Chapter 3

Minimal English: The Science Behind It

Cliff Goddard

As described in Chap. 2, the Minimal English project has emerged from, and in a sense rests upon, the findings of a program of linguistic research known as NSM (Natural Semantic Metalanguage). NSM is a wide-ranging approach to semantics, that is, the systematic study of meaning in languages. Unlike many forms of contemporary linguistics, especially the generative linguistics or biolinguistics of Noam Chomsky, the NSM approach places words and meaning at the very centre of language study (cf. Goddard and Wierzbicka 2014). Since its early days more than four decades ago, the NSM approach has been much interested in cross-linguistic semantics, that is, comparing and contrasting how different languages construe and package meaning differently in their vocabulary and grammar. The goal of this chapter is to summarize and review those aspects of NSM research which are most relevant to Minimal English. Some readers may prefer simplyto browse the chapter and return to it later, as necessary.

3.1. Sketch History and Main Ideas Behind NSM Approach

3.1.1 Paraphrase using simple, easily understood words

The NSM approach starts with two very simple ideas. The first idea is that if we want to explain words and their meanings, we have no choice but to do so using words themselves. As Hanks (2007: 6) put it: 'There is no metalanguage other than words themselves (in one language or another) – or at a pinch derivatives of word such as logical symbols – for expressing thoughts about words'. The second idea is that, this being the case, we need to proceed very carefully to avoid the danger of getting 'tangled up in words'. In particular, we need to ensure that our explanations (or, explications) of word meanings are as clear as possible and that they are not undermined by definitional circularity (i.e. going around in 'vicious circles'). The clarity principle was already clear to Aristotle (1928), who enunciated it in the section 'On definitions' in his *Topica* (6.4.141ª25). Aristotle stressed that to be effective any definition must be phrased in terms that are 'prior and more intelligible', that is, simpler and easier to understand. It would seem to follow that the ideal metalanguage for the purpose of semantics would consist of words which are simple and easy to understand.

People familiar with linguistics may feel a little uneasy at this point. Is it not true that many linguists, and many dictionaries for that matter, are prone to using technical terms or sophisticated expressions to describe meanings? For example, they may say that part of the meaning of the verb *to promise* is 'to undertake an obligation' or that part of the meaning of the verb *to cut (something)* is 'to bring about a separation in material integrity'. The unfortunate truth is that many linguists are either oblivious to metalanguage issues[1] or believe that they are free to employ technical formulas, often called 'formalisms', without anchoring them in ordinary natural language. As John Lyons once observed, however: '[A]ny formalism is parasitic upon the ordinary everyday use of language, in that it must be understood intuitively on the basis of ordinary language' (Lyons 1977: 12).

At this point, we can appreciate the designation 'Natural Semantic Metalanguage approach'. It is an approach to describing meanings based on using simple words of natural language. Describing the meaning of a word by using other, simpler words sounds a lot like paraphrase, of course; and in a very real sense paraphrase is indeed the main technique behind NSM semantics. But to be clear, the key criterion is not simply paraphrase, but more precisely, paraphrase without circularity. In this respect, NSM semantics makes a radical break with traditional lexicography,[2] which also relies (imperfectly) on paraphrase in natural language. To make this more concrete, consider how the Collins Cobuild English Language Dictionary (1991) defines the word part via piece, piece via part and bit, and bit via piece and part, as follows:

> *part* – a part of something is one of the pieces that make up an object.
> *piece* – a piece is a bit or part of something that has been broken off, or cut off.
> *bit* – in informal English, a bit of something is also a small piece of it
> *bit* – in fairly informal English, you can refer to a particular part of something […]
>   as a particular *bit* of it.

The set of definitions is confusing and self-referring. At first sight it may seem that the *Cobuild* dictionary succeeds in paraphrasing all three words—*part*, *piece*, and *bit*—but this result is possible only because of the circularity. As a matter of fact, according to NSM research it is impossible to paraphrase the word 'part' (in contexts like 'a person's head is part of this person's body' and 'this thing has two parts') without circularity, which means that 'part' has to be recognized as a semantic prime, that is, as an irreducible or 'rock bottom' meaning. On the other hand, using 'part' and other elements, it is possible to produce satisfying paraphrases of the words *piece* and *bit* (cf. Goddard 2010).

This brings us to the fundamental notion behind NSM semantics, namely, the concept of semantic primes.

3.1.2 Semantic primes: The atoms (or elements) of meaning and thought

Assuming it is possible to fully paraphrase word meanings in terms of simpler meanings, it would seem to be a matter of logical necessity that 'at the end of the line', so to speak, there are some ultimately simple meanings which cannot be paraphrased at all. The logic leading to this conclusion was well known to seventeenth-century philosophers, including Descartes, Pascal, Arnauld, Locke and Leibniz. A classic quotation comes from Arnauld and Nicole's *Logic, or the Art of Thinking*.

> I say it would impossible to define every word. For in order to define a word it is necessary to use other words designating the idea we want to connect to the word being defined. And if we again wished to define the words used to explain that word, we would need still others, and so on to infinity. Consequently, we necessarily have to stop at primitive terms which are undefined. (Arnauld and Nicole 1662)

The seventeenth-century philosophers also saw clearly that semantic primes (as we term them these days) are not just 'in' language, but also 'in' our minds, that is, they are concepts. They called them 'simple ideas' and spoke of them as the fundamentals of 'human understanding'. A clear formulation comes from Gottfried Wilhelm Leibniz (1646-1716): 'If nothing could be understood in itself, nothing at all could ever be understood' (Leibniz 1903: 430).

Among his seventeenth-century contemporaries, Leibniz stands out as the true precursor of modern systematic semantics (Wierzbicka 2001) on account of his conviction that simple ideas underlying human thought could be discovered by systematic lexical analysis. For decades he experimented with definitions of numerous words,[3] trying to establish 'the catalogue of those concepts which can be understood by themselves, and from whose combinations our other ideas arise' (Leibniz 1903: 430). He anticipated that they might be relatively few in number ('because nature usually achieves as much as possible with as few elements as possible') and universal in their reach. As George Steiner ([1992]1975: 75) put it in his classic work After Babel: 'Leibniz was profoundly interested in the possibilities of a universal semantic system, immediately accessible to all men, … grounded in the very architecture of human reason and […] independent of all local variation'.

For various reasons, Leibniz was unable to complete this task in his lifetime. One reason no doubt was the huge number of other projects – practical, philosophical, mathematical, diplomatic – which occupied this remarkable polymath. He also lacked collaborators. His nineteenth-century editor Gerhard wrote: 'There are a great many manuscripts among Leibniz's papers which deal with

elementary concepts and definitions.... Convinced that he would never be able to complete the great task alone, how often he yearned for help!' But in another sense, the time was not right for it. Linguistics as a discipline did not yet exist and relatively little information was available about the languages of the world. After this death, no-one progressed Leibniz's interest in empirical universal semantics and the program was lost.

3.1.3 Sketch history of the NSM approach

Fast-forwarding to the 20th century,[4] the Leibnizian project re-surfaced in the 1960s with the Polish linguist Andrzej Bogusławski (1966, 1970, and other works). Like Leibniz, he advocated seeking out the elementary units of thought through empirical investigations into natural language. The project was taken up in earnest by a young Polish linguist: Anna Wierzbicka. Wierzbicka's first landmark publication in English was Semantic Primitives (1972). The following year she moved to Australia and took up a post in the Australian National University, where she has been based ever since. Renowned for its typological- descriptive studies and for field work, the ANU Department of Linguistics provided an especially conducive environment for the subsequent development of the NSM theory. In the following three and half decades, Wierzbicka was joined by an increasing number of like- minded colleagues, the theory underwent successive waves of development and empirical testing, and experienced a gradual upward trajectory in terms of reception and influence. The historical developments can be arranged into five periods, as described below. (Needless to say, no short overview like this can be complete or entirely accurate, and it is also the case that NSM discovery process has depended in part on accidents of history, personalities, and a degree of luck.)

1972 to mid-1980s This was the early development period, from the publication of *Semantic Primitives* (1972), through to *Lingua Mentalis* (1980), and into the mid-1980s. The term Natural Semantic Metalanguage was not yet in use, so the metalanguage version is sometimes referred to as Pre-NSM. Over this period only 13-14 semantic primitives were proposed. Wierzbicka was virtually the sole author and was widely seen as brilliant but iconoclastic. Valuable analytical work was done on emotion terms, body parts, stage of life terms, speech acts, cases, and 'concrete' vocabulary (Wierzbicka 1985). Due to the dominance of generative linguistics, these were very tough times for a meaning-based approach to language.

Mid-1980s to mid-1990s This roughly 10-year period can be called the 'expanding set' phase, following Wierzbicka (1989). The prime inventory grew to 37 in *Semantic and Lexical Universals* (1994). A couple of the first-generation semantic primitives ('world' and 'imagine') were withdrawn, as evidence had shown them to be not cross-linguistically viable. A further 20-odd

primes were announced in *Semantics: Primes and Universals* (1996), bringing the total to 55. This version of the metalanguage is sometimes termed NSM1. The expansion of the prime inventory was driven from two directions: on the one hand, there was the need to improve the power and transparency of the metalanguage in the areas of time, space, causation, quantification, and logical relations; on the other hand, as more empirical- descriptive work was done, it often became apparent that adjustments were needed to the prime inventory or to the grammar. For detailed accounts, see Wierzbicka (1989, 1996: Ch2), also Goddard (1989, 2008).

This period saw increasing emphasis on cross-translatability and universals. The first major cross-linguistic survey of semantic primes appeared, the edited collection *Semantic and Lexical Universals* (Goddard and Wierzbicka, eds. 1994). Described as 'monumental' by one reviewer, it included data on 17 languages in addition to English, only 1 of which (French) was European. Another major work was Wierzbicka's *Semantics: Culture and* Cognition (1992). The first textbook to offer a thorough grounding in NSM appeared, Goddard's (1996) *Semantic Analysis: A Practical Introduction* (revised and re-issued in 2011a).

Late-1990s to 2007 The prime inventory expanded modestly to 62 primes over these years. This version of the metalanguage is sometimes termed NSM2. The key theoretical and cross-linguistic volume, *Meaning and Universal Grammar* (Goddard and Wierzbicka, eds., two vols, 2002) is situated in the middle of the period. As indicated by the title, there was increasing attention to universals of syntax and to 'whole metalanguage' studies (cf. Goddard 2002b).

An important publication towards the end of the period was Wierzbicka's (2006) *English: Meaning and Culture*, which argued comprehensively that the lexicon and grammar of English continues to bear the imprint of its own highly specific cultural history. A collective volume on semantic primes in the Romance languages (Peeters ed. 2006) broke new ground in its detailed treatment of semantic primes and their grammar in a well-known language family. Many other significant publications appeared in this period, particularly on the cross- linguistic semantics of emotions. Wierzbicka's (1999) *Emotions Across Languages and Cultures* led the way, followed by two collective volumes (Harkins and Wierzbicka eds. 2001; Enfield and Wierzbicka eds. 2002) that together presented work on 15 languages, by a similar number of authors. These years also saw the development and consolidation of the theory of cultural scripts (also termed ethnopragmatics), the pragmatic 'sister theory' of NSM semantics (cf. Wierzbicka 2003; Goddard and Wierzbicka eds. 2004; Goddard ed. 2004; Goddard ed. 2006).

2008-2014 This period is book-ended by the edited collection *Cross-Linguistic Semantics* (Goddard ed., 2008) and by the co-authored *Words and Meanings* (Goddard and Wierzbicka 2014). There was

a slight expansion of the prime inventory to 65 primes. Two new primes (BE (SOMEONE/SOMETHING) and LITTLE~FEW) were added, and an 'old' prime DON'T WANT was restored to the inventory. This version of the metalanguage is sometimes termed NSM3. The main trend over this recent period was intensified attention to the theory of semantic molecules, as described further in section 3.4. There were some simplifications to metalanguage syntax, for example, dropping the 'that-complement' construction for THINK. In lexical semantic practice, the theory of semantic templates was developed further, especially in relation to verbs.

In terms of descriptive work, two notable historical semantic studies appeared: Bromhead's (2009), *The Reign of Truth and Faith*, about aspects of seventeeth-century English, and Wierzbicka's (2010) massive *Experience, Evidence, Sense: The Hidden Cultural Legacy of English.* Monographs appeared on Russian and Danish (Gladkova 2010, Levisen 2012, respectively). There were several collective projects, such as an edited collection on 'memory' concepts (Amberber ed. 2007) with half a dozen NSM chapters and a special issue on 'social cognition' (Goddard ed. 2013).

2014 onwards Recent years have seen only small adjustments to the NSM metalanguage, such as the change from HAVE to MINE as the semantic prime underlying ownership and possession (Goddard and Wierzbicka 2016a). Some fine-tuning and re-arrangement of syntactic frames has been implemented, principally in relation to KNOW and CAN. This version of the metalanguage has been designated NSM4. The theory of lexicosyntactic molecules began to be developed. There were two key monographs in 2014-15: Goddard and Wierzbicka's *Words and Meanings* and Wierzbicka's *Imprisoned in English*, as well as several collective volumes (Peeters ed. 2015; Goddard and Ye eds., 2015; Ye ed. 2017). Major works appeared on the semantics of Chinese musical concepts (Tien 2015) and the culture of Singapore English (Wong 2014). 2014 can also be identified as the beginning of the development process of Minimal English (see section 3.5).

It will be apparent from this review that there is a vibrant international NSM research community. It is still strongest in Australia, but there are advanced practitioners in Poland, Denmark, Israel, Ireland, Singapore, and Finland, among other countries, and of course they are all interconnected via internet.[5] Languages that have been intensively studied using NSM methods include: Arabic, Bislama (Vanuatu), Cantonese, East Cree (Canada), Danish, English, Ewe (Ghana), Finnish, French, Hebrew, Italian, Japanese, Korean, Koromu (Papua New Guinea), Mandarin, Polish, Russian, Spanish, and Vietnamese. In terms of geographical distribution and language families, there are many gaps, but there is also a great deal of diversity among these languages—six language families including languages from every continent, and many different

cultural types.

3.2.   A Closer Look at Semantic Primes: Simple Universal Words

As we know, a semantic prime is a word-meaning or concept which cannot be paraphrased (decomposed) in any simpler terms. Semantic primes are expected to be expressible by words or phrases in all languages, that is, to be lexical universals. These twin criteria mean that only a small number of words are possible semantic primes, because the vast majority of words can readily be shown to be semantically complex and/or language-specific (cf. Goddard 2001). There is also a third consideration: taken as a whole, the metalanguage of semantic primes is intended to enable reductive paraphrase of the entire vocabulary and grammar of the language at large, that is, it is intended to be comprehensive.

3.2.1   How were primes identified in the first place?

It is impossible to show that a concept cannot be further decomposed except by trying to decompose it and failing; and it is impossible to show that a concept *can* be decomposed except by trying to decompose it – and succeeding. (Wierzbicka 1980: 10)

We will briefly consider two semantic primes: GOOD and SAY. How could one possibly paraphrase, using simpler words, the meaning of *good* in contexts like (a) *This is good,* (b) *Something good happened*, (c) *People say many good things good about him*. It just seems impossible. Terms such as 'positive evaluation', 'valued', or 'pleasing' are obviously more complex than 'good' (and also highly language-specific).

The only plausible route would be to try to decompose 'good' in terms of actual or potential 'desirability'; for example, by saying that 'this is good' means 'I want this' or 'people want this', but this doesn't work out for several reasons. Most importantly, to say that something is 'good' presents the evaluation in an objective mode, not as the desire of any specific person, or even of people in general. The difficulty of finding a satisfactory paraphrase makes GOOD a candidate for a semantic prime. Furthermore, GOOD is clearly required for explicating innumerable lexical items such as *nice, tasty, kind, happy, and pretty,* and for grammatical constructions such as benefactives. Upon checking in a range of languages, one finds that all languages appear to have a word which can express the meaning GOOD. (Obviously, this does not mean that different cultures share the same views about what kind of things are good.)

Now consider the word *say*, in everyday contexts like these: (a) *She said something to me*, (b) *What did she say?,* (c) *She said "I don't want to do it"*. How could one paraphrase *say* in these contexts, using simpler words? It would be no good to use 'verbally express', since 'express' and

'verbally' are intuitively more complex and more language-specific than SAY. The only plausible route appears to be via 'do', 'want', and 'know'; for example, 'X said something to Y' = 'X did something, because X wanted Y to know something'. But it fails because the right-hand side of the equation could be satisfied by many actions which are non-verbal and not symbolic.

There are numerous lexical items whose meanings are based on SAY– most notably, speech-act verbs, that is, words like *promise*, *praise*, *apologize*. In some languages, there are grammatical meanings which involve SAY; for example, certain evidential particles. Do all languages have lexical equivalents to SAY? The answer appears to be yes, although in some languages the word for SAY can also express another meaning, such as DO or WANT. In such languages, there is always (as far as we know) language-internal evidence for polysemy (cf. section 3.3.1).

3.2.2 *How it is that primes can be related to one another?*

In the early days, it was assumed that primes should be 'conceptually independent' from another, but it has since been realized that this is not so. In fact, semantic primes can be intuitively related to one another in a variety of ways, without this implying that any of them can be paraphrased in terms of the others. Consider pairs of 'opposites' like GOOD and BAD, for example. It is impossible to define either word in terms of the other (certainly, *good ≠ not bad*, *bad ≠ not good),* yet obviously the meanings GOOD and BAD are closely related –even to the point that there is an implicational relationship between the two; that is, 'good' implies 'not bad', 'bad' implies 'not good'.

Similarly, the inventory of primes includes several pairs of converses, such as BEFORE and AFTER, and ABOVE and BELOW. From a strictly logical point of view, 'A happened after B' is equivalent, to 'B happened before A'; and mutatis mutandis for ABOVE and BELOW.[6] However, as pointed out by Wierzbicka (1996: 110) 'a natural language has its own ("natural") logic and in this "natural logic" BEFORE and AFTER, and ABOVE and BELOW are not mutually equivalent'. She continued:

> In both pairs each element is linked with one particular point of view, and in human communication a difference in point of view may be as important as a difference between two predicates. For example, while past and future tenses both situate events in time with reference to the present moment, past tenses do so in terms of the concept BEFORE, whereas future tenses do so in terms of the concept AFTER.

Facts like these show that not all semantic relationships are of a compositional nature, that is, there can be affiliations between meanings which are not reducible to any specifiable common component. As an example of a different kind, consider the primes IF, CAN and

MAYBE. Certainly one feels that they all have something in common (compared with, say, TWO and BODY), but it seems to be impossible to paraphrase any of them. In NSM theory, such affiliations are called 'non-compositional relationships' (Goddard 2002a: 26-31).

3.2.3   A review of the semantic primes, in 12 groups

We will now briefly run through the 65 semantic primes, using the 12 groupings used in Chap. 2. A slightly more elaborate set of groupings, accompanied by linguistic labels, is often used in technical linguistic publications. For reference, it is given in Appendix A. There are also other ways of grouping primes, according to their shared grammatical properties; for example, we could regard GOOD, BAD, BIG, SMALL, NEAR, FAR, A LONGTIME, and A SHORT TIME as falling into a single group – on the basis that all of them can be modified by VERY. Likewise, I YOU, HERE and NOW can be seen as forming a grouping(linguists term them deictics) because they are all tied to the speech situation and because they are all resistant to modification.

- **I~ME, YOU, SOMEONE, SOMETHING~THING, PEOPLE, BODY, KIND, PART**

This group represents what traditional grammar calls 'substantive' words. The presence of I and YOU as semantic primes implies that first-person and second-person orientations are integral to human thinking and cannot be reduced to third-person formulations. Note that there is a clear distinction drawn between SOMEONE and SOMETHING. This might like seem simple common sense, but in linguistics and logic, scholars often use abstract or technical terms such as 'participants', 'arguments', or 'entities', which gloss over this distinction. PEOPLE also appears as a semantic prime. As far as I know, no other system of semantic analysis has any analogue of this preeminently 'social' semantic unit. The presence of BODY as a semantic prime is another distinctive feature of NSM semantics. It provides a unique hinge for 'embodiment' semantics. PARTS and KINDS are both 'relational' in the sense that are always connected to something else; for example, 'many parts of something', 'many kinds of something', and so on. Few would deny that that parts (partonomy) and kinds (taxonomy) are essential ingredients in human thinking.

   Why, it may be asked, are there so few nouns in this division of the inventory of semantic primes? Where are words for various universals of human experience, such 'sun', 'water', 'hands', and 'mouth? The answer is that these important notions are not semantically simple and are therefore not semantic primes; rather, they are universal semantic molecules (see section 4).

- **THIS, THE SAME, OTHER~ELSE**
- **ONE, TWO, MUCH~MANY, LITTLE~FEW, SOME, ALL**

The words in these two groups combine can directly with the substantives (except for I and YOU) to form expressions like 'this someone' (= 'this person', 'he/she'), 'something else', 'the same kind', 'many people', 'all these things', and so on. As well, the semantic prime THIS can stand alone in various contexts, for example, 'this is bad', 'it is like this: ...', 'I don't want this'.

- **GOOD, BAD, BIG, SMALL**

These are the only indefinable, that is, semantically simple, descriptors. GOOD and BAD, the semantic lynchpins of 'evaluation', are very versatile, appearing in contexts such as, for example, 'something good happened', 'this someone did something very bad', 'I want to do something good for you', 'if you do this, people can think something bad about you'.

Other important physical properties, for example, 'hard', 'long', 'round' are semantic molecules.

- **THINK, KNOW, WANT, DON'T WANT, FEEL, SEE, HEAR**

There is a fairly rich inventory of mental primes, some of which (especially, THINK) are largely reserved for use with human subjects. These have important implications for naïve 'theories of mind'. One matter that calls for comment concerns WANT and DON'T WANT. In English, as in many (not all) languages, their exponents are morphologically related to one another, and DON'T WANT is also morphologically related to NOT (negation). However, the prime DON'T WANT (as in 'I don't want this', usually voiced by young children simply as *No!)* represents 'rejection' (metaphorically, pushing something away), not the simple negation of wanting. The relationship between WANT and DON'T WANT is non-compositional, akin to the relationship between opposites or converses.

- **SAY, WORDS, TRUE**

The semantic prime SAY is, obviously, crucial to human communication. To SAY something in its basic sense, however, is not necessarily tied to WORDS. One can SAY ('express') something using gestures, and one can easily imagine someone saying something with a look or tone of voice. Of course, the primes SAY and WORDS are closely related (normally, people say things with words; see section 3.3 below), but there appears to be a basic distinction between the expressions 'say something', which is focussed on the content of what of what is said, and 'say some words', which is focussed on the form, so speak. For discussion, see Wierzbicka (1996: 107-8); Goddard (2011b). As for TRUE, it is the most specialized of all the semantic primes, in the sense that it can apply only to something someone says (or, perhaps, thinks): 'it is true' or 'it is not true'.

• **HAPPEN, DO, MOVE**

These are the 'dynamic' primes, underlying notions of events, actions and movement. They are all time-bound concepts: something happens at some time, someone does something at some time, something moves in a place at some time.

• **BE (SOMEWHERE), BE (SOMEONE/SOMETHING), THERE IS**

It is widely accepted in linguistics that the 'be of location', for example, in a sentence like 'She is in Sydney', is a different element to the 'be of identity', for example, in a sentence like 'She is my mother'. Many languages have two different words. NSM linguists propose that the 'be of location' also allows us to speak (and think) about 'being <u>with</u> someone'. As for THERE IS (roughly, existence), again, many languages have a separate word for this, for example, Spanish *hay*. Despite the formal overlaps in English and in many other languages, it seems clear that none of these concepts can be satisfactorily explicated in terms of the others.

• **(IS) MINE**

This prime underlies concepts of ownership and 'possession' (alienable possession, in linguistic parlance), which can all be sourced back to the key expression 'it is mine' (Goddard and Wierzbicka 2016a).

• **LIVE, DIE**

Presumably few would question the plausibility of LIVE as a semantic prime. It is less obvious why we need DIE as well, but the basic idea is that DIE represents a deeper and more existential event than simply coming to the end of one's life. Separating LIVE and DIE is also necessary to provide conceptual ingredients for various notions of life after death.

• **TIME~WHEN, NOW, BEFORE, AFTER, A LONG TIME, A SHORT TIME, FOR SOME TIME, MOMENT**
• **PLACE~WHERE, HERE, ABOVE, BELOW, FAR, NEAR, SIDE, INSIDE, TOUCH**

These two groups of primes are connected with time and place, respectively. They have a similar number of members, and there are obvious analogies or parallels between them. Both groups have a substantive-like member. One can speak, for example, of something happening 'at this time' and 'in this place' (or, 'at the same time', 'in the same place', etc.) Both groups have a deictic element: NOW and HERE, respectively. Both include converses (e.g. BEFORE and AFTER, ABOVE and BELOW) and 'dimensional' expressions (A LONG TIME and A SHORT TIME, FAR and NEAR). The time group has FOR SOME TIME and MOMENT, which are necessary for capturing ideas of

duration and 'momentariness'. The place group includes INSIDE (containment) and TOUCH (contact), which are both about physical objects, as much as place per se.

   • **NOT, CAN, MAYBE, IF, BECAUSE, VERY, MORE, LIKE~AS**

From a grammatical point of view, this is a rather disparate collection of items, albeit they all express notions which can be seen as broadly 'logical' in some aspect. Few would dispute that negation (NOT) is a fundamental of human language and thought. The NSM system recognizes no less than three primes (CAN, MAYBE, IF) connected with possibility and hypothetical thinking. BECAUSE is the natural language anchor for notions about causality. Broadly speaking, VERY and MORE are about intensification and 'increase', though they are quite different in their syntax and in how they are utilized in explications. Finally, LIKE (as in 'it happened like this', and 'someone like me') is one of the most versatile and multi- functional of all primes.

3.2.3 Four complications (lexical polysemy, formal complexity, allolexy, portmanteau expressions)

It is useful at this point to introduce the term 'exponent' to designate an expression in a given language that expresses a semantic prime. We can now say that the relationship between a semantic prime and its exponent(s) are not necessarily, or even typically, one-to-one. The exponent of a prime can have more than one meaning, that is, it can have other meanings in addition to the semantically primitive one, and, vice versa, a prime can have more than one exponent. Misunderstanding on these points is rampant among many under- informed critics of NSM.

Lexical polysemy Exponents of primes can be polysemous, with different secondary meanings in different languages. For example, the English word *feel* and the Malay word *rasa* have the same primary sense, namely, the prime FEEL, but the English word *feel* has a secondary meaning related to 'touching' which is not shared by the Malay word, while Malay *rasa* has a secondary meaning 'taste' which is not shared by English *feel*. Another example, which will be closer to home for many readers, concerns the prime DO. In many languages, including European ones, exponents of DO can also mean something like 'make'. So, when it is said that a given language has exponents for the semantic prime FEEL or DO, for example, this is not the same as saying that these words are 'the same' in all respects as their English equivalents.

    After more than 20-years study of the exponents of semantic primes across diverse languages, NSM researchers have built up a stock of knowledge about recurrent patterns of polysemy. Certain polysemies are known to be common across languages. A selection is given in Table 3.1.[7]

Table 3.1. Selected common polysemies of exponents of semantic primes (after Goddard 2012, data from studies in Goddard and Wierzbicka 1994, 2002; Peeters 2006; Goddard 2008; and Gladkova 2010).

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 (*palya<u>ni</u>*)
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 (*kòòn¹*), 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*)

Formal complexity Exponents of primes in different languages can be formally complex in various ways. For example, in English the words SOMEONE and INSIDE are morphologically complex, but their meanings are not composed from the meanings of the morphological 'bits' in question. That is, the meaning SOMEONE does not equal 'some + one'; the meaning INSIDE does not equal 'in + side'. In meaning terms, SOMEONE and INSIDE are indivisible.

Exponents of primes can also be phrasemes, that is, fixed multi-word expressions, such as the English expressions *a long time* and *a short time*. It is not hard to discern that in these expressions the words 'long' and 'short' do not have their normal meanings. Despite the formal complexity, NSM researchers argue that these composite forms each express a single, unitary meaning. Exponents of primes can also be bound morphemes. For example, in Yankunytjatjara, BECAUSE is expressed by the suffix *–nguru*.

Allolexy In many languages, including English, there are multiple exponents of a single prime which occur in different contexts. In linguistic terminology, the prefix *allo-* indicates the existence of alternative versions of items belonging to various categories; for example, allomorphs are alternative forms of a given morpheme, allophones are alternative forms of a given phoneme. Following this terminology, alternative word-forms for a given semantic prime are called 'allolexes'. Patterns of allolexy are different in different languages, sometimes for no apparent

reason. For example, in English the semantic prime OTHER has a postnominal allolex *else*, which is normally used with indefinite expressions; compare *another place* vs. *somewhere else, another thing* vs. *something else*, and so on. Spanish and German, on the other hand, get by perfectly well without any word comparable to *else*. They use the same word for OTHER (Spanish *otro*, German *andere(r)*) with indefinites and with ordinary nouns. Other instances of allolexy are connected with the grammatical character of a language. For example, German is a language with grammatical gender, so its word for OTHER has different forms (*andere*, *anderer*) depending on the gender of the noun it is modifying. Many languages have systems of 'case', meaning that words appear in different forms depending on their grammatical role in a sentence. English has a remnant of this in its alternation between the forms *I* (used before a verb) and *me* (elsewhere). As a final example, English NOT has the verbal allolexes *don't* and *doesn't* (with the 'dummy verb' *do*), which are used to negate most verbs; for example, 'I don't know' = 'I not know'. The point is that all these formal variants are purely cosmetic so far as meaning is concerned.[8]

How can one be sure that a particular formal variation is an instance of allolexy? As usual in NSM semantics, the test is paraphrase. Two forms can be allolexes only if there is no paraphrasable difference between them. Another useful benchmark is what happens when NSM sentences in one language are transposed into another language. Naturally, all the grammatical trappings in the original language are lost. For example, if NSM based on Russian, which is a highly inflectional language, is transposed into Chinese, a great amount of formal, allolexical variation (gender, case, person agreement on verbs) is lost.

This raises the question of the different tense forms of verbs, for example, the difference between English 'This often happens' and 'This often happened'. In everyday English, this difference (*happens* vs. *happened*) is not allolexy because it is meaning-bearing: the form *happened* (past tense) conveys that the event took place 'at some time before (now)'.[9] This means that when using NSM based on English (or any other language which has verbal tense), one has to be careful that particular tense forms do not 'smuggle in' implicit semantic content. The solution is to ensure that any past tense use of a verb, for example, is accompanied by a phrase such as 'at some time before' which puts the semantic content of the tense ending into words. This makes the meaning explicit and cross- translatable even into languages which lack verbal tense.

Portmanteau expressions A single word (or affix, or phraseme) can express a combination of semantic primes. Portmanteaus with NOT are fairly common. A simple example is English CAN'T, combining CAN and NOT. Many languages have several portmanteaus that involve NOT. Portmanteaus with LIKE are also common. For example, German has a portmanteau *so* for 'like this'; likewise Yankunytjatjara has *alatji*. In some languages certain combinations of semantic

primes are expressible only via a portmanteau; for example, Cantonese has aportmanteau *móuh* of NOT and THERE IS; Polish has the portmanteau expression *dawno (temu)* for A LONG TIME BEFORE. This is not a problem, provided that native speakers can clearly recognize what is going on.

In English NSM, and in Minimal English, we commonly use the portmanteau expressions 'often' (= at many times) and 'sometimes' (= at some times).

3.3.   Grammar Matters: Universal Grammar Versus Language-Specific Grammar

The term 'grammar' can refer to different aspects of language structure: both to combinations of words and the order in which they occur (linguists call this syntax) andto the changing forms that words take in different grammatical contexts (linguists call this morphology, which also includes how words can spawn other, related words by changes in form). In the end, the simplest idea is that grammar is about how words can be combined — about what can go together with what.

Obviously, the metalanguage of semantic primes has its own syntax, and it is (or, ought to be) universal in the sense that the same principles of combination should apply in any language. For example, it should be possible, in any language, to combine the primes I, NOT and KNOW to order to say the equivalent of 'I don't know', and to combine PEOPLE, ALL, and DIE to say the equivalent of 'all people die'. For a more complex example, it should be possible in all languages to combine IF, DO, SAY, CAN, BAD and other primes, to express a complex meaning like 'if you do this, people can say something bad about you'. Notice that the order of the words and phrases doesn't matter. Different languages put words in different orders. All that matters is whether the same range of combinations is possible in all languages. As described earlier, NSM researchers have been investigating the grammar of semantic primes for many years, and a great deal is now known about it.

It turns out that the best way forward is not to try to state very general 'rules' that would apply across the whole metalanguage, but rather to look at each semantic prime individually, asking: How does this prime 'work' grammatically? How can it combine with other primesto form phrases and sentences? From this point of view, it can be said that every prime has itsown mini-grammar. Some primes, such as DIE and TRUE, have a very limited range of grammatical options. Others have a basic or minimal frame, but also allow several extended grammatical options. For example, using the semantic prime DO we can say not only that 'someone does something (at some time)', but also that 'someone does something <u>to</u> someoneelse'. Using the semantic prime SAY, we can say not only that 'someone said something (at some time)' but that 'someone said something <u>about</u> someone else'. We expect that in all languages the words for DO and SAY (or their

allolexes) can be used to say such things, even though the formal expression will be different. English makes use of propositions like 'to' and 'about', but other languages may use other kinds of grammatical words, or case endings, or more exotic devices such as serial verb constructions, to get the same effect. It doesn't matter so long as the meanings expressed are the same.

Extended grammatical options are often referred to in the technical literature as grammatical frames or as valency options. Table 3.2 shows some of the key grammatical frames for DO and SAY, which actually have the most elaborate array of possibilities of all the semantic primes. It shows five frames for DO and six frames for SAY. The claim is that all these frames have their equivalents in other languages, be it Chinese, Finnish, Russian, Vietnamese, or whatever.

Table 3.2 Selected grammatical frames for DO and SAY

someone DOES something
someone DOES something to someone else
someone DOES something to something
someone DOES something with something
someone DOES something good for someone else

someone SAYS something
someone SAYS something to someone else
someone SAYS something about something
someone SAYS something like this: " "
someone SAYS something with some words
someone SAYS a word to someone else

Over the years, there has been a lot of fine-tuning of the posited grammar of semantic primes, in response to accumulating evidence about languages and/or as a result of deeper conceptual analysis. In recent years, NSM researchers have been using a visual display in the form of a Chart of NSM Semantic Primes, downloadable from the NSM Homepage and updated from time to time. Each prime has a cell of its own, in which certain key grammatical combinations or frames are listed. Sometimes the exponent of a prime allows grammatical extensions which look like additional valency options, but which are actually complex language-specific constructions. For example, in English the verb *do* can be used to talk about 'doing something <u>about</u> something', but this construction expresses a complex meaning (involving thinking about a situation, wanting it to be different, and doing something because of it), which, furthermore, lacks equivalents in many languages. Likewise, in English one can speak about 'feeling <u>that</u> such-and-such', e.g. *I feel (that)*

we should go now, but this construction is a kind of amalgam of FEEL and THINK, lacking equivalents in many other languages. Even the 'think (that) ...' construction which sounds so basic in English, lacks equivalents in many languages; hence, in NSM instead of saying *I thought that she was at home*, for example, one says instead 'I thought like this: she is at home'.

These examples underline the point that NSM is not just a vocabulary list, but truly a mini-language. To use NSM properly, one has to learn and comply with the allowable grammar, as well confine oneself to the allowable words. The same applies to Minimal English, though perhaps not so strictly (some compromise on strict NSM grammar may be acceptable, provided it does not interfere with translatability.) See Chap. 2. There are many other aspects of NSM grammar that cannot be covered here for reasons of space.

3.4. Semantic Molecules, Other Lexical Universals, and Near-Universals

3.4.1 *What are semantic molecules?*

Semantic molecules are non-primitive, but still relatively simple, meanings which function alongside semantic primes as building blocks of meaning. In NSM explications, they are marked with the notation [m] (Goddard and Wierzbicka 2014, Goddard 2016). Semantic molecules can be explicated into primes, so there is no danger of circularity and no compromise of the reductive principle. The notion is similar to that of 'intermediate expressions' in the Moscow Semantic School (Zholkovsky 1964; 2000; Mel'čuk 1989, 2012). The idea of semantic molecules was implicit in Wierzbicka's mid-1980s works on concrete vocabulary (Wierzbicka 1985), which employed defining vocabularies of several hundred words. It was already apparent that plausible explications of many words from these domains could not be achieved solely in terms of semantic primes, unlike with words from abstract domains such as emotions, values, speech-acts, and discourse particles.

Explications for animal words like *cat*, *dog*, *horse*, and *mouse*, for example, require the use of body-part terms such as 'head', 'legs', 'mouth', 'teeth', 'tail', 'eyes', and 'ears', and also descriptor words such as 'long', 'round', and 'sharp'. Verbs like 'kill' and 'eat' are also needed to describe characteristic habits of these creatures. Explications for folk-biological terms need words such as 'fish', 'bird' and 'creature' as semantic molecules (for example, it would be a very implausible explication for *salmon* and *trout* that did not include the information that they are 'fish'). Explications for artefact words (such as *cup* and *knife*) require the nouns 'hands', 'fingers' and 'mouth', among others, as well as verbs like 'hold', 'eat', and 'drink'. Many body-part words require descriptors like 'long' and 'round', like *legs* and *eyes*, respectively, and some require physical quality words such as 'hard', e.g. *teeth*, *bones*. The reader will already be noticing that

many of the same putative molecules are needed across different areas of the lexicon.

The principle was clear enough: certain complex terms were needed as 'concept building' elements (Wierzbicka 1991). The expression 'semantic molecule' was being used in 1996 (Goddard 1996). It was also clear that considerable empirical work would be needed, across different lexical fields, to figure out the optimal inventory of semantic molecules, to see how they were arranged into hierarchies or levels of complexity, and to explore the relative roles of universal or near-universal vs. language-specific semantic molecules. Broadly speaking, systematic theoretical work on molecules got going in earnest only in the twenty-first century. The remainder of this section summarizes our current understandings in this field.

One important result is that there are 'molecules within molecules' or, to put it another way, there are chains of semantic dependency between words of different levels of complexity. This principle is depicted schematically in Fig. 3.1 for three 'top-level' domains, which we can loosely call artefacts for eating and drinking, names for celestial bodies, and kin terms, respectively. Each level depends, semantically, on the level beneath it — with the whole thing resting ultimately upon the bedrock of semantic primes. For example, looking at the left-most diagram, 'mouth' and 'hands' can be explicated solely in terms of semantic primes. Using these words in combination with semantic primes, it is possible to explicate the concept of 'water' (Goddard 2016). The concept 'drink' in turn, depends partly on 'water' (along with many other elements, of course) and moving one level higher, the concept of 'cup' depends partly on 'drink'.
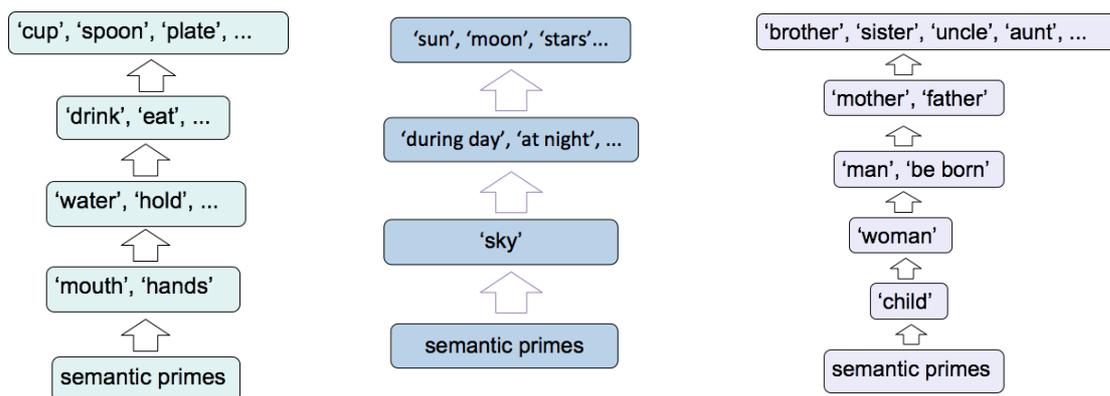
Fig. 3.1 Diagrams illustrating how complex concepts are successively built up from simpler ones, in three domains.

In similar fashion, the middle diagram in Fig. 3.1 depicts the claim that 'sky' is a foundational molecule for many environmental concepts, including 'during the day' and 'at night'. Obviously, both these concepts also depend heavily on expressions fashioned purely from semantic primes such as 'at a time when …', and 'people can/can't see things well from some time'. 'During the day' and 'at night' in turn are conceptual ingredients in concepts such as 'sun', 'moon' and 'stars' (Goddard 2016; cf. Chap. 8, this volume). Likewise, in the domain of kin terms, 'child' (or, 'children') is a basic semantic molecule which is conceptually prior even to 'woman' and 'man'. To build into an account of kinship, the molecule 'be born' is also indispensable but with these molecules in hand the most basic kinship concepts, namely, 'mother' and 'father', can be explicated, and these in turn open the way for more complex, and more language-specific, terms such as 'brother' and 'sister' to be explicated (Wierzbicka 2016).

Needless to say, many of the details of these claims are non-obvious. For example, it may be asked: How do we know that 'during the day' is simpler than 'sun'? Couldn't it be the other way around? And why is 'woman' regarded as semantically simpler than 'man'? The conventional view in linguistics either puts them on a par, or else sees 'man' (or, 'male') as the relatively simpler term. To questions like these, the answer is that these claims have emerged from painstaking conceptual analysis, that is, experimenting with different approaches and with many alternative formulations (Goddard and Wierzbicka 2014: Ch 2). It is not possible, or necessary, in this chapter to go over all the arguments and evidence for individual analyses. The curious or sceptical reader can consult the relevant literature.

Following from the exposition so far, it is clear that semantic molecules enable great compression of semantic complexity, even though this complexity is 'disguised by its being encapsulated and telescoped into lexical units embedded one in the other, like a set of Russian dolls' (Wierzbicka 2009a). Though this observation has great theoretical importance for conceptual semantics, for the purpose of Minimal English this kind of 'semantic compression' hardly matters. The more important take-away message is that there is a stock of basic, and probably universal, semantic molecules that can be taken for granted and incorporated into Minimal English. And, equally important, that the membership of this second tranche of 'safe words' (in addition to the first tranche, the semantic primes) does not depend on anyone's immediate and untested intuitions, but rather rests on a body of serious analytical work.

In the next section, we will overview the proposed inventory of universal molecules, before moving to the topic of 'approximately universal' molecules, and then to highly language-specific semantic molecules. Before that, it is well to make one further point to correct a possible mis-impression that may be gained from the layout of Fig. 3.1. Contrary to the look of this diagram, it is not the case that semantic molecules are restricted to particular 'silos of meaning'. A molecule

such as 'water', for example, is not only necessary for explicating the concept of drinking (basically, it helps get to the notion 'liquid', that is, something like water). It is obviously necessary for explicating dozens or hundreds of ordinary concepts, such as 'rain', 'river', 'lake', 'sea', 'tears', as well asbeing a conceptual constituent of verbs like 'pour' and 'flow', and the names of other'liquids' such as *milk*, *oil*, *tea*, *petrol*, etc. Likewise, 'children' (and 'child') not only helps explicate 'women', 'be born' and many other obvious 'child words', such as 'play', 'toy' and 'school', it may turn up in unexpected places, for example, in the explication for English words like 'tender' (in the sense of 'tender feelings', which incorporates a reference to how one may feel towards a small child) and 'cute'. In short, all or most of the inventory of proposed universal molecules are extremely prolific or versatile.[10]

3.4.2 *Universal semantic molecules*

As mentioned several times, current research suggests that some semantic molecules arelikely to be found in all or most languages. Examples are given in the list below (about 60 in number) and discussed briefly in their respective groups. Aside from its importance to cross- cultural communication, the likely existence of certain semantic molecules in all languages has profound implications for our understanding of human cognition and knowledge.

• **'Body-part' words: hands, mouth, eyes, head, ears, nose, face, legs, teeth, fingers, breasts, skin, bones, blood**

Notice that not all English body-parts are included in this list. Words like 'shoulders', 'wrist', and even 'feet' and 'neck', are absent because they are known to be language specific. It may well be, of course, that such words still play a part in concept- building in languages which have them; for example, 'feet' is a necessary molecule for explicating the English verb *walk* (Goddard et al. 2016).

• **'Biosocial' words, that is, stage of life and kin terms: children, men, women, be born, mother, father, wife, husband**

Again, it bears noting that many seemingly ordinary English words are absent from this list on account of non-translatability; for example, 'baby', 'boy', 'girl' are absent from the stage of life list, and 'brother' and 'sister' are absent from the kin list.

• **'Biological' world: creature, grow (in ground)**

It seems that perhaps only these two words have a reasonable claim to strict universality. 'Creature' underlies all words for animals, fish, birds, and other life-forms,[11] and 'grow(in

the ground') underlies the world of 'plants'. There are also many other near-universal or approximately universal molecules in this area (see next section).

- **'Physical qualities': long, round, flat, hard, soft, sharp, smooth, heavy**

Such words are essential for characterizing all manner of real-world things and their parts, including most body parts. It seems likely that these meanings are present in all languages, once polysemy is taken into account. Obviously words like 'long' 'hard' and 'sharp' have several meanings. The meanings we are interested are found in expressions such as 'something long', 'something hard', and 'something sharp'.

- **'Position and orientation in space': be on (something), at the top, at the bottom, in front, around**

This group of words are to do with where something is in relation to a physical object, someone's body, and so on. Except for 'be on (something)', whose meaning has to do with touching and stability of position, the words in this group involve where something is in relation to something else. They depend on semantic primes ABOVE, BELOW, or ON ONE SIDE, together with other elements.

- **'Materials': wood, stone**

These may well be the only two materials which are universally attested and available across all geographical zones. The language-specific molecules in this grouping are, of course, much more diverse, including words like 'paper', 'glass', and 'metal'.

- **'Environmental' words: the sky, the sun, the ground, the earth; during the day, at night; water, fire**

This is grouping falls into three clusters. They roughly correlate with places ('sky', 'sun', 'ground', 'on earth'), with times ('during the day', 'at night'; notice that this pair of expressions is adverbial), and with natural elements ('water', 'fire').

- **'Time periods, natural units of time': day, year**

'Day' in this sense is noun-like, as in expressions like 'one day after'. The cross-linguistic status of 'year' is still uncertain.

- **'Knowing and naming': know (someone), be called**

The idea that we 'know' certain people (and in some cases, 'know them well') is an essential

ingredient in many concepts and cultural scripts concerned with human interactions and relationships.[12] As for 'be called', this semantic molecule underlies naming, that is, linking a person, place or thing with an identifying word.

- **'Doing': hold, sit, lie, stand, sleep; play, laugh, sing; make, kill**

Before going through this list, it is interesting to observe the absence of various verbs that are often assumed to represent universals, such as 'hit', 'eat', and 'go'. Some of these figure in later sections of this chapter, but in the present section we are concerned with plausible candidates for universal status, and it is known that verbs like 'hit', 'eat', and 'go' are quite variable cross-linguistically.

Most of the better candidates, such as those listed, are fairly simple in their semantic structure, though 'kill' and 'make' (in the sense of 'making a tool, etc.') are the only ones that may be composed purely of semantic primes. 'Hold (something in the hand)' is a plausible candidate on account of the importance of the hands in using tools of all kinds. The three 'body posture' verbs ('sit', 'lie', 'stand') involve what can be thought of as canonical positions of rest or readiness of the human body. 'Sleep' (in the sense 'be sleeping') includes 'lie' and 'at night' in its meaning structure — not as necessary conditions, but in terms of a prototype ('as people often do when it is like this: ...').

'Play' (as in 'children were playing outside'), 'laugh' (as in 'she was laughing'), and 'sing' (as in 'they were singing') are all activity verbs, that is, their basic senses all imply duration, and they are all connected with potential good feelings in people.

'Kill (someone)' is a fairly plausible semantic molecule, reflecting the role of violence in human life. The status of 'make' is not so clear. Certainly it is often and rightly said that humans are the 'tool makers' par excellence, but it is also true that 'make verbs' in the world's languages are very polysemous and can have widely different ranges of use. The current hypothesis is that there is a shared meaning, but the details are not perfectly clear at this point (cf. Goddard and Wierzbicka 2016b), so the proposal that 'make' is a universal semantic molecule is still tentative.

In the next two sections, we turn to semantic molecules which are language-specific to a greater or lesser extent: from 'approximate' universals (section 3.4.3) to highly culture-specific molecules (section 3.4.4).

3.4.3 Language-specific but 'approximate' universal semantic molecules

Despite the inherent interest and importance of universal semantic molecules, it is incontrovertible that there are many aspects of shared human experience which are not associated with any strictly universal words. Consider 'eat' and 'drink', for example. All people eat and drink, as we

say in English, but it is a fact that not all languages have verbsthat precisely match English 'eat' and 'drink' in their meanings. There are languages like Kalam (Wierzbicka 2009b), which have one main verb to cover both activities, as we see them in English, and there are languages like Mandarin Chinese and Shanghainese (Ye 2010) where the semantic boundary between the near-equivalents to 'eat' and 'drink' is drawn slightly differently to English. Nonetheless, it can hardly be doubted that in any language the nearest equivalents to 'eat' and 'drink' are important semantic molecules. 'Eat' and 'drink', then, are examples of what we are terming 'approximate' universal molecules (the term is somewhat oxymoronic, to be sure).

So-called "life form" words (Berlin 1992), such as 'bird', 'fish' and 'tree', furnish a second example. On the one hand, a reliable body of linguistic anthropology and ethnobiology indicates that comparable words do not have exactly the same meanings in all languages. On the other hand, it seems likely that all languages have some such terms and, equally importantly, that in any language there will be numerous words for different kinds of bird, different kinds of fish, different kinds of tree.

Presumably in many cases, it will be possible to identify a common core of meaning that isstable across languages, despite the variability in other respects. For example, with 'eat' and 'drink' the common core evidently turns on the idea of someone doing something to something as people often do, because of wanting some of it to be first inside the mouth and then somewhere else inside the body (cf. Goddard 2011a: 278-285). How 'eat and drink' words differ across languages concerns the mechanics of the process, which is connected with the kinds of foodstuff being consumed (roughly, how solid and/or liquid it is). With 'bird' and 'fish', the common core presumably turns on the idea of creatures of many kinds that belong in the air and in the water, respectively, and have some sharedbodily features, such as wings or fins (cf. Wierzbicka 1985; Goddard 2011: Ch 7).

At the present time, it seems reasonable to propose the following short list of possible 'approximate' universal molecules. They are a disparate set. In the biological sphere: 'seeds', 'grass', 'flower', 'egg', and 'dog' (because dogs are nearly universal animals in people's lives). In the area of human life and social relations, perhaps 'house' and 'family' would havesome claim to being 'approximate' universal molecules, remembering that this hypothesis only posits the existence of a shared semantic core and thus allows for some semanticvariability.

In the environmental department, it may be that words like 'river', 'sea', and 'mountain' have some claim to being approximate universals, despite the known variability in ethnogeographical vocabulary (Bromhead 2011, 2013). The putative common core of the concepts 'river' and 'sea' is that they are places where is a lot of water, with 'river' additionally implying that the water is moving (flowing) and that such a place, that is, a river, has two sides, and 'sea' additionally

implying something like 'vastness'. As for 'mountain', the putative common core would be that of a very elevated place, that is, the place at the top of it is (or, is thought of as) very far above the surrounding land below. One wonders about words like 'jungle' or 'forest', which might have a common core in terms of there being a great many trees in one place, such that when people are there, they can't see very far because of it.

In relation to 'foods', it is possible that 'flour' and 'sweet' (and perhaps 'meat') might qualify as semantic molecules. The rationale for 'flour' is that grinding is a very widespread, if not universal, technology and that the meaning of the verb 'grind' appears to require something like 'flour' as a semantic molecule (grinding is, of course, typically used not only for grain but for other purposes as well). As for 'sweet', it appears to be a very widespread semantic molecule in words for kinds of fruits, as well as human made delicacies.

3.4.4 Other possibly universal or near-universal words

Are there words whose meanings are universal or approximately universal, but which are not semantic molecules, that is, do not function as conceptual building blocks for other word meanings? It seems logically possible. If it were true, it would be of considerable interest to linguists, particularly to semanticists and lexical typologists. It is not particularly important from the point of view of Minimal English, however, where the main consideration is translatability.

I will just mention a few possibilities here. 'Moon' and 'stars' seem like reasonable candidates for universal status, yet they do not seem to be particularly important as semantic molecules (though 'moon' may be part of the concept of 'month'). Also in the environmental department, one may ask whether approximations to 'east', 'west', 'north', and 'south' exist in all or most languages. Among body-parts, 'brain' and 'heart' seem likely candidates. If there is a universal 'tool', with functional equivalents across all cultures, it would arguably be 'knife', yet it too does not seem to be very important as a semantic molecule. Among 'doing' words, it may be that words like the following are widely attested, at least approximately, in the world's languages, without being semantic molecules: 'go/went', 'take (someone somewhere)', and 'learn'. With a few exceptions, relatively little research has yet been done into the cross-linguistic semantics of most of these words.

3.4.5 Language/culture-specific semantic molecules

In English and in other contemporary languages, and presumably in all languages, there are many semantic molecules which are tied to local aspects of geography, history or culture. This does not necessarily make them any less important to the lexical structure of their respective languages

or to habitual ways of thinking in their respective cultures.

The most obvious examples are words connected with the environment, such as 'island' or 'snow'; staple foods, such as 'rice', 'wheat', 'corn', 'yams', and 'plantain'; and words for large and important animals, such as 'horse', 'cow', 'buffalo', 'moose', or 'caribou'. There are materials that have existed in some, but not all, parts of the world for centuries, such as 'gold', 'iron', 'glass', 'paper', 'leather', 'cloth', and 'wool'; and similarly, items of traditional technology and transport, such as 'wheel', 'road', and 'boat'. There is also an abundance of similar words that belong to the industrial and information ages: 'new' materials such as 'metal', 'rubber', 'plastic'; 'new' kinds of transport and technology, such as 'car', 'plane', 'train', 'engine'; 'gun'; 'telephone', 'radio', 'television', and (perhaps) 'computer'.

In relation to people's ways of living and social organization, words like 'village', 'city', and 'country' are important language-specific molecules in the modern world; social institutions such as 'school', 'hospital' and 'the law'; and words for professions such as 'doctor', 'teacher', and 'soldier'. The related words 'read', 'write', and 'book' are hugely important semantic molecules in many concepts related to modern life, not to mention 'money' and 'God'. Needless to say, many of these words and concepts, albeit language/culture-specific, are not confined to a single language or culture, but are found across a broad cultural area.

The language specificity of some molecules may not be obvious to ordinary speakers of a language, especially if they are deeply ingrained in the lexicon and in everyday ways of thinking. Three pertinent examples from English and other European languages are the words 'colour', 'number', and the pair 'brother' and 'sister'. 'Colour' arguably functions as a semantic molecule in the meaning of words like *red*, *blue*, *green*, and so on. Wierzbicka (2008) contends that many non-European languages lack 'colour words' in the true sense, because their visual descriptors do not involve any comparable molecule. The semantic molecule 'number' forms part of the meaning of many quantitative and measurement terms of English and similar languages, as well as underpinning the number system (Goddard 2009. Finally, Wierzbicka (2016) argues that 'brother' and 'sister', though emphatically not lexical universals (unlike 'mother' and 'father', for example), function as semantic molecules in the kinship vocabulary of many European languages, for example, in the meanings of words like *uncle* and *aunt*.

Needless to say, the discussion in this section has not been anything like comprehensive. The details of the molecule theory are still being worked through, and refinements and adjustments are ongoing. Compared with 40 years developing the core NSM of semantic primes, as yet only 10–15 years has been devoted to semantic molecules. There is still much to be learnt.

3.4.6 *A parting note on other useful international words*

As mentioned in Chap. 2, there are plenty of other useful, or even essential, words for Minimal English which are not semantic molecules (or, at least, not particularly important ones) and which are not universal or near-universal either, taking the broad sweep of human history and human geography into account. Rather, they are needed in Minimal English because they are important to international discourse in the modern world. They include words to do with countries, government and migration, such as 'government', 'border','passport', and 'vote'; words that are important for talking about health issues, such as 'medicine', 'mosquitoes', 'flies', 'salt', and 'sugar'; conflict and security, such as 'war' and 'bomb'; words needed to talk about climate change and pollution, such as 'electricity', 'oil', 'coal', 'petrol', and 'air'; natural disasters, such as 'flood', 'storm', and 'earthquake'; and cultural words such as 'newspaper', 'film', 'photo', 'internet', and 'music'.

3.5 From NSM to Minimal English

As mentioned, the Minimal English project began to take shape in 2013, impelled by Anna Wierzbicka's sense that the NSM research community had accumulated enough knowledge and experience about semantic variation and cross-translatability that it was now practical to adapt NSM into a user-friendly tool for thinking and communicating outside the confines of Anglo English (cf. Wierzbicka 2014: Ch 14).

It would be difficult to summarize the development process of Minimal English so far, and the precise details are hardly important. A few key observations will suffice. The process began with consultations and discussions among linguists in the NSM research community, led by Anna Wierzbicka and Cliff Goddard. The starting point was the empirical findings in hand, so to speak, about semantic primes and molecules, and about shared syntax. As for what extensions were necessary to this fundamental core, decisions were guided by an inductive approach, that is, Wierzbicka, Goddard, and others began to experiment with crafting texts in Minimal English for various purposes and in different domains. Bert Peeters, Marie-Odile Junker, and Gian Marco Farese all made notable contributions to these efforts. When additional non-universal words and expressions were needed, efforts were made to identify the most translatable ways to go. In this process it was significant that the NSM research community comprises a diversity of scholars with rich experience in cross- linguistic semantics, whose importance can perhaps be likened to the importance of 'clinical experience' in medicine. Some of the early Minimal English texts eventually developed into texts that appear in chapters in this volume: the ethics scripts in Wierzbicka's Chap. 6, and the 'popular science' texts in Chaps. 8 and 9. Wierzbicka began an

ambitious project now titled 'The Story of God and People' (cf. Wierzbicka 2017, forthcoming).

In 2015 a Symposium titled 'Global English, Minimal English' was held at the Australian National University, with the goal of bringing the idea of Minimal English to the attention of scholars and professionals across many fields. The uptake and interest were very promising, as shown by a number of the chapters in this book. This development process — inductive and practice-based — is continuing; for example, with the collaboration with narrative medicine (cf. Chap. 11), several projects in cross-cultural and language education (cf. Peeters ed. forthcoming; Sadow 2015, 2016; Tully 2016) and literacy development and translation practice in the Pacific (cf. Hill et al. 2016).

The aim of the present chapter has been to give the non-specialist a sense of the knowledge base upon which Minimal English rests.

Appendix A: Semantic Primes, English Exponents (after Goddard and Wierzbicka 2014)

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

Notes: • Exponents of primes can be polysemous, i.e. they can have other, additional meanings • Exponents of primes may be words, bound morphemes, or phrasemes • They can be formally complex • They can have language-specific combinatorial variants (allolexes, indicated with ~) • Each prime has well-specified syntactic (combinatorial) properties.

Notes

1 The only other linguists who clearly acknowledge the importance of semantic metalanguage, and have done so since the mid-1960s, are the Moscow Semantic School, whose leading figures are Jurij Apresjan and Igor Mel'čuk. In his book *Systematic Lexicography*, Apresjan (2000: 224) explains: '[T]he metalanguage of lexicography is a sub-language of the object language, comprising a relatively small and unified vocabulary and syntax. The basis of this metalanguage is semantic primitives. With the aid of the metalanguage, complex semantic units of the object language (grammatical as well as lexical) are reduced to a fixed structure of semantic primitives by a process of hierarchical breakdown.'

2 In some ways, the NSM approach can be seen as following the tradition of lexicography, but, by taking a much more systematic approach, striving to avoid the many pitfalls of conventional lexicography (Wierzbicka 1996: Ch 9; Goddard in press/2017). Aside from circularity and the 'pseudo-defining' of relatively simple words in terms of more complex ones, perhaps the most common fault of conventional dictionary-making is not distinguishing invariant meaning from contextual enrichment, leading to the postulation of unlikely dozens of 'senses'. Another common flaw is reliance on open-ended terms such as usu. (usually) or esp. (especially), and the so-called lexicographer's crutch, the all-purpose 'etc.'

3 Written mainly in Latin, they were compiled by his editor, after Leibniz's death, as his "Table de definitions" (see Leibniz 1903: 437-510.

4 If one had to nominate a single benchmark in the intervening period, that distinction would fall to Wilhelm von Humboldt (1767-1835), one of the founding figures of modern comparative linguistics and anthropological linguistics. Humboldt tended to be more interested in culture-unique words than in universal words, on account of his interest in folk (ethno-national) cultures. '[E]ach language draws a circle around the people to whom it adheres which it is possible for the individual to escape only by stepping into a different one.' Even so, he also wrote: 'To be sure, a midpoint, around which all languages revolve, can be sought and really found, and this midpoint should always be kept in mind in the comparative study of languages, both in the grammar and lexicon'.

5 Established NSM researchers, roughly arranged into chronological waves, include: (i) Bert Peeters, Felix Ameka, Jean Harkins, Cliff Goddard, (ii) Zhengdao Ye, Jock Wong, Anna Gladkova, Rie Hasada, Deborah Hill, Kyung-Joo Yoon, Marie-Odile Junker, (iii) Carsten Levisen, Maria Auxiliadora Barrios Rodriguez, Sandy Habib, Adrian Tien, Radoslava Trnavac. Recently completed PhDs include Lien Huong Vo, Helen Leung, and Gian Marco Farese.

6 To be precise, one would first have to expand 'A happened after B' to 'B happened at some time, A happened after'. Then, it could be claimed that the latter is equivalent to 'A happened at some time, B happened before'.

7 When the exponent of a prime is polysemous, the intended primitive meaning can always be made perfectly clear by reference to key grammatical frames (see section 3.3) in which it occurs, but from which the other meanings are excluded.

8 To give an example which will be less familiar to many readers, there are plenty of languages in which the prime SOMETHING has two or more exponents. A typical pattern is for one exponent to be used across a wide range of contexts, while another occurs only

when SOMETHING is the complement of a particular verb, such as DO or SAY. Persian (Farsi) is one such language. The main exponent of SOMETHING IS *čiz(i),* but as the complement of DO, the allolex *kār* (lit. 'work') is used (Arab 2016).

9 This formulation, using 'now', is an over-simplification, as the real implicit reference point is the time of speaking, that is, 'when I say this now'.

10 Equally it is interesting to note that there is much variation in how many molecules (if any) are required in explications. Many explications require no molecules at all and others require only one or two. Explications for common verbs of "doing and happening" (Goddard and Wierzbicka 2009, 2016b) usually require a handful. Explications for concrete nouns, on the other hand, require significantly more. Wierzbicka's (2015b) explications for English *spoon* and Chinese *tangchí,* for example, use about 15 molecules per explication.

11 Some might wonder: what about 'animal'? Actually, this word is much more English-specific than 'creature' and much less cross-translatable. Even in languages as close to English as Danish and German, the normal translation equivalents for *animal* (Danish *dyr*, German *Tier*) are closer to 'creature', insofar as they are used freely about spiders, fish, etc.

12 In English and many languages, this meaning exists as a distinct sense of the exponent of KNOW, but it is well known that many other languages have a distinct verb for 'know (someone)', e.g. German *kennen*, French *connaître*. NSM researchers once believed that the two verbs in French and German, for example were allolexes but it has become clear that 'know (someone)' can be explicated, as shown (Wierzbicka in press).

 I know this someone:

 I know some things about this someone

 because I was with him/her for some time before

 because of this, I can think like this: "he/she is like this"

References

Amberber, Mengistu, ed. 2007. *The Language of Memory in a Cross-Linguistic Perspective.* Amsterdam: John Benjamins.

Apresjan, Jurij. 2000. *Systematic Lexicography*. Translated by Kevin Windle. Oxford: Oxford University Press.

Arab, Reza. 2016. Farsi (Persian) exponents of semantic primes. [www.griffith.edu.au/humanities-languages/school-humanities-languages-social-science/research/natural-semantic-metalanguage-homepage/downloads]

Aristotle. 1928. Of Definitions. Extracts from *Topica* Book VI, Ch 1–4. In *The Works of Aristotle*., ed. W.D. Ross. London: Oxford University Press.

Arnauld, A., and P. Nicole. 1996 [1662]. *Logic or the Art of Thinking*. Trans. J. Buroker. Cambridge: Cambridge University Press.

Berlin, Brent. 1992. *Ethnobiological Classification*. Princeton, N.J.: Princeton UniversityPress.

Bogusławski, Andrzej. 1966. *Semantyczne pojęcie liczebnika i jego morfologia w języku rosyjskim*, Wrocław: Ossolineum.

Boguslawski, Andrzej. 1970. On semantic primitives and meaningfulness. In *Sign, Language and Culture*, ed. A. J. Greimas, Roman Jakobsen, and M. A. Mayenowa, 143-152. The Hague: Mouton.

Bromhead, Helen. 2009. *The Reign of Truth and Faith: Epistemic expressions in 16th and 17th century English*. Berlin: Mouton de Gruyter.

Bromhead, Helen. 2011. Ethnogeographical categories in English and Pitjantjatjara/Yankunytjatjara. *Language Sciences* 33(1): 58-75.

Bromhead, Helen. 2013. *Mountains, Rivers, Billabongs: Ethnogeographical categorization in cross-linguistic perspective.* PhD thesis, Australian National University.

Collins Cobuild English Language Dictionary. 1991.

Enfield, N.J., and Anna Wierzbicka, eds. 2002. *The Body in Description of Emotion.* Special Issue of *Pragmatics and Cognition* 10(1/2).

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

Goddard, Cliff. 1989. Issues in Natural Semantic Metalanguage. *Quaderni di Semantica* 10(1): 51-64.

Goddard, Cliff. 1994. Semantic Theory and Semantic Universals. In *Semantic and Lexical Universals – Theory and Empirical Findings*, ed. Cliff Goddard and Anna Wierzbicka, 7-30. Amsterdam: John Benjamins.

Goddard, Cliff. 1996. *Semantic Analysis: A Practical Introduction.* [First edition] Oxford: Oxford University Press.

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

Goddard, Cliff. 2002a. The search for the shared semantic core of all languages. In *Meaning and Universal Grammar — Theory and Empirical Findings, Volume I,* ed. Cliff Goddard and Anna Wierzbicka, 5-41. Amsterdam: John Benjamins.

Goddard, Cliff. 2002b. The on-going development of the NSM research program. In *Meaning and Universal Grammar – Theory and Empirical Findings, Volume II*, ed. Cliff Goddard and Anna Wierzbicka, 257-300. Amsterdam: John Benjamins.

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. 2010. *A piece of cheese, a grain of sand*: The semantics of mass nouns and unitizers. In *Kinds, Things and Stuff. Mass Terms and Generics,* ed. Francis Jeffry Pelletier, 132-165. New York: Oxford University Press.

Goddard, Cliff. 2011. *Semantic Analysis – A Practical Introduction.* [Revised 2nd edition] Oxford: Oxford University Press.

Goddard, Cliff. 2012. Semantic primes, semantic molecules, semantic templates: Key concepts in the NSM approach to lexical typology. *Linguistics* 50(3): 711–743.

Goddard, Cliff, ed. 2013. *Semantics and/in Social Cognition.* Special Issue of *Australian Journal of Linguistics* 33(3).

Goddard, Cliff. 2016. Semantic molecules and their role in NSM lexical definitions. *Cahiers de lexicologie* 2016/4: 13-36.

Goddard, Cliff. In press/2017. The future of lexicography: NSM perspectives. In *International Handbook of Modern Lexis and Lexicography*, ed. Patrick Hanks and Gilles- Maurice de Schryver. Springer.

Goddard, Cliff, and Anna Wierzbicka, eds. 1994. *Semantic and Lexical Universals – Theory and Empirical Findings*. Amsterdam: John Benjamins.

Goddard, Cliff, and Anna Wierzbicka, eds. 2002. *Meaning and Universal Grammar – Theory and Empirical Findings*. [Two volumes]. Amsterdam: John Benjamins.

Goddard, Cliff, and Anna Wierzbicka, eds. 2004. Cultural Scripts. Special Issue of *Intercultural Pragmatics* 1(2).

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

Goddard, Cliff, and Anna Wierzbicka. 2014. *Words and Meanings: Lexical Semantics Across Domains, Languages and Cultures*. Oxford: Oxford University Press.

Goddard, Cliff, and Wierzbicka, A. 2016a. 'It's mine!'. Re-thinking the conceptual semantics of "possession" through NSM. *Language Sciences* 56: 93-104.

Goddard, Cliff, and Wierzbicka, Anna. 2016b. Explicating the English lexicon of "doing" and "happening". *Functions of Language* 23(2): 214-256.

Goddard, Cliff, and Zhengdao Ye, eds. 2015. *"Happiness" and "Pain" Across Languages and Cultures*. Amsterdam: John Benjamins.

Goddard, Cliff, Wierzbicka, Anna and Wong, Jock. 2016. "Walking" and "running" in English and German: The conceptual semantics of verbs of human locomotion. *Review of Cognitive Linguistics* 14(2), 303-336.

Harkins, Jean, and Anna Wierzbicka., eds. 2001. *Emotions in Crosslinguistic Perspective*. Berlin: Mouton de Gruyter.

Hanks, Patrick W. 2007. General Introduction. In *Lexicology: Critical Concepts in Linguistics*, ed. Patrick W. Hanks, Vol I, 1–23. Abingdon: Routledge.

Hill, Deborah, Gabriel Ropovono, and Matilda Matala. 2016. *Tangavulu u'unu ni Longgugi* [10 Longgu Stories]. University of Canberra.

Leibniz, G. 1903. *Opuscules et fragments inedits de Leibniz*. Ed. Louis Couturat. Paris.

Levisen, Carsten. 2012. *Cultural Semantics and Social Cognition. A case study on the Danish universe of meaning*. Berlin: Mouton de Gruyter.

Lyons, John. 1977. *Semantics*. Cambridge: Cambridge University Press.

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

Mel'čuk, Igor. 2012. *Semantics: From Meaning to Text*. Amsterdam: John Benjamins.

Peeters, Bert, ed. 2006. *Semantic Primes and Universal Grammar: Empirical evidence from the Romance Languages*. Amsterdam: John Benjamins.

Peeters, Bert, ed. 2015. *Language and Cultural Values: Adventures in applied ethnolinguistics*, Special issue of *International Journal of Language and Culture* 2(2).

Peeters, Bert, ed. Forthcoming/2018. "Words on the Move: English Key Word Exports Captured in Minimal English". Special Issue, *Multilingua*.

Sadow, Lauren. 2015. Cultural Scripts in ESL Classrooms. Paper presented at the Annual Applied Linguistics Association Australia Conference, Adelaide, Australia.

Sadow, Lauren. 2016. Teaching Invisible Culture in ESL Classrooms. Paper presented at National University of Singapore Centre for English Language Communication Symposium, Singapore.

Steiner, George. 1992 [1975]. *After Babel*. Oxford University Press.

Tien, Adrian. 2015. *The Semantics of Chinese Music. Analysing Selected Chinese Musical Concepts.* Amsterdam: John Benjamins

Tully, Alex. 2016. Minimal English Paraphrasing in Language Teaching. MA Thesis, Australian National University.

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

Wierzbicka, Anna. 1980. *Lingua Mentalis: The Semantics of Natural Language*. Sydney: Academic.

Wierzbicka, Anna. 1985. *Lexicography and Conceptual Analysis.* Ann Arbor: Karoma.

Wierzbicka, Anna. 1987. *English Speech Act Verbs: A Semantic Dictionary*. Sydney: Academic Press.

Wierzbicka, Anna. 1988. *The Semantics of Grammar*. Amsterdam: John Benjamins.

Wierzbicka, Anna. 1989. Semantic primitives – The expanding set. *Quderni di Semantica* X(2): 309-332.

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*. New York: Oxford University Press.

Wierzbicka, Anna. 1996. *Semantics: Primes and Universals*. New York: Oxford University Press.

Wierzbicka, Anna. 2001. Leibnizian linguistics. In *Perspectives on Semantics, Pragmatics, and Discourse: A Festschrift for Ferenc Kiefer*, ed. I. Kenesei, and R. Harnish, 229-253. Amsterdam: John Benjamins

Wierzbicka, Anna. 2003. *Cross-Cultural Pragmatics: The Semantics of Human Interaction.* Berlin: Mouton de Gruyter. (Expanded 2nd edition).

Wierzbicka, Anna. 2006. *English: Meaning and Culture.* New York: Oxford University Press.

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

Wierzbicka, Anna. 2009a. The theory of the mental lexicon. *HSK (Handbücher zur Sprach und Kommunikationswissenschaft. Handbook of Linguistics and Communication Science) "Slavic Languages",* 848–863. Berlin: Mouton de Gruyter.

Wierzbicka, Anna. 2009b. All people eat and drink: Does this mean that 'eat' and 'drink' are universal human concepts? In *The Linguistics of Eating and Drinking*, ed. John Newman, 65–

89. Amsterdam: John Benjamins.

Wierzbicka, Anna. 2010. *Experience, Evidence and Sense: The hidden Cultural Legacy of English*. New York: Oxford University Press.

Wierzbicka, Anna. 2014. *Imprisoned in English: The Hazards of English as a Default Language*. New York: Oxford University Press.

Wierzbicka, Anna. 2015. The idea of a 'spoon': semantics, prehistory, and cultural logic. *Language Sciences* 47: 66–83.

Wierzbicka, Anna. 2016. Back to 'Mother' and 'Father': Overcoming the Eurocentrism of Kinship Studies through Eight Lexical Universals. *Current Anthropology* 57(4): 408-429.

Wierzbicka, Anna. In press. I KNOW: A human universal. In *Epistemology for the Rest of the World*, ed. Masaharu Mizumoto, Eric McCready, Jason Stanley, and Stephen Stich.

Wong, Jock O. 2014. *The Culture of Singapore English*. Cambridge: Cambridge University Press.

Ye, Zhengdao. 2010. Eating and drinking in Mandarin and Shanghainese: A lexical- conceptual analysis. In *ASCS09: Proceedings of the 9th Conference of the Australasian Society for Cognitive Science,* ed. E. Christensen, E. Schier, and J. Sutton, 375–83. Sydney: Macquarie Centre for Cognitive Science.

Ye, Zhengdao, ed. 2017. *The Semantics of Nouns*. Oxford: Oxford University Press.

Zholkovsky, A. 1974 [1964]. The vocabulary of purposeful activity. In *Machine Translation and Applied Linguistics*, ed. Ju Rozencveig, vol I, 197-234. Frankfurt am Main: Athenion [First Published in Russian 1964].