Language Sciences 33 (2011) 40–57

Contents lists available at ScienceDirect

Language Sciences

journal homepage: www.elsevier.com/locate/langsci

The lexical semantics of *language* (with special reference to *words*)

Cliff Goddard *

School of Behavioural, Cognitive and Social Sciences, University of New England, Armidale, NSW 2351, Australia

ARTICLE INFO

Article history:
Received 11 September 2008
Received in revised form 27 March 2010
Accepted 29 March 2010

Keywords:
Cross-linguistic semantics
Language concept
Lexical semantics
NSM
Polysemy
Semantic primes

ABSTRACT

Language can be regarded as one of the key words of English, as well as the foundational term of the discourse of linguistics. It is well to remember, however, that the concept of *a language* lacks precise semantic equivalents in many languages. This study presents a semantic-lexicographic analysis of several meanings of the word *language* in contemporary English, using the Natural Semantic Metalanguage method of semantic description (Wierzbicka, 1996, 1997; Goddard, 1998, 2008). The study is similar in scope and approach to an earlier study (Goddard, 2005) of the word *culture*, which resembles *language* in several important respects. One distinctive aspect of the explications for *language* is their reliance on the proposed semantic prime WORDS, which is discussed at some length. Though primarily focused on English, the study makes reference to Yankunytjatjara, Chinese, and Russian, among other languages.

1. Introduction

This study is first and foremost an investigation of the lexical semantics of various senses of the word *language* in ordinary contemporary English, conducted using the Natural Semantic Metalanguage (NSM) approach to semantic analysis (Wierzbicka, 1996; Goddard and Wierzbicka, 1994, 2002; Peeters, 2006; Goddard, 1998,2008; and other works). It is a companion paper to an earlier study of the lexical semantics of *culture* (Goddard, 2005); as we will see, there are important parallels in the semantic structure of the two concepts. As far as I know, no previous study has sought to bring the methods of linguistic semantics to bear on the word *language*. In addition to its primary goal, the study has several secondary goals: to shed light on the culture-historical positioning of the "language concept"; to clarify the NSM claim that WORDS (a crucial element in the proposed semantic structure of *language*) is a semantic universal and to refute certain overstated claims to the contrary; and to consider the utility and pitfalls of *language* as a key word of contemporary Anglophone discourse.

The general assumptions of the NSM approach are reasonably well-known. It is a conceptualist approach to meaning, whose method of description is reductive paraphrase. Its primary tool is a highly constrained vocabulary of 63 semantic primes (such as SOMEONE, SOMETHING, PEOPLE, DO, SAY, WANT, KNOW, GOOD, BAD, BECAUSE, CAN, and others), which have been arrived at after a lengthy program of semantic research, beginning with Wierzbicka (1972). The current inventory of semantic primes is displayed in full in Appendix A, in their English versions. Evidence suggests that semantic primes are present as lexical units, i.e. as discrete meanings of words or word-like elements (bound morphemes or phrasemes), in all languages; and furthermore, that they can be combined into phrases and sentences following combinatorial rules which are shared across all languages. Together, the semantic primes and their rules of combination constitute a kind of "mini-language" which is an ideal tool for semantic-conceptual analysis: hence the term "natural semantic metalanguage". Versions of this metalanguage have been documented for French, Spanish, Russian, Chinese, Polish, Korean, Japanese, Mbula (PNG), Amharic, East Cree, and a variety of other languages, in various NSM publications (Goddard and Wierzbicka, 2002; Peeters, 2006; Goddard, 2008).

* Tel.: +61 2 6773 3309; fax: +61 2 6773 3748.
E-mail address: cgoddard@une.edu.au

0388-0001/$ - see front matter © 2010 Elsevier Ltd. All rights reserved.
doi:10.1016/j.langsci.2010.03.003

The goal of an NSM lexical-semantic analysis is to arrive at a well-evidenced reductive paraphrase (an "explication") framed exclusively in semantic primes for each discrete sense of the word (or other expression) under consideration.[1] I will argue that the word *language* has five main senses in contemporary English. Section 2 deals with the count noun sense of the word, according to which *English* and *Greek*, for example, are two different *languages*$_1$. Section 3 takes up the other main meanings of the word: *language*$_2$, roughly 'language in general' (a mass noun), as in expressions like *the origins of language*; *language*$_3$, roughly 'word usage', as in expressions like *bad language* and *Shakespeare's language*; *language*$_4$, roughly 'a specialised way of speaking', as in expressions like *the language of science*; and *language*$_5$, roughly 'an expressive medium', as in expressions such as *the language of music*. For each of these senses I will propose and seek to justify a semantic explication phrased exclusively in the metalanguage of semantic primes. Naturally, the constrained metalanguage means that NSM explications turn out to be much lengthier than other modes of semantic description, and, as one would expect, the restricted choice of words gives them an unusual stylistic quality; but by relying on a small inventory of simple cross-translatable meanings, the NSM approach eliminates the circularity and obscurity that plague most dictionary definitions (and most scholarly discussions), enables an extremely fine-grained resolution of meaning, and safeguards against terminological ethnocentrism.

2. Language$_1$

In Sections 2.1 and 2.2, I provide some culture-historical and lexicographic perspectives on the meaning *language*$_1$, after which, in Section 2.3, I present and justify a reductive paraphrase explication. Section 2.4 reviews the status of this English concept in cross-linguistic perspective.

2.1. A brief culture-historical perspective

Linguistic historiographers and language historians have established that the concept of *a language*, as we know it in English, is linked with the social processes of language standardisation; and from the beginning, language standardisation was linked with territories and nations. A benchmark date is 1492. Not only was this the year of Christopher Columbus's famous voyage, it was also the year of Antonio de Nebrija's standardising grammar of Spanish. His *Gramática Castellana* was presented to the King of Spain as a way in which "to aggrandize the things of our nation" (Joseph, 2002, p. 3). Needless to say, there was (and still is) a great deal of variation in ways of speaking in different parts of the Iberian peninsula, but standardisation was based on the view that the continuum of geographical and social variation was a manifestation of imperfect approximations to a single "correct" form. Subsequently, the writing and promulgation of grammars formed an integral part of the nation building projects of post-Renaissance Europe, bringing with them the notion of multiple "languages" as distinct entities: hence the term "language making" used by some commentators (Harris, 1980). In short, the concept of *a language* is a cultural product of post-Renaissance Europe. As Haugen (1972) observed in a classic study, this basic insight was even recognised by some during the Renaissance itself. He quotes George Puttenham (1589), who wrote in his book *The Arte of English Poesie*: "After a speach is fully fashioned to the common understanding, and accepted by consent of a whole country and nation, it is called a language" (Haugen, 1972, p. 241). Conversely, it has also been argued that "print-languages" have been one of the foundational elements in the development of national consciousness; cf. Anderson (1991, pp. 43–46).

I raise these points simply to establish the culture-historical specificity of the "*language* concept", rather than with a view to condemning or dismissing it as a myth or a fiction. I agree with Pennycook (1994, p. 117), who is highly critical of the "language concept", that it is important to "historicize our view of linguistics ... [as] the legacy of very particular political and cultural circumstances in Europe"; but on the other hand, the culture-historical specificity and "constructedness" of a particular concept does not necessarily make it useless or insidious, provided it is well understood and handled with care (see Section 5).

2.2. Language$_1$ in lexicographic perspective

The meaning of a word, it is often said, reveals itself in the company it keeps. In linguistic terms, frequent collocations provide clues to semantic content.[2] To begin with, therefore, we will review the most frequent collocations of the meaning *language*$_1$ in a large corpus of the English language, namely, Collins Wordbanks Online: English. Although this exercise is an instructive one, it is also important to bear in mind that corpus techniques have their limitations. Any given corpus has certain local peculiarities; Wordbanks Online is skewed towards the English of the United Kingdom and this is evident in some of the collocation frequency figures. Furthermore, even a rather large corpus (Wordbanks has over 450 million words) does not

[1] In addition to semantic primes, some NSM explications make use of "semantic molecules", i.e. certain relatively simple but non-primitive word-meanings (themselves decomposable into primes) which function as units in the semantic structure of other concepts (Goddard, 2010). No semantic molecules are used in the present study.

[2] In linguistics, this saying is often sourced to Firth (1957). In law, there is a well-known principle of interpretation (or canon of construction) *Noscitur a sociis* ('a word is known by the company it keeps'), normally interpreted to mean that when a word is ambiguous, its meaning may be determined by reference to the rest of the statute.

necessarily contain all expressions for a given word that are subjectively "familiar", let alone all the expressions that are interpretable. Importantly too, semantic argumentation can depend on linguistic facts, such as the possibility of contrasting different meanings of the same word, the existence of different antonyms or derivatives, the non-acceptability of certain combinations, the presence of entailments and implications, that cannot be read off from corpora.

The set of expressions listed in (1) below includes all the relevant examples from the top 50 collocation candidates from Wordbanks Online: English, augmented by a selection of similar examples from the second 50 most common collocation partners. (Function words such as *the*, *a*, and *of* have been excluded from consideration.) All these words have *T*-scores above 5.0, meaning that they are highly correlated with *language* (a score above 2.0 is statistically significant).

The examples have been grouped to reflect their semantic content. The examples in (a) show *language*$_1$ in combination with an adjective derived from a proper name designating a country or region. (Relatedly, the adjective that describes the country, region or (in some cases) the people, can usually stand alone as a noun, i.e. as a language name; for example, *Chinese*, *German*, *Greek*, *Maori*, *Swahili*, and so on.) The examples in (b) are likewise based on regions (places). Those in (c) evoke countries. The various expressions in (d) reflect links between languages and collectivities of people, while those in (e) reflect the fact that it is understood that there is a diversity of languages. Those in (f) reflect the fact that languages can be learnt in adulthood, and our knowledge that languages of the past were different from those of the present day.

(1) a. *English language, French language, Irish language, Arabic language, Welsh language, Chinese language, Spanish language, German language, Maori language, Greek language, Kurdish language, Japanese language, Russian language, Italian language*
b. *local language(s), European languages, African languages, Indian languages, Asian languages, regional language(s)*
c. *foreign language, official language*
d. *own language, their language(s), our language(s), native language(s), first language, indigenous languages*
e. *different language(s), same language, common language, several languages, another language, many languages, two languages*
f. *second language, new language, modern language, ancient language*

The expressions in (2) show *language*$_1$ with the most common following words, according to Wordbanks Online. Once again, the examples all have *T*-scores above 5. Those in (a) are connected with language learning, while those in (b) are connected with communication problems due to language differences. Because these are compounds, they appear in stem form, i.e. they are not subject to plural marking. Nonetheless, paraphrase equivalences make it clear that the relevant meaning in the compound is *language*$_1$; for example, 'language teaching' ≈ 'teaching of languages (or, a language)'; 'language skills' ≈ 'skills with a language (or, with languages)'.

(2) a. *language skills, language learning, language teaching, language class(es), language schools, language training, language courses, language lessons, language teachers, language centres*
b. *language barrier, language difficulties, language problems*

Expressions combining the word *language(s)* with the verbs *speak*, *learn*, *teach*, and *know* are also extremely common. For example, *speak + language* (occurring in the range of three words before *language*) has a *T* score of 38.4. The comparable figures for *learn*, *teach*, and *know* in this range are also very high: 26.5, 11.1 and 10.6, respectively.

In summary, the collocational profile of *language*$_1$ links it with countries and other places where many people live, with collectivities of people, with diversity, with learning and knowing, with saying (speaking) and with communication (roughly, saying what one wants to say effectively, so that others can understand one). All these links are significant in terms of the semantic structure of *language*$_1$.

2.3. Explicating the concept of "a language$_1$"

As we have just seen, the *language* concept originated in association with places and peoples, and linguistic usage continues to evidence these links, both in collocational profile and in the way that language names are systematically related to the names of countries and to the names of people who live in these countries. Explication [A] below is consistent with this connection between the names of languages, on the one hand, and place-names and the names of peoples, on the other. The coordinated references to places and to kinds of people living in these places also account for the links between the concepts of *language* and *culture* – because the "*culture* concept" evokes a similar set of ideas; cf. Goddard (2005).[3]

[3] As explicated in Goddard (2005), the concept of *culture* (in the relevant sense) shares a nearly identical set of (d) components. It too is predicated on the assumption that (roughly speaking) there are different kinds of people living in different parts of the world. In the case of *culture*, the subsequent components refer to people in these different places doing things and thinking about things in their own specific ways. This is obviously congruent, in both form and content, with the corresponding components in explication [A], which concern people in various places saying things using their own specific words. Not surprisingly, the collocations *language(s) and culture(s)* are extremely frequent in English corpora; in Wordbanks Online they are the most frequent expressions of the form *language(s) and X*.

[A] *language*$_1$, e.g. *a language, different languages*
a. something of one kind
b. people can say what this something is with the word *language*
c. people can say something about something with this word when they think like this:
d. ''there are people of many kinds,
these people live in many places, people of one kind in one place, people of another kind in another place
e. when people in one of these places want to say something to other people in this place,
they say it in one way, they say it with words of one kind
f. other people in this place can know what these people want to say, because they know these words
g. when people in another place want to say something, they say it in another way, they say it with words of another kind"

Component (a) of the explication says that *a language* is 'something of one kind'. The expression 'something of one kind' reflects the fact that *language* is a count noun, and, like most other count nouns, represents individual things as instances of a 'kind' (or class). Component (b) is shared by innumerable other common nouns, simply stating the status of the word form in question (i.e. in this case, *language*) as an identifying label (Goddard and Wierzbicka, submitted for publication). Component (c) states that people can say something about something with this word (i.e. use the word, so to speak, as a ''discourse tool'') when they are thinking in terms of a certain mental model.[4] The bulk of the explication spells out the content of the mental model. Component (d) lays the basis in terms of some assumptions about people and places; in particular, the notion that certain 'kinds of people' live in certain places: 'people of one kind in one place, people of another kind in another place'. There is of course essentialism here (cf. Gelman, 2005): in the assumption that there are people of various 'kinds' who belong in particular places (normally, their country or named region). Component (e) conveys the idea that people in these individual places have their own distinctive ways of speaking: that people in a given place express themselves 'in one way', using 'words of one kind'. Component (f) states that this mode of expression can be an effective means of conveying what these people want to say because 'other people in this place ... know these words'. Component (g) reflects the diversity of languages. The pivotal role in the explication of components based on SAY and saying 'with words' is consistent with the high-frequency of collocations involving *speak(ing) a language*, and the component with KNOW is connected with the collocations involving *know(ing) a language*, *learn(ing) a language*, *teach(ing) a language*, and related expressions.[5]

Two aspects of the concept as explicated that could seem problematical are as follows. First, the reference in component (d) to 'people of many kinds' living in different places represents an essentialisation and ''territorialisation'' of social units. Second, although nothing in the explication rules out adaptation and change, they are certainly not provided for explicitly, and the overall phrasing suggests a static uniformity. These aspects are no longer problematical when one recalls that explication [A] is intended to pull apart the portfolio of assumptions which underlie an English folk concept, viz. *language*, showing how it intertwines a particular set of views about relationships between people, places, words, and ways of speaking. It is not intended in any sense as a model of ''reality''. That it cannot be regarded as a faithful picture of reality is clear from the perennial difficulties discussed by linguists under the rubric of the ''language/dialect'' question (e.g. Haugen, 1972; Simpson, 1994).

2.4. "*Language*$_1$" in cross-linguistic perspective

Many languages and cultures of the world go without any word or lexicalised meaning answering precisely to the English *language* concept, as explicated in [A]. Though space prohibits a detailed survey here, it will be useful to look at two examples from very different cultural settings – Central Australia and China.

Yankunytjatjara (Central Australia) has no distinct word for 'language'. The nearest candidate is the polysemous word *wangka*, which can mean 'word(s)', or 'talk, way of speaking', or 'voice' (Goddard, 1996a). Many languages of the world are reported to exhibit an identical or similar polysemy, which is hardly surprising given the importance of words in human communication. Many dictionaries of such languages list the word 'language' as a possible English gloss, along with 'speech', 'talk', and 'way of speaking', but in my view this is most unlikely to be semantically accurate. In the case of Yankunytjatjara, while the word *wangka* can be used to designate what would in English be termed different 'languages' (e.g. *wangka Ingkilitji* 'English'), it can equally well be used to designate any 'way of speaking' – including what linguists would term mutually intelligible dialects of a single language (e.g. *wangka Yankunytjatjara*, *wangka Pitjantjatjara*), speech styles (e.g. *wangka kiti-kiti* 'indirect speech'), and other kinds of registers or ''lects'', e.g. *tjitjiku wangka* 'children's speech'. There is no language-internal evidence for polysemy between these usages, all of which can be adequately (albeit roughly) glossed as 'way of speaking'.[6]

[4] Component (c) is framed in such a way that it does not necessarily imply that the speaker believes in the reality or validity of the model. The model represents a way of thinking about something.

[5] It should be noted that if explication [A] is valid, then the expressions *programming language(s)* and *computer language(s)*, referring to ''languages'' such as Fortran, Perl and C, must be regarded as fixed expressions, in which the word *language* expresses an extended meaning based on *language*$_1$, but not identical to it. This follows because computer languages are not about saying things to 'other people'. This seems to be the correct outcome. From a non-specialist point of view, a *computer language* is not a *language*.

[6] In traditional Western Desert society, *yankunytjatjara* was only one of a set of terms for referring to different regional ways of speaking, by way of identifying a distinctive common word of the variety in question. Many of these terms used the suffix *-tjara* 'having': *yankunytja-tjara,* for example, is based on the word *yankunytja* 'come, go', which contrasts with *pitjantja* 'come' in a nearby variety. Similarly, the pair *mula-tjara* and *maṯu-tjara* are based on alternative words for 'true'. In the original system it is questionable whether such labels were proper nouns. In recent times, however, social and demographic changes have led to a certain degree of standardisation, and the use of the capital letter in contemporary orthography reflects this (cf. Goddard, 2004a).

It is arguable that the "*language* concept" could not arise in traditional Aboriginal societies because one of its ideological underpinnings – the assumption that distinct 'kinds of people' belong in particular regions – was not present. It is true that, as Rumsey (1993, pp. 194–195) puts it, Aboriginal people "all around Australia [say] that they think of the land as divided up into more-or-less clearly bounded regions", and it is also true that each of these regions is associated with one or more labels, such as Yankunytjatjara, Warlpiri, or Wiradjuri, which denote certain distinctive words and ways of speaking. But as Rumsey (1993) also emphasises, the Aboriginal system differs radically from the European way of thinking in that, in Aboriginal thinking, geographical areas and ways of speaking are directly linked, i.e. certain words and ways of speaking inherently belong with certain geographical areas. Consequently, whenever people are in a particular region, they ought to speak using those certain words and ways of speaking – simply and solely because they are in that region. In traditional Aboriginal thinking, there was no concept of a particular 'kind of people' (such as "the Yankunytjatjara", "the Wiradjuri", "the Warlpiri") who occupy the land, on the one hand, and who speak the 'language', on the other.

Aboriginal ideologies of social organisation and land tenure are constructed on completely different principles, which typically do not map onto the idealised words-region distribution in any clear fashion (Merlan, 1981; Harvey, 2002). In legal hearings into Native Title in Australia, the Eurocentric assumption that indigenous labels can be validly used to co-designate entities of three different kinds (a group of people (the claimants), their 'language', and the 'territory' under claim) has often proved deeply problematical.

Shifting to a very different cultural context, in Modern Standard Chinese also there is no word which precisely matches English *language*.[7] The closest colloquial candidate is *huà* 'talk, speech' (cf. *shuōhuà* [lit. say *huà*] 'to talk'); but again this word is closer to 'talk, speech, way of speaking' than to 'language', and it has a correspondingly wider range of use. In particular, it is used to talk about the various regional ways of speaking across China. For example, one can ask: *Nǐ shuō shá/shěnme dìfāng de huà?* [you say what place NOM speech] or *Nǐ shuō de shì shá/shěnme dìfāng de huà?* [you say NOM be what speech]; and such questions can be answered with expressions such as *Guangdong huà* 'Guangdong speech, i.e. Cantonese', *Shanghai huà* 'Shanghainese', *Beijing huà* 'Beijing dialect', *Suzhou huà* 'Suzhou dialect'. As is well-known, many of these regional ways of speaking are mutually unintelligible, while others are mutually intelligible. To inquire of a foreigner 'What language do you speak?', one could ask: *Nǐ shuō de shì nǎguó huà?* [you say NOM what.country speech], but the emphasis is very much on the spoken word. Possible answers have the same form, e.g. *Hánguó huà* [Korea speech] or *Rìběn huà* [Japan speech].

Perhaps equally unexpected from an English-speaker's point of view is the fact that there are several different Chinese terms for "Chinese". Using *huà*, one can speak of *Zhōngguó huà* 'China speech' or *Putonghuà* 'the common tongue', referring to the standardised national lingua franca of China. There is also the term *hanyu* (roughly) 'Han language', which is becoming a standard term used in the growing trade of teaching Chinese as a foreign language. The first element in this expression (*Han*) is an ethnic national term, referring to native Chinese, while its second element (*yu* or *yuyan*) is a technical term not often used in ordinary speech. (Nowadays, influenced by written Chinese, one could ask a question formed on the bound morpheme *-yu*, e.g. *nǎguó yu* [what.country yu] and receive an answer like *Déyu* 'German'.)

Finally and importantly, there is the word *Zhōngwén*, which refers to the written traditional language of China – the language of five thousand years of Chinese culture and civilisation. In this expression, *Zhōng* (lit. middle) may be understood to designate China, and *wén* to designate something like "text/written language" (Ye 2007, p. 158), with richer associations of culture and history than conveyed by the English word *language*. It is possible to apply the word *wén* to languages other than Chinese, e.g. *Yīngwén* 'English', *Déwén* 'German', provided that such associations of literacy and high culture are satisfied (Ye, 2007).

In the case of China, it can be argued that the preconditions for the European "*language* concept" do not exist, but for very different reasons than for Aboriginal Australia. Chinese ways of thinking about "Chinese language" are predicated on the historical and geographical unity of China. The different ways of speaking in different places in China are all viewed as different manifestations of a single unique entity (*Zhōngwén*) constituted by Chinese words-in-writing, i.e. by Chinese characters, which has endured over many centuries.[8] Mutual intelligibility, regional distinctiveness, the existence of different kinds of people in different places: ideas like these do not enter into the cultural–conceptual configuration.

3. Other meanings of *language*

Aside from *language*$_1$, the word *language* has at least four other productive polysemic meanings, as well as figuring in a number of fixed expressions with somewhat idiosyncratic meanings.[9] For linguists, the most salient of the productive meanings is no doubt that found in expressions such as *the origins of language* and *the language faculty*. I will deal with this next, labelling it *language*$_2$. It should be noted, however, that in terms of frequency, another meaning (roughly, 'word usage', e.g. *bad language*, *Shakespeare's language*), is much more significant. I will designate this as *language*$_3$ and deal with it in the subsequent section. A fourth meaning *language*$_4$ (roughly, 'a specialised way of speaking', e.g. *the language of science*) is dealt with next, followed by a fifth meaning: *language*$_5$ (roughly, 'an expressive medium', e.g. *the language of music*). From a grammatical point of view, these four additional meanings are all clearly distinct from *language*$_1$ because they are "mass" (non-count) nouns.

[7] I thank Zhengdao Ye for providing the information in the following paragraphs.

[8] A term like 'Chinese script' is not really an appropriate way to represent indigenous Chinese ideas, because the term 'script' presupposes the primacy of speech.

[9] One such expression is *speaking different languages*, in the sense of (roughly) 'being unable to understand each other', and its converse, *speaking the same language*. Needless to say, these expressions are closely related in meaning to *language*$_1$.

(I will not deal with *language* as it is used in educational parlance, where it means roughly, knowledge about and skills with the English language, e.g. *language and literature*.)

Since the main focus of this section is lexical-semantic analysis of polysemy, a few words on the NSM approach to polysemy are in order. The NSM approach upholds the traditional "definitional" approach (Geeraerts, 1994; Wierzbicka, 1996; Goddard, 2000). In general, one assumes to begin with that any given word or expression has but a single meaning, and attempts to state it in a clear and predictive fashion, in the form of a translatable reductive paraphrase (explication). Polysemy is only posited when it appears to be impossible to accommodate the full range of use under a single explication. The next hypothesis is that there are two distinct meanings, and attempts are made to state both in a clear and predictive fashion, and so the process goes, until the full range of application of the word can be captured within the specified range of senses. The paraphrase analysis procedure allows us not only to detect polysemy, but also to understand it, because it enables us to "see" and to compare the relevant meanings in detail. Needless to say, this procedure can only be implemented if the principles of good definition are followed. The NSM requirements that meanings be stated as reductive paraphrases, and that their validity be testable by substitution, provide clear guidelines to limit the ad hoc postulation of polysemy. Usually these procedures converge on the same results that would be expected on the basis of standard lexicological tests for detecting polysemy (cf. Cruse, 1986). These include the existence of different syntactic properties attaching to the different senses, the possibility of directly contrasting the two meanings without contradiction, the existence of different antonyms for the putatively different senses, the existence of derived or compounded forms accepting only one of the senses, different entailments and implications, and so. To these traditional tests, I would add the existence of different "collocational profiles" or collocational preferences for the different senses.[10]

In the following treatment I will draw on these lexicographic tests, as appropriate, to help establish the distinctness of the several different senses of *language*. The primary argument for recognising polysemy, however, remains the need for distinct explications to accommodate different uses of the word in question.

3.1. *Language*$_2$: roughly, 'language in general'

The meaning *language*$_2$ can be seen as a more abstract or more generalised counterpart of *language*$_1$. Though the concept is, in a sense, "semi-scientific", it is nevertheless an ordinary meaning embedded in the lexicon of the English language, employed in expressions such as *human language* and *the origins of language*. See also the examples in (3).

(3) a. *Language is the single most defining characteristic of our species.*
b. *Linguistics is the study of language.*
c. *It is the interdependence of hearing loss and language that is important.*

This meaning is not particularly common in Wordbanks Online, but it is clearly different in meaning and in grammar to *language*$_1$. In terms of collocational preferences, it is notable, firstly, that *language*$_2$ generally appears without any modifier; and, secondly, that the range of modifiers it does take is highly restricted: essentially, either the word *human* (i.e. in the expression *human language*) or adjectives describing the modality by which words are expressed; as in expressions such as *spoken language*, *oral language*, and *written language*.

Language$_2$ can be explicated as in [B].

[B] *language*$_2$ (e.g. *the origin of language*):
a. something
b. people can say what this something is with the word *language*
c. people can say something about something with this word when they think like this:
d. "when people want to say something to other people, they can say it with words
e. other people can know what these people want to say, because they know these words"

Notice that the meaning structure represented in [B] does not involve any 'kinds' of people or any links between such kinds and the places where they live; rather, by abstracting away from these details, it comes out as a statement about 'people' in general. As such, it lends itself to comparisons with "non-people", in particular, with animals; hence, the existence of examples such as *Do dolphins have language?* At first blush one might think that explication [B] would predict that such a sentence is contradictory, since it would seem that the scenario depicted in components (d) and (e) cannot apply to dolphins. However, it is a characteristic of the "abstract noun" semantic structure of an explication like [B] that it lends itself to extended uses (Goddard and Wierzbicka, submitted for publication). In the case of this sentence, the overall interpretation is roughly as follows: "Is it the case that dolphins are like people, inasmuch as when people want to say something to other people, they can say it with words (etc.)".

[10] Though the concept of a "collocational profile" (also termed collocation profile or lexical profile) is well known in corpus linguistics, it is usually applied to a lexeme as whole, rather than to an individual sense (or lexical unit) of a given lexeme (cf. Stubbs, 2001, pp 85–96), and consequently there is no effort made to correlate the collocational preferences with the specifics of individual senses. The approach adopted in the present study, which uses collocational profiles as evidence for sense differentiation, was developed initially by Wierzbicka (2009), though see Teubert and Čermáková (2004, pp. 151–156) for a partial precedent.

Many languages do not have any lexicalised meaning which even approximately matches that of English *language*$_2$, while in others which do have such a meaning it is expressed by a different word to that which matches *language*$_1$. Spanish, for example, distinguishes lexically between *lengua* (roughly) 'language$_1$' (e.g. *la lengua española* 'the Spanish language', *lengua materna* 'mother tongue') and *lenguaje* (roughly) 'language$_2$' (e.g. *lenguaje humano* 'human language'). The Spanish version of an English sentence such as 'Linguists study languages$_1$ to better understand the nature of language$_2$' uses both: *Los lingüistas estudian las lenguas para entender mejor la naturaleza del lenguaje*.[11]

Two fixed expressions closely related to *language*$_2$ are *body languag*e and *sign language* (in its naïve "mass noun" sense, referring to the use of improvised gestures to convey a message). *Body language* is not language at all, in any literal sense, both because it does not involve words and because it designates an unconscious expression of attitudes and feelings, rather than deliberate communication, as implied by *language*$_2$. *Sign language* (as in an example like *Research is showing that if you use sign language with babies, they learn to speak earlier and are more verbal*) is intentional, but it is distinct from *language*$_2$ because it does not involve words. (Needless to say, this statement does not apply to the linguistic concept of *signed languages*, referring to languages such as British Sign Language, Auslan, etc. This is clearly an instance of *language*$_1$, i.e. *language* in its count noun sense.)

3.2. *Language*$_3$: roughly, 'word usage'

Like *language*$_2$, this meaning shows a notable parallelism in structure to *language*$_1$. Essentially, *language*$_3$ refers to how someone expresses something in a distinctive way through their words and "wording". It is very common. The examples in (4) include all the *language*$_3$ collocations from the first 100 collocation candidates listed in Wordbanks Online, augmented by a selection of phrases of the relevant types from other sources. Aside from combinations with adjectives and other prenominal modifiers, such as *bad/foul language* and *Obama's language*, the meaning *language*$_3$ can also occur with a modifying prepositional phrase, e.g. *the language of the resolution*, *the language of hate*.

From a semantic point of view, the modifiers (whether adjectival, phrasal, or, indeed, clausal) fall into four groups, exemplified in (4a)–(4d). They can characterise the kind of words and wording used, as in (4a); identify a certain person or quasi-person as the originator of the words and wording, as in (4b); identify a document, as in (4c); or they can characterise a motive or illocutionary function, as in (4d) (e.g. *the language of hate* means, roughly speaking, 'how someone expresses themselves in words when they want to incite hate'). In terms of collocational profile, these modifiers are all notably different to those found with *language*$_1$.

(4) a. *bad language, foul language, abusive language, strong language, plain language, sexist language, simple language, colourful language, everyday language, obscene language, offensive language, vulgar language, racist language, technical language, poetic language, accessible language, language everyone can understand*

b. *Shakespeare's language, Obama's language, the language of Washington*

c. *the language of the resolution, the language of the novel, the language of the text, the language of the legislation*

d. *the language of hate, the language of war, the language of compromise, the language of persuasion*[12]

Some sentence examples follow. Notice that, as shown in (5a) and (5b) examples, *language*$_3$ can (and frequently does) occur outside of any of the common modifier collocations listed above. So far as verbs are concerned, *language*$_3$ tends to correlate with verbs like *speak*, *express* and *use*. (The strong association between the verb *use* and *language*$_3$ provides an interesting point of contrast with *language*$_4$ (roughly) 'a specialised way of speaking', to be dealt with the next section.)

(5) a. *Perhaps this explains the ultra-cautious language being used by the WHO this time around.*
b. *There's no other way to describe the language used by Secretary Powell other than to say that it was misleading.*
c. *Speaking the language of war will not bring a solution.*

Quite often *language*$_3$, without any modifier, is used in a sense close to 'bad language'; for example, in sentences like *Don't use that language to me*, and in the fixed expression *Mind your language*.

Explication [C] is based around the idea – given in component (d) – that someone can go about saying something in a particular way, with certain words as opposed to others, because they want to do so.

[C] *language*$_3$ (e.g. *bad language, Shakespeare's language, the language of the resolution*):
a. something
b. people can say what this something is with the word *language*
c. people can say something about something with this word when they think like this:
d. "when someone wants to say something to other people,
this someone can say it with some words, not with other words,
because this someone wants to say it in this way"

[11] I do not want to imply that Spanish *lengua* means exactly the same as English *language*$_1$, if only because Spanish has a second (more informal) word *idioma* for (roughly) 'spoken language, way of speaking'; *¿Qué idioma se habla en australia?* 'What language do they speak in Australia?'.

[12] Similar in form, but in reality a fixed expression is *the language of love* (as in *Music is the language of love*); cf. *the language of the heart*.

The collocational patterns shown in (4a)–(4d) above are consistent with explication [C] because they provide further specification or amplification of the basic "way-of-saying-things, using-certain-words" meaning articulated in explication [C].

3.3. Language$_4$: roughly, 'a specialised way of speaking'

The fourth meaning to be dealt with here designates a specialised vocabulary and way of speaking, something like the linguistic concept of a register, whose use is (or is considered to be) confined to discourse in a particular domain. In contradistinction to the *language$_3$* meaning, it is not concerned with one's choice of words (a position that assumes the possibility of alternative forms of expression). The idea is rather that discourse on a certain topic or in a certain domain is carried out in a more-or-less stable and distinctive fashion, which must be known in order to be understood.

In this meaning, it possible to speak of particular words as belonging to or being part of a *language$_4$*, as in the naturally-occurring examples in (6).

(6) a. *Numbers are part of the language of science.*
b. *[S]in, prayer, and contrition are part of the language of Christian religion.*
c. *Togo's new generation of protesters have started to dismiss the language of democratic struggle – people-power, grassroots solidarity and all that.*

Language$_4$ also occurs with prepositional phrases, such as those shown below. The nouns in these phrases identify a domain, either an occupational domain, as in (7a), or a social domain, as in (7b).

(7) a. *the language of diplomacy, the language of politics, the language of the law.*
b. *the language of the playground, the language of the locker-room.*

In either case, the *language$_4$* meaning carries an implication of something like exclusivity. There is the sense that unless one knows the specialised way of speaking, one would not be able to understand or participate in the relevant kind of discourse. Following from this, it is notable that (unlike as with *language$_3$*), it makes sense in relation to *language$_4$* to speak, in a literal way, about learning and knowing such *language$_4$*; for example, to speak about learning *the language of science* or learning *the language of diplomacy*.

In many cases, expressions with *language$_4$* can occur in two parallel forms: with a prepositional phrase, as in the examples above, or with a corresponding adjective that identifies a domain of discourse, e.g. *scientific language*, *religious language*, *mathematical language*, *and legal language*, *diplomatic language*. In terms of collocational profile, it will be apparent that these modifiers are quite different in character from those of *language$_3$*. Furthermore, some modifier types which are very common with *language$_3$*, such as evaluative and modal-like adjectives (e.g. *bad language$_3$*, *difficult language$_3$*), are seldom if ever used with *language$_4$*. In terms of grammatical properties, *language$_4$* is clearly distinct from *language$_3$* in that it seldom if ever appears independently, i.e. without a modifier (either propositional or adjectival).

For *language$_4$*, I would propose the explication in [D]. The (d) component indicates the existence of a particular way of speaking, including 'words of one kind' that are commonly used when speaking of 'things of one kind'. The component in (e) conveys the sense of exclusivity that attaches to this sense of the word *language*.

[D] *language$_4$* (e.g. *the language of science*)
a. something
b. people can say what this something is with the word *language*
c. people can say something about something with this word when they think like this:
d. ''when someone wants to say something about things of one kind to other people,
at many times this someone says it in one way, this someone says it with words of one kind
e. other people can't know what this someone wants to say if these other people don't know these words"

It is important at this point to note that many phrases involving *language* and a prepositional phrase with *of* are ambiguous. For example, a phrase such as *the language of photography* can have the meaning just discussed, i.e. a special lexicon and way of speaking about photography, or it can refer to photography as an "expressive medium", as a way of expressing things. The latter represents a fifth meaning of *language*, to which we will now turn.

3.4. Language$_5$: roughly, 'an expressive medium'

I will designate as *language$_5$* the meaning exemplified in expressions such as: *the language of photography, the language of architecture, the language of clothes, the language of music*. Some sentence examples of *language$_5$* are the following:

(8) a. *The language of art consists of things like color, line, shape, space, value, and texture.*
b. *Use the "Language of Flowers" to create your own "secret" message for the one you love.*
c. *The Language of the Fan. In the past, hand fans were used not only as cooling instruments, but also as convenient communication devices, mainly for transmitting more or less furtive love messages.*[13]

[13] Example (8a) comes from the website of the Tweed Museum of Art: www.duluth.umn.edu/tma/collections/language/index.html. Example (8c) is the heading and first sentence from an online article: www.ideco.com/fans2/language.htm.

Especially in intellectual commentaries, art critiques, and the like, *language*$_5$ can appear with certain adjectival modifiers, as in (9). The adjectives indicate a domain of expressive activity or practice.

(9) *musical language, architectural language, visual language*

Like *language*$_4$ in this respect, *language*$_5$ seldom if ever appears independently, i.e. without a modifying *of*-phrase or an adjective.

I propose the explication in [E] for *language*$_5$. Essentially, it captures the notion that doing certain things with flowers (music, art, etc.) can allow someone to express themselves to other people, in a way analogous to how they can say things to other people with words. That is, people can use flowers, (music, art, etc.) as or like a language. It would be fair to call this meaning "figurative" inasmuch as component (e) contains the element LIKE, drawing an explicit comparison with how people can say things with words.

[E] *language*$_5$ *(of things of this kind) (flowers, music, art,* etc.):
a. something
b. people can say what this something is with the word *language*
c. people can say something about something with this word when they think like this:
d. ''when people do something with things of this kind (flowers, music, art, etc.),
if they want, they can say something to other people in this way
e. like people can say something to other people with words"

3.5. Review and discussion

I have proposed a set of five inter-related explications for different senses of the word *language*. These five posited meanings (together with a handful of fixed expressions, as noted in the footnotes) appear to cover the great majority of the attested uses of *language* in Wordbanks Online.[14] Of course I do not say this on the basis of having examined all the attestations – over 40,000 for the singular form alone. I have, however, examined several random samples of 250 attestations each (generated by the Wordbanks software) and in each case there were only a handful of examples that could not be allocated to one of the meanings or fixed expressions dealt with so far.

Some readers without an interest in systematic linguistic semantics may have found the simple vocabulary and unusual phrasing of the explications [A]–[E] hard-going, so it may be worthwhile to recapitulate some of the merits of the reductive paraphrase approach. Because the explications are phrased in a controlled, highly restricted vocabulary, they avoid definitional circularity and obscurity. By the same token, they enable a maximally explicit and detailed conceptual analysis. They are free of flaws such as the use of disjunction ('or') and other non-explicit "lexicographic crutches" (such as the use of 'etc.' and 'esp.') that mar so many conventional dictionary definitions. On account of their fine-grained quality, the explications bring into clear view the semantic affiliations between the various polysemic meanings, to a much greater extent than would be possible with conventional dictionary definitions. Likewise, the details of individual explications provide an explanatory rationale for the varying collocational preferences of each sense. Finally, although the explications are presented in the English version of NSM, the vocabulary and phrasing is not tied to English in any essential way; i.e. the explications can be transposed without loss or distortion of meaning into the corresponding NSMs of other languages. To illustrate this point, Appendix B contains Russian and Chinese versions of explication [A].

4. WORDS as a universal semantic foundation for 'language' and related concepts

A key element in explications [A]–[E] is the semantic prime WORDS. It was first advanced as a semantic prime by Wierzbicka (1996, pp. 107–108), with further explanation in Wierzbicka (2002, pp. 99–101). The following discussion builds on and amends these contributions in certain respects.[15] After a brief introduction, I will structure the discussion around Dixon and Aikhenvald's (2002) arguments against WORDS as a universally lexicalised meaning.

4.1. The status of WORDS in the NSM metalanguage

The basic idea is that while one can talk about people saying things without any reference to WORDS, there are some contexts in which a reference to WORDS appears to be essential to the intended meaning; for example, the contrast between (10a) and (10b) below turns precisely on whether or not words are specifically at issue. Further, it would appear to be conceptually necessary for various purposes to be able to distinguish between the content of what is said, and the form – i.e. the words – in which it is expressed. The notion of paraphrase (fundamental to semantic analysis) and the notion of translation both turn on the idea that one can say the same thing in or with other words. NSM theory postulates that one can express the equivalent of an English sentence like (11) in all languages.

[14] For interest, the sense differentiation proposed here corresponds rather closely to that of Collins Cobuild dictionary. The main difference is that Cobuild does not recognise the "figurative" use designated as *language*$_5$; but on the other hand it recognises an additional sense in relation to 'computer languages' and the like.

[15] Wierzbicka (2002) maintained that an essential property of the semantic prime WORDS was that it could refer not only to "individual words" but also to an utterance as a whole, seen from the point of view of how it is expressed. This now seems to be incorrect. While the exponent of WORDS in many languages does indeed have an extended meaning corresponding to 'what (someone) said at one time with words', e.g. *His words moved me*, *God's word*, etc., such usages are better analysed as polysemic extensions, rather than as a use of WORDS in its semantically primitive sense.

Table 1
Dictionary definitions of word.

Oxford English Dictionary (1989)	a combination of vocal sounds, or one such sound, used in a language to express an idea … and constituting an ultimate minimal element of speech having a meaning as such.
Collins Cobuild [accessed online at http://dictionary.reverso.net/english-cobuild/word, 04/01/2010]	A **word** is a single unit of language that can be represented in writing or speech. In English, a word has a space on either side of it when it is written.

(10) a. *He/she said one word.*
b. *He/she said one thing.*

(11) *He/she said the same thing in/with other words.*

Diverse important speech categories and speech practices also crucially involve the notion of WORDS, including names and naming (words for identifying persons, places, etc.; cf. Goddard and Wierzbicka (2008)), counting (words for keeping track of 'how many' things there are somewhere; cf. Goddard (2009a)), speech formulae, e.g. *Goodbye* and *How do you do?* (cf. Ameka (2006)), honorific words (cf. Yoon (2004)) and magical formulae (roughly, saying certain words to cause certain things to happen). The notion of figurative speech (metaphor, etc.) also implies recognising a potential contrast between what a speaker says with certain words and what the words say (Goddard, 2004b).

Not only is the notion of WORDS indispensable for satisfactory explications in many domains, of equal importance is its resistance to non-circular definition. To illustrate the point, consider the dictionary definitions presented in Table 1.

Both definitions have a distinctly scholarly and counter-intuitive tone, and neither could be substituted into ordinary sentences that use the word *word*. Specifically, the OED's fails both because the expression 'ultimate minimal unit' is obscure,[16] and more importantly, because the word 'speech' itself implies 'words', thus making the definition implicitly circular. The same faults vitiate the Collins Cobuild definition: the expression 'single unit' is obscure, and using the word 'language' in the definition is circular, because (as we have seen) the meaning of 'language' itself depends on 'words'. Furthermore, on account of their complex and language-specific wording, neither definition could be transposed freely across languages without distortion. On the NSM view, it is literally impossible to reductively paraphrase the meaning of 'words', as used in sentences like those in (10a) and (11).

Having said this, it is crucial to add that exponents of WORDS are frequently polysemous (as are the exponents of many other semantic primes). Even in English, the word *word* has extended meanings that go beyond the semantically primitive meaning; for example, in fixed and semi-fixed expressions such as: *a word of warning*, *a kind word* (or *kind words*), *to have a word with someone*, *to have the last word*, *to put in a good word*, *to get/bring word of something*, and *to give one's word*. All these expressions refer to someone saying something (usually something brief) about something: they are not examples of WORDS in its semantically primitive sense. In Biblical language (i.e. in the characteristic phraseology of the English Bible) and theological language generally, the word *words* is used even more freely to mean, in essence, what someone has said; for example, in phrases such as *to hear (speak, etc.) someone's words*. The examples in (12) come from New International Version of the Bible, while those in (13) are from the King James Bible.

(12) a. *Adah and Zillah, listen to me; wives of Lamech, hear my words:* … (Genesis 4:23)
b. *When Esau heard his father's words, he burst out with a loud and bitter cry* … (Genesis 42:16)
c. *When Moses went and told the people all of the Lord's words and laws, they responded with one voice, "Everything the Lord has said we will do*" (Exodus 20:1)

(13) a. … *So likewise ye, unless you utter by the tongue words easy to be understood, how shall it be known what is spoken?* (Corinthians 14:9)
b. … *the words that I speak unto you, they are spirit, they are life* (John 6:63)

Presumably these and similar usages (such as the expression *the word of God* (or *God's word*) to designate the Bible, in whole or in part) originate in the Greek and Hebrew versions of the New and Old Testaments that pre-date the English translation. They are highly naturalised for those well familiar with the Christian Bible and must be counted as part of the English language, even though they would hardly be acceptable in everyday conversational English (as one can see from the oddity of comparable expressions such as *?I listened to George's words*, *?John remembered Max's words*, and the like).[17] But to repeat: these are not instances of the universal semantic prime WORDS, they are extended polysemic meanings.

Failure to adequately take account of polysemy has led some linguists to deny the universality of WORDS as a lexicalised meaning. Dixon and Aikhenvald (2002, p. 2) reject Wierzbicka's (1996) claim that WORD is a universal semantic prime on the basis of the fact that Russian *slovo*, Old English *word*, and French *mot* have extended ranges of application, beyond that

[16] Probably the intuition behind expressions like 'minimal unit of (speech)' and 'single unit of (language)' is connected with the ideas of 'one thing' and 'part'. On this interpretation, the dictionaries are saying that a 'word' is 'one thing' that is 'part of speech' (or 'part of language').
[17] In some contexts, especially when auditory impression or precise semantic content is involved, such sentences do occur, e.g. *Her words echoed in his ears*.

of *word* in contemporary English. But the fact that these words have extended ranges of application by no means disqualifies them as exponents of the proposed NSM prime (see Appendix B for a Russian explication using *slovo* WORD).

Dixon and Aikhenvald (2002) raise two additional arguments against the universality of WORD, one based on claims about the languages of small non-literate ("tribal") cultures and the other based on facts about Chinese. I will consider these in turn.

4.2. WORDS as a lexicalised meaning in the languages of small non-literate cultures

There is no doubt that attitudes towards words vary greatly across languages and cultures, depending above all on whether or not a given culture has the institution of literacy, and if so, how widely it is distributed across the speech community, what particular forms it takes, how it is associated with other cultural traditions, and so on (Ong, 2002). All these things will influence people's attitudes towards and beliefs about words. Acknowledging this, however, is a far cry from accepting that such languages lack the concept of words altogether, even to the extent that they lack any lexicalised meaning corresponding to the proposed semantic prime WORDS.

This, however, is precisely Dixon and Aikhenvald's (2002, p. 3) claim: "The vast majority of languages spoken by small tribal groups (from a few hundred to a few thousand speakers) have a lexeme meaning '(proper) name' but *none* have the meaning 'word'" [emphasis added]. They state that this claim can be substantiated in respect of "many languages of Australia, ... Amazonia ... and New Guinea with which we, or our colleagues, are familiar", though they mention by name only two languages, the Australian language Arrernte and the Amazonian language Jarawara. On the face of it, their claim is manifestly incorrect. In relation to Australian languages, I have consulted a number of dictionaries and linguists, and have yet to find a single Australian language (including Arrernte[18]) which does not have a plausible candidate exponent for the NSM prime WORDS. The only complicating factor is that (as in English) the candidate exponents typically have additional meanings as well. For example, they may express meanings that are glossed by the dictionary-makers as 'way of speaking', 'language', 'talk', 'message', 'story' or 'voice'.

The Yankunytjatjara language of Central Australia is a case in point, and I will use it to illustrate the issues involved. The Yankunytjatjara noun *wangka* has at least three meanings: (i) 'word', (ii) 'talk, way of speaking' (a general meaning that subsumes glosses such as 'language', 'dialect', 'speech style'; cf. Section 2.4), and (iii) 'voice'. How then can one establish that 'word' is a discrete meaning? One way is to refer to contexts where only this interpretation would make sense; for example (14) and (15) below.

(14) *Paluṟu wangka kutju wangka-ngu.*[19]
3sg.NOM word one say-PAST
'He/she said one word (≠ he said one thing)'.

(15) *Piṟan-tu wangka-pai 'soakage', ka ngana-mpa wangka 'raalpa'.*
white-ERG say-CHAR 'soakage' CONTR 1PL-POSS word 'raalpa'.
'White people say 'soakage', but our word (is) 'raalpa'.' (Goddard, 1996a, p. 216)

In both the examples above, *wangka* is used to designate an individual spoken word. *Wangka* can also be used with the pluralising word *tjuṯa*. As one would expect, *wangka tjuṯa* is equivalent to English *words*. An expression like *iṯi-ku wangka tjuṯa* [baby-POSS word PL], for example, corresponds to "baby words", i.e. distinctive words that are used by young children (analogous to English words like *mummy* and *choo-choo).* As expected, the expression *wangka tjuṯa* can be used to talk about saying things 'in other words', as in (16).

(16) *Paluṟu palupuṟunypa wangka-ngu palu wangka kutjupa tjuṯa-tjara-ngku.*
3sg.ERG the.same say-PAST but words other PL-having-ERG
'He/she said the same (thing), but in other words'.

On NSM assumptions, examples like these establish 'words' as a discrete meaning for *wangka*,[20] because it would appear to be impossible to bring them under a single reductive paraphrase that would also embrace *wangka* in its alternative meanings of 'way of speaking' and 'voice'. Furthermore, Yankunytjatjara speakers perceive a clear contrast between the available interpretations of potentially ambiguous expressions. The expression *wangka kutjara* [wangka two], for example, can mean 'two words', or 'two ways of speaking', or 'two voices'.

Table 2 below gives a sample of dictionary entries for potential exponents of WORDS in a number of other Australian languages. In all cases, the authors of the dictionaries include 'word' as one of the possible glosses of the indigenous language term, in addition to other meanings or glosses, as indicated.

[18] Dixon and Aikhenvald (2002, p. 3) state without qualification that Arrernte does not have a lexeme for 'word', and they refer the reader to Henderson's (2002) chapter in the same volume. What Henderson (2002, p. 100) actually says is that "there is no Arrernte term which *exclusively* picks out a word-level unit" (emphasis added).

[19] Note that the form *wankga-* also exists as a verb-stem, with the meaning SAY.

[20] In some contexts, especially with written words, the words *ini* 'name' and *tjukurpa* 'story, meaning' can be used to refer to words; for example: *Ini/tjukurpa nyangatja wangka ka-na kulila* [name/story this say and-I listen] 'Say this word for me to listen to'. However, there are decisive contexts in which all three words (i.e. *wangka*, *ini*, *tjukurpa*) can be used in and these contexts, it is clearly *wangka* that conveys the meaning 'word'. Compare example (14), for instance, with: *Paluru tjukurpa kutju wangka-ngu* [3sgnom "story" one say-past] 'He/she said one thing'.

Table 2
Candidates for exponents of WORDS in a selection of Australian languages.

Language and source	Candidate exponent	Meanings or glosses given in the source dictionary (bolding added)
Arrernte (Henderson and Dobson (1994): 134)	*angkentye*	1. language 2. a general term for languages 3. an Aboriginal language, especially Arrernte 4. the kind of words someone uses; way of talking, language 5a. message, news, story 6. a **word**, phrase or other piece of language 7. someone who talks in a particular way, a speaker of a language 8. voice. 9. sound or noise that something usually makes
Bundjalung (Sharpe 1992, p. 225)	*ŋuyay*	**word**, language, speech
Kayardild (Evans 1992, p. 68)	*kangka*	1. **word**, talk 2. language 3. voice 4. characteristic sound
Ngan'gityemerri Reid and McTaggart, 2009, p. 96	*ngan'gi*	1. language, a particular variety of speech associated with some country 2. **word**, story, message, news 3. Mass. Any church related meeting or service
Warlpiri (Laughren, Hale et al., 2006)	*yimi*	speech, talk, utterance, verbal **word**, sentence, text; language; story, tale, account, narrative, information, news
Wik-Mungkan (Kilham et al., 1986, p. 252)	*wik*	**word**, talk, speech, language, bird or animal call, noise, meeting to make decision

Comparable data is more difficult to access for the languages of Papua New Guinea, but in three languages from which such information is available, good potential exponents for WORDS appear to be as follows: Yimas *pia-/-mpwi* (Foley p.c.), Ku Waru *ung/ing* (Merlan and Rumsey, 1991, p. 347), Mangaaba-Mbula *sua* (Bugenhagen, 2002, p. 18). The following selection of examples can be added from languages from North America: Mohawk (Iroquian) *weńn*, Ojibwa (Algonquin) *kidwin* (Rhodes, 1985, p. 207), Sm'algyax (Salish) *algya̲x* (*Sm'algya̲x Learners' Dictionary*, 1999), East Cree (Algonquin) *ayimuwin* (Junker, 2008).[21]

Given that all these languages are spoken by "small tribal groups (of a few hundred to a few thousand speakers)", these examples are sufficient to refute Dixon and Aihkenvald's (2002) unqualified assertion that all "small tribal languages" lack any word for 'words'. More detailed lexicographic evidence would be required before one could conclude that the apparent exponents can be used in all the ways predicted for NSM semantic prime WORDS, but the overall prospectus seems positive.

4.3. WORDS in Chinese

Dixon and Aikhenvald (2002) adduce one further example of a language in which the status of WORDS is problematical. That language is Chinese, and the situation in Chinese is indeed extremely interesting and instructive. The normal, common Chinese word for WORDS is *zì*, but on account of the tremendous salience of Chinese character writing, *zì* is strongly associated with written words. Chinese people tend to think of "oral words" as pronunciations of *zì*, in the sense of written characters. There is another Chinese word for 'words', namely *cí*, but this is a technical term, something like 'word-form' in the linguistic sense. It is *zì* which is overwhelmingly dominant in Chinese cultural ideas about "language". As the eminent Chinese linguist and sociolinguist Chao (1968) put it in a classic passage, *zì* is the "sociological unit" of the language:

> ... that type of unit, intermediate in size between a phoneme and a sentence, which the general, nonlinguistic public is conscious of, talks about, has an everyday term for, and is practically concerned with in various ways ... Thus it has all the social features of the common small change of every day speech which one would call a 'word' in English. (Chao, 1968, p. 136; cited Dixon and Aikhenvald, 2002)

Does Chinese, then, provide an example of a language in which there is no exponent of WORDS in the general sense, unrelated to writing? This question has been considered in two NSM treatments, Chappell (2002, pp. 276–278) and Ye (2007, pp. 155–159). Both observe that despite the psychological and cultural prominence of the written (graphemic) word, *zì* can – and is – also used freely to refer to spoken words, as in the following examples:

(17)	*Yíge*	*zì*	*yíge*	*zì*	*de*	*mànmànde*	*shuō.*
	one:CL	word	one:CL	word	LIG	slowly.RDP.ADV	say
	'Please speak slowly, one word after another.'						
(18)	*Tā*	*xiànzà*	*néng*	*shuō*	*dānge*	*de*	*zì*
	3SG	now	can	speak	single:CL	LIG	words
	'She can now say single words'. (about the speaker's one-year old daughter).						

As Ye (2007, p. 157) notes, there are plenty of fixed expressions which point to the "spoken *zì*"; e.g. *tǔzì* [utter *zì*] 'pronounce words correctly or in the traditional way; articulate, pronounce', and set phrases such as *tǔzìqīngchǔ* [utter-*zì*-clearly] 'enunciate clearly for describing clear pronunciation' and *zìzhèngqiānyuán* [*zì*-standard-tone-round] 'sing/speak with clear and rich

[21] A number of these languages, including Ngan'gityemerri (Australia), Yimas (PNG) and the North American languages, are strongly polysynthetic. As noted in a previous study (Goddard, 2001), there seems to be no particular difficulty in locating potential exponents of words in polysynthetic languages, despite the tremendous difference in the kinds of words that can be found in them.

tone' for describing a theatrical verbal experience. A Chinese Yahoo search [www.yahoo.com.cn] conducted by Ye brought up more than ten million examples of "saying *zì*". Chappell (2002, p. 276) also notes that NSM expressions like 'saying the same thing in different words' present no problem in Chinese. In view of data such as these, Chappell (2002) and Ye (2007) concur in concluding that *zì* is the exponent of the NSM prime WORDS in Chinese.[22] To demonstrate this point, Zhengdao Ye has kindly provided a Chinese version of explication [A] for *language*$_1$. It is presented in Appendix B.

It should be sufficiently clear at this point that the status of the concept of WORDS is starkly different to that of *language*. The "*language* concept" is a complex one, culturally constructed and lexicalised in only a subset of the world's languages. WORDS, on the other hand, is a universal human concept, which (along with other semantic primes) can enable us to break down the *language* concept into simple and cross-translatable terms.

5. Discussion and concluding remarks

With the exception of the excursus into WORDS in the preceding section, this study has been a semantic-lexicographic inquiry into a single language-specific word, namely, the English word *language*. From a purely semantic point of view, the exercise adds a contribution to the study of abstract nouns, a relatively under-developed area of lexical semantics (cf. Goddard and Wierzbicka, submitted for publication),[23] and provides an instructive case study in lexical polysemy. Its broader interest of course turns on the pivotal role that the particular noun *language* plays (and has played) in English-language discourse, as a key word in public discussion about world affairs, immigration, education, and a host of other topics, in scholarly discourse in the arts and humanities, and most particularly in disciplines such as linguistics, communication studies, and lexicography, among others.

In recent times, various critics have voiced objections to the "*language* concept". Some of these objections concern the everyday folk concept of *language* (in some respects echoing a similar critique of the "*culture* concept" in anthropology; cf. Duranti, 1997; Kuper, 1999; Shweder, 2001). Others concern technical ideas about the nature of language – ideas belonging to the discourse and conceptual framework of linguistics as a discipline; for example, that a language is a "code", or a "system", or a "set of sentences" (cf. Harris, 1980; Love, 2004, 2007). Sometimes objections of both kinds are raised at the same time, with little or no distinction being drawn between them. I would like to close with some reflections about those criticisms that pertain to the everyday folk concept of *language*.

The semantic exegesis undertaken in the present study, crystallised in explication [A], is consistent with criticisms of *language*$_1$ that highlight its implicit essentialism and reification, and its historical links with territoriality, nationalism and ethnicity. These characteristics, and the overall culture-specificity of the concept, mean that the "*language* concept" must indeed be handled with care.

On the other hand, I would not go all the way with critics who condemn the *language* concept unreservedly for perpetrating a "myth" or a "false ontology", and call for it to be abandoned in favour of more flexible, inclusive or dynamic notions such as *communication*, *discourse*, *interaction*, *communicative practice*, *multimodal semiotics* or "*languaging*" (e.g. Harris, 1980, 1981; Davis and Taylor, 1990; Makoni and Pennycook, 2006). I have two main reasons. First, just because a nominal category concept lacks a real-world referent does not necessarily make it "pernicious", as Pennycook (2006) describes *language*. Many nominal categories (perhaps all abstract ones) lack real-world referents (think of *consciousness*, *identity*, *rights*), and this does not prevent them from fulfilling a valuable function as topics of discourse without necessarily carrying with them any ontological commitments.[24]

Second, many critics of the "*language* concept" seem oblivious to the fact that their preferred alternative concepts are no less semantically complex and language-specific than *language* itself. Just as the semantic content of *language* frames and shapes any discourse in which it is a key term, highlighting certain considerations and backgrounding others, so too do terms such as *communication*, *discourse*, and *interaction*, each in its own way. For example, while *communication* (Goddard, 2009b) is more inclusive and process-oriented than *language*, the everyday meaning of *communication* does not carry with it any acknowledgement of the important role of WORDS in the process, nor does it tip the attention towards cultural, regional or historical variability. On the contrary, it is all too easy to speak of "*human* communication" in general terms, and in the process to elide the tremendous differences in communication practices across languages and cultures. The same goes for other general terms, such as *interaction* and *discourse*. They are likely to tip our attention away from geographical and cultural differences in ways of speaking and in "ways with words". It is hard to speak about *languages* (in the plural), however, without some allowance for cultural differences and particularities.

In the end, however, none of the terms *language*, *communication*, *discourse*, *interaction* or *semiotics* provide a suitable conceptual basis for the human sciences, because all of these terms are language-specific and culture-bound, and because they

[22] Ye (2006, p. 239) takes the analysis of *zì* a step further in proposing that it has a distinct second sense *zì*$_2$, roughly 'character word', for which she provides a formal NSM explication.

[23] Previous NSM studies of individual abstract nouns include Wierzbicka (2004) on *consciousness*, Wierzbicka (2006a) on *experience*, Wierzbicka (2006b) on *dialogue*, Goddard (1996b) on *enterprise*, Goddard (2005) on *culture*, and Goddard (2009b) on *communication*. An important addition to this literature will come with the publication of Wierzbicka (2010) – a book-length study of *evidence*, *experience*, and *sense*.

[24] Further, to the extent that the word *language* comes with essentialist ontological commitments, it would seem that these are more or less dissolved when the word is used in the expression *language varieties* (presumably based on *language*$_3$), which sociolinguists have used for decades as a neutral cover term for dialects, registers, styles, and the like.

are too conceptually complex to allow for optimal formulation of testable hypotheses. The more promising approach would be to re-frame our fundamental research questions and hypotheses in terms of semantic primes, such as SAY, WORDS, DO, WANT, KNOW, and others (Wierzbicka, 2005). This would serve to detach the terminology from the grip of any single language and at the same time greatly advance conceptual clarity. Though this project is beyond the scope of the present study, I would hope that it has made a start towards this goal.

Acknowledgements

My thanks go to Anna Wierzbicka, Dorothea Cogill, Catherine Travis, Elizabeth Ellis, Anna Gladkova and Zhengdao Ye for helpful discussions and information. I am particularly grateful to Anna Gladkova and Zhengdao Ye for their contributions in respect of Russian and Chinese, respectively. Thanks to Paul Eckert for input regarding Pitjantjatjara and Yankunytjatjara. Earlier versions of parts of this paper were presented at Australex 2004 at Sydney University, at ConCom05 'Conceptualising Human Communication' at the University of New England in December 2005, and at a seminar in the Department of Linguistics at Göteborg University, Sweden, in July 2007. I thank the participants of all these gatherings for their helpful comments. Several anonymous reviewers for *Language Sciences* provided comments that have helped me improve the paper in various ways. This research was supported by the Australian Research Council.

Appendix A. Semantic primes – English exponents, grouped into categories

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

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

Appendix B. Explications for *language$_1$* in Russian and in Chinese

The following Russian and Chinese versions for explication [A] (with interlinear glossing) have been provided by Anna Gladkova and Zhengdao Ye, respectively. For convenience, the English version of the explication is presented again below.

[A] *language$_1$*, e.g. *a language, different languages*

a. something of one kind
b. people can say what this something is with the word *language*
c. people can say something about something with this word when they think like this:

d. ''there are people of many kinds,
these people live in many places, people of one kind in one place,
people of another kind in another place
e. when people in one of these places want to say something to other people in this place,

(*continued on next page*)

they say it in one way, they say it with words of one kind
f. other people in this place can know what these people want to say, because they know these words
g. when people in another place want to say something, they say it in another way, they say it with words of another kind"

Russian. One can see the word *slovo* WORD in instrumental form *slovom*, in components (b) and (c); and in plural instrumental form *slovami*, in components (e)–(f). Note, incidentally, that this explication employs two allolexes (*rod* and *vid*) as Russian exponents of KIND. See Gladkova (2007, 2010) for a thorough treatment of Russian NSM.

[A1] Russian explication for *language*$_1$:
a. čto-to odnogo roda
b. ljudi mogut skazat', čto ėto slovom *language*
c. ljudi mogut skazat' čto-to o čem-to ėtim slovom, kogda oni dumajut tak:
d. ''est' ljudi mnogix vidov
ėti ljudi živut vo mnogix mestax, ljudi odnogo vida v odnom meste,
ljudi drugogo vida v drugom meste
e. kogda ljudi v odnom iz ėtix mest xotjat skazat' čto-to drugim ljudjam v ėtom meste,
oni govorjat ėto odnim obrazom, oni govorjat ėto slovami odnogo roda
f. drugie ljudi v ėtom meste mogut znat', čto ėti ljudi xotjat skazat'
potomu čto oni znajut ėti slova
g. kogda ljudi v drugom meste xotjat skazat' čto-to, oni govorjat ėto drugim obrazom
oni govorjat ėto slovami drugogo roda"

[A1] Russian explication for *language*$_1$ (Russian)

a.	čto-to	odnogo	roda				
	something.NOM	one.MASC.SG.GEN	kind.MASC.SG.GEN				
b.	ljudi	mogut	skazat',	čto	ėto		
	people.PL.NOM	can.3PL.PRES	say.INF	what	it		
	slovom	*language*					
	word.NEUT.SG.INST	*language*					
c.	ljudi	mogut	skazat'	čto-to	o		
	people.PL.NOM	can.3PL.PRES	say.INF	something.SG.ACC	about		
	čem-to	ėtim	slovom				
	something.PREP.SG	this.MASC.SG.INST	word.NEUT.SG.INST				
	kogda oni	dumajut	tak:				
	when they.3PL	think.3PL.PRES	like.this.ADV				
d.	''est'	ljudi	mnogix	vidov			
	be.PRES	people.PL.NOM	many.PL.GEN	kind.PL.GEN			
	ėti	ljudi	živut	vo mnogix	mestax,		
	this.PL.NOM	people.PL.NOM	live.3PL.PRES	in many.PREP.PL	place.PREP.PL		
	ljudi	odnogo	vida	vo			
	people.PL.NOM	one.MASC.SG.GEN	kind.MASC.SG.GEN	in			
	odnom	meste,	ljudi	drugogo			
	one.PREP.MASC.SG	place.PREP.MASC.SG	people.PL.NOM	another.MASC.SG.GEN			
	vida	v	drugom	meste			
	kind.MASC.SG.GEN	in	another.PREP.MASC.SG	place.PREP.MASC.SG			
e.	kogda	ljudi	v	odnom	iz	ėtix	mest
	when	people.PL.NOM	in	one.PREP.SG	out.of	this.PL.GEN	place.PL.GEN
	xotjat	skazat'	čto-to	drugim	ljudjam		
	want.3PL.PRES	say.INF	something.SG.ACC	other.PL.DAT	people.PL.DAT		
	v	ėtom	meste,	oni	govorjat	ėto	
	in	this.PREP.MASC.SG	place.PREP.MASC.SG	they.3PL	say.3PL.PRES	it	
	odnim	obrazom,	oni		govorjat	ėto	
	one.SG.INST	way.SG.INST	they.3PL		say.3PL.PRES	it	
	slovami	odnogo	roda				
	word.PL.INST	one.MASC.SG.GEN	kind.MASC.SG.GEN				
f.	drugie	ljudi	v	ėtom	meste	mogut	
	other.PL.NOM	people.PL.NOM	in	this.PREP.MASC.SG	place.PREP.MASC.SG	can.3PL.PRES	

	znat',	čto	ėti	ljudi	xotjat		
	know.INF	what	this.PL.NOM	people.PL.NOM	want.3PL.PRES		
	skazat'	potomu	čto	oni	znajut	ėti	slova
	say.INF	because	that	they.3PL	know.3PL.PRES	this.PL.ACC	word.PL.ACC
g.	kogda	ljudi	v	drugom	meste		
	when	people.PL.NOM	in	another.PREP.MASC.SG	place.PREP.MASC.SG		
	xotjat	skazat'	čto-to,	oni	govorjat	ėto	
	want.3PL.PRES	say.INF	something.SG.ACC	they.3PL	say.3PL.PRES	it	
	drugim	obrazom,	oni		govorjat	ėto	
	other.MASC.SG.INST	way.MASC.SG.INST	they.3PL		say.3PL.PRES	it	
	slovami	drugogo	roda"				
	word.PL.INST	other.MASC.SG.GEN	kind.MASC.SG.GEN				

Chinese. One can see the Chinese exponent of WORDS *zì* in components (b) and (c), and again in components (e)–(f). See Ye (2006) and Chappel (2002) for treatments of Chinese NSM.

[A2] Chinese explication for *language*$_1$:
a. yìzhǒng dōngxi
b. rénmen néng yòng 'language' zhège zì shuō zhè dōngxi shì shěnme
c. dāng rénmen zhèyàng xiǎng shí, rénmen néng yòng zhège zì jiù mǒxiē dōngxi xiǎng shuōxie shěnme:
d. ''yǒu hěnduō zhǒng rén
zhèxie rén shēnghuó zài hěnduō dìfāng, yìzhǒng rén zài yíge dìfāng,
qítā lìng yìzhǒng rén zài qítā ling yíge dìfāng
e. dāng qízhōng yíge dìfāng de rén xiǎngyào duì zhège dìfāng de qítā rén shuōxie shěnme shí,
tāmen yòng yìzhǒng fāngshì shuō, yòng yìzhǒng zì shuō
f. zhège dìfāng de qítā rén néng zhīdào tāmen xiǎng shuō de, yīnwèi tāmen zhīdào zhèxie zì"
g. dāng lìng yíge dìfāng de rénmen xiǎngyào shuōxie shěnme,
tāmen yòng lìng yìzhǒng fāngshì shuō, yòng lìng yìzhǒng zì"

[A2] Chinese explication for *language*$_1$ (Chinese)

a.	yìzhǒng	dōngxi									
	one.kind	thing									
b.	rénmen	néng yòng	'language'	zhège		zì	shuō	zhè	dōngxi shì	shěnme	
	people	can do.with	language	this.CL		word	say	this	thing be	what	
c.	dāng	rénmen	zhèyàng	xiǎng	shí,	rénmen	néng	yòng	zhège	zì	
	when	people	this.way	think	when	people	can	do.with	this.CL	word	
	jiù	mǒxiē	dōngxi	xiǎng	shuōxie	shěnme:					
	about	some	thing	think	say.CL	something					
d.	''yǒu	hěnduō	zhǒng	rén							
	there.is	many	kind	people							
	zhèxie	rén	shēnghuó	zài	hěnduō	dìfāng,					
	these	people	live	loc	many	places					
	yìzhǒng	rén	zài	yíge	dìfāng,	qítā	ling	yìzhǒng	rén		
	one.kind	people	LOC	one.CL	place	other	another	one.kind	people		
	zài	qítā	ling	yíge	dìfāng						
	LOC	other	another	one.CL	place						
e.	dāng	qízhōng	yíge	dìfāng	de	rén	xiǎngyào				
	when	which.among	one.CL	place	LIG	people	think.want				
	duì	zhège	dìfāng	de	qítā	rén	shuōxie	shěnme	shí,		
	towards	this.CL	place	LIG	other	people	say.CL	something	be		
	tāmen	yòng	yìzhǒng	fāngshì	shuō,	yòng	yìzhǒng	zì	shuō		
	they	do.with	one.kind	way	say,	do.with	one.kind	word	say		
f.	zhège	dìfāng	de	qítā	rén	néng	zhīdào	tāmen	xiǎng	shuō	de,
	this.CL	place	LIG	other	people	can	know	they	want	say	LIG
	yīnwèi	tāmen	zhīdào	zhèxie	zì"						
	because	they	know	these	words						

(*continued on next page*)

g.	dāng	lìng	yíge	dìfāng	de	rénmen	xiǎngyào	shuōxie	shěnme,	
	when	another	one.CL	place	LIG	people	think.want	say.CL	something	
	tāmen	yòng	lìng	yìzhǒng	fāngshì	shuō,	yòng	lìng	yìzhǒng	zì"
	they	do.with	another	one.kind	way	say	do.with	another	one.kind	words

B.1. Interlinear glosses

3 third person, ACC accusative, ADV adverb, CHAR characteristic, CL classifier, CONTR contrastive, DAT dative, ERG ergative, GEN genitive, INF infinitive, INST instrumental, LIG ligature, LOC locative, MASC masculine, NEUT neuter, NOM nominative, PAST past tense, PL plural, POSS possessive, PRES present tense, PREP prepositional, PRT particle, RDP reduplication, SG singular.

References

Ameka, F.K., 2006. "When I die, don't cry": The ethnopragmatics of "gratitude" in West African languages. In: Goddard, C. (Ed.), Ethnopragmatics: Understanding Discourse in Cultural Context. Mouton de Gruyter, Berlin, pp. 231–266.

Anderson, B., 1991[1983]. Imagined Communities: Reflections on the Origins and Spread of Nationalism. [Revised and extended edition]. Verso, London/New York.

Bugenhagen, R.D., 2002. The syntax of semantic primes in Mangaaba-Mbula. In: Goddard, C., Wierzbicka, A. (Eds.), Meaning and Universal Grammar – Theory and Empirical Findings. John Benjamins, Amsterdam, pp. 1–64.

Chao, Y.R., 1968. A Grammar of Spoken Chinese. University of California Press, Berkeley/Los Angeles.

Chappell, H., 2002. The universal syntax of semantic primes in Mandarin Chinese. In: Goddard, C., Wierzbicka, A. (Eds.), Meaning and Universal Grammar – Theory and Empirical Findings. John Benjamins, Amsterdam, pp. 243–322.

Cruse, D.A., 1986. Lexical Semantics. Cambridge University Press, Cambridge.

Davis, H.G., Taylor, T.J. (Eds.), 1990. Redefining Linguistics. Routledge, London.

Dixon, R.M.W., Aikhenvald, A.Y., 2002. Word: a typological framework. In: Dixon, R.M.W., Aikhenvald, A.Y. (Eds.), Word: A Cross-linguistic Typology. Cambridge University Press, Cambridge, pp. 1–41.

Duranti, A., 1997. Linguistic Anthropology. Cambridge University Press, Cambridge.

Evans, N., 1992. Kayardild Dictionary and Thesaurus: A Vocabulary of the Language of the Bentinck Islanders. University of Melbourne, Department of Linguistics and Language Studies, North-West Queensland, Parkville.

Firth, J.R., 1957. Papers in Linguistics 1934–1951. Longman, London.

Geeraerts, D., 1994. Polysemy. In: Asher, R.E., Simpson, J.M.Y. (Eds.), The Encyclopedia of Language and Linguistics. Pergamon Press, Oxford/New York, pp. 3227–3228.

Gelman, S.A., 2005. The Essential Child: Origins of Essentialism in Everyday Thought. Oxford University Press, New York.

Gladkova, A., 2007. Russian Emotions, Attitudes and Values: Selected Topics in Cultural Semantics. Ph.D. Thesis, Australian National University.

Gladkova, A., 2010. Russkaja kul'turnaja semantika: ėmocii, cennosti, žiznennye ustanovki [Russian cultural semantics: emotions, values, attitudes.] Languages of Slavic Cultures, Moscow. (in Russian), in press.

Goddard, C., 1996a. Pitjantjatjara/Yankunytjatjara to English Dictionary, revised second ed. IAD Press, Alice Springs.

Goddard, C., 1996b. Can linguists help judges know what they mean? Linguistic semantics in the courtroom. Forensic Linguistics 3 (2), 250–272.

Goddard, C., 1998. Semantic Analysis: A Practical Introduction. Oxford University Press, Oxford.

Goddard, C., 2000. Polysemy: a problem of definition. In: Polysemy: Theoretical and Computational Approaches. Oxford University Press, Oxford, pp. 129–151.

Goddard, C., 2001. Lexico-semantic universals: a critical overview. Linguistic Typology 5 (1), 1–66.

Goddard, C., 2004a. Linguistic Report to the De Rose Hill Native Title Hearing. Adelaide, SA.

Goddard, C., 2004b. The ethnopragmatics and semantics of 'active metaphors'. Journal of Pragmatics 36, 1211–1230.

Goddard, C., 2005. The lexical semantics of *culture*. Language Sciences 27 (1), 51–73.

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

Goddard, C., 2009b. The 'communication concept' and the 'language concept' in everyday English. Australian Journal of Linguistics 29 (1), 11–25.

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

Goddard, C. (Ed.), 2008. Cross-Linguistic Semantics. John Benjamins, Amsterdam.

Goddard, C., Wierzbicka, A., 2008. New semantic primes and new syntactic frames: "Specificational be" and "abstract this/it". In: Goddard, C. (Ed.), Cross-Linguistic Semantics. John Benjamins, Amsterdam, pp. 35–58.

Goddard, C., Wierzbicka, A., submitted for publication. The semantics of "abstract nouns": Modern NSM analyses and the thinking of John Locke and Jeremy Bentham.

Goddard, C., Wierzbicka, A. (Eds.), 1994. Semantic and Lexical Universals – Theory and Empirical Findings. John Benjamins, Amsterdam.

Goddard, C., Wierzbicka, A. (Eds.), 2002. Meaning and Universal Grammar – Theory and Empirical Findings, vol. I and II. John Benjamins, Amsterdam.

Harris, R., 1980. The Language-Makers. Duckworth, London.

Harris, R., 1981. The Language Myth. Duckworth, London.

Harvey, M., 2002. Land tenure and naming systems in Aboriginal Australia. The Australian Journal of Anthropology 13 (1), 23–44.

Haugen, E., 1972. Dialect, language, nation. In: Dil, A.S. (Ed.), The Ecology of Language. Stanford University Press, Stanford, CA, pp. 237–254.

Henderson, J., 2002. The word in Eastern/Central Arrernte. In: Dixon, R.M.W., Aikhenvald, A.Y. (Eds.), Word: A Cross-linguistic Typology. Cambridge University Press, Cambridge, pp. 100–124.

Henderson, J., Dobson, V., 1994. Eastern and Central Arrernte to English Dictionary. IAD Press, Alice Springs.

Joseph, J.E., 2002. From Whitney to Chomsky: Essays in the History of American Linguistics. John Benjamins, Amsterdam.

Junker, M-O., 2008. Semantic primes and their grammar in a polysynthetic language: East Cree. In: Goddard, C. (Ed.), Cross-Linguistic Semantics. John Benjamins, Amsterdam, pp. 165–206.

Kilham, C., Pamulkan, M., Pootchemunka, J., Wolmby, T., 1986. Dictionary and Source Book of the Wik-Mungkan Language. Summer Institute of Linguistics, Darwin.

Kuper, A., 1999. Culture: The Anthropologists' Account. Harvard University Press, Cambridge, MA.

Laughren, M., Hale, K. Warlpiri Lexicography Group, 2006. Warlpiri-English Encyclopaedic Dictionary. Electronic files. University of Queensland, St. Lucia, Brisbane.

Love, N., 2004. Cognition and the language myth. Language Sciences 26 (6), 525–544.

Love, N., 2007. Are languages digital codes? Language Sciences 29 (5), 690–709.

Makoni, S., Pennycook, A. (Eds.), 2006. Disinventing and Reconstituting Languages. Multilingual Matters, Clevedon.

Merlan, F., 1981. Land, language and social identity in Aboriginal Australia. Mankind 13, 133–148.
Merlan, F., Rumsey, A., 1991. Ku Waru: language and segmentary politics in the Western Nebilyer Valley. Papua New Guinea. Cambridge University Press, Cambridge.
Ong, W.J., 2002. Orality and Literacy. Routledge, London/New York.
Peeters, B. (Ed.), 2006. Semantic Primes and Universal Grammar: Empirical Evidence from the Romance Languages. John Benjamins, Amsterdam.
Pennycook, A., 1994. The Cultural Politics of English as an International Language. Longman, London.
Pennycook, A., 2006. Postmodernism in language policy. In: Ricento, T. (Ed.), An Introduction to Language Policy: Theory and Method. Blackwell, Malden, MA, pp. 60–74.
Reid, N., McTaggart, P., 2009. Ngan'gi Dictionary. Australian Linguistics Press, Armidale.
Rhodes, R.A., 1985. Eastern Ojibwa-Chippewa-Ottawa Dictionary. Berlin, Mouton de Gruyter.
Rumsey, A., 1993. Language and territoriality in Aboriginal Australia. In: Walsh, M., Yallop, C. (Eds.), Language and Culture in Aboriginal Australia. Aboriginal Studies Press, Canberra, pp. 191–206.
Sharpe, M., 1992. Dictionary of Western Bundjalung. University of New England, Armidale.
Shweder, R.A., 2001. Culture: contemporary views. In: Smelser, N.J., Baltes, P.B. (Eds.), International Encyclopedia of the Social and Behaviorial Sciences, vol. 5. Pergamon Press, Oxford, pp. 3151–3158.
Simpson, J.M.Y., 1994. Language. In: Asher, R.E., Simpson, J.M.Y. (Eds.), Encyclopedia of Language and Linguistics. Pergamon Press, Oxford/New York.
Sm'algyax Learners' Dictionary, 1999. English-Sm'algyax Version. Ts'msyeen Sm'algyax Authority and First Nations Education Council (District 52). April 1999 Draft.
Stubbs, M., 2001. Words and Phrases, Corpus Studies of Lexical Semantics. Blackwell, Oxford.
Teubert, W., Čermáková, A., 2004. Directions in corpus linguistics. In: Halliday, M.A.K., Teubert, W., Yallop, C., Čermáková, A. (Eds.), Lexicology and Corpus Linguistics: An Introduction. Continuum, London, pp. 113–166.
Wierzbicka, A., 1972. Semantic Primitives. Athenäum Verlag, Frankfurt.
Wierzbicka, A., 1996. Semantics: Primes and Universals. Oxford University Press, Oxford.
Wierzbicka, A., 1997. Understanding Cultures through their Keywords. Oxford University Press, Oxford.
Wierzbicka, A., 2002. Semantic primes and universal grammar in Polish. In: Goddard, C., Wierzbicka, A. (Eds.), Meaning and Universal Grammar – Theory and Empirical Findings. John Benjamins, Amsterdam, pp. 65–144.
Wierzbicka, A., 2004. Conceptual primes in human languages and their analogues in animal communication and cognition. Language Sciences 26 (5), 413–441.
Wierzbicka, A., 2005. Empirical universals of language as a basis for the study of other human universals and as a tool for exploring cross-cultural differences. Ethos 33 (2), 256–291.
Wierzbicka, A., 2006a. ''Experience" in John Searle's account of the mind: brain, mind, and Anglo culture. Intercultural Pragmatics 3 (3), 241–255.
Wierzbicka, A., 2006b. The concept of 'dialogue' in cross-linguistic and cross-cultural perspective. Discourse Studies 8 (5), 675–703.
Wierzbicka, A., 2009. Exploring English phraseology with two tools: NSM semantic methodology and Google. Journal of English Linguistics 37 (2), 101–129.
Wierzbicka, A., in press. Evidence, Experience, and Sense: The Hidden Cultural Legacy of English. Oxford University Press, New York.
Ye, Z., 2006. Ways of Meaning, Ways of Life: A Semantic Approach to Chinese Ethnopsychology. Ph.D. Thesis. The Australian National University.
Ye, Z., 2007. 'Memorisation', learning and cultural cognition: The notion of *bèi* ('auditory memorisation') in the written Chinese tradition. In: Amberber, M. (Ed.), The Language of Memory in a Cross-Linguistic Perspective. John Benjamins, Amsterdam, pp. 139–180.
Yoon, K.-J., 2004. Not just words: Korean social models and the use of honorifics. Intercultural Pragmatics 1 (2), 189–210.