Lexico-semantic universals: A critical overview

CLIFF GODDARD

Abstract

Are there any word-meanings which are absolute and precise lexico-semantic universals, and if so, what kind of meanings are they? This paper assesses the status, in a diverse range of languages, of about 100 meanings which have been proposed by various scholars (linguists, anthropologists, psychologists) as potential universals. Examples include: 'I', 'this', 'one', 'big', 'good', 'true', 'sweet', 'hot', 'man', 'mother', 'tree', 'water', 'sun', 'wind', 'ear', 'say', 'do', 'go', 'sit', 'eat', 'give', 'die', 'maybe', 'because'. Though relatively small, the sample is variegated enough to justify the preliminary conclusion that the semantic primes proposed by Wierzbicka (1996) and colleagues are much stronger contenders for universal status than are terms designating natural phenomena, body parts, concrete objects, and other putative experiential or cultural universals.

Keywords: *basic vocabulary, lexical universals, lexicon, Natural Semantic Metalanguage, polysemy, semantic primes, semantics*

1. Introduction

1.1. *Preliminaries: Definitions and approaches*

This paper addresses the following questions: Are there any word-meanings which are universal, i.e., which are found in all languages? If there are, what kind of meanings are they? Are they, for example, based on universals of experience, or on cultural universals, or are they sui generis, belonging to a domain of conceptual semantics? Needless to say, the scope and nature of lexico-semantic universals – if there are any – is of profound theoretical importance to lexical typology, lexicology, and semantic methodology in general (see Section 3).

It is necessary at the onset to establish some terminology, and for this purpose I propose to adopt some of Cruse's (1986: 77–78) definitions. The key an-

Linguistic Typology 5 (2001), 1–65 1430–0532/2001/005-01

alytical construct is "lexical unit" – defined as the pairing of a single specifiable sense (meaning) with a lexical form; cf. also Mel'čuk (1989).[1] The concept of "lexical unit" is not to be identified either with "lexeme" (a family of lexical units) or with "lexical form" (a family of word forms which differ only in respect of inflection). A polysemous word is a lexeme which consists of more than one lexical unit. A lexical form need not be formally monomorphemic: it may be a compound or derived word, or a phraseme. The main concern of this paper can therefore be re-stated as follows: Are there any meanings which exist as the senses of lexical units in all languages? I will refer to meanings which putatively satisfy this description as (putative) "lexico-semantic universals", or "universally lexicalised meanings", or more simply, as "lexical universals".

One issue which immediately arises is what standard of precision it is reasonable to adopt. In discussions of lexical universals there is a tendency to speak at a level of approximate, rather than precise, semantic identity. For example, it is sometimes said, in the wake of Berlin & Kay (1969), that all languages have words for 'black' and 'white' but this statement can only be valid at an approximate level, because in languages with only a small number of abstract colour words, the words which cover black and white also cover many other hues.

Many suggestions about approximate lexical universals are closely linked with proposals, scattered through the anthropological literature, about cultural universals. D. E. Brown (1991) is a useful summary. Aside from various universals concerning language, he nominates things like: use of non-verbal communication, making of tools and shelter, use of fire, some form of marriage, a system of kinship, sexual regulations and standards of sexual modesty, division of labour, some system of law and sanctions, some standards of etiquette and hospitality, some concept of personal belongings, religious or supernatural beliefs, rituals, bodily adornment, medicines and healing practices, dancing and music, aesthetic standards, and many others. Ideas about approximate lexical universals also come from generalisations about the physical environment, and from inherent properties and capabilities of the human body.

To deal adequately with a reasonable sample of possible approximate lexical universals would be impossible within the confines of a journal article. Furthermore, precise lexical universals (if they exist) are surely more interesting than

1. Mel'čuk (1988: 170) identifies a similar concept of "lexical unit" (though he uses the term "vocable" rather than "lexeme"). He furthermore enunciates a principle whereby one of the lexical units of a polysemous word can be identified as the basic lexical unit: "The basic lexical unit of a vocable is the lexical unit which has a semantic bridge [= is semantically linked directly] with the majority of other lexical units of the vocable". The concept of semantic bridging is also necessary to establish the integrity of a lexeme (vocable), in the sense of distinguishing between polysemy (a single lexeme with significantly related meanings) as opposed to homonymy (distinct lexemes with identical forms and unrelated meanings).

approximate ones. For these reasons, in this paper we will be concerned with the search for precise lexical universals. This search obviously presupposes that we have a principled and practical method of semantic description. Here we adopt reductive paraphrase in natural language as the basic method of specifying meanings. That is, a "meaning" of an expression will be regarded as a paraphrase, framed in semantically simpler terms than the original expression, which is substitutable without change of meaning into all contexts in which the original expression can be used. This is the foundational postulate of the Natural Semantic Metalanguage (NSM) approach to semantic analysis, originated by Anna Wierzbicka (1972, 1980, 1992, 1996, among other works; Goddard 1998a). It is also followed, broadly speaking, by the Meaning-Text Model of Igor Mel'čuk and colleagues (cf. Mel'čuk 1988, 1989).[2] The postulate implies the existence, in all languages, of a finite set of indefinable expressions (words, bound morphemes, phrasemes). The meanings of these indefinable expressions, which represent the terminal elements of language-internal semantic analysis, are known as "semantic primes". About 60 semantic primes have been proposed (Wierzbicka 1996). They are listed in Section 1.2 below.

Wierzbicka and colleagues hypothesise that the semantic primes of all languages coincide. This is certainly the strongest claim about universally lexicalised meanings to be found in the contemporary linguistic literature, and the specifics of this claim (i.e., the universality or otherwise of the proposed NSM primes) will be dealt with in the body of the paper. But even if one grants that semantic primes (if they exist) have a plausible claim to being universals, it

2. Concerning the decision to adopt the reductive paraphrase approach, it could be asked: Doesn't this decision necessarily bias the outcome of the study? What would happen if, for example, one adopted a prototype approach instead? Given that any empirical claim is relative to the theoretical framework in which it is located, I freely agree that the reductive paraphrase approach brings with it certain perspectives, constraints, and procedures, which impart their stamp on this study. It would hardly be practical here, however, to conduct a parallel investigation in terms of several alternative frameworks, or to engage in metatheoretical debate on the merits of alternative frameworks (cf. Goddard 1998b). The reductive paraphrase approach has the merit of setting rather stringent standards of semantic identity, simply because within this framework the claim that two semantically similar lexemes have (or do not have) precisely "the same meaning" can be operationalised in a very clear fashion: Is it or is it not possible to identify a PARAPHRASABLE meaning difference, i.e., to identify a semantic component which can be stated in discrete, reductive terms and which belongs to one lexeme but not to the other? As I hope to show, using this criterion we can often identify small and subtle meaning differences between semantic near-equivalents in different languages, and thus overturn many potential candidates for the status of precise lexical universal. Within a prototype approach, such as that explored, for example, in Newman (1996), the question of PRECISE semantic identity of senses across languages does not arise in a particularly sharp fashion. A prototype approach seems more compatible, and more comfortable, with claims about approximate lexical universals, than with claims about precise lexical universals.

does not follow that semantic primes are the only possible lexical universals. There may be certain meanings which, though complex, represent universals of human experience, and are thus plausible candidates for lexicalisation in all or most languages. For example, Wierzbicka (1992) has argued that all languages have a lexical unit whose meaning corresponds to the primary meaning of the English word 'mother', even though she does not regard this meaning as a semantic prime (see below). The 100 or so potential lexical universals considered in this article are about evenly divided between proposed primes and other terms.

As to data, it is unfortunately not possible to rely on conventional dictionaries. Often they do not provide the necessary level of language-internal semantic analysis, either because they do not characterise semantic invariants of meaning in a sufficiently clear or predictive fashion, or because they do not distinguish adequately between polysemy and semantic generality. Dictionaries of less well known languages may be particularly prone to these difficulties, but even major monolingual dictionaries of English and other European languages are not immune. Before reliable conclusions about lexical universals can be drawn, it is fundamental that polysemy be recognised as appropriate, and distinguished from semantic generality, because otherwise we cannot even establish the number and nature of lexical units belonging to each lexeme. (To take a simple example, one could not conclude that 'sun' is not a lexical universal solely from the fact that in various languages a single lexeme corresponds both to English *sun* and *day*. Language-internal analysis is needed to establish whether or not such a lexeme consists of two lexical units, one of which has a meaning corresponding to 'sun'.)

It is not possible here to discuss in detail the criteria for recognising polysemy, except to say that in general terms I adopt the traditional "definitional" approach (Geeraerts 1994); i.e., an expression has a single meaning if and only if it is possible to formulate a single reductive paraphrase explication which is predictive of its full range of distribution. An expression is polysemous if this is not possible, and two (or more) distinct explications are required. Formal indications of polysemy include the putatively distinct meanings having different syntactic properties, or having different antonyms, or only one of the senses participating in a derivational process (cf. Apresjan 1972, 1974 [1992]; Mel'čuk 1988; Wierzbicka 1996: 242–244, 270–274).

Given the unsuitability of data from dictionaries, it may be asked whether we are in a position to make any real progress on identifying universally lexicalised meanings. If we confine ourselves to high quality linguistic information, as I have tried to do in this paper, we will not be able to sample anything more than a very small fraction of the 5,000 or so languages of the world. It turns out, however, that this limitation is not as profound as it seems, because, as we will see, most proposed lexical universals fail even on a small

sample of geographically dispersed languages. The sample of languages referred to in this paper is an opportunistic one. It consists primarily of languages which I know personally and languages which I have been able to discuss with native speakers or colleagues who have done intensive semantic-lexicographical research: Yankunytjatjara (Australia), Malay, Lao, Cantonese, Kalam (Papua New Guinea), Japanese, Polish, and Russian (see acknowledgements). These languages are mostly from Europe, East Asia, Papua New Guinea, and Australia; but at least seven different languages families are represented. A wide variety of other languages are also referred to in relation to specific points.

1.2. *Divergent approaches: Swadesh, Wierzbicka, and Brown*

As a starting point for a survey of possible lexico-semantic universals, we can contrast Morris Swadesh's (1972) list of 100 Basic Vocabulary items with Anna Wierzbicka's (1996) list of 60-odd proposed semantic primes. The purposes and origins of these two lists are, of course, very different. Swadesh was chiefly interested in identifying a sample of stable, lexicalised meanings which could be used for lexicostatistical analysis. His approach was purely pragmatic and inductive. He adjusted his initial hypotheses several times as a result of pre-testing before proposing the 100 items listed in Table 1, which he describes as designating "universal and simple things, qualities, and activities, which depend to the least possible degree on the particular environment and cultural state of the group". They include "pronouns, some quantitative concepts, parts and simple activities of the body, movements, and some general qualities of size, colour, and so on" (Swadesh 1972: 275). Swadesh recognised that languages may not have precise semantic equivalents to these items.

Swadesh does not ascribe any significance to his numbering system, but to some extent it must reflect a hypothesis about the durability and universality of the items. It can be no coincidence that almost all of the first dozen-and-a-half prove to be relatively strong candidates as universally lexicalised meanings. Surprisingly perhaps, in view of their different origins, most of the early Swadesh items also occur on Wierzbicka's list of universal semantic primes.

Wierzbicka's proposals differ markedly from those of Swadesh in that they are derived entirely from semantic analysis. The items listed in Table 2 are, according to Wierzbicka (1996), meanings which are both incapable of further (non-circular) definition, and indispensable for adequate paraphrase analysis of the lexicon at large. The "double-barrelled" items, such as *something/thing*, *someone/person*, and *when/time*, indicate meanings which, in English, are expressed by variant forms (allolexes) in different grammatical contexts. It is claimed that every one of the proposed primes exists as the meaning of a lex-

Table 1. *100-item Basic Vocabulary List (Swadesh 1972)*

1.	I	26.	root	51.	breasts	76.	rain
2.	you	27.	bark	52.	heart	77.	stone
3.	we	28.	skin	53.	liver	78.	sand
4.	this	29.	flesh	54.	drink	79.	earth
5.	that	30.	blood	55.	eat	80.	cloud
6.	who	31.	bone	56.	bite	81.	smoke
7.	what	32.	grease	57.	see	82.	fire
8.	not	33.	egg	58.	hear	83.	ash
9.	all	34.	horn	59.	know	84.	burn
10.	many	35.	tail	60.	sleep	85.	path
11.	one	36.	feather	61.	die	86.	mountain
12.	two	37.	hair	62.	kill	87.	red
13.	big	38.	head	63.	swim	88.	green
14.	long	39.	ear	64.	fly	89.	yellow
15.	small	40.	eye	65.	walk	90.	white
16.	woman	41.	nose	66.	come	91.	black
17.	man	42.	mouth	67.	lie	92.	night
18.	person	43.	tooth	68.	sit	93.	hot
19.	fish	44.	tongue	69.	stand	94.	cold
20.	bird	45.	claw	70.	give	95.	full
21.	dog	46.	foot	71.	say	96.	new
22.	louse	47.	knee	72.	sun	97.	good
23.	tree	48.	hand	73.	moon	98.	round
24.	seed	49.	belly	74.	star	99.	dry
25.	leaf	50.	neck	75.	water	100.	name

ical unit in all languages, and, furthermore, that the meaning correspondence across languages is not merely approximate, but precise.[3]

Wierzbicka acknowledges that a list like that in Table 2 is insufficient, in itself, to identify the intended meanings because many of these English expressions are polysemous. For this reason each prime is provided with a set

3. The groupings and labels in Table 2 are not intended to bear any great theoretical weight. Their status in the table, and also in the organisation of the present paper, is purely expository. No claim is implied that these groupings divide the meta-lexicon into grammatically self-contained domains or fields. The studies in Goddard & Wierzbicka (eds.) (1994), Goddard (ed.) (1997), and Goddard & Wierzbicka (eds.) (in press) amply document that individual languages may allocate morphosyntactic properties across the meta-lexicon in diverse and cross-cutting ways.

Table 2. *Proposed semantic primes (after Wierzbicka 1996, with more recently proposed items in parentheses)*

Substantives:	I, you, someone/person, something/thing, (body)
Determiners:	this, the same, other
Quantifiers:	one, two, some, all, many/much
Evaluators:	good, bad
Descriptors:	big, small
Mental predicates:	think, know, want, feel, see, hear
Speech:	say, word, (true)
Actions, events, movement:	do, happen, move, (touch)
Existence and possession:	there is, (have)
Life and death:	live, die
Time:	when/time, now, before, after, a long time, a short time, for some time, (moment)
Space:	where/place, here, above, below, far, near, side, inside
"Logical" concepts:	not, maybe, can, because, if
Intensifier, Augmentor:	very, more
Taxonomy, partonomy:	kind of, part of
Similarity:	like

of "canonical contexts", i.e., a set of sentences or sentence fragments exemplifying its characteristic grammatical (combinatorial) contexts. For example, to ascertain which sense of English *know* is intended as a prime one can refer to canonical contexts such as *I don't know where he is* and *This person knows something about you*. (Equivalents of semantic primes may have different morphosyntactic characteristics, and hence belong to different "parts of speech", in different languages, without this necessarily disturbing their canonical combinatorial properties.)

The claim that semantic primes are universal has been empirically investigated in a collection of studies published as *Semantic and Lexical Universals* (Goddard & Wierzbicka (eds.) 1994). In these studies, contributors investigated whether 39 proposed semantic primes (the full number at that time) were to be found "embodied", as it were, in lexical units in a varied range of languages. The sample was typologically and genetically diverse, containing only one European language, French. The other languages involved were as follows: Acehnese (Indonesia), Arrernte (Australia), Ewe (Ghana), Japanese, Longgu (Solomon Islands), Kalam (Papua New Guinea), Kayardild (Australia), Mandarin Chinese, Mangap-Mbula/Mangaaba-Mbula (Papua New

Guinea), three Misumalpan languages (Nicaragua), Samoan, Thai, Yankunytjatjara (Australia). Subsequent work (Goddard (ed.) 1997) has examined aspects of the combinatorial syntax of primes in Japanese, Longgu, French, and Hawaiian Creole English. A set of comprehensive studies of Lao, Malay, Mandarin Chinese, Mangaaba-Mbula, Spanish, and Polish is presently underway (Goddard & Wierzbicka (eds.) in press).

Finally, mention should be made of the crosslinguistic lexical surveys of Cecil H. Brown, Stanley R. Witkowski, and colleagues (Witkowski & Brown 1978, C. H. Brown 1984, 1989). Inspired by the work of Berlin & Kay (1969) these scholars have mounted a series of studies into regularities in body-part nomenclature, folk biology, weather terminology, and various other lexical fields. The relevance of these surveys to the existence of precise lexical universals is, however, rather limited; first, by the fact that the investigators were primarily interested in finding implicational universals of lexical field structure, rather than precise lexical universals, and second, by their reliance on data from bilingual dictionaries, which, as noted earlier, is generally not of a suitable semantic standard to support the identification of precise lexical universals.

2. A survey of potential lexico-semantic universals

We can now begin a survey of potential universally lexicalised meanings, including both proposed semantic primes and other, presumably non-prime, meanings such as those identified in the work of Swadesh and others. For convenience, the full list is provided in Table 3. To anticipate, we will find that although there are some strong contenders of both types, the proposed primes fare much better than their putatively "cultural universal" competitors. Connected with this is the fact that, paradoxical as it may seem from some points of view, the best candidates for lexical universals tend to have a rather "abstract" and general character, rather than designating concrete objects or natural phenomena.

2.1. *Substantives*

2.1.1. *Substantives (I, you, something/what/thing, someone/who).* Typologists generally take for granted the universality of the singular pronouns 'I' and 'you'. The only substantial issue concerns the situation in languages like Thai and Japanese in which there are several different pronouns in each category, the choice between them signalling some kind of interpersonal message to do with solidarity, social distance, or the like. This has been discussed by Diller (1994) and Onishi (1994), who both conclude that, initial impressions notwithstanding, it is possible to identify primary, semantically unmarked equivalents for 'I' and 'you' in Thai and Japanese, respectively. It is sometimes claimed

Table 3. *Meanings considered in this study*

SUBSTANTIVES	
2.1.1	Substantives: 'I', 'you', 'something/what/thing', 'someone/who'
2.1.2	Relational substantives: 'part of', 'kind of'
2.1.3	Social categories and relationships: 'people', 'man', 'woman', 'child', 'mother', 'father'
2.1.4	The body, body parts, and products: 'head', 'eye', 'nose', 'ear', 'hand', 'shit', 'piss'
2.1.5	Life-forms: 'tree', 'wood', 'bird', 'fish'
2.1.6	Cosmological features: 'sun', 'moon', 'night', 'day'
2.1.7	Other features of the natural world: 'water', 'fire', 'rock', 'mountain', 'wind'
SPECIFIERS	
2.2.1	Determiners: 'this', 'the same', 'other'
2.2.2	Quantifiers: 'one', 'two', 'much/many', 'all'
ATTRIBUTES AND INTENSIFIER	
2.3	Attributes: 'good', 'bad', 'big', 'small', 'old'
2.4	Intensifier: 'very'
PREDICATES	
2.5.1	Events, actions, motion: 'happen', 'do', 'move', 'hit', 'break', 'go', 'come'
2.5.2	Mental predicates: 'think', 'know', 'want'
2.5.3	Speech: 'say', 'word', 'true'
2.5.4	Life and death: 'live', 'die', 'kill'
2.5.5	Existence and possession: 'there is', 'make', 'have', 'give'
2.5.6	Sensations and emotions: 'see', 'hear', 'feel'
2.5.7	Bodily postures and activities: 'sit', 'lie', 'stand', 'eat', 'drink'
2.5.8	Sensations and emotions: 'hot', 'cold', 'hungry', 'thirsty', 'sweet', 'fear', 'anger'
THE DOMAIN OF TIME	
2.6.1	Deictic and categorical temporal meanings: 'now', 'when/time'
2.6.2	Sequence: 'before', 'after'
2.6.3	Duration: 'a long time', 'a short time', 'for some time'
THE DOMAIN OF SPACE	
2.7.1	Deictic and categorical spatial meanings: 'here', 'where/place'
2.7.2	Vertical dimension: 'above', 'below'
2.7.3	Laterality: 'on one side', 'left', 'right'
2.7.4	Topological relations: 'inside', 'on'
LOGICAL CONCEPTS	
2.8.1	Negation: 'not'
2.8.2	Possibility and potentiality: 'maybe', 'can'
2.8.3	Causation: 'because'
2.8.4	Conditional: 'if'
SIMILARITY	
2.9	Similarity: 'like'

that English *you* is indeterminate as to number, but in reality the word is polysemous between *you* (SG) and *you* (PL), as shown by the contrast between the reflexives *yourself* and *yourselves* (cf. Wierzbicka 1976, Goddard 1995).

Almost all languages appear to have separate words for 'someone' and 'something' (cf. Haspelmath 1997). Sometimes the same words are used as interrogatives or as so-called "knowledge complements", i.e., in constructions like *I don't know who did it* or *I know what happened*, as with Acehnese *soe* 'who/someone' and *peue* 'what/something' (Durie et al. 1994). More commonly, one set of forms will be morphologically basic and the other(s) will be built upon it. Occasionally the expressions for 'someone' and 'something' are phrasemes, as with Kayardild *ngaaka dangkaa* 'someone' and *ngaaka thungalda* 'something' (Evans 1994). In some polysynthetic languages, the equivalents of 'something' and 'someone' are bound morphemes. For example, in Koasati (Louisiana) *na:si-* 'something' and *a:ti-* 'someone' are normally bound morphemes appearing as the first element of a verbal word (Kimball 1985: 106, 135–139).

Very occasionally, it may appear that the same expression is used to cover both 'someone' (who) and 'something' (what). This is the case in Lithuanian where the relevant form is *kas*. But Tatjana Bulygina (Anna Wierzbicka, personal communication) argues that *kas* is polysemous: kas_1 'someone', kas_2 'something'. Her arguments include the fact that kas_1 and kas_2 have different agreement patterns; for example, *Kas linksmas, tas mielas* 'who(ever) is cheerful is nice' (adjectives with masculine agreement) does not equal *Kas linksma, tai miela* 'what(ever) is cheerful is nice' (adjectives with feminine/neuter agreement). The two meanings also exhibit different behaviour in the genitive case. Kas_2 'something' has a single genitive form *ko*, which is used for the full range of functions of genitive case. Kas_1 'someone', however, has a special genitive form *kieno* 'whose, by whom', which is used to denote possession and to mark the subject of a passive verbal construction, with *ko* being used for other genitive functions (cf. Dambriūnas et al. 1966: 285). Another language which at first seems to lack a lexical distinction between 'someone' and 'something' is Wambaya (Australia). The same stem *gayini* is used for both, but the distinction is made by choice of different gender suffixes (Nordlinger 1998: 120–122): with the inanimate gender suffix *gayini* means 'something', with an animate gender suffix it means 'someone' (the masculine animate being used when the actual gender of the referent is unknown).

In some languages (e.g., Arrernte, Samoan, Yankunytjatjara), the word for 'someone' is identical in form with the word for 'other', but because the syntactic properties of 'someone' and 'other' are so different ('someone' being a substantive and 'other' a specifier) it is usually straightforward to establish polysemy on language-internal grounds. For example, Yankunytjatjara *kutjupa* means 'other' when it is adnominal (e.g., *kungka kutjupa* 'another woman') and

'someone' when it is the head of an NP in its own right, as in *Kutjupa-ngku katingu* 'Someone (ERG) took (it)'. It might be thought that the latter usage is elliptical, with an implied head noun such as *a<u>n</u>angu* 'people/person', but this analysis is not viable since *kutjupa(ngku)* 'someone' can be used to refer to non-human beings, such as the Christian God, who could never be referred to as an *a<u>n</u>angu* '(human) person'.

In some languages, the word for 'something (what)' is polysemous, expressing also the meaning 'part of', when it appears in an appropriate grammatical context. This is dealt with in Section 2.1.2 below.

2.1.2. *Relational substantives (part of, kind of).* Linguists seem to agree that the part-whole relationship is fundamental to the vocabulary structure of all languages (cf. partonomic relationships such as *hand*/*arm)*. But though the element 'part (of)' is postulated to be a semantic prime, there are languages which do not have a unique word for 'part'. In such cases, the meaning 'part (of)' is typically expressed by means of the word for 'something', 'thing', or 'what', used in a grammatical construction associated with "possession". This can be illustrated from Yankunytjatjara and Mangaaba-Mbula, which are unrelated languages (Goddard 1994, Bugenhagen in press). In these languages 'part' belongs to a lexical unit of the same lexeme as 'something' (presumably no coincidence, since a part of something is itself a something).

(1) Yankunytjatjara
Puntu kutju, palu kutjupa-kutjupa tju<u>t</u>a-tjara.
body one but something many-having
'(It is) one body, but with many parts.'

(2) Mangaaba-Mbula
Iti tomtom na koroŋŋa-nda boozo, kumbu-ndu, nama-nda, …
we person given thing-our many leg-our head-our
'We people, our parts are many: legs, heads, …'

As with partonomy, so with taxonomy. Linguists and cognitive anthropologists seem to agree that taxonomy is a basic principle of lexical organisation, especially in the realm of living things (cf. Berlin 1992). All languages have hierarchies of designations which specify that certain individually named animals and plants are 'kinds' of some higher level "life forms"; for example, a *sparrow* is a kind of *bird*, a *trout* is a kind of *fish*, an *oak* is a kind of *tree*. There is less agreement about whether all languages contain a lexical unit which could serve to articulate the nature of this arrangement, i.e., whether all languages have a lexical unit meaning 'kind (of)'. All languages examined in the *Semantic and*

Lexical Universals survey did contain such a lexical unit, but the situation is made complex by the fact that in several languages this unit belonged to a polysemous lexeme. For example, in Yankunytjatjara the meaning 'kind of' is expressed by a lexeme *ini* which can also mean 'name'; e.g., *Wayuṯa kuka ini kutjupa* 'the possum is another kind of meat-animal' (Goddard 1994). In Kayardild 'kind of' is expressed by *minyi*, which can also mean 'colour' (Evans 1994).

2.1.3. *Social categories and relationships (people, man, woman, child, mother, father).* It seems that all languages have a lexical item one of whose meanings corresponds to that of English *people* in non-specific uses like 'people think ...', 'people say ...', and so on. In European languages this word is often a collective term, unrelated to the word for an individual human being; for example, French *les gens*, Russian *ljudi*. In languages without obligatory number marking, the term for 'people' can usually be used to refer to a single individual. This is the case, for example, with Japanese *hito*, Yankunytjatjara *aṉangu*, and Malay *orang*.

In some languages the term for 'people' is formally complex, e.g., Kalam *bin-b* (lit., woman-man) (Pawley 1993). Sometimes the expression for 'people' appears to be a pluralised version of the term for 'someone'; for example, Mangaaba-Mbula *zin tomtom* (lit., plural 'someone'). The Mangaaba-Mbula expression is not semantically the sum of its parts, however, because *tomtom* 'someone' can refer to beings other than humans, whereas *zin tomtom* cannot.

'Man' and 'woman' are not proposed semantic primes, but they nevertheless have some claim to being universal meanings.[4] It is true that their nearest equivalents are often conflated with "social" information (e.g., rank or age-level), but it can often be shown that such lexemes are polysemous, with both general and specific meanings. In Pitjantjatjara, for example, there are two words for 'woman', *kungka* and *minyma*, depending on the age and associated seniority of the individual. Further, *wati* 'man' normally refers only to initiated men. However, it is arguable that in both cases there is polysemy, with *wati* and *minyma* also having more general meanings as 'man' and 'woman', respectively (cf. Goddard 1996a). One piece of evidence for this conclusion is that *wati* and *minyma* are used when it is necessary to speak about men and women in general, without regard to their social status (for example, to discuss sexually transmitted diseases or to discuss ritual or economic division of labour). Both words are also routinely used about non-Aboriginal men and women.

4. Actually the plural counterparts, i.e., 'men' and 'women', may be semantically more basic than the singulars 'man' and 'woman', because it is easier initially to explicate 'men' and 'women' as social categories, i.e., as two 'kinds of people'.

A potential counter-example of a different kind is found in Japanese, where instead of single words for 'man' and 'woman' there are phrasemes of the form 'male person' and 'female person': *otoko no hito* and *onna no hito*. Normally, however, these Japanese expressions are not used to refer to children, and as far as I know, they have the same referential range as English *man* and *woman*. Until and unless some specific differences are identified, the claim that 'man' and 'woman' are lexical universals remains viable. Cantonese at first appears to have similar phrasemic equivalents to 'man' and 'woman' but on closer examination there is evidence that the forms are lexical compounds. For example, in *neoi*^{5}*jan*^{2} 'woman (lit., female person)' the second morpheme *jan*^{2} has rising tone, but as a separate word *jan*^{4} 'person' has low-falling tone.[5] This kind of tone change indicates that the combination has been lexicalised (cf. Matthews & Yip 1994: 26).

'Child' may also have some claim to universal status, despite apparent arguments to the contrary. The nature of these counter-arguments, and the reasons why they are not decisive, can both be illustrated from Spanish. First, there may be no gender-neutral term answering to 'child', as in Spanish where *niño* and *niña* designate a male child and female child, respectively. Nevertheless, the Spanish masculine plural *niños* can be used about children generally, and even in singular contexts *niño* is the default choice when gender is unknown, or known but irrelevant, e.g., *Para ella es dificil conseguir trabajo porque tiene un niño* 'It is difficult for her to get work because (she) has a child'. Second, the nearest equivalents to 'child' may not have the same range of use (as *child* in English). After about 11 or 12 years of age the Spanish words *muchacho* and *muchacha*, roughly 'boy' and 'girl', are called for instead of *niño* and *niña*. However, this difference may have to do with different cultural attitudes as to when young people become able to look after themselves (assuming that this is one of the key semantic factors at play), rather than with the meaning of the word for 'child' itself. It is not possible to treat this issue adequately in the space available here.

Kinship is probably the most intensively studied of all cultural phenomena. According to most experts, the biological link between mother and child provides an essential "linking principle" in all kin systems, raising the question of whether 'mother' could be a universally lexicalised meaning. This proposal is not necessarily upset by the existence of "classificatory" kin terminology because there may be language-internal grounds for recognising polysemy in such languages (cf. Wierzbicka 1992). To illustrate the kind of evidence and arguments involved, consider Yankunytjatjara, in which the word *nguny-*

5. Cantonese words are here transcribed according to the Romanisation recommended by the Linguistic Society of Hong Kong (http://www.hku.hk/linguist/lshk/), which differs from that adopted by Matthews & Yip (1994).

tju can be used not only for one's mother, but also for one's mother's sisters and mother's female cousins (among others). Even so, there are several arguments in favour of recognising polysemy, i.e., the existence of two lexical units: *ngunytju*$_1$ 'mother', and an extended sense *ngunytju*$_2$. First, the expression *ngunytju mula* (*mula* 'true') can only refer to one's biological mother; if *ngunytju* did not have a distinct sense as "biological *ngunytju*" it is hard to see how *ngunytju mula* could have this sense. Second, unlike other kin terminology, the word *ngunytju* is routinely used about animals, where there is no question of classificatory extensions being involved and where the biological basis for the usage seems particularly clear. A third, more theoretical, argument is that it is not possible to state the meaning(s) of *ngunytju* at all unless one is allowed to employ the meaning 'mother' (Scheffler 1978). This is because in order to enunciate the principles by which certain individuals can be reckoned as *ngunytju* (in its extended sense) one must refer to the relationship of biological motherhood, which is the focal or "logically prior" sense of the term *ngunytju*.

In other languages, there may be additional arguments for a polysemy analysis. For example, in addition to its words *yabu(ndi)* for mother (including classificatory mother) and *ŋuman(ndi)* for father (including classificatory father), the Dyirbal language has *jarraga* and *galŋan* to refer to non-classificatory mothers and fathers, respectively (Dixon 1989). Walmajarri, in addition to *ngamaji* 'mother' (including classificatory mother), has a special term *kumpurru* for foster-mother. The same language also has the words *parnmarn* 'a man's actual mother-in-law' and *karntiya* 'a woman's actual son-in-law' (Richards & Hudson 1990). If the concept of 'mother' (i.e., actual biological mother) is salient enough that it figures as a semantic component of other word-meanings, this at least suggests that 'mother' is likely to exist as the meaning of a lexical unit in its own right.

Provisionally then, we can say that 'mother', in its biological sense, has a reasonable claim to the status of a universally lexicalised meaning. Wierzbicka (1992) contends that the same logic applies to the meaning 'father', but as noted by Foley (1997) this conclusion does not seem as secure because the role of father is more open to social factors than that of mother. In some languages (Yankunytjatjara, for one) the person normally denoted by the term nearest in meaning to 'father' is not necessarily the begetter, but the mother's husband at the time of the child's birth or conception. Foley cites data from Gough (1959, 1961) to the effect that in Nayar society (southern India), one's kin relations on the male side are reckoned exclusively through the man (termed *appan* "father") with whom one's mother formed a "permanent union" in a special premenstrual ceremony, though this man is usually not the biological father.

2.1.4. *The body, body parts (head, eye, nose, ear, hand) and products (shit, piss).* There is a substantial literature in linguistics and anthropology about body-parts and about cultural construals of the body, so it comes as something of a surprise that little attention has been directed to the question of whether 'body' exists as the meaning of a lexical unit in all languages. Early surveys of body-part nomenclature (C. H. Brown 1976, Andersen 1978) claimed that the meaning 'body' is indeed universally lexicalised, but such claims have also been disputed.

Unfortunately many of the counter-claims are to be found in the anthropological literature, where either insufficient data is given to decide whether or not polysemy is involved, or the data given is at best equivocal. For example, Wilkins (1996a) refers to claims by anthropologist Gilbert Lewis (1974) that in the Papuan language Gnau "there is no single word corresponding to English 'body' ". According to Wilkins, Lewis says that the lexeme *matil* 'human' must be used when speaking about the body, and that this word must be distinguished for sex in its singular form. Wilkins continues: "Lewis makes it clear that there is no polysemy involved here, it is 'human being', not 'body', which is referred to when *matil* is used". I do not agree with Wilkins' assessment on this point. What Lewis says is that *matil* is an adjective, and that to make it into a singular noun a suffix is needed – a suffix which must be either masculine (*-den*) or feminine (*-da*). But Lewis immediately gives data which suggests that the formally masculine word is not necessarily semantically "masculine", i.e., confined to males (Lewis 1974: 52):

> If the Gnau wish to speak of the remains buried to rot when a man or woman dies – the material element, the body or corpse – they say it is the *matilden* that they must bury [...] *Matilden* then is the word they use where we might say 'body', and it must necessarily distinguish sex in its singular form. So the phrasing of my request for a list of body parts had therefore to be: *djisapeg wobla wobla beiya matilden* "you tell me part part of human being (male)"; and *matilden* was understood in this context to mean 'body': the distinctive aspects of a woman's anatomy were included in the lists I was given without prompting or altering the question.

Wilkins (1996a) charts shifts and overlaps between the meanings 'body' and 'person' (cf. English *somebody*), and 'body' and 'skin' in various languages,[6] but it seems to be generally necessary, for independent reasons, to recognise

6. This may be a good place to observe that semantic simplicity and "stability" of formal realisation across time are not necessarily correlated, particularly in the case of semantic primes such as 'body', which tend to be expressed by polysemous lexemes. Not only can there be frequent and recurrent semantic shifts (cf. Wilkins 1996a), it is known that some languages have borrowed terms for semantic primes, presumably replacing earlier indigenous words. In Malay, for example, the lexemes for 'feel' and 'think' (*rasa* and *fikir*) have been borrowed

polysemy in such cases. This can be illustrated from a language that Wilkins has documented in detail, namely, Arrernte (Wilkins 1996b). The Arrernte word *tyerrtye* can mean, inter alia, either 'people/person' or 'body'. The existence of a separate meaning 'body' is clear in sentences like the following: (i) *Ayenge welheme tyerrtye urinpe-arle-irrerlenge* 'I feel my body getting hot (i.e., I'm getting a fever)' (ii) *Tyerrtye ikwerenhenge, intel-tnye ikwerenge anteme kemirreke thipe kngerrepenhe anteme* 'From his body, from where he lay dead, there then arose a large bird'. As it happens, Arrernte has also a monosemous word for 'body', namely, *arlke* (Henderson & Dobson 1994), so the existence of a lexical unit with this meaning does not hinge on the analysis of *tyerrtye*. In nearby Yankunytjatjara, the situation is similar. The word for 'people/person' *a̲na̲ngu* can also mean 'body', but there is also another word *puntu* for 'body'.

Coming now to individual body-parts, we will first consider words for 'eye', 'ear', and 'nose', which, according to Andersen (1978), strongly tend to be formally unanalyzable. It may be that all three are lexical universals, notwithstanding that lexemes for all of them often have other, polysemic meanings. The most common polysemies seem to be that both 'eye' and 'nose' can sometimes also mean 'face'. As well, the word for 'eye' can sometimes also mean 'seed' (occasionally the lexeme for 'eye' is formally complex; e.g., "seed-face" in some Mayan languages). It is not uncommon for the word for 'nose' also to mean 'point, tip', as with Yankunytjatjara *mulya* which can mean 'nose' or 'tip', and also 'face'. It is well known that in different languages words for 'nose' can be extended in different ways to refer to features of animal physiology. For example, Yankunytjatjara *mulya* and Russian *nos* may refer to the beak of a bird, Japanese *hana* may refer to the trunk of an elephant. However, these extended uses do not necessarily present a challenge to the status of 'nose' as a lexical universal in view of the fact that there are strong arguments for recognising polysemy in such cases.

Occasionally there may be two terms for 'eye'. Timberlake (1993: 881) says that in nineteenth century Russian the words *glaz* and *oko* were still both in use (the former has now displaced the latter except in poetry and songs), but with different meanings: *glaza* refers to eyes as "instruments of physical perception, with which one merely reads or looks", whereas with one's *oči*, one

---

from Sanskrit and Arabic, respectively. The independence of form-meaning stability and semantic simplicity is nicely illustrated by studies of the cognation rates (and by implication, retention rates) for meanings from Swadesh's Basic Vocabulary List. Dyen et al. (1967) found that the five most stable items among Austronesian languages were: 'five', 'two', 'eye', 'we', and 'louse'. Pawley (1997) found the three most stable items in Trans New Guinea languages were: 'eat', 'louse', and 'I'. The only term which ranks in the top five in both families is 'louse', which (though it may perhaps be an approximate lexical universal) obviously has no claim to semantic simplicity.

"gazes actively or reflects a sad thought". Illustrating by reference to Puškin's *Evgenij Onegin*, Timberlake points out that: "In an identical collocation, the insensitive general does not take his [. . .] *glaza* from Tat'jana, but this perceptive heroine does not take her [. . .] *oči* from Onegin". Presumably *oko* is the more semantically complex and language-specific of the two words, on account of its "cognitive" dimension.

At first blush, 'ear' appears to have little claim to the status of a universal meaning, due to the fact there are languages which apparently "divide" the concept behind the English word *ear*. In Spanish, for example, *oido* designates the 'organ of hearing', essentially the inner ear; it is etymologically related to the verb *oir* 'hear'. But *oreja* designates the outer, visible parts of the ears only; it is a physical thing. For example, if one follows the expression *me duele* 'it hurts (lit., to-me hurts)' with *oreja*, the message is that someone must have hit me, or something else has happened to my outer ear; but using the word *oido* implies that I may have a middle ear infection, or something similar, which interferes with my hearing. On the other hand, it can be argued that English *ear* is polysemous, with two senses corresponding closely to the Spanish words. Some evidence for this proposal comes from the fact that when *ear* appears in some compound words it carries only one of the two (putative) meanings. Especially telling in this regard is the word *earache*, which can only refer to a pain in the inner ear (even though an ache in the outer ear is possible). Notice also that the word *ear-piercing* has two meanings, one linked with each of the (putative) meanings of *ear* (i.e., 'terribly loud' vs. 'piercing of the outer ear for the purpose of attaching earrings').

The situation with 'hand' is not clear. It is well known that in many – perhaps most – languages the word for 'hand' can also be used to refer to all or part of the arm (cf. C. H. Brown 1989). Words like these (e.g., Russian *ruka*) are often assumed to be semantically general over hand and arm – but there are reasons to suspect polysemy. First, it is not as easy as it might seem to actually state a suitably general meaning. Second, there are derivations and collocations which seem to be based specifically on the meaning 'hand'; for example, *rukovički* 'mittens', *požat' ruku* 'to shake (lit., squeeze) someone's (DAT) hand', *ručka* hand-DIM 'little hands'. In view of these facts, it is too early to discount the possibility that 'hand' is a universally lexicalised meaning.

Waste products of the human body, such as 'shit' and 'piss', though universal referents, are not lexical universals, because in many – but not all – languages the words are "loaded" with evaluational meaning. For example, there is no neutral term in English corresponding to Yankunytjatjara *kuna*. Whether we choose to translate *kuna* as *shit*, *poo*, *dung*, *excrement*, or *faeces* we are always adding in an element of a social-evaluational nature which is absent from *kuna* itself. Mutatis mutandis, the same applies to 'piss'; as it does also to "taboo" body-parts such as 'penis' and 'vagina'.

2.1.5. *Life-forms (tree, wood, bird, fish).* There has been a good deal of work aimed at uncovering universals in the organisation and evolution of ethnobiological nomenclature. On the basis of this work, which involves mass comparisons of large numbers of languages, it has been claimed (C. H. Brown 1984) that certain so-called "life-form" terms are extremely common, if not absolutely universal. This applies in particular to the term 'tree' (from the botanical realm) and to 'bird' and 'fish' (from the zoological realm). When looked at closely, however, it is apparent that, at best, these words can only lay claim to the status of approximate lexical universals. English, for example, recognises several botanical life-form words (*tree*, *bush*, *vine*, etc.), with the result that English *tree* is much narrower in its referential range than the nearest term in a language with fewer life-form categories, such as Yankunytjatjara, whose main botanical life-form term *puṉu* takes in trees, bushes, vines, grasses, and fungi. *Puṉu* can be glossed as 'living thing which grows out of the ground', whereas English *tree* includes this component plus the specification that the thing in question have a 'trunk'. Clearly the semantic match between *puṉu* and *tree* is not precise.

Given that words for 'tree' often have a related sense 'wood' – C. H. Brown (1984: 60) cites the figure of two thirds of the world's languages – it might be suggested that 'wood' is a possible candidate for a universally lexicalised meaning. This proposal fares better in relation to English and Yankunytjatjara, because *puṉu* indeed has a second sense approximating 'wood'. Admittedly, firewood is not normally referred to as *puṉu* but as *waṟu* (also 'fire'), but it is at least possible to refer to firewood as *puṉu*. However there are European languages which lack an exact equivalent to 'wood', i.e., a word which can be used indifferently about the hard stuff from trees, regardless of the function to which it is put. For example, Polish *drzewo* refers to the "material" which comes from trees (and *drzewo* can also mean 'tree'), but one could not use *drzewo* to refer to a pile of firewood; *drwa* (PL) 'pieces of firewood' would be used instead.[7]

Similar facts to those which disqualify 'tree' as a true lexical universal also apply to 'bird' and 'fish'. At best these are approximate universals (which is to say, not true universals at all). C. H. Brown (1984) lists 24 languages in which the "bird" category is extended to bats, and 21 languages in which the "bird" category excludes some wild flying birds, domestic birds, or large flightless birds. As for 'fish', as C. H. Brown (1984) acknowledges, it is not

7. Another Polish word *drewno* can also be used to refer to 'wood' as "material", but it has a slightly technical ring to it and is less common than *drzewo*. Also, *drewno* is most comfortably used when one is speaking about making things such as craft items, by hand; one could not speak of a house, for example, made of *drewno* (though to speak of a house made of *drzewo* is perfectly normal).

difficult to find languages in which the nearest word for 'fish' extends also to other "fish-shaped" marine creatures, such as sharks and dolphins (e.g., Malay *ikan*), and even, in some cases, to other creatures which live in the water, such as eels, turtles, frogs, crocodiles, and crustaceans. For example, Chrau (Vietnam) includes eels in its "fish" category *ca*, Tzeltal includes crustaceans in its "fish" category *čay*, and Lao includes some kinds of turtle (as well as dolphins, sharks, whales, and eels) in its "fish" category *paa*^{3}. There are also languages from regions which have very few fish species, and where, consequently, there is no word at all answering to 'fish'. Brown cites Kyaka Enga and Ndumba, both from the New Guinea highlands. Conversely, there are languages like Kayardild (Nicholas Evans 1992, personal communication) which "split" fish into two categories: *yakuri* 'bony, scaly fish' and *wanku* 'cartilaginous, scale-less fish'.

2.1.6. *Cosmological features (sun, moon, night, day).* In many languages, the words for 'sun' and 'moon' can also be used to designate the temporal units 'day' and 'month', respectively; e.g., Tagalog *araw* 'sun, day' and *buwan* 'moon, month'. This kind of polysemy, however frequent, does not impeach the viability of 'sun' and 'moon' as universally lexicalised meanings. There are, however, languages which have been reported to have more than one word for 'sun'; for example, Nyawaygi (Australia) *bujira* 'sun low in the sky' and *jula* 'hot sun overhead' (Dixon 1980: 104).

As for 'day' and 'night', it is necessary to indicate which sense of these terms are intended as possible universal meanings. Is it the sense which describes the ambient conditions (e.g., *I went during the day/at night*) or the "unit of time" sense (e.g., *two days/nights*)? In some languages, separate words are necessary in these two uses. For example, Yankunytjatjara *kaḻaḻa* 'by day, in the daylight', *tjintu* (kutjara) '(two) days'; Malay *siang* 'by day', *(dua) hari* 'two days', Lao *vên*^{2} 'daylight, by day', *mùù*^{4} 'day (unit)'. It seems that the "unit" sense of 'day' is more likely to be a universal than the ambient sense. So far I am not aware of any counter-example.[8]

On the other hand, there are languages in which the word closest in meaning to 'night' is not normally used in a "unit" sense at all. For example, in Polish it would sound very odd to say the equivalent of 'three nights'; the same applies to Yankunytjatjara. This suggests that if 'night' is to have any claim to being

8. Whorf (1956: 40) claimed that the Hopi word for 'day' (*taala*) could not be enumerated with cardinal numbers, though he allowed that similar meanings could be conveyed by the use of ordinals; e.g., instead of saying 'after three days', one would say 'on the fourth day'. Malotki (1983) has shown that Whorf's claim stands in need of amendment in various respects. Even if Whorf were correct, however, it is unclear whether this would force us to conclude that *taala* means something different to 'day'.

a lexical universal, it is the "ambient" sense of the word (e.g., in a phrase like 'at night') which should be our focus of interest. Polish, however, can also be used to furnish counter-evidence to this proposal. The significant fact is that the Polish word *wieczór* (roughly) 'evening' differs from its nearest English counterpart. Whereas English *evening* is viewed as the first part of the night, *wieczór* is viewed as, so to speak, the last part of the day. The Polish *noc* (roughly) 'night' doesn't start till the *wieczór* is over; in other words, *noc* starts later than English *night*. Differences like this suggest that there is no sense in which 'night' is a precise lexical universal.

2.1.7. *Other features of the natural world (water, fire, rock, mountain, wind).* Surprising as it may seem to English speakers, 'water' is probably not a universal lexical unit. Japanese has two words (*mizu* and *yu*) for 'water', with *yu* (often with an honorific prefix *o*-) being reserved for hot water (Suzuki 1978: 51–52). *Mizu* cannot be used about hot water. Furthermore, combining the adjective *atsui* 'hot' with *mizu* sounds unnatural – Suzuki calls it "self-contradictory" – though there is no such restriction in relation to other liquids, e.g., *atsui miruku* 'hot milk' (cf. Wierzbicka 1996: 229). These facts imply that *mizu* and *yu* both have a reference to temperature built into their meanings.

The claim that 'fire' is a universal lexicalised meaning seems quite plausible at first; however, Russian may present a counter-example. As well as *ogon'*, normally glossed as 'fire', Russian has another common word *kostër*, for a fire which is lit outside usually for the purpose of keeping people warm. (*Kostër* is normally glossed as 'bonfire', but *bonfire* is really quite different both in meaning and in stylistic effect.) The significant fact is that a *kostër* cannot be referred to, in Russian, as an *ogon'* (or as a kind of *ogon'*), though it would definitely qualify as *fire* in English.

Foley (1997: 35) expresses confidence that a noun corresponding to 'rock' is "a predetermined category" in the vocabularies of all languages but this statement does not hold up if we understand it to mean that all languages have a lexical unit meaning precisely the same as English *rock*. Polish *skała*, for example, is normally glossed as 'rock' but it is quite different to English *rock*. *Skała* is used for big rocks set into the ground, and could not be used to refer to a rock the size of (say) a loaf of bread.

Swadesh had one natural geographical feature on his list of 100 Basic Vocabulary items, namely, 'mountain', but it is not a viable lexical universal. English *mountain* contains a specification that the feature referred to is very big (very high) and difficult to climb; smaller, easier to climb, features are known as *hills*. But many languages do not lexicalise this distinction, e.g., Yankunytjatjara *apu*, Tagalog *bundok*. There are also languages which have no word resembling 'mountain' because the locale simply does not have any mountains. Nida (1947: 135) uses the Mayan language Yucatec to illustrate this point: "the

Maya country is extremely flat, except for some slight knolls, rising perhaps 100 feet at the most [. . .] A *muul* 'a low hill' is the closest equivalent which the Maya language possesses". Comparable examples could no doubt be found in the atoll languages of Micronesia and the Pacific.

C. H. Brown (1989) has claimed that 'wind' is a lexical universal, but this claim seems doubtful, at least if we are interested in precise semantic identities across languages. Nida (1947: 159) says that: "In some languages no one word will be found for 'wind'. There may be several types of winds, e.g., 'zephyrs', 'tornadoes', 'hot winds off the desert', or 'freezing winds', but there may be no general word for 'wind' ". Unfortunately, however, Nida does not give any examples.

2.2. *Specifiers (determiners and quantifiers)*

2.2.1. *Determiners (this, the same, other).* Linguists rightly expect all languages to have at least one demonstrative word. Usually, of course, languages have several such words, organised into a paradigm structured according to various semantic parameters, but in all such systems it seems possible to identify one demonstrative which is semantically basic, and to claim that this basic demonstrative has the same meaning as 'this'. For example, French *ce*/*cette* is a "solo" demonstrative, English *this* contrasts with one other element (*that*), Yankunytjatjara *nyanga* contrasts with two other elements (*pala* and *nyara*). But even so, one can maintain that *ce*/*cette*, *this*, and *nyanga* have the same meaning, in the sense that it is not possible to state, in the form of a paraphrase, any difference between them. The existence of different numbers and kinds of contrasting elements brings with it differences in patterns of usage (but differences in usage do not always entail paraphrasable differences in meaning).

Evidence suggests that all languages have a lexical unit meaning 'the same', though the same evidence makes it clear that its range of use differs somewhat from language to language. In particular, in some languages, e.g., Samoan (Mosel 1994), the expression for 'the same' can only be used as a predicate complement; i.e., to say that two people 'do the same' or 'think the same', but not to say that 'the same person' came. In these languages, however, the word for 'one' often functions to fulfil the nominal specifier work of 'the same'. It may well be, therefore, that closer analysis would reveal that in such cases the word for 'one' can function as an allolex of 'the same'. Occasionally, a single lexeme serves to express both meanings 'the same' and 'one'. This occurs in Acehnese, for example, with the form *sa*, but there is language-internal evidence for polysemy (the different meanings are associated with different allomorphy patterns).

On available evidence, it appears that 'other' is a lexical universal. Some languages have more than one word which can serve as equivalents to English

other, but on closer examination it turns out that one of these terms is more basic. For example, Longgu (Solomon Islands) has two "*other*" words, *ve'ete* and *lou*, but *ve'ete* can probably be decomposed as 'not the same' (or perhaps 'not of the same kind'). On the other hand, *lou* can be used in contexts like 'you and two other people', where this line of analysis will not work (Hill 1994, Wierzbicka 1994: 471).

2.2.2. *Quantifiers (one, two, much/many, all).* No convincing counter-examples have been reported to the claim that 'one' and 'two' are lexical universals (though the word for 'one' sometimes is polysemous, also meaning 'only'). Of course, one has to recognise that words for 'one' and 'two' may have different morphosyntactic properties (i.e., belong to different word-classes) in different languages. For example, in Cayuga they are formally verbs, carrying verbal affixes which cross-reference the person of the head; so, to say the equivalent of 'two men', one says, in effect, 'man+be-two' (Sasse 1996). It also has to be recognised that a language's having terms for 'one' and 'two' does not necessarily mean that the speakers of the language employ these words for "counting". Many cultures lack the institution of serial counting (one, two, three, four, etc.). Some have no words for numbers higher than three or four. (In some Sinitic languages, e.g., Cantonese, a different word is used for 'two' when it is used in counting, as opposed to specification.)

It seems likely that all languages have at least one word with the meaning 'many' and/or 'much'. Even languages with a very small inventory of numerals are invariably reported to have terms for 'one', 'two', and 'many'. The unresolved question is whether 'many' and 'much' are properly regarded as separate meanings, as suggested by the existence of distinct words in many languages aside from English (e.g., Thai *lǎ:y* 'many', *mâ:k* 'much'), or whether they are contextual variants of a single meaning, as suggested by the fact that in many other languages a single word is used in both functions (for example, Mandarin *xǔduō*, Arrernte *arunthe*, French *beaucoup*).

There are also languages in which there are overlaps in form. For example, in German *viele* is 'many' (*viele Leute* 'many people') but *viel* is 'much' (*viel Bier* 'much beer'); the two words also have somewhat different case-marking patterns. Yet another possibility is for there to be a single form which manifests slightly different combinatorial syntax depending on its meaning. For example, Malay *banyak* can be combined with *sangat* 'very' when it means 'many' (e.g., *sangat banyak kucing* 'very many cats'), but not when it means 'much' (e.g., **sangat banyak air* 'very much water'). The fact that the same restriction applies in English perhaps suggests that 'many' and 'much' should be regarded as separate meanings.

Despite occasional claims that certain "primitive" languages lack the resources for making "absolute generalisations", no reputable linguistic descrip-

tion has reported a language which lacks an equivalent for 'all'. On the other hand, it is clear that words for 'all' vary somewhat in their syntax from language to language. In particular, it seems that in some languages 'all' has an "adverbial" syntax, i.e., it does combine directly with nouns. For example, in Japanese *minna* 'all' is a verbal modifier (Onishi 1994); in Mandarin Chinese *dōu* 'all' is an adverb (Chappell 1994); see Evans (1995) for a close discussion of "adverbial quantifiers" (A-quantifiers) in Mayali.

2.3. *Attributes (good, bad, big, small, old)*

All languages appear to have terms for the proposed semantic primes 'good' and 'bad' (Goddard & Wierzbicka (eds.) 1994, cf. Hill 1987, Dixon 1982); for a contrary view on 'bad', see Myhill (1996). It appears that it is possible in all languages to express meanings involving 'good' and 'bad' in both attributive and in predicative frames; i.e., to say the equivalents of things like 'do something good/bad' and 'be good/bad'.

It must be acknowledged that expressions for 'good' and 'bad' do not have exactly the same ranges of use in different languages. Obviously, one source of such disparities is simply that different cultures regard different things as good and bad (and for different reasons). But not all differences can be explained in this way. In particular, there may be idiosyncratic restrictions on combinations of these terms with semantically complex words, e.g., in French one can speak of *mauvais temps* 'bad weather', but not of **bon temps* 'good weather'. It is also notable that in some cultural contexts people generally prefer not to label events, actions, or other people as 'bad', preferring in most cases to use a milder expression such as 'not good'. This is the case in Malay, for example (Goddard in press a).

Terms for 'big' and 'small', which are also among the posited semantic primes, also appear to exist in all languages (Goddard & Wierzbicka (eds.) 1994; cf. Dixon 1982). In some languages, the term for 'big' shares a form with 'much/many' or with 'very'; but in these cases the existence of different grammatical properties makes it necessary to recognise polysemy.

There is at least one non-prime "attributive" meaning which may have some claim to universal status – namely, 'old'. In English, the word *old* is polysemous, contrasting both with 'young' and with 'new'. It has been suggested that all languages have at least one "age" term (Dixon 1982). Given the social significance of old people in most societies, it might be thought that 'old' is the best candidate for a lexical universal of "age". Presumably the meaning is quite simple; to say that person X 'is old' is to say that person X has 'lived for a long time'. However, given the fact that in many cultures age is associated with status, age-standing is likely to be conflated with other social meanings, particularly gender. Thus, in Pitjantjatjara to speak of an 'old' man one uses

the term *tjilpi*, whereas to speak of an 'old' woman one uses the term *pampa*. As far as I know there is no single term which can be used "generically", as it were, to refer to 'old people'. One has to say 'old men (and) old women' or 'grandparents (and) grandmothers'. This suggests that 'old' is not a universally lexicalised meaning.

2.4. *Intensifier (very)*

On present evidence, it appears that all languages have an intensifying word with the same meaning as English 'very', which can combine with words like 'big' and 'good'. In some languages the same word can combine with verbs, but in others a variant word or phraseme is needed (e.g., *very much* in English, or French *beaucoup*, as opposed to *très*). In some languages, the form for 'very' is the same as, or overlaps with, that for 'much'. This is the case in Lao (a single form *laaj*^{3} for both) and Samoan (two different words for 'very', one of which, *tele*, is identical with 'much'). In these cases, however, the combinatorial differences require a polysemy analysis.

2.5. *Predicates*

2.5.1. *Events, actions, motion (break, hit, happen, do, move, go, come).* Most complex events are easily disconfirmed as lexical universals. For example, 'break' has no hope of attaining this status because some languages differentiate different kinds of breaking. For example, Malay has *putus* 'break in two', *patah* 'break but not sever', and *patah* 'break into many pieces, smash'. Cantonese has, among others: *zing*^{2}*laan*^{6} (roughly) 'break into pieces', *zing*^{2}*tyun*^{5} 'cut in two' (both employing *zing*^{2} 'make'), *daa*^{2}*laan*^{6} 'smash into pieces (*daa*^{2} 'hit', implying forceful impact on a fragile object), *ngaau*^{2}*tyun*^{5} 'break into two by bending' (*ngaau*^{2} 'bend'). Similarly, 'hit' is not a universally lexicalised meaning, because some languages have different verbs for different kinds of event which are all lumped together in English as 'hitting'. For example, Yankunytjatjara has *a<u>t</u><u>u</u>ni* 'hit with a stone', *rungka<u>n</u>i* 'hit with a piece of wood', *punganyi* 'hit with hand'. Dyirbal has several words for hitting depending on the kind of implement involved, e.g., *bijin* 'hit with a rounded object, e.g., a clenched fist, a stone', *bunjun* 'hit with a long flexible object, e.g., with the flat of the hand, with a belt or a bramble' (Dixon 1982: 60).

The most plausible candidates for universal "event" meanings are more general in nature; specifically, the proposed semantic primes 'happen' and 'do'. A good deal of crosslinguistic work has been done on this question (Goddard & Wierzbicka (eds.) 1994), and it appears that both meanings have a strong claim to universality, once various instances of polysemy are properly understood. In many languages, the word for 'happen' has a secondary meaning approximating 'appear' or 'arrive'. The Australian language Yolngu Matha

has both (Cooke & Goddard 1997). In the Djambarrpuyngu dialect, 'happen' is expressed by *malŋ'thu-*, which can also mean 'appear'; in the Gumatj dialect, it is expressed by *buna-*, which can also mean 'arrive'. The supposition that these words are monosemous is hardly credible, in view of examples like those in (3a) and (3b) below, which have an abstract substantive phrase such as *ŋula=nhä manymak*/*ŋamakurru* 'something good' as subject.

(3) a. Djambarrpuyngu dialect
ŋula=nhä manymak malŋ'thu-rr ŋarra-k
something good "appear"-3 I-DAT
'Something good happened to me.'

b. Gumatj dialect
ŋula=nhä ŋamakurru ŋarra-ku buna-na
something good I-DAT "arrive"-3
'Something good happened to me.'

The polysemy is not difficult to understand once we see that 'appearing' and 'arriving' both involve something 'happening' in a place, after which something or someone is in the place in question. In the case of 'appearing', there is presumably an additional component involving being 'able to see' something, and in the case of 'arriving' there is an additional component involving prior motion. This kind of polysemy also recurs in various languages, including Mangaaba-Mbula, Ewe, and French. (An older variety of English had a similar use for 'happen', cf. sentences like *He happened upon me just as I was reading the letter*.)

A more complex example of polysemy is found in the Papuan language Kalam, where a single stem *g-* expresses not only 'happen', but also 'do' and 'feel'. Let us consider first the 'do/happen' polysemy, analogues of which turn up sporadically in languages around the world. There are certain intransitive contexts in which *g-* can only mean 'happen', as in (4a). The distinctive thing about this sentence seems to be that its subject is *ak* 'this'. There are certain contexts where *g-* can only mean 'do', as in (4b) and (4c). The distinctive thing about (4b) is the reciprocal particle *pen* 'the same in return', while the distinctive thing about (4c) is the sequence of verbs, with *ap* 'come' serialised with *gpan* 'you.did'.

(4) a. *Mñab nb ak ned wagn ak g gek*
country such this first origin the happen it.happened
mñab Aytol-jl alym.
country Aytol-jl down.there
'The place where this originally happened was down there at Aytol-jl.' (Pawley 1994: 408)

b. *Pen* *gpan* *gpin.*
reciprocally did.you did.I
'I did the same as you.' (Pawley 1994: 404)

c. *Gos* *etp* *agi* *ap* *kun* *gpan?*
thought what having.thought come such did.you
'Whatever were you thinking that you came and did that?' (Pawley 1994: 396)

On the other hand, there are sentences like (5) which are ambiguous between 'happen' and 'do'.

(5) *Tap* *etp* *gp?*
thing what 3SG.happen/do
'What's happened?' Or: 'What has he done?' (Pawley 1994: 408)

The analytical question posed by examples like these is whether it is necessary to adopt a polysemy interpretation (as the presentation so far has assumed), or whether it is possible to sustain a monosemy interpretation. One could perhaps claim that *g-* has a unitary meaning in (4a–c) which is Kalam-specific and simply cannot be stated in English, and that the apparent differences (from an English perspective) are imposed by the lexico-grammatical contexts. But if *g-* has a unitary meaning, then (5) cannot be ambiguous after all – which seems unlikely in the light of Pawley's description. One salient difference between a 'happen' interpretation and a 'do' interpretation concerns "aboutness" (cf. Sasse 1987). A question like 'What happened' is not ABOUT any particular person, whereas 'What did he/she do?' is about a specific person. To claim that *g-* has a unitary meaning in (5) is to claim that this difference does not exist in Kalam, and that a Kalam speaker would be completely indifferent to it. If, as I assume, this is not the case, a interesting question arises: How could a Kalam speaker identify the distinct meanings involved, given that both are expressed by the same lexical form? Presumably – by reference to other, unambiguous, sentences. The Kalam speaker can say: "it can mean the same as *g-* in example (4a) [i.e., 'happen'], or it can mean the same as *g-* in example (4b) [i.e., 'do']".

As for the meaning 'feel', it is readily separated from the other two meanings of *g-* because it is found in a distinctive "experiencer construction". This takes the form: Experiencer–Condition–Verb+TENSE+3SG. The experiencer appears as a noun or free pronoun with objective case-marking as in (6).

(6) *Yp* *tep* *g-p.*
me good feel-3SG.PERF
'I feel good.'

Coming now to the question of the universality of 'do', the first thing to note is that a great many languages either have a unique lexical form for this

meaning or a lexical form which is polysemous between 'do' and 'make' (see Section 2.5.5). The equivalents of sentences such as 'What did you do?' and 'X did something bad to Y' are perfectly straightforward and unambiguous in many – probably most – languages. Wierzbicka and colleagues in the NSM framework claim that 'do' is a universal semantic prime. It must be acknowledged, however, that this is something of a controversial position. Though many linguists seem happy enough to accept the concept of "agency" as a fundamental linguistic notion, the suggestion that this concept can be anchored in a universal lexicalised meaning often meets with stiff resistance. To be sure, there are complications in some languages. We have just seen one example, namely, polysemy between 'do' and 'happen' (as in Kalam). Another recurrent polysemy is between 'do' and 'say'; but again, there are invariably good language-internal grounds to recognise polysemy (see Section 2.5.3 for a discussion of Samoan).

Another difficulty concerns languages which maintain a strict transitivity distinction in their verbal morphosyntax. Many Australian languages are often described in these terms. All verbs in such languages, it is said, are either strictly transitive (selecting an ergative case subject) or strictly intransitive (selecting a nominative case subject). Since 'do' can be either monovalent (as in 'X did something') or bivalent (as in 'X did something to Y'), it may be asked how the meaning 'do' be expressed in a uniform way in such a system? As far as one can tell from the available evidence, languages with so-called strict transitivity distinctions tend to use the formally transitive verb even in semantically unmarked contexts. For example, in Yankunytjatjara (Goddard 1994) the meaning 'do' is expressed by a lexical unit of the verb *palyaṉi*, which is formed, via zero-derivation, from the root *palya* 'good'. *Palyaṉi*, which can also mean 'make' or 'fix', is a formally transitive verb, selecting an ergative case subject; but when used to mean 'do' it does not require an object argument, as illustrated in (7a–b). Notice that in (7b) the word *kura* 'bad' occurs as the complement of *palyaṉi*, showing that the formal make-up of the verb (i.e., its root *palya* 'good') is not semantically "active".

(7) a. *Nganaṉa puṯu iritilpi alaṯji palyaṉingi.*
1PL in.vain for.ages like.this do.PAST.IMPRF
'We've already been doing (like) that for ages.'

b. *Ka wati tjilpi-ngku tjinguṟu kura palyanyangka* …
and man old-ERG if bad do.NMZ.LOC
'For if an old man does something bad …' (1 Timothy 5: 1)

Though linguists generally assume that any language will contain some "verbs of motion", and that "motion" is, in some sense, a universal linguistic category, it is unlikely that any complex verb of motion will be a true lexical universal, given that it is well established that languages differ greatly

in the patterns by which semantic specifications of other kinds (e.g., manner, means, path) are conflated with motion (Talmy 1985). If any motion verb is a viable lexical universal, the best candidate is probably the simplest motion verb of all, namely, 'move', which has been proposed as a semantic prime (Wierzbicka 1996) in contexts such as 'something moved in the bushes' or 'I was so cold/frightened, I couldn't move'. Given that 'move' is not the kind of word which normally figures in basic vocabulary lists, it is surprising that there is often no difficulty in finding equivalents to it, in these contexts, in languages from around the world; for example, Malay *bergerak*, Yankunytjatjara *yurinyi*, Kalam *am-*. It has reported that some languages distinguish obligatorily between moving-without-change-of-location and moving-with-change-of-location, e.g., Lao *nêng*^{3} 'move (not from place to place)' and *ñaaj*^{4} 'move from place to place' (Enfield in press). It is not clear whether reports of this kind challenge the universal status of 'move'. The matter requires further investigation.

'Go' and 'come' are certainly not viable as universally lexicalised meanings. Even within Europe, there are plenty of languages which do not have exact semantic equivalents for 'go'. For example, German has two everyday words for "translational motion": *gehen* (roughly) 'go on foot' and *fahren* 'go, not on foot (e.g., in a vehicle)'. A similar situation is found in Polish, which has *iść* 'move from one place to another on foot' and *jechać* 'move from one place to another in a vehicle'. In Polish, furthermore, these verbs can occur either in the imperfective form or with various perfective prefixes, such as *po-* and *od-*. For example, *X szedl* (imperfective) means roughly 'X was walking', *X po-szedl* means roughly 'X went', and *X od-szedl z miejsca A* 'X left (from) place A'. In other words, different verbs will be used for 'was going along', 'went', and 'went from place-A': *szedl*, *po-szedl*, and *od-szedl*, if we are on foot; and *jechał*, *po-jechał*, and *od-jechał*, if we are in a vehicle.

As far as 'come' is concerned, the situation is even worse. One of the most distinctive things about English *come* is its ability to support so-called "deictic projection". It is well known that apparent equivalents of *come* in other languages (e.g., French *venir*, Italian *venire*, Japanese *kuru*, Malay *datang*) do not allow deictic projection as freely as *come*. For example, using Japanese *kuru* one cannot "project" even to an addressee's location; so to say the equivalent of *I'm coming (to you)* in Japanese, one must say the (near) equivalent of *I'm going (to you)*. This situation is often attributed to (unexplained) pragmatic differences in the range of deictic projection of different languages, but it is much more likely that there are real semantic differences between the apparent equivalents for 'come' in different languages (cf. Wilkins & Hill 1996, Goddard 1997b).

2.5.2. *Mental predicates (think, know, want).* As mentioned earlier, English *know* can be used in several distinct contexts ('know something/know that', 'know someone', 'know how') which in many other languages require different words. The meaning claimed to be a semantic prime is found in contexts such as 'I don't know where he is' and 'She knows you said something about her'. 'Know' in this sense (German *wissen*, French *savoir*, Malay *tahu*, Polish *wiedzieć*) has a strong claim to being a universally lexicalised meaning. In some languages, the relevant expression is formally complex, but not in ways which challenge the semantic integrity of the meaning 'know' itself. For example, in Japanese 'know' is *sit-te iru* which consists of the verb *siru* with the linking suffix *-te* and the auxiliary *iru*. Since *siru* by itself means 'learn, come to know' it might be thought that *sit-te iru* means 'be in the state of having come to know', but as Onishi (1994: 368) points out, resultative readings of this kind are possible only in highly specific contexts. Usually *sitte iru* simply indicates a state. The best interpretation is therefore that *sitte iru* is polysemous, with its stative meaning being fully equivalent to 'know'. (Onishi notes that additional evidence of the polysemy is to be found in the fact that under negation the two meanings (resultative and stative) can be contrasted, cf. Kuno (1983: 109–116).)

Various anthropologists in the "primitive thought" tradition, including Hallpike (1979), have claimed that some indigenous languages do not draw any distinction between mental predicates like 'think' and 'know', and perceptual ones like 'hear' or 'listen'. But, where data is available, this claim turns out to be a confusion based on failure to understand polysemy. For example, in English one can say *I see what you mean*, but no-one concludes from this that *see* is semantically general over visual perception and understanding. The correct conclusion, which can be supported by a range of syntactic as well as semantic evidence, is that *see* has two distinct meanings (just as French *entendre* has two distinct meanings 'hear' and 'understand'). What holds for English and French holds also for "exotic" languages such as Pitjantjatjara, despite the claims of Bain (1992: 86) that: "There is no way to differentiate the concept of thinking, listening, and heeding in Pitjantjatjara. The same verb *kulini* does duty for all". When looked at closely it turns out that it is not difficult to differentiate the three senses ('think', 'listen', 'heed') on language-internal grounds, since each of the senses just mentioned has a distinct syntactic frame, from which the others are excluded (cf. Goddard 1991, Evans & Wilkins 2000).

In some languages, the expression for 'think' is formally complex. For example, in Kayardild the word for 'think' is *marral-marutha* (formally, 'ear-put') (Evans 1994). In Kalam, the expression for 'think' is *gos nŋ-*, where *nŋ-* is the stem meaning 'know'; furthermore, *nŋ-* is found in several other "cognitive" verbs such as *wdn nŋ-* 'see' (*wdn* 'eye'), *tumd nŋ-* 'hear' (*tumd* 'ear'), and

wsn nŋ- ‘dream’ (*wsn* ‘sleep’). Despite the fact that the presence of *nŋ-* in these various formations appears to be indexing some kind of semantic affiliation (a family resemblance) between them, it seems impossible to assign a uniform meaning to *nŋ-* which would, in combination with the other elements, produce the requisite meanings. In particular, *gos nŋ-* ‘think’ cannot be regarded as a semantic compound of ‘know’ and any other element, if only because ‘think’ does not imply ‘know’ (for more on Kalam *nŋ-*, see Section 2.5.6 below).

One language which has been reported to lack an expression meaning ‘know’ is Yidiny, though Dixon (1991: 263) notes that there is a Yidiny verb *ngannga-n*, which has ‘don’t know (about)’ as one of its meanings. One possibility which bears investigation is that ‘know’ in Yidiny is a lexical unit of the verb *binanga-l*, whose chief meaning is ‘hear, listen to’. Polysemy between ‘hear’ and ‘know’ is known to occur in other Aboriginal languages, and there are some usages in Dixon’s (1991) texts which suggest that this may be true of Yidiny also. For example, in one Dreaming story an ancestral being gives the following explanation for why he has given names to the places along his route of travel: *garru binangalna bulmba wanyaja galing*. Dixon translates this clause as ‘so that by-and-by [people] can listen to [and remember the sequence of place-names along a route and remember] where the places are’. This translation, which alludes to the Aboriginal navigational practice of mentally reciting the sequence of names along a Dreaming “track”, preserves the interpretation of *binangalna* (an inflected form of *binanga-l*) as ‘hear, listen to’. However, it would be more straightforward to gloss the clause as ‘so that by-and-by [people] can KNOW where the places are’. In my view, the issue remains open. One key piece of information is whether one could, in Yidiny, explain the meaning of *ngannga-n* ‘don’t know’ in terms of negation and *binanga-l*.

As for ‘want’, evidence assembled in the *Semantic and Lexical Universals* studies, and subsequently in Harkins’ (1995) crosslinguistic study of desiderative expressions, strongly suggests that ‘want’ is a universally lexicalised meaning, though the situation is frequently complicated by language-specific patterns of polysemy which mean that the lexemes in question do not have identical ranges of use. For example, the range of use of the Yankunytjatjara equivalent of ‘want’ *mukuringanyi* does not correspond to that of English *want*, because *mukuringanyi* can also mean ‘like, be fond of (a person)’ (Goddard 1991). The two meanings each have a distinctive frame: ‘want’ is associated with a clausal complement, e.g., ‘want to do’, ‘want something to happen’, whereas ‘like, be fond of’ takes an NP complement in purposive case. Similar polysemies involving ‘want’ are found in many languages. For example, Spanish *querer* not only has the meaning ‘want’, but also a secondary meaning, roughly ‘like, desire, fancy (a person)’. Polysemy between ‘want’ and ‘seek’ is also fairly common, being found, for example, in Ewe (Ameka 1994) and in Misumalpan languages such as Ulwa (Hale 1994).

One language which is reported not to have a lexical equivalent of 'want' is Kayardild. According to Evans (1994), the Kayardild potential and desiderative verbal inflections both include 'want' as a semantic component (along with other elements), but there is no way of isolating 'want' in lexical form. It seems possible, however, that the Kayardild word *janija* may be polysemous between 'seek' and 'want'. This has been suggested (Kenneth Hale, personal communication) in relation to the cognate word *janijani* in the closely related language Lardil (cf. Ngakulmungan Kangka Leman 1997: 117).

Some languages use distinct lexical forms for 'want' in different syntactic contexts. For example, the Japanese desiderative suffix *-tai* is the normal Japanese equivalent of 'want' in an "equi" complement clause, which has an embedded structure, as in (8a). With non-equi complement clauses, however, 'want' is expressed by a distinct adjectival word *hosii*. The complement clause is marked by the subordinating suffix *-te*, as in (8b). The examples are from Onishi (1997: 224), who also describes and discusses some additional complexities concerning alternative case-marking possibilities within the complement clause.

(8) a. *Ore wa soko no iki-tai.*
I TOP [there LOC go]-want
'I want to go there.'

b. *Ore wa sore oki-te hosii.*
I TOP that happen-CONJ want
'I want that to happen.'

Despite the different allolexes, and the different syntactic structures, there seems to be no specifiable semantic difference between *-tai* and *hosii*, so that they have to be analysed as members of the same lexical unit.

2.5.3. *Speech (say, word, true).* In many languages there is a unique lexical form meaning 'say', as in English. But there are languages in which the word for 'say' can also express the meanings 'do' (as in Samoan) or 'want' (as in Kalam and Mangaaba-Mbula). It goes without saying, of course, that polysemy should never be postulated without language-internal evidence. As an example of the kind of evidence which requires a polysemy analysis, consider the situation with Samoan *fai* (Mosel 1994). Its two meanings are associated with different morphosyntactic properties. *Fai* 'say' is a non-ergative verb, selecting an absolutive subject, as in (9a). *Fai* 'do', on the other hand, selects an ergative subject, as in (9b) and (9c). As well, *fai* 'do' often occurs in the so-called long (suffixed) form *fai=a*, which is usual when an ergative verb is preceded by a pronoun, even when *fai* 'do' is used in a non-transitive frame, as in (9c).

(9) a. *Ona toe fai atu lea 'o le fafine, "Se …"*
then again say DIR then ABS the woman friend
'Then the woman said again, "Friend, …" ' (Mosel 1987: 459)

b. *… 'ua fa'apênâ lava ona fai e le tama.*
PERF like.this EMPH that do ERG the youth
'… the youth did it like this.' (Mosel 1987: 122)

c. *'O ai na faia?*
PRES who PAST do?
'Who did it?'

Another common polysemy in which 'say' is involved concerns 'making sounds' of various kinds. For example, the Yankunytjatjara word *wangkanyi* 'say' can also be used to designate a bird or insect making its characteristic sound (though in this case, the verb *wangkanyi* selects a nominative subject, rather than an ergative subject, as with 'say'). A similar situation is found in Thai and in Mandarin (Diller 1994, Chappell 1994). Another common pattern is illustrated in Kalam, where the stem *ag-* 'say' is also found in various compounds which designate different kinds of sound-making activities (e.g., *mnm ag-* 'speak', *mukbel ag-* 'belch'). However, since it is possible to 'say' something without making any sounds, for example, by means of signed messages, the expression for 'say' (in any language) cannot be decomposed into a combination of 'making sounds' and some other elements.

It has recently been suggested that 'word' is a semantic prime, and hence a possible candidate as universally lexicalised meaning (Wierzbicka 1996), not in the sense of a discrete individual unit, but in a vaguer sense as something that can be heard, and that can be used to express a message (cf. Polish *słowo*, rather than *wyraz*). The viability of this proposal is hard to gauge at present (though it is interesting to note that Swadesh included at least one related metalinguistic item on his 100-item list, namely, 'name'). It is known that in many languages the word for 'word' is polysemous, and can also carry meanings such as 'talk', 'language', 'story', and 'message'. In some languages the word for 'word' is morphologically related to 'say' or to 'mouth'; for example, Malay *kaka-kata* 'words' is a reduplication of *kata* 'say'; Lao *kham*^{3}*.sap*^{2} is morphologically a compound of *kham*^{3} 'mouthful' and *sap*^{2} 'word' (a loan from Sanskrit).

One question which comes easily to mind is whether polysynthetic (or "polypersonal") languages would have an item meaning 'word', given that words in polysynthetic languages are often multimorphemic, and can often convey a self-contained meaning equivalent to that of a whole sentence in English. In those languages which I have been able to check, a term for 'word' does exist, though as one might expect of polysynthetic languges, sometimes this term is a bound item. For example, Mohawk (Iroquoian) *weńn-*, Ojibwa (Algonquin) *kidwin* (Rhodes 1985: 207), Ngan'gityemerri (Daly, Australia) *ngan'gi*

(Nicholas Reid, personal communication), Yimas (Papuan) *pia-/-mpwi* (William Foley, personal communication), Sm'algyax (Tsimshian) *algayax̲* (Tonya Stebbins, personal communication).

As with 'word', so too with another recently proposed prime 'true'. All languages looked at so far appear to have a word which includes 'true' as one of its meanings, but the sample is too small to be certain whether 'true' will stand up as a universal. One language which initially presents a problem is Russian, which has two such words – *pravda* and *istina*. *Pravda* is usually used for practical and/or mundane matters, while *istina* is used when there is a suggestion of "deeper" significance (cf. Mondry & Taylor 1992). As this description suggests, it appears that *pravda* corresponds to 'true', while *istina* involves additional components (for example, 'it would be good if people could know this').

2.5.4. *Life and death (live, die, kill).* Wierzbicka (1996) has argued that 'live' (in the sense found in examples like *This person lived for a long time*) is a semantic prime. It is well known that some languages, e.g., German, distinguish between the prime sense of 'live' (*leben*) and a meaning 'live in a place, dwell in' (*wohnen*). In many languages the meaning 'live in a place' is conveyed by the same verb as expresses the meaning 'stay, remain (somewhere)'; for example, in Malay the prime 'live' is *hidup*, and *tinggal* can mean either 'live in a place' or 'stay'. In other languages, the meaning 'live' is conveyed by a lexeme which can also mean 'there is' or 'sit'.

At present, no counter-examples are known to the proposition that 'die' is a universally lexicalised meaning. It is notable that in many cultures it is usual to use indirect or euphemistic expressions to refer to people dying. For example, in Yankunytjatjara the verb *wiyaringu* 'finished' would normally be used when referring to the death of a human being, in place of *ilungu* 'died'. Interestingly, the converse effect is found in Polish. Essentially, *umzreć*' 'die' can be used only of people. To speak of animals dying one uses the word *zdechnąć*, and even this word is not really suitable for animals of all kinds, since it has a distinctly "non-sentimental", or even mildly pejorative, tone, which would make it an insensitive choice to describe the death of, for example, a pet (Wierzbicka in press). The significance of usage restrictions like these remains to be clarified.

Though 'kill' is not a proposed semantic prime, it may be a universally lexicalised meaning. In saying this, I do not mean to deny the well known fact that expressions corresponding to 'kill' are often formally complex. For example, Turkish 'kill' is *öl-dür-*, which is formally a causative version of *öl-* 'die'; in the Papuan language Watam the expression for 'kill' consists of two verbs in a close-knit construction, *mo-* 'do' plus *minik-* 'die' (Foley 1997: 34). But, as has often been noted in discussions of the semantics of causatives, 'kill' does not mean precisely 'cause to die'; if only because 'kill' implies more di-

rect causation than 'cause to die'. (Incidentally, in Turkish *öl-dür-* is formally UNlike other, genuinely semantic causatives, in as much as it allows application of a further causative suffix; i.e., one can form the word *öl-dür-dür-* 'cause to kill', whereas most other "double causatives" are impossible.)

It is known that in many languages one could not use the lexeme for 'kill' in the same broad range of contexts as English *kill*. For example, in Polish, Yankunytjatjara, and no doubt many other languages, one could not speak of a bullet or an explosion "killing" someone, or someone "being killed" in an accident or in battle. But it is arguable on independent grounds (cf. Goddard 1998a: 279–281, Parsons 1994: Ch. 6) that English *kill* is polysemous, with a primary sense ($kill_1$) requiring a human (or human-like) agent (a "do-er"), and a secondary sense ($kill_2$) in which the causal trigger is a "localised event". (One piece of evidence for this is that only in the agentive sense can *kill* take an instrument argument: *John killed the cat with a knife*, but **The explosion killed John with a piece of flying rock*.) Clearly, only the posited $kill_1$ meaning has any chance of universality.

In some Australian languages, a single word can be used for both 'kill' and 'hit', but this does not necessarily mean that a language lacks a lexical unit meaning 'kill'. In Arrernte, for example, *atwene* is ambiguous in this way, but there are two fixed expressions *irrerlknge atwene* and *tetye atwene* (*irrerlknge* 'dead', *tetye* 'to death'), which do unambiguously mean 'kill' (Henderson & Dobson 1994: 323). In Yankunytjatjara *punganyi* 'hit' can also be used to mean 'kill by hitting', especially when speaking about animals; but there is another verb *iluntantanyi* 'kill' (a causative formation based on *ilu-* 'die'). In Yidiny the verb *bunja-n* can be used to mean either 'hit to inflict injury' or 'kill'. There is no other lexical form for 'kill' in Yidiny, so when Dixon (1991: 255) says that *bunja-n* appears to have a unitary meaning, i.e., 'hit in such a way that it could result in death', the implication is that this language lacks a lexical unit meaning 'kill'. However, Dixon notes that there is a derived noun *bunjaybunjay* 'killer', which may be a small piece of evidence in favour of recognising polysemy for *bunja-n*. The matter requires more investigation.

2.5.5. *Existence and possession (there is, have, make, give).* It is well known that English *there is* has peculiar language-specific properties (such as the fact that the word *there* functions as a syntactic "subject" for some purposes, such as tag-question formation). In other languages 'there is' may be expressed by a monolexemic verb, such as *hay* (*haber*) in Spanish. Often, existence is expressed by the same verb which serves as a copula, or by a verb which can also mean 'stand', or 'live', or 'give'. In some languages 'there is' is expressed not by a verb at all but by a particle, or even by a determiner-like element; for example, in Tolai (Mosel 1984) existential meanings are expressed by means of the definite article *a* in a verbless sentence.

It appears that in all languages the meaning 'there is' can take a "locus" argument (Wierzbicka 1996), as in sentences like 'There is someone IN THIS PLACE' and 'There is no-one HERE'. This helps explain the intimate relationship between existential and locative constructions in many languages. Misunderstanding of the tie-ups between locational and existential expressions sometimes leads to assertions that a particular language lacks an existential word altogether. This point can be illustrated from Yankunytjatjara, where, as in many Australian languages, posture verbs are routinely used to designate the location of particular kinds of entities. Thus, in normal circumstances, to assert that 'X is over there' one would choose the verb *ngaṟanyi* 'stand' for a tree, *nyinanyi* 'sit' for a person whose actual posture is unknown or irrelevant, *ngarinyi* 'lie' for water, and *pupanyi* 'crouch' for a hut. It looks like there is no general verb of existence. When the difference between existence and location is taken into account, however, this conclusion can be shown to be incorrect. In fact *ngaṟanyi* functions as the general existential verb, in which function it can be predicated of any kind of entity, as shown in (10). Existential *ngaṟanyi* is also used in relation to "abstract" subjects, for example, to speak of events taking place, customs or rights applying, or seasons being current. In other words, *ngaṟanyi* is clearly polysemous between the meanings 'be standing' and 'there is'.

(10) *Puṉu/aṉangu/kapi/waḷi wiya ngaṟa-ngi palula aṟa-ngka.*
tree/people/water/house not there.is-PAST that time-LOC
'There were no trees/people/water/houses at that time.'

The lexeme which expresses the meaning 'there is' quite commonly also functions to express "alienable possession" (i.e., 'have'). This is particularly common in languages of Southeast Asia and East Asia. Formal overlaps between existence, possession, and location have been noted for many years (cf. Lyons 1967). However, the meanings 'there is' and 'have' are kept distinct by the fact that they occur in different grammatical frames: 'there is' is monovalent and can take a "locus" argument, whereas 'have' is bivalent, takes a personal substantive as subject, and generally does not allow a locational adjunct (though this is marginally possible in English). These features can be illustrated from Malay where *ada* expresses both existence and possession.

(11) a. *Ada dua ekor lembu (di padang tu).*
ada (=there are) two CLF cattle (in field that)
'There are two cattle in the field.'
b. *Orang ini ada dua ekor lembu.*
person this *ada* (=has) two CLF cattle
'This person has two cattle.'

It is not uncommon for ‘have’ (alienable possession) to be expressed by a copula verb in a special construction. For example, in Russian the equivalent of ‘He had a book’ is expressed in a construction of the form “at-him was book”. Comparable constructions are common in South Asian languages. That is, the possessor appears in the guise of a locative and the verb is the copula (which, incidentally, is normally omitted in Russian in the present tense). In view of the distinctive morphosyntax of the possessive construction, it is appropriate to recognise a distinct lexical unit of the copula verb with the meaning ‘have’. For a comprehensive survey of how ‘have’ is expressed crosslinguistically, see Verhaar (ed.) (1967–1973).

As humans are quintessentially “tool-making” creatures, it might be conjectured that all languages would have a lexical unit corresponding to ‘make’. It is well known that in many languages one uses the same lexical form for both ‘do’ and ‘make’ (e.g., Malay *buat*, Yankunytjatjara *palyaṉi*, Miskitu *daukaia*, Arrernte *mpware-*, Kalam *g-*), but it seems likely that there are language-internal grounds for positing polysemy in such cases. In particular, the ‘make’ meaning would always be compatible with a “material source” argument (i.e., one could ask the equivalent of ‘What did you make it from?’), whereas the ‘do’ meaning would not be compatible with this kind of oblique argument. One would expect native speakers to clearly recognise that some sentences, e.g., those of the form ‘What did you do/make?’, are ambiguous between the two meanings. Polysemy between ‘do’ and ‘make’ makes sense on the interpretation that the meaning of ‘make’ involves ‘do’ along with other elements, including – crucially – ‘there is’. ‘Making’ something (e.g., a Y) involves purposefully ‘doing something’ with some material with a certain goal in mind, a goal which can be roughly stated as: ‘if I do this, there will be a Y here; I want this’. Some languages have specialised “manufacture” verbs which are normally used to speak of making particular tools and implements, but this does not necessarily mean that a general “verb of making” is absent from the language. So far, I know of no counter-example to the proposal that ‘make’ is a universally lexicalised meaning.

Is ‘give’ a universal lexical unit, notwithstanding the fact that – like ‘make’ – it is unlikely to be a semantic prime? Certainly “giving”, in a broad sense, has a good claim to being a universal of human behaviour. In Newman’s (1997: vii) terms, ‘giving’ arguably combines “experiential basicness” with “internal semantic complexity”. He notes that basic vocabulary lists almost always include ‘give’, that it is one of the first action verbs understood by children acquiring English as a first language, and that a morpheme meaning ‘give’ (or something similar) is found even in restricted small vocabularies of languages like Kalam. Admittedly, a few languages are reported to lack an overt lexical form corresponding to ‘give’. For example, in Amele (Roberts 1987, 1997) affixes which normally attach to a verb stem appear instead as a string, as shown in (12).

This situation could be interpreted as the meaning 'give' being indicated by a zero-morpheme, an interpretation entertained by Roberts (1987). However, in later work Roberts (1997: 25–27) adduces evidence that the indirect object agreement morphology, e.g., the element *ut-* 3SG.IO in the example, is actually functioning as a "relexicalised verb stem".

(12) *Dana uqa=na mel sigin ut-i-a.*
man 3SG=of boy knife 3SG.IO-3SG.SBJ-TODP
'The man gave his son the knife.'

Presumably, "giving", in a broad sense, is related to the proposed semantic prime 'have' (as well as to 'do'). In some languages this is morphologically transparent, e.g., Ainu where 'give' is *kore-e* have-causative (Shibatani 1990: 48). Dixon (1982) includes GIVING as one of his fundamental "semantic types", implying that every language has at least one "verb of giving". The question is, however: Does every language have one "verb of giving" which has precisely the same meaning in all languages? One factor which suggests otherwise is the familiar fact that "giving" can be seen from (at least) two perspectives: from the point of view of the recipient, as in the English construction with the recipient as grammatical object, i.e., *X gave person-Y thing-Z*, or from the point of view of the thing transferred, as in the English construction with the thing as direct object and the recipient as an oblique, i.e., *X gave thing-Z to person-Y*. Arguably these two kinds of construction encode meanings something like those in (13a–b) below.

(13) a. 'give' scenario with recipient (person-Y) as direct object:
X had thing-Z
X wanted person-Y to have it
because of this, X did something
after this, person-Y had thing-Z

b. 'give' scenario with thing transferred (thing-Z) as direct object:
X had thing-Z
X wanted person-Y to have it
because of this, X did something to thing-Z
after this, person-Y had thing-Z

Many languages allow both types of construction, with one or the other usually appearing as the more grammatically "basic" of the pair but there are also languages in which only one construction type is found. For example, French and Maori only have the "thing as direct object" pattern, whereas Ojibwa and Tzotzil only have the "recipient as direct object" pattern (cf. Dryer 1986). If one were to assume that the different construction types express different construals (even in languages with only one construction type), these facts would

militate against the universality of any single 'give' meaning. Admittedly, such an assumption would require independent justification.

A further complication is the existence, in some languages, of separate verbal stems for 'give' depending on the number or nature of items given. Foley (1986: 129) reports that in the Papuan language Kiwai many verbal stems, including the stem for 'give', begin with the vowel *i* to indicate a nonsingular "undergoer". Compare: *agiwai* 'give one (thing)', *iagiwai* 'give some (things)'. Some other Papuan languages have suppletive number-specific variants of this kind, e.g., Barai *m-* 'give one thing', *vaj-* 'give many things' (citations from Newman (1996: 17–19)). Athapaskan languages commonly have a set of verb stems for 'give' which typically encode eight or nine animacy/shape/consistency distinctions in the object, among other things. In a detailed description of Chipewyan, Rice (1997) refers to the "near absence of either a neutral or generic" verb of 'giving' in Athapaskan languages. She reflects that "languages differ greatly on which aspects of transfer events they encode as well as on any other aspects of these fundamental human interactions which might also get elaborated" (Rice 1997: 109). At the moment, then, the indications are that 'give' is likely to be only an approximate lexical universal.

2.5.6. *Sensations and emotions (see, hear, feel).* I know of no convincing counter-evidence to the proposal (Wierzbicka 1996) that 'see' and 'hear' are universally lexicalised meanings, provided that we allow, first, that the lexical forms need not be formally monomorphemic, and, second, that expressions for 'see' and 'hear' are sometimes polysemous with other "cognitive" meanings, such as 'know', 'think', and 'understand' (the example of Yankunytjatjara *kulini*, which is polysemous between 'hear' and 'think', has already been mentioned). In saying this, however, I do not wish to deny that there are some confusing and unresolved issues in certain languages.

The nature of these issues can be illustrated in a particularly sharp form with Kalam (Pawley 1993, 1994). In this language, a single verb *nŋ-* is often used in contexts where in English one would have to choose between 'know', 'see', and 'hear'. Example (14a) shows *nŋ-* used as both 'know' and as 'see' in a single sentence. When *nŋ-* corresponds to 'see' it is possible to substitute in its place the phraseme *wdn nŋ-* (*wdn* 'eye'); and similarly, when it corresponds to 'hear' it is possible to substitute *tumd nŋ-* (*tumd* 'ear'). These substitutions are "marked" in distributional terms, but they are perfectly grammatical and semantically normal (Pawley, personal communication). No "expanded" version is possible when *nŋ-* corresponds to 'know', as in the second instance in (14a), or in a sentence like (14b).

(14)	a.	*Tap*	*ebap*	*nŋbin*	*ak,*	*mey*	*tap*	*kun*	*etp*	*ak*
		thing	one	I.see	that	that	thing	such	what	that

ma-nŋin.
not-I.know
'I see something, but I don't know what kind of thing it is.' (Pawley 1994: 391)

b. *Cn tap kun ak tap tmey ak nŋbun.*
we thing such this thing bad this we.know
'We know that this sort of thing is bad.' (Pawley 1994: 394)

Initially the Kalam data seem to be compatible with several different interpretations. First, it might be proposed that *nŋ-* means 'know', and that the phrasemes *wdn nŋ-* 'see' and *tumd nŋ-* 'hear' are semantic compounds, i.e., that *wdn nŋ-* is 'know by the eyes' and *tumd nŋ-* is 'know by the ears'. This is not an altogether unattractive analysis (cf. Wierzbicka 1980) but there are two problems with it, which for simplicity I will illustrate solely by reference to 'see'. First, it leaves us having to further decompose *wdn* 'eye' without recourse to the meaning 'see' – if we are to avoid circularity, that is. The only way to do this is to rely entirely on the physical characteristics of the eyes; and this leads to complex and counter-intuitive results. Second, there is the fact that 'seeing' does not always lead to 'knowing'. One can see a mirage or a hallucination, for example, and not know anything as a result. It is for reasons like these that Wierzbicka (1996) abandoned her own earlier attempts to explicate 'see' as '(come to) know through the eyes'.

A second proposal would be that *nŋ-* does not mean 'know' but 'perceive'. Thus *wdn nŋ-* would mean 'perceive by the eyes'. This proposal runs into the problem that the meaning of *nŋ-* in contexts like (14b) above cannot be stated as 'perceive': in such contexts it seems to mean 'know'; and 'know' is not the same as 'perceive'. A third proposal would be that the meaning of *nŋ-* cannot be matched with the meaning of any English word, but rather has a Kalam-specific meaning – midway, as it were, between 'know' and 'perceive'. This is Pawley's preferred interpretation. This position is difficult to refute. It may indeed be empirically unfalsifiable. One problem with this position is that it offers no explanation for the fact that *nŋ-* does appear to mean precisely 'know' when it is used with a "sentential" complement clause, as in (14b). Adopting the position that *nŋ-* has a Kalam-specific meaning entails accepting semantic incommensurability between Kalam and English (and, probably, most languages of the world). Such a radical conclusion, it could be argued, should not be endorsed until all other possible avenues of explanation are exhausted.

A fourth, and final, proposal, is that *nŋ-* is polysemous, with three discrete and separable meanings: 'know', 'see', and 'hear'. To assess the credibility of this position on Kalam-internal grounds, some additional data would be required, bearing on the possibility that each of three putatively separate meanings has certain distinctive syntactic properties. It would be relevant to know,

for example, how sentences like the following could be expressed in Kalam: (i) 'X knows something about you', (ii) 'X saw some people in this place', (iii) 'X heard these words'. These sentences embody hypotheses about distinctive syntactic (combinatorial) characteristics of the proposed semantic primes 'know', 'see', and 'hear' (cf. Goddard & Wierzbicka (eds.) in press). The implication of (i) is that the meaning 'know' is compatible with an optional "topic" argument (realised in English by means of the preposition *about*). It is expected that 'see' and 'hear' do not possess this option on a universal basis (admittedly, in English one can *hear something about someone*, but this is an English-specific construction which is decomposable in terms of 'hear' and 'know about'). The implication of (ii) is that the meaning 'see' is compatible with a "locus" expression (such as *in this place*) in its complement. Presumably nothing comparable is possible for 'know' or 'hear'. Finally, the implication of (iii) is that the meaning 'hear' can take an expression involving *words* as its complement. In my estimation, the available data on Kalam is not sufficient to decide between the rival proposals: a uniform but Kalam-specific meaning for *nŋ-* vs. three-way polysemy justified on syntactic evidence.

A final challenge to the status of 'see' and 'hear' as universal meanings is that many languages do not make a lexical distinction between 'see' and 'look', or between 'hear' and 'listen', i.e., between the proposed primes and corresponding "active" meanings (which presumably involve 'do' in addition to the basic prime). In such languages, the range of use of the nearest equivalent to 'see' or 'hear' is typically broader than that of its English counterpart; for example, Yankunytjatjara *nyanganyi* 'see' has a broader range of use than English *see*. This does not necessarily mean, however, that there is a paraphrasable semantic difference between *nyanganyi* and *see*. As far as I know, it is impossible to state any such difference. The point is that the mere existence, in English, of a semantically "active" counterpart of *see*, namely *look*, does not necessarily mean that English *see* is semantically specified as "inactive" (i.e., that it contains a component involving 'not doing'): it may simply be semantically unspecified in this respect. Comparisons with English *see* are complicated, however, by some other unusual properties of the English word. For example, English *see* cannot be combined readily with durative expressions (**I saw it for a long time*) or used to express "imperfective" meanings (the sentence *I see your house*, for example, suggests a momentary event). To express comparable durative or imperfective meanings the locution *can see* is used (e.g., *I could see it for a long time, I can see your house*). Space does not permit us to pursue these matters in detail here.

The proposed semantic prime 'feel' is neutral to the distinction between "emotion" and "sensation", and is the common foundation, so to speak, for terms of both these kinds. Emotion terms (*sad*, *angry*, *excited*, etc.) involve 'feel' in combination with cognitive verbs such as 'think' and 'want' (and other

elements). Sensation terms (*hungry*, *cold*, *itchy*, etc.) involve 'feel' in combination with 'want' and 'body' (and other elements). Another lexical field involving 'feel' is that of "tactile" words (*rough*, *smooth*, *sharp*, etc.) which involve similar components, along with 'touch'.

Perhaps because of its frequent compounding with elements of other kinds, lexemes for 'feel' are often polysemous and/or formally complex. For example, English *feel* can convey not only the primitive meaning, but also a "cognitive" meaning (e.g., *He felt it was wrong*), and, in a different syntactic frame, a "touch-related" meaning (e.g., *She felt his pulse*). In Malay, *rasa* 'feel' can also mean 'taste', as well as conveying a similar cognitive meaning to that found in English (presumably involving a combination of 'feel' and 'think'). Similar polysemies involving 'feel' can be found in Ewe, Lao, French, and Mandarin Chinese.

To illustrate some of the formal complexities found with terms for 'feel', one can consider the following three formations, from Yolngu Matha (Australia), Lao, and Ewe, respectively: (i) *dhäkay-ŋänha* 'taste-hear' (Yolngu Matha; Cooke & Goddard 1997), (ii) *huu*^{4}*.sùk*^{2} 'know-[unknown root]' (Lao; Enfield in press), (iii) *se sese le-lãme* 'hear feeling in-body' (Ewe; Ameka 1994). It is also notable that in various languages the lexeme for 'feel' is identical to a body-part word, typically the word for 'liver', 'heart', or 'stomach', or less commonly a more general term for 'insides', as with Mangaaba-Mbula *lele-*. For example, in Yankunytjatjara (Goddard 1994), the expression for 'I feel good' is *Ngayulu tjuni palya*, which can be glossed word-for-word as 'I stomach (=feel) good'. Sentences like this have often been interpreted as figurative, and no doubt in some languages they are – when there is another language-internal way of stating the literal meaning. For example, Malay has numerous emotion-related expressions employing *hati* 'liver' (Goddard in press c), and English has a fair number employing *heart*. In both Malay and English, however, there are other words (*rasa* and *feel*, respectively) which express the prime 'feel' pure and simple, as it were. Not so in Yankunytjatjara, where locutions with *tjuni* are the plainest and simplest way of expressing the meaning.

2.5.7. *Bodily postures (sit, lie, stand) and activities (eat, drink).* 'Sit' is not a universally lexicalised meaning, if only because in some languages the verb which covers sitting (in the English sense) also covers squatting on one's haunches (without the bottom touching the ground). This applies to the Lao word *nang*^{1} and the Tagalog word *upo*. In the case of Lao, the squatting posture is probably the prototypical "sitting" position. The universal status of 'stand' can probably be disconfirmed on the basis of European (Romance) languages. In French, for example, there are two expressions which are near-equivalents to English *be standing*, both based on the adverb *debout* 'being upright'. The expression *être debout* 'be *debout*' merely describes one's posi-

tion, whereas the more common expression *se tenir debout* '(lit.) reflexive hold *debout*' conveys the idea of an activity (something akin to "holding" or maintaining a posture). It seems that either French expression conveys a slightly different meaning to that conveyed by English verb *stand*. So far, I know of no concrete counter-evidence to the proposition that 'lie' (be lying) is a universally lexicalised meaning. However, it has been suggested to me that some languages with highly elaborate systems of "positional" marking (such as Mayan languages) may insist on encoding the difference between, say, 'lying prone', 'lying on the side', and 'lying face up'.

Aside from postures, the most plausible candidates for universals in the domain of body actions are probably 'eat' and 'drink', but their claim to universal status is doubtful. Some languages do not have separate words for 'eat' and 'drink'; for example, Kalam *ñŋ-*, Yimas *am-*, Warlpiri *ngarni* can all be glossed 'ingest, consume' (Pawley 1993, Foley 1997, Laughren & Hale forthcoming). In some cases these languages have phrasemes roughly corresponding to 'eat' and 'drink'; e.g., Kalam *tap ñŋ-* 'food consume' (Pawley 1993: 107), but the Kalam phraseme does not mean precisely the same as English *eat*, if only because *eat* can be used about ingesting solids other than food (e.g., *The baby was eating sand*).

Even when a language has a word which is usually glossed as 'drink', the semantic correspondence need not be precise. For example, Japanese *nomu* can be used not only about drinking water, tea, coffee, etc. but also for swallowing solid items such as pins and rings, and for smoking a cigarette. Suzuki (1978: 17–19) argues that *nomu* means 'to introduce something into one's body without chewing it'. He also notes that rice is normally something to *taberu* 'eat', but if a fish bone is stuck in someone's throat, one says in Japanese, "You should *nomu* some rice". The universality of 'eat' and 'drink' is also challenged by the fact that some languages distinguish between these activities as done by humans (or, in a human-like fashion) as opposed to animals, e.g., German *essen* vs. *fressen*.

2.5.8. *Sensations (hot, cold, hungry, thirsty, sweet) and emotions (fear, anger, etc.).* It seems obvious that all languages must have words to refer to the kinds of sensations indicated by the English words *hot* and *cold*. The question is: Can we find in all languages lexical units with PRECISELY the same meanings as the English terms or is the correlation only approximate? The evidence favours the latter position. First, English draws a distinction between *hot* and *warm*, and between *cool* and *cold* (if something is *warm*, it is not *hot*; if it is *cool* it is not *cold*). In languages which lack these distinctions, the nearest equivalents to *hot* and *cold* have a broader range (e.g., Yankunytjatjara *waṟu* 'hot, warm', *wari* 'cold, cool'). Second, languages may make finer distinctions even in the realm of high (or low) temperatures. For example, Russian distinguishes *gorjačij*

(roughly) 'hot perceived by touch' from *žarkij* (roughly) 'hot perceived via the air' (Koptjevskaja-Tamm & Rakhilina 1999).

So far I am not aware of any language which lacks an equivalent to 'hungry' or 'thirsty'. It is not necessarily a counter-example to the universality of 'hungry' and 'thirsty' that in German, for example, these meanings are usuallly expressed by nouns in combination with *haben*. The expressions *Hunger haben* and *Durst haben* can be regarded as lexical units, equivalent to 'be hungry' and 'be thirsty', regardless of the formal differences. Nor is it a counter-example that in some languages the relevant expressions are formally complex, either consisting of complex words (e.g., Yankunytjatjara *anymatjara*, where *-tjara* 'having') or of a phraseme (e.g., Yankunytjatjara *kapi ilu* lit. 'water die').

"Basic tastes" like 'sweet', 'bitter', 'sour', and 'salty' may have some claim to universal status, on account of the fact that receptor cells for these tastes are localised in particular regions of the human tongue. The best candidate is probably 'sweet', given that sweet foods are supposed to be universally valued by human beings, but it can be ruled out as a precise lexical universal on the evidence of Cantonese and Japanese. The nearest Cantonese equivalent *tim*^{4} can be used not only about sweet food, but about other nice-tasting food and drinks generally. The nearest Japanese equivalent *amai* can be used not only about sweet tastes but also about "mild" tastes (Backhouse 1994). One language which is known not to have any word even approximating 'sweet' is Kayardild, traditionally spoken on isolated Bentinck Island in the Gulf of Carpentaria (Australia). The Kayardild traditionally lacked any sweet foods, even honey. Today, the word *rirrk* 'fat, grease, rich food' is used for honey as well (Nicholas Evans, personal communication). There is little chance that any of the other taste sensations are lexical universals, given that 'bitter', 'sour', and other "bad tastes" are all covered by one word *kumarlpa* in Warlpiri (Laughren & Hale forthcoming).

Moving now to emotions, there is a school of thought in psychology, associated primarily with Paul Ekman (e.g., 1992, 1993), which holds that there is a small set of physiologically in-built "basic emotions", such as fear, anger, sadness, disgust, surprise, and joy. It is often claimed, in connection with this proposal, that all (or most) languages have words for these basic emotions. However, a substantial body of anthropological and semantic evidence indicates that at best there is only an approximate match across languages between the meanings of basic emotion terms (cf. Russell 1991, Wierzbicka 1986, 1992, Goddard 1996b, 1997c, Harkins 1996). Some of them are known to have no equivalents, not even near-equivalents, in particular languages; for example, Yankunytjatjara has no near-equivalent to 'disgust', Tahitian has no near-equivalent to 'sadness'. In other cases, it can be demonstrated that an apparent equivalence is only approximate. This is easiest in languages which have more than one near-equivalent for a particular putative basic emotion, e.g., Yankuny-

tjatjara *pikaringanyi* and *mirpan̲arinyi* for 'angry', Malay *terkejut* and *terperanjat* for 'surprise', German *Angst* and *Furcht* for 'fear'. Even when there is a single near-equivalent, semantic differences can often be detected; for example, it can be shown that Russian *grust'* is not precisely identical with *sadness*, that Italian *rabbia* is not precisely identical to *anger*, that Malay *malu* is not precisely identical to *ashamed*, and so on.

2.6. *The domain of time*

It is frequently asserted, in connection with the Sapir-Whorf hypothesis, that there are languages which have a "fundamentally different concept of time" to that of English. Whorf (1956) set the tone with his celebrated assertions about Hopi being a "timeless language", but many other writers have stated that the concept of time can vary greatly between languages and cultures. According to anthropologists such as Claude Lévi-Strauss, Marshall Sahlins, and Clifford Geertz, many tribal peoples have cyclical or atemporal cosmologies which are very different to those of the Western world-view.

That cultures differ greatly in the extent and manner in which they elaborate basic temporal notions, and in the functional role which temporal notions play in the culture, is beyond doubt. Contemporary "Anglo" culture, in particular, displays a veritable obsession with time, as many observers have noted. It has invented numerous methods of measuring and arranging times (clocks, calendars, schedules, etc.), and employs these for regulating and ordering countless aspects of everyday life. But cultural elaboration is one thing, and semantic fundamentals are another. Differences in the cultural construction of history, ritual, and myth, do not necessarily reach right down to the bedrock linguistic encoding of temporal notions in everyday talk (cf. Keesing 1994). This point is perhaps best illustrated with Hopi, which, contrary to the implications of Whorf's claims, certainly has equivalents to the proposed semantic primes of time (cf. Malotki 1983, Goddard 1997b).

2.6.1. *Deictic and "categorical" temporal meanings (now, when/time).* All languages appear to have a term with the meaning 'now'; e.g., French *maintenant*, Malay *sekarang*, Cantonese *ji*^{4}*gaa*^{1}. Sometimes this word is polysemous, having also the meaning 'today', as with Yankunytjatjara *kuwari* or Hopi *pu'*. Sometimes it is polysemous, also having the meaning 'this' (and/or 'here').

As far as we know, it is possible in all languages to ask the equivalent of 'When did it happen?', and to receive an answer like 'It happened at this time'. Of course, one would not expect to find (in any language) words which are equivalent in every way to the English word *time*, with an identical range of polysemic meanings and uses; for example, its uses as an abstract noun (e.g.,

We didn't have time, Time flies, Times have changed), its role in phrasemes such as *a long time*, in compounds such as *lunchtime*, and so on. These English-specific usages do not represent examples of the proposed semantic prime 'when/time', which need only occur in a narrow range of basic, and putatively universal combinations, such as 'I know when it happened', 'it happened at this time', and 'they did it at the same time'. This can be illustrated from Hopi (Malotki 1983). The equivalent to 'when', both as an indefinite and as an interrogative, is *hisat*, as shown in (15a). Morphologically this word is analysable as a question formative *hi-* (much like English *wh-*) and *-sat* 'time'. In particular, *-sat* 'time' can combine with the demonstrative *yàa-* 'this' to form the expression *yàa-sat* 'at this time' as shown in (15b). An allomorph of *-sat*, namely *-saq*, combines with the Hopi equivalent of 'the same' *suu-/sú-*, to form expressions meaning 'at the same time', as in (15c).

(15) a. *Pam hisat nima?*
that when go.home
'When did he go home?' (Malotki 1983: 305)

b. *Taavok yàa-sat=haqam ay nu' tsöng-moki.*
yesterday this-time=APPROX ASSR I hunger-die
'Yesterday at about this time I got really hungry.' (Malotki 1983: 146)

c. *Pam sú-'inùu-saq nakwsu.*
that the.same-I-time start.out
'He started out at the same time as I.' (Malotki 1983: 144)

In some languages there are two words which appear to correspond with 'time' (in the appropriate sense). This is the case in some Sinitic languages, including Cantonese (*si*^{4}*hau*^{6} vs. *si*^{4}*gaan*^{3}) and Mandarin (*shíhou* vs. *shíjiān*). When looked at closely, however, it turns out that one of the words is semantically complex, implying a "length" or "segment" of time, whereas the other is vague, i.e., semantically unspecified (Chappell 1994: 135–136, Tong et al. 1997). The Cantonese and Mandarin equivalents of 'time' in its most basic sense are *si*^{4}*hau*^{6} and *shíhou*, respectively. The examples below illustrate Cantonese constructions.

(16) a. *Go*^{2} *go*^{3} *si*^{4}*hau*^{6}, *ngo*^{5} *mi*^{1}*dou*^{1} *m*^{4} *zi*^{1}*dou*^{3}.
that CLF time I anything not know
'At that time, I didn't know anything (about it).'

b. *Tung*^{4}*jat*^{1} *si*^{4}*hau*^{6} *faat*^{3}*sang*^{1}.
the.same time happen
'(It) happened at the same time.'

The status of "frequency" expressions such as 'once' (one time), 'twice' (two times), and 'often' (many times), remains unclear. In English, the same

lexical form (i.e., *time*) is used for saying when something happened, and for saying how often it happened, an alignment which is not by any means unique. For example, Yankunytjatjara *a̲ra* is used in both contexts: *a̲ra nyangangka* 'at this time', *kutjupa a̲ra* 'two times'. But in many – perhaps most – languages, a different word is used in these two contexts, as illustrated in (17) with examples from Cantonese, Malay, Lao, and French.

(17)			
	Cantonese:	*si*^{4}*hau*^{6}	*ci*^{4}
	Malay:	*masa*	*kali*
	Lao:	*tòòn*^{3}	*thùa*^{1}
	French:	*moment*	*fois*
		'[at this] time'	'[two] times'

It has been suggested (Tong et al. 1997: 248) that "frequency-time" may be a distinct concept – a different prime – to "when-time", but more research is needed on this question.

2.6.2. *Sequence (before, after).* It appears that semantic equivalents to 'before' and 'after' can be identified in all languages. In Hopi, they appear to be the postpositions *-pyeve* and *-ngk*, respectively (though the temporal particles *angwu* 'beforehand' and *ason* 'later, afterwards' are closely related). As in many languages, the Hopi expressions for 'before' and 'after' also have spatial meanings – 'going ahead of' and 'following', respectively.

(18) a. *Pam put hinhin a-pyeve tìi-ti-wa.*
that that somewhat he-before child-CAUS-PASS
'He was born a little bit before him.' (Malotki 1983: 107)

b. *Puma pay hìisap itamu-ngk öki.*
those ASSR a.short.time we-after arrive
'They arrived a little after us.' (Malotki 1983: 144)

A potentially confusing instance of polysemy occurs in languages in which the expression meaning 'before' can also be used to mean 'first'. This is the case, for example, with the Samoan word *muamua* (Mosel 1994). Furthermore, Samoan *muamua* is a verb, normally glossed as 'come/go first, be first'; cf. sentences like *Muamua 'oe, 'ae mulimuli a'u* 'Go first, and I will follow'. It might seem, then, that the nearest Samoan equivalent to English 'before' is not an exact equivalent at all, but, when examined more closely, this idea does not stand up. We can see this by looking at a context which concerns a series of three or more events; for example, the birth of three siblings, A, B, and C. In this kind of situation it makes sense to say that A was born 'first' (i.e., before all the others), but it doesn't make sense to say that B was born 'first' in respect of C. But in fact Samoan *muamua* can still be used in such a situation, to specify

the relative ordering of events B and C. Faced with this situation, it is tempting to respond that *muamua* still really means 'first', but in a relative sense: 'first with respect to C'. However, this interpretation isn't really coherent. What is the difference in meaning between 'first with respect to C' and 'before C'? Surely the answer is: No difference at all.

It is commonly reported that the words for 'before' and 'after' are polysemous, expressing also locational or motional meanings such as 'in front' or 'ahead', and 'back' or 'following', respectively. In many cases it can shown that lexemes for 'before' and 'after' are etymologically derived from body-part nouns or verbs of motion. Such formal overlaps and recurrent patterns of semantic change certainly reflect the fact that there are intimate links – both semantic and experiential – between time, space, and motion. The structure of the human body, which can move forwards much more easily than it can move backwards, is responsible for establishing a number of correspondences of this type. But the formal overlaps between exponents of time, space, and motion, no matter how motivated they may be, do not mean that temporal meanings can be reduced to spatial or motional meanings.

Bohnemeyer (1998a, b) has claimed that Yucatec Maya has no lexical equivalents to 'after' and 'before'. Communicating about temporal sequencing is achieved, he argues, by implicatures generated largely from aspectual or phase verbs (Phrasenaktionsarten) such as *ts'o'k* 'end, finish, cease' and *ho'p'* 'begin, start', in combination with verbal aspect. Thus, though (19a) strongly implicates a certain temporal sequence, as in gloss (ii), its true semantic structure is better indicated as in gloss (i). Bohnemeyer rejects a suggestion from Wierzbicka that in such uses *ts'o'k'* 'end' is functioning literally as an exponent of 'after', pointing to the fact that if the order of clauses is reversed, as in (19b), the interpretation of the temporal sequence also changes – a result that one would not expect if *ts'o'k'* simply meant 'after'.[9]

(19) a. *Pedro-e'* *káa* *h* *ts'o'k* *u*
Pedro-TOP káa PFV end(B.3SG) A.3
ts'íib-t-ik *le* *kàarta-o'* *káa* *t-u*
write-APP-INC(B.3SG) DEF letter-DIST káa PFV-A.3
ts'u'ts'-ah *hun-p'éel* *chamal.*
suck-CMP(B.3SG) one-CLF.INAN cigarette
(i) 'Pedro, when he had finished writing the letter, he smoked a cigarette.'
(ii) 'After Pedro wrote the letter, he smoked a cigarette.'

9. For expository clarity I have slightly varied the presentation of Bohnemeyer's (1998a: 255) free glosses for those examples.

b. *Pedro-e'* *káa* *t-u* *ts'u'ts'-ah* *hun-p'éel*
Pedro-TOP káa PFV-A.3 suck-CMP(B.3SG) one-CLF.INAN
chamal-e'. *káa* *h* *ts'o'k* *u*
cigarette.TOP káa PFV end(B.3SG) A.3
ts'íib-t-ik *le* *kàarta-o'.*
write-APP-INC(B.3SG) DEF letter-DIST
'Pedro, he smoked a cigarette, and finished writing the letter.'

However, things are not quite as clear-cut as they may seem. First, the phase verb *ts'o'k* is not in the same grammatical environment in (19a) and (19b): in the latter sentence it is in the main clause, whereas in the former it is in an adjoined subordinate clause, roughly approximating the status of an adverbial clause. Second, as Bohnemeyer himself stresses, the phasal verb *ts'o'k* differs strikingly from its putative near-equivalents (such as 'end' and 'finish') in Indo-European languages in that it is readily compatible with punctual events. Thus, there is nothing unusual in Yucatec Maya in combining *ts'o'k* with a punctual verb such as *ah* 'wake up', as in (20), though would-be "literal" glosses as in (i) are semantically odd in English. This phenomenon is all the more striking since it is not shared by other egressive phase verbs in Yucatec Maya, such as *ch'éen* 'stop' and *xúul* 'stop', which would be ungrammatical if substituted for *ts'o'k* in (20).

(20) *Káa* *h* *ts'o'k* *inw* *ah-al-e',* *káa* *h*
káa PFV end(B.3.SG) A.1SG awaken-INC-TOP káa PFV
líik'-en *uk'ul.*
lift.ACAUS-B.1SG drink.ATP
(i) '(When) I finished waking up, I rose to have breakfast.'
(ii) 'After I woke up, I rose to have breakfast.'

Bohnemeyer (1998a: 270) comments:

> *Ts'o'k* must be assumed to represent a type of phasal operator unattested and probably unparalleled in Indo-European languages. The terminal boundary it imposes on the interpretation of the target event in discourse is purely referential in nature and simply ignores the lexical event structure properties of the embedded verb. This underlines the status of *ts'o'k* as an operator of temporal coherence rather than merely a lexical verb, which in turn may lend some plausibility in retrospect to Wierzbicka's suggestion [...] that *ts'o'k* represents a temporal operator – to be precise, a connective – rather than a phase verb.

Ts'o'k, and its "ingressive" counterpart *ho'p*, also have some grammatical properties not shared by other phase verbs in the language. It is therefore possible – in my view, likely – that *ts'o'k* and *ho'p* will turn out to be poly-

semous, functioning both as lexical phase verbs and as temporal operators, i.e., semantic equivalents to 'after' and 'before'.

2.6.3. *Duration (a long time, a short time, for some time).* On presently available evidence, no language is known to lack a lexical unit with the meaning 'a long time'. Interestingly, the expressions are not typically phrasemes (as in English), but formally simple expressions such as Malay *lama*, Russian *dolgo*, Lao *don*^{3}. An interesting example of polysemy is furnished by Hopi, where the word *hisat* not only means 'when' (as mentioned above) but also 'a long time', as in (21).

(21) *Nu'* *pay* *hisat* *tsoo-tsong-ngwu.*
I ASSR a.long.time RDP-pipe-HAB
'I've been smoking for a long time.' (Malotki 1993: 155)

Further evidence that *hisat* can mean 'a long time' comes from other formations. For example, when adjectivalised by suffix *-wa*, the resulting word *hisatwa* means 'old/ancient'; and *hisat* also enters into various compounds with nouns, again meaning 'old/ancient', e.g., *hisat-himo* 'old things', *hisat-sinom* 'the old people [of long ago]' (Malotki 1983: 159–160).

It is not clear whether 'a short time' is a universally lexicalised meaning. No definitive counter-example is yet known, but it is known that equivalents to 'a short time' are often not morphosyntactically parallel to those for 'a long time'. For example, in Cantonese 'a long time' is an adverbial expression *(hou*^{2}*) noi*^{6}, but 'a short time' is expressed by a *jat*^{1}*zan*^{6}, where *jat*^{1} is 'one' and *zan*^{6} is a specialised classifier associated with brief events (e.g., *jat*^{1} *zan*^{6} *fung*^{1} one CLF wind 'a burst of wind', *gei*^{1} *zan*^{6} *jue*^{5} few CLF rain 'a few showers'). Furthermore, expressions for 'a short time' are sometimes subject to combinatorial restrictions not shared by 'a long time'. To illustrate again from Cantonese, *noi*^{6} 'a long time' can be modified by *hou*^{2} 'very' or *fei*^{1}*soeng*^{4} 'extremely', but it is impossible to combine these intensifiers with *jat*^{1}*zan*^{6}. The status of the "indefinite" durational expression 'for some time' is also unclear, though many languages are known to have such an expression, which is often glossed in dictionaries as 'for a while'.

2.7. *The domain of space*

Till recently it was assumed that the linguistic encoding of space would not vary greatly from language to language, on account of the presumed universal physiological basis for spatial perception. Lately however, a rich new body of crosslinguistic data has shown that there is "astonishing variation between languages in the way they express and schematicize space" (Sinha 1995: 7). So much so that some linguists are now suggesting that languages may dif-

fer "fundamentally" in their ways of describing spatial information (cf. Foley 1997). However, despite striking differences in the semantic elaboration of spatial semantics, and in the grammatical systems employed to express spatial meanings, it appears that there are a number of strong candidates for the status of universally lexicalised meanings in the spatial domain.

2.7.1. *Deictic and "categorical" spatial meanings (here, where/place).* All languages appear to have a term with the meaning 'here'. Quite commonly, the forms involved are related to the form for 'this' (e.g., Malay *sini* 'here', *ini* 'this'). Sometimes they may appear to be compositionally derived from 'this' and 'place', as with Cantonese *ni*^{1}*dou*^{6} (*ni*^{1} 'this', *dou*^{6} 'place). Evidence that *ni*^{1}*dou*^{6} is fully lexicalised with the meaning 'here' is provided by sentences like (22) below, which would hardly make sense as 'this place is far from this place'.[10] The Cantonese situation of one-word compound *ni*^{1}*dou*^{6} 'here' vs. two-word phrase *ni*^{1} *dou*^{6} 'this place' can be compared with the situation of English *something* vs. *some thing*.

(22) *Ni*^{1} *dou*^{6} *lei*^{4} *ni*^{1}*dou*^{6} *hou*^{2} *jyun*^{5}.
this place from here DEG far
'This place (e.g., Gundaroo) is far from here.'

It appears that 'where' is a universal meaning, at least as an interrogative; that is, that in all languages one can say the exact equivalent of 'Where is X?' It also appears to be true that all languages have lexical units which indicate a corresponding indefinite sense. Sometimes these are separate lexemes (as with English *where* and *somewhere*), but in many languages the same lexeme is used in both interrogative and indefinite uses. It is highly probable that all languages have an expression corresponding in meaning to 'place', at least in "specified" contexts such as 'this place' and 'the same place'. It has been claimed, however, that in some languages the "place word" can only be used about "socially significant places", i.e., about places linked to 'people'. Harkins & Wilkins (1994: 299–300) make this claim about Arrernte *pmere*.

2.7.2. *Vertical dimension (above, below).* As far as I know, all languages have words which express the spatial relational meanings 'above' and 'below', in contexts such as 'X is above Y', 'Y is below X'. It is well known that in many languages the lexemes involved are "noun-like", e.g., they bear nominal affixes which cross-reference the person/number of the relatum. In such languages, the expressions for 'above' and 'below' are usually designated as

10. It is true that a sentence like 'this place is far from this place' can make sense if supplemented by deictic gestures; but the Cantonese example does not require any such gestures.

"locational nouns". There are also languages in which the spatial relational terms are assimilated to "verb-like" morphosyntax. These facts do not necessarily impeach the claim of 'above' and 'below' to the status of lexical universals. For example, despite the fact that in Tzeltal *-ajk'ol-al* 'above' and *-alan-il* 'below' normally take "possessive" cross-referencing prefixes (e.g., 3SG *s-/y-*), there does not appear to be any specifiable semantic difference between, for example, Tzeltal *ta y-ajk'ol-al na* and English *above (the) house*; or between Tzeltal *ta y-anil-il mexa* and English *below (the) table* (cf. P. Brown 1994 and personal communication, Levinson 1994).

In some languages, polysemy may confuse the situation. One common pattern is for lexemes with relational meanings 'above' and 'below' also to have "substantive" meanings such as 'top' and 'bottom', respectively; or for them also to have positional meanings such as 'high' and 'low', respectively. In languages I have examined from this point of view, language-internal semantic analysis shows that a polysemy analysis is necessary, and that the explicitly relational meanings (i.e., 'above' and 'below') are semantically prior to the other meanings (cf. Hill & Goddard 1997).

One phenomenon which could upset the claim to universality of 'above' and 'below' (and several other proposed lexical universals) is the possibility that in some languages these meanings are "distributed" across several lexical items. This is dealt with in Section 2.7.4 below.

2.7.3. *Laterality (on (one) side, left, right).* All languages appear to have an expression meaning 'on (one) side of', which can appear in a frame like 'X is on (this/that) side of Y'. Often there are two or more such words, some of them also incorporating a reference to nearness (e.g., English *beside*).

It is well known that in many languages (e.g., in Australia, Papua New Guinea, Mesoamerica) the words for 'left' and 'right' are not used to indicate direction or orientation, and that in such languages people use an "absolute" frame of reference, employing terms like 'north' or 'south'. It appears, however, that all languages have words which can be used to distinguish one hand from the other. This applies for example to Tzeltal, a language which has been much written about in connection with its preference for absolute frames of reference; cf. Tzeltal *xin* 'left', e.g., *xin k'ab(al)* 'left hand', and *wa'el* 'right', e.g., *wa'el k'ab(al)* right hand' (cf. Levinson & Brown 1994). Some languages, e.g., Tzeltal's better studied neighbour Tzotzil, are reported to have only a term for 'left-hand' (in this case *tz'et*), with the right-hand being designated as the 'true' or 'correct' hand *batz'il k'obtik*, but this is probably a case of polysemy, not substantially different to the polysemy of *right* in English.

In some languages there are terms which literally mean 'left-hand' and 'right-hand' (and/or 'left-handed' or 'right-handed'), but which are not normally used as modifiers about other body-parts, such as legs, or eyes, or ears. This is the

case with Yankunytjatjara *tjampu* and *waku*, and Arrernte *alyenge* and *akwe-arratye*, which mean 'left-hand, left-handed' and 'right-hand, right-handed', respectively. It would seem, therefore, that while all languages have the resources for distinguishing between 'left' and 'right' (by locutions such as 'on the same side as the right-hand'), it would not be correct to say that 'left' and 'right', as general modifiers, are precise universals.

2.7.4. *Topological relations (on, inside).* It is clear that 'on' (support and contact) is not a universally lexicalised meaning. Many languages lack any element which can be used indifferently in the three situations described by English *on the table*, *on the wall*, and *on the ceiling*. For example, in German one uses *auf* for the first of these (support from below), but *an* for the other two. Also, in many languages, e.g., Italian, Malay, and Cantonese, there is no differentiation between 'X is on Y' (as in 'on the table') and 'X is above Y'.

The situation is very different in regard to 'inside'. Most languages seem to have a reasonably straightforward equivalent, although as mentioned above, this equivalent may be morphosyntactically noun-like or verb-like. It is, of course, not a counter-example to the universality of 'inside' to observe that the Spanish preposition *en* can be used in both "in-situations" and "on-situations", because there is a clear Spanish equivalent of 'inside' in the expression *dentro de* 'inside of'. Nor is it necessarily a problem if the relational meaning 'inside' happens to be expressed by a word which can also express a more concrete, nominal meaning, such as 'stomach, belly', as in Mixtec and some other Mesoamerican languages, cf. Brugman & Macaulay (1986), Langacker (1999). In such cases, there are usually language-internal arguments in favour of a polysemy analysis. That is, it can be established that 'inside' exists as the meaning of a distinct lexical unit.

A more serious objection to the universal status of 'inside' is posed by reports that in some languages, particularly Australian languages, the word for 'inside' is also used to cover the meaning 'under' and, often, 'low' and 'below', as well (Evans & Wilkins 1995). For example, in Arrernte a single term *kwene* covers both 'inside' and 'below' (as well as 'low'): *X Y-locative kwene* is the Arrernte equivalent for both 'X is inside Y' and for 'X is below Y'. However, there is some language-internal evidence for polysemy. For example, when the allative suffix *-akerle* is added, the resulting word *kwene-akerle* can only mean 'downwards' – it cannot mean 'into' or 'inwards'. Also, the inchoative motion verb *kwene-irreme* can only mean 'go down, duck down' (not 'go inside'); and there is a derived form *kwenengenenge* which is glossed as 'right in, completely inside' (Henderson & Dobson 1994). However, the situation requires more research.

Another challenge to the universality of 'inside' – and of various other spatial meanings – is posed by the fact that in some languages the meaning is

apparently "distributed" across two (or more) lexical items. For example, in Ewe (Ameka 1995) spatial adjuncts normally require both a postposition and a preposition. Examples (23a–c) show that static locational relationships are signalled by the "general" preposition *le* (glossed for the time being as 'at') together with a postposition (e.g., *me*, *té*, or *gbɔ*) which indicates the specific nature of the spatial arrangement. Preposition *le* has evolved from the locational verb *le* 'be somewhere'. The postpositions have evolved from nouns, often from body-part nouns.

(23) a. *É-fle agbalẽ lá* [*le fiásé me*].
3SG-buy book the at shop inside
'She bought the book in the shop.'
b. *Me-kpɔ-e* [*le máŋgɔ-tí té*] *kpɔ.*
1SG-see-3SG at mango-tree under PFV
'I once saw it under a mango tree.'
c. *É-vá dze* [*le Atsú gbɔ*].
3SG-come land at Atsu near
'It (a fly) came and landed near Atsú.'

It can be argued (cf. Ameka 1995) that there is a semantic "division of labour" between the two adpositions, the preposition indicating that a locative relationship exists and the postposition designating a substantive meaning ("place as entity"). If so, then neither adposition could be regarded as semantically equivalent to monolexemic (and putatively prime) meanings such as 'inside', 'below', and 'near'. The situation is not as straightforward as the description so far would suggest, however. In fact preposition *le* is required only when a static locational phrase occurs as a syntactic adjunct, but not when the locational phrase occurs in predicative function, either with the verb *le* 'to be somewhere' as in (24a), or with spatial predicates such as *mlɔ* 'lie', as in (24b).

(24) a. *X le Y me / gbɔ.*
X be.at.PRES Y inside / near
'X is inside/near Y.'
b. *Dadi lá mlɔ* [*aba dzí*].
cat the lie mat surface
'The cat is lying on the mat.'

Furthermore, *le* is a multi-purpose preposition, being found not only with (static) locative adjuncts, but also with adjunct phrases of other kinds, including temporal, causal, degree, and several other semantic roles (Ameka 1995: 160). Consequently it is possible that preposition *le* is best analysed as a syntactic marker of adjunct status, having an indexical function, rather than expressing

a specifiable (paraphrasable) meaning. In this case, the Ewe postposition *me* could be regarded as semantic equivalent to 'inside'. The matter requires further investigation.

2.8. *Logical concepts*

2.8.1. *Negation (not).* Linguists generally accept that negation is a fundamental – and universal – element of human language. Some languages lack any interjection corresponding to *no*; for example, to deny or refuse something in the Amazonian language Jarawara one must repeat the predicate, combining it with a clausal negator (R. M. W. Dixon, personal communication). It does appear, however, that all languages have at least one clausal negator, corresponding semantically to 'not'.

The universality of 'not' is not contested by the fact that some languages have a series of specialised portmanteau negators, in addition to a semantically simple negator. An outstanding example is Minnan (Taiwan), which has no less than eight negative words (Chappell 1994). As well as the simple negator/negative adverb *bo*^{24} 'not' (which can also function as a verb, meaning 'there is not' and 'not have'), the other Minnan negators are *m*^{22} 'don't, not want to', *bue*^{22} 'can't', *be*^{22} 'not yet', *mai*^{24} 'don't do', *bian*^{53} 'no need', *boai*^{21} 'not want to', and *mmo*^{53} 'better not'. The standing of *bo*^{24} as the simplest negator is fairly clear. It is the element used for simple clausal negation, e.g., to express meanings such as 'It isn't big' or 'He didn't go', and it is used to form yes/no questions (by simply being placed at the end of a statement). It is also much more common in texts and has a wider range of uses than the other negators. It seems highly likely that the other negators can be analysed as portmanteaux of *bo*^{24} 'not' along with other elements, such as 'want', 'can', 'good', 'before', and 'do' (Hilary Chappell, personal communication).

2.8.2. *Possibility and potentiality (maybe, can).* Wierzbicka (1996) has proposed that 'can' and 'maybe' are universal semantic primes. At first sight, they appear rather similar, but closer examination shows that the two elements have quite different properties. Bolinger (1989) distinguishes "extrinsic possibility" ('maybe') from "intrinsic potentiality" ('can'). This difference correlates with a widespread contrast in the formal means by which 'can' and 'maybe' are expressed in languages: 'can' tends to manifest as a verb, verbal inflection, or verbal particle, whereas 'maybe' is typically realised as a sentence particle. Regardless of their formal realisation, the key fact which indicates that 'can' and 'maybe' are distinct is that they can both occur in the same clause. For example: *I can't do this. Maybe someone else can do it.* This is the case even when, as occasionally happens, the two meanings are expressed by the same lexical form, as in Polish:

(25) *Ona nie może tego zrobić, może ktoś inny może.*
she not can this do maybe someone else can
'She can't do this; maybe someone else can.'

Despite the tradition of distinguishing different "kinds" of 'can' (e.g., the 'can of ability' vs. the 'can of possibility'), indifference to such artificial dichotomies is one of the hallmarks of 'can'. In language after language a modal element is reported which is vague in this way, i.e., which can be used both with personal subjects, as in 'I can/can't move' and 'You can't do things like this', and with non-personal subjects as in 'Something bad can happen' and 'This thing can move'. To illustrate some of these properties, consider the following examples from two non-European languages. In Mangaaba-Mbula (Papua New Guinea), 'can' is expressed by a co-verb *-rao* (Bugenhagen in press). In Northern Pomo (California) it is expressed by a clitic *male*, which O'Connor (1992: 51–53) labels "potential modality". In both languages it can be found with personal and with non-personal subjects, as shown by the (a) and (b) examples, respectively, below.

(26) Mangaaba-Mbula

a. *Ni i-rao i-wit koroŋ tana som.*
NOM.3SG 3SG-can 3SG-lift thing that not
'He can't lift that thing.'

b. *Mbulu sanan-ŋa-na sa i-rao*
event be.bad-NMZ-GEN.3SG non-REF 3SG-can
i-pet pio na som.
3SG-happen REF.1SG give not
'Nothing bad can happen to me.'

(27) Northern Pomo

a. *Moːw k'o haynam mac'aː-nha male.*
3SM.A not stick break-NEG can
'He can't break the stick.'

b. *Xanam muː thinda padim-ʔa male.*
water.SPEC DEM EVID swim-PASS can
'(Someone) can swim across the river here.'

2.8.3. *Causation (because).* No convincing counter-example is known to the claim that all languages have a lexical unit with the meaning 'because'. In many languages there is an unambiguous connective analogous to English *because*; for example, *yīnwie* in Mandarin Chinese, *tana* in Mangaaba-Mbula, *'ona* in Samoan. In some languages, the lexeme expressing the meaning 'because' is morphosyntactically nominal, as with Acehnese *kareuna* or Kalam *juj*. It is known that some languages do not have a lexeme uniquely dedicated

to expressing 'because'. Commonly, 'because' is the meaning of a lexical unit of a lexeme which can also express "locational source", i.e., the starting-point of motion. This is the case in the Australian languages Yankunytjatjara and Arrernte, for example, where the ablative case-markers *-nguṟu* and *-nge*, respectively, are used to indicate 'because'.

2.8.4. *Conditional (if).* Various authors have claimed that particular languages do not distinguish between 'if' and 'when' (both claimed to be semantic primes by Wierzbicka 1996). Taking German as an example, Reilly (1986) says that the same word *wenn* is used both for 'when', as in *When Clare comes home, we'll have lunch*, and for English 'if', as in *If Clare comes home, we'll have lunch*. But as Wierzbicka (1996: 191) points out, German does lexically distinguish between 'if' (*wenn*) and 'when' (*als*, *wann*) – except in subordinate clauses referring to future events, such as those quoted by Reilly. In relation to future events, *wenn* is polysemous, and means either 'if' or 'when'. The polysemy is particularly clear in view of the fact that both meanings can be contrasted in one sentence (small capitals indicate contrastive stress): *Wenn er kommt* – WENN *er kommt, werde ich ihn sehen* 'when he comes – IF he comes, I'll see him'.

A slightly different situation obtains in Japanese, a language which has also been claimed not to distinguish 'if' from 'when'. It is true that Japanese constructions employing conjunctive *-ba* can be (and commonly are) used in both temporal and conditional contexts. As Hasada (1997) points out, however, if the particle *moshi* is employed, the *-ba* construction becomes unambiguously conditional; thus, *moshi* can be regarded as the Japanese equivalent of 'if'.

Various indigenous languages (particularly Australian Aboriginal languages) have also been claimed not to have an equivalent to 'if', but, where claims of this kind can be checked, they have never stood up (cf. McConvell 1991). It is worth noting, however, that in some Australian Aboriginal languages 'if' is expressed by the same the form which (as a sentence adverb or particle) expresses 'maybe'. The Arrernte particle *peke* and the Yankunytjatjara particle *tjinguṟu* are both polysemous in this way.

2.9. *Similarity (like)*

'Like' is a strong candidate as a universally lexicalised meaning. I know of no language which does not have at least one word or bound morpheme which expresses this meaning (in contexts such as 'someone like me' or 'do it like this'). From a morphosyntactic point of view, the lexical element meaning 'like' is sometimes verbal; as with Samoan *pei* 'be like' (Mosel 1994). In addition to a simple term for 'like', some languages have additional portmanteau

words for semantic combinations such as ‘like this’, ‘do (it) like this’, and ‘say (it) like this’.

This concludes the present review of potential lexico-semantic universals.

3. Conclusions and discussion

As stated at the onset, it is impossible to proclaim with absolute certainty that any meaning is attested as the meaning of a lexical unit in all languages. The sample of languages for which we are able to obtain information and analysis of the necessary quality is too small. This does not mean, however, that we reach the end of our survey without any firm conclusions. First, we know that only a very small set of meanings have any chance at all of being universals. From even a small sample of languages it is clear that many impressionistically “basic” items of English vocabulary (such as *go*, *water*, and *eat*) lack exact equivalents in other languages.

Second, we may conclude that universality and semantic simplicity are closely linked. The best candidates for the status of universal meanings are overwhelmingly to be found within the set of proposed semantic primes. To see this, one needs only to consider the fact that of the 48 NON-prime candidates for universal status reviewed in this paper, only the following seem to have much hope: ‘man’, ‘woman’, ‘child’, ‘mother’, ‘head’, ‘eye’, ‘ear’, ‘nose’, ‘hand’, ‘day’, ‘kill’, and ‘make’. In general, however, doubts remain about the universality of these non-prime meanings, doubts which are exacerbated by their semantic complexity. There is always the possibility that apparent equivalents in different languages may differ slightly in their underlying semantic configuration. Indeed, it might seem unlikely that ANY complex meaning configuration – no matter how solidly based in human experience – will be present in precisely the same shape (i.e., identical in every single detail) in all languages. On the other hand, of the 60 or so semantic primes proposed in the latest NSM work, at least 40 can be regarded as relatively secure candidates for universal status. (There is no definitive evidence against any of the other 20-odd proposed primes, but there has not been sufficient crosslinguistic work to establish their credibility either.)

Substantives:	I, you, someone, people, something/thing
Determiners and quantifiers:	this, the same, one, two, all, much/many
Attributes:	good, bad, big, small
Mental and speech predicates:	think, know, want, see, hear, say
Actions, events, states, existence:	do, happen, live, die, there is
Time:	when/time, now, before, after, a long time
Space:	where/place, here, above, below, inside
“Logical” concepts:	not, maybe, because, if
Other:	like, very

Third, it appears that the best candidates for universally lexicalised meanings are not based – in any obvious way – on universals of experience, or on universal features of the environment, or on cultural universals. Rather they represent CONCEPTUAL universals, and as such they tend to have a rather general or "schematic" character. This is not to say that meanings like 'I', 'you', 'this', 'good', 'do', 'think', 'see', and so on are in any sense "remote" from ordinary experience. On the contrary, they are everyday concepts, grounded in simple everyday words. Arguably, concepts like these suffuse, and, in a sense, help constitute, the lived experience of human beings.

With renewed research attention being focused on the question of semantic universals, and more and better descriptive accounts of languages now becoming available, it is perhaps not too much to hope that the next decade may see the establishment of a comprehensive inventory of lexico-semantic universals. This would constitute a substantial advance in linguistic typology for several reasons. Most obviously, it would represent a foundational advance in lexical typology in the sense of a "theory of vocabulary structure" (cf. Lehrer 1992). On any account, lexical typology is surely at present the "poor cousin" among typological studies. I am taking it for granted, of course, that typology includes not only establishing patterns of systematic co-variation across languages, but also establishing what is constant, i.e., universal, across languages. If only because universals place a limit on diversity, a finding about universals IS a finding about typology.

In my view, however, there would also be a deeper theoretical significance to the discovery of a robust inventory of lexico-semantic universals, which derives from what one might call the logic of typological investigation. The point is that any typological framework, i.e., a framework which enables us to identify and order the variability across languages, necessarily presupposes descriptive parameters which are constant and language-neutral, in the sense of not depending on the vagaries of any individual language. More simply, to describe and compare any set of things, one must have some terms, some *tertium comparationis*, which are stable and equally applicable across the entire set of things being compared. Consider the example of syntactic typology. It requires a framework of putative syntactic universals, which might include concepts such as 'noun' (or noun-phrase), 'verb', 'clause', 'subject', 'object', 'transitive', 'passive', 'causative', 'relative clause', and so on. Such terms constitute a descriptive metalanguage of syntactic typology. The details are, of course, controversial, and so they should be, for they "set the frame" for the entire venture. If they are faulty – for example, if they are language-specific, unnecessarily vague, or conceptually incoherent – then the entire venture of syntactic typology is compromised. The importance of establishing an inventory of lexico-semantic universals is that one could use this inventory as a base upon which to construct a semantic framework for lexical typology; for ex-

ample, to construct a theory of lexical domains, to investigate systematically language-specific patterns of lexicalisation (in the sense of Talmy (1985)), to map out implicational universals in the lexicon, to investigate recurrent patterns of lexical polysemy, and so on.

The establishment of a firm baseline of lexico-semantic universals would be important for grammatical typology as well, in several ways. At the simplest level, it is apparent that many recurrent grammatical phenomena are or can be semantically-based (cf. Goddard 1997a, Wierzbicka 1998, 1999). It is also possible that lexico-semantic universals may bring with them, as it were, a substantial slab of universal syntax, because presumably they have inherent syntactic properties which will be manifest in all languages (though with different formal realisations). For example, if it turns out that 'want', 'know', and 'think' are lexico-semantic universals then presumably this will carry implications about the universality of certain types of clausal complementation. Similarly, if 'can' and 'maybe' turn out to be lexico-semantic universals this will presumably carry implications about the internal structure of a simple clause (since 'can' can be seen as a predicate operator and 'maybe' as an "external" clause operator). The idea that there can be linkages, at the very deepest level, between lexis and syntax deserves much fuller treatment than is possible here (cf. Goddard & Wierzbicka (eds.) in press).

More broadly, the connection between lexical typology and grammatical typology is well characterised by Behrens & Sasse's (1997) expression "the interweavement of lexical and grammatical structuring". Writing in opposition to conventional "subsystem typologies", Behrens & Sasse urge that typology be re-construed as the study of a composite lexico-grammar. "Viewed in the context of comparative linguistic research", they write, "the concept of lexico-grammar leads to the assumption that we can expect, in different languages, quite divergent patterns of interactions between lexicon and grammar, and these divergences are of great typological significance. It is therefore proposed that lexical semantics and its repercussions on grammar be assigned a central role in typological investigations" (Behrens & Sasse 1997: 2). In this perspective, it can be seen that the identity and nature of lexico-semantic universals holds potentially profound implications for linguistic typology at large.

Received: 14 January 2000 *University of New England*
Revised: 30 May 2000

Correspondence address School of Languages, Cultures and Linguistics, University of New England, Armidale, NSW, 2351, Australia; e-mail: cgoddard@metz.une.edu.au

Acknowledgments I would like to thank Lyle Campbell, Nick Enfield, and Anna Wierzbicka for comments and criticisms which have helped me improve earlier versions of this paper. My thanks are also due to three reviewers for

Linguistic Typology, and to two reviewers for the journal *Lexicology* (unfortunately, now discontinued). I would also like to thank the following people for information and discussion of particular languages: Penelope Brown (Tzeltal), Timothy Curnow (Spanish), Nick Enfield (Lao), Nicholas Evans (Kayardild), Raquel Flores (Tagalog), Kumie Fujimori (Japanese), Annie Langlois (French), Tony Liddicoat (French), Judy Ho (Cantonese), Mee Wun Lee (Cantonese), Beat Lehmann (German), Andrew Pawley (Kalam), Catherine Travis (Spanish), Anna Wierzbicka (Polish, Russian), David Wilkins (Arrernte). Needless to say, they are not responsible for errors of fact or interpretation which remain. Data on Yankunytjatjara and Malay comes from my own work. A condensed version of part of this paper appears as Goddard (in press b).

Abbreviations 1, 2, 3 1st, 2nd, 3rd person, A cross reference set A, ABS absolutive, APP applicative, APPROX approximation, ASSR assertive, ATP antipassive, B cross-reference set B, CAUS causative, CLF classifier, CMP completive, CONJ conjunctive suffix, DAT dative, DEF definite, DEG degree word, DIM diminutive, DIR directional, DIST distal deixis, EMPH emphatic, ERG ergative, EVID evidential, FOC focus, GEN genitive, HAB habitual, INAN inanimate, INC incompletive, IO indirect object, LOC locative, NEG negative, NMZ nominaliser, NOM nominative, O object, PASS passive, PAST past tense, PAST.IMPF past imperfective, PERF perfective, PL plural, PRES present tense, PFV perfective, RDP reduplication, REF referent case, SBJ subject, SG singular, SPEC specifier, TODP today's past tense, TOP topic. The boundary symbol = designates a clitic juncture in Amele and Hopi in examples (12), (15), (17), and (20). In example (2) it designates a juncture between two word-like elements in Yolngu Matha which form a fixed phrasal expression.

With minor exceptions, interlinear glosses are essentially as given by the original authors.

References

Andersen, Elaine S. (1978). Lexical universals of body-part terminology. In Joseph H. Greenberg (ed.), *Universals of Human Language*, Volume 3: *Word Structure*, 335–368. Stanford: Stanford University Press.

Ameka, Felix (1994). Ewe. In Goddard & Wierzbicka (eds.) 1994, 57–86.

— (1995). The linguistic construction of space in Ewe. *Cognitive Linguistics* 6: 139–181.

Apresjan, Yuri D. (1972). Regular polysemy. *Linguistics* 124: 5–32.

— (1974 [1992]). *Leksičeskaja semantika: Sinonimečeskie sredstva jazyka*. Moskva: Nauka. [English translation: *Lexical Semantics: User's Guide to Contemporary Russian Vocabulary.* Ann Arbor: Karoma.]

Backhouse, Anthony E. (1994). *The Lexical Field of Taste: A Semantic Study of Japanese Taste Terms*. Cambridge: Cambridge University Press.

Bain, Margaret (1992). *The Aboriginal-White Encounter in Australia: Towards Better Communication*. Darwin: Summer Institute of Linguistics.

Behrens, Leila & Hans-Jürgen Sasse (1997). Lexical typology: A programmatic sketch. *Arbeitspapiere des Instituts für Sprachwissenschaft der Universität zu Köln*, Neue Folge 30.

Berlin, Brent (1992). *Ethnobiological Classification: Principles of Categorization of Plants and Animals in Traditional Society*. Princeton: Princeton University Press.

Berlin, Brent & Paul Kay (1969). *Basic Color Terms: Their Universality and Evolution*. Berkeley: University of California Press.

Bohnemeyer, Jürgen (1998a). Temporal reference from a radical pragmatics point of view: Why Yucatec Maya does not need to express 'after' and 'before'. *Cognitive Linguistics* 9: 239–282.

— (1998b). Time relations in discourse: Evidence from a comparative approach to Yukatek Maya. Doctoral dissertation, Katholieke Universiteit Brabant, Tilburg.

Bolinger, Dwight (1989). Extrinsic possibility and intrinsic potentiality: 7 on *may* and *can* + 1. *Journal of Pragmatics* 13: 1–23.

Brown, Cecil H. (1976). General principles of human anatomical partonomy and speculations on the growth of partonomic nomenclature. *American Ethnologist* 3: 400–424.

— (1984). *Language and Living Things*. New Brunswick, N.J.: Rutgers University Press.

— (1989). Universal constraints on polysemy and overt marking. *Quaderni di Semantica* 10: 33–50.

Brown, Donald E. (1991). *Human Universals*. Philadelphia: Temple University Press.

Brown, Penelope (1994). The INs and ONs of Tzeltal locative expressions: The semantics of static descriptions of location. *Linguistics* 32: 743–790.

Brugman, Claudia & Monica Macaulay (1986). Interacting semantic systems: Mixtec expressions of location. *Berkeley Linguistics Society* 12: 315–327.

Bugenhagen, Robert D. (in press). The syntax of universal semantic primes in Mangaaba-Mbula. In Goddard & Wierzbicka (eds.) in press.

Chappell, Hilary (1994). A typology of negative verbs of existence and possession in Sinitic languages. In Hajime Kitamura, T. Nishida, & Y. Nagano (eds.), *Current Issues in Sino-Tibetan Linguistics*, 303–311. Osaka: Museum of Ethnology, Organizing Committee of the 26th International Conference on Sino-Tibetan Languages and Linguistics.

Cooke, Michael & Cliff Goddard (1997). Semantic primes in Yolngu Matha. Unpublished manuscript.

Cruse, D. A. (1986). *Lexical Semantics*. Cambridge: Cambridge University Press.

Dambriūnas, Leonardas, Antanas Klimas, & William R. Schmalstieg (1966). *Introduction to Modern Lithuanian*. Brooklyn, N.Y.: Franciscan Fathers Press.

Diller, Anthony (1994). Thai. In Goddard & Wierzbicka (eds.) 1994, 149–170.

Dixon, R. M. W. (1980). *The Languages of Australia*. Cambridge: Cambridge University Press.

— (1982). *Where Have All the Adjectives Gone? and Other Essays in Semantics and Syntax*. Berlin: Mouton.

— (1989). The Dyirbal kinship system. *Oceania* 59: 245–268.

— (1991). *Words of our Country*. St. Lucia: University of Queensland Press.

Dryer, Matthew S. (1986). Primary objects, secondary objects, and antidative. *Language* 62: 808–845.

Durie, Mark, Bukhari Daud, & Mawardi Hasan (1994). Acehnese. In Goddard & Wierzbicka (eds.) 1994, 171–201.

Dyen, Isidore, A. T. James, & J. W. L. Cole (1967). Language divergence and estimated retention rate. *Language* 43: 150–171.

Ekman, Paul (1992). An argument for basic emotions. *Cognition and Emotion* 6: 169–200.

— (1993). Facial expression and emotion. *American Psychologist* 48: 384–392.

Enfield, Nick (in press). The syntax of natural semantic metalanguage expressions in Lao. In Goddard & Wierzbicka (eds.) in press.

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

— (1994). Kayardild. In Goddard & Wierzbicka (eds.) 1994, 203–228.

— (1995). A-Quantifiers and scope in Mayali. In Emmon Bach, Eloise Jelinek, Angelika Kratzer, & Barbara H. Partee (eds.), *Quantification in Natural Languages*, 207–270. Dordrecht: Kluwer.

Evans, Nicholas & David P. Wilkins (1995). The inside story of under: A strange case of regular polysemy in Australian languages. Seminar paper, Cognitive Anthropology Research Group, Max Planck Institute for Psycholinguistics, Nijmegen.

— (2000). In the mind's ear: The semantic extensions of perception verbs in Australian languages. *Language* 76: 546–592.

Foley, William A. (1986). *The Languages of New Guinea*. Cambridge: Cambridge University Press.

— (1997). *Anthropological Linguistics*. Oxford: Blackwell.

Geeraerts, Dirk (1994). Polysemy. In R. E. Asher & J. M. L. Simpson (eds.), *The Encyclopedia of Language and Linguistics*, 3227–3228. Oxford: Pergamon.

Goddard, Cliff (1991). Testing the translatability of semantic primitives into an Australian Aboriginal language. *Anthropological Linguistics* 33: 31–56.

— (1994). Lexical primitives in Yankunytjatjara. In Goddard & Wierzbicka (eds.) 1994, 229–262.

— (1995). Who are *We*? The natural semantics of pronouns. *Language Sciences* 17: 99–121.

— (1996a). *Pitjantjatjara/Yankunytjatjara to English Dictionary*. Revised 2nd edition. Alice Springs: IAD Press.

— (1996b). The "social emotions" of Malay (Bahasa Melayu). *Ethos* 24: 426–464.

— (1997a). Grammatical categories and semantic primes. *Australian Journal of Linguistics* 17: 1–43.

— (1997b). The semantics of coming and going. *Pragmatics* 7: 147–162.

— (1997c). Contrastive semantics and cultural psychology: 'Surprise' in Malay and English. *Culture & Psychology* 3: 153–181.

— (1998a). *Semantic Analysis: A Practical Introduction*. Oxford: Oxford University Press.

— (1998b). Bad arguments against semantic primitives. *Theoretical Linguistics* 24: 129–156.

— (in press a). The syntax of universal semantic primes in Malay (Bahasa Melayu). In Goddard & Wierzbicka (eds.) in press.

— (in press b). Universal units in the lexicon. In Martin Haspelmath, Ekkehard König, Wulf Oesterreicher, & Wolfgang Raible (eds.), *Language Typology and Linguistic Universals: An International Handbook.* Berlin: de Gruyter.

— (in press c). *Hati*: A key word in the Malay vocabulary of emotion. In Jean Harkins & Anna Wierzbicka (eds.), *Emotions in Cross-linguistic Perspective.* New York: Oxford University Press.

Goddard, Cliff (ed.) (1997). *Studies in the Syntax of Universal Semantic Primitives*. Special Issue of *Language Sciences* 19-3.

Goddard, Cliff & Anna Wierzbicka (eds.) (1994). *Semantic and Lexical Universals: Theory and Empirical Findings*. Amsterdam: Benjamins.

— (in press). *Meaning and Universal Grammar*. Amsterdam: Benjamins.

Gough, Kathleen (1959). The Nayars and the definition of marriage. *Journal of the Royal Anthropological Institute* 89: 23–34.

— (1961). Variation in matrilineal systems. In David M. Schneider & Kathleen Gough (eds.), *Matrilineal Kinship*, 70–84. Berkeley: University of California Press.

Hale, Ken (1994). Preliminary observations on lexical and semantic primitives in the Misumalpan languages of Nicaragua. In Goddard & Wierzbicka (eds.) 1994, 263–284.

Hallpike, Christopher Robert (1979). *The Foundations of Primitive Thought*. Oxford: Clarendon.

Harkins, Jean (1995). Desire in language and thought: A cross-linguistic and cross-cultural study. Ph.D. dissertation, Australian National University.

— (1996). Linguistic and cultural differences in concepts of shame. In David Parker, Rosamund Dalziell, & Iain Wright (eds.), *Shame and the Modern Self*, 84–96. Melbourne: Australian Scholarly Publishing.

Harkins, Jean & David Wilkins (1994). Mparntwe Arrernte and the search for lexical universals. In Goddard & Wierzbicka (eds.) 1994, 285–310.

Hasada, Rie (1997). Conditionals and counterfactuals in Japanese. *Language Sciences* 19: 277–288.

Haspelmath, Martin (1997). *Indefinite Pronouns*. Oxford: Clarendon.

Henderson, John & Veronica Dobson (1994). *Eastern and Central Arrernte to English Dictionary*. Alice Springs: IAD Press.

Hill, Deborah (1987). A cross-linguistic study of value-judgement terms. M.A. thesis, Australian National University.

— (1994). Longgu. In Goddard & Wierzbicka (eds.) 1994, 311–330.

Hill, Deborah & Cliff Goddard (1997). Spatial terms, polysemy and possession in Longgu (Solomon Islands). *Language Sciences* 19: 263–276.

Keesing, Roger M. (1994). Radical cultural difference: Anthropology's myth? In Martin Pütz (ed.), *Language Contact and Language Conflict*, 3–24. Amsterdam: Benjamins.

Kimball, Geoffrey D. (1985). A descriptive grammar of Koasati. Ph.D. dissertation, Tulane University.

Koptjevskaja-Tamm, Maria & Ekaterina V. Rakhilina (1999). Prototypes in lexical typology: Temperature terms in Russian and Swedish. Paper presented at the Sixth International Cognitive Linguistics Conference, Stockholm, July 1999.

Kuno, Susumu (1983). *Sin-Nihonbunpoo-Kenkyuu*. [New Research on Japanese Grammar.] Tokyo: Taisyuukan Syoten.

Langacker, Ronald (1999). A study in unified diversity: English and Mixtec locatives. *Rask* 9/10: 215–256.

Laughren, Mary & Kenneth Hale (forthcoming). *Warlpiri Dictionary*. Electronic version available, Department of English and Linguistics, University of Queensland.

Lehrer, Adrienne (1992). A theory of vocabulary structure: Retrospective and prospectives. In Martin Pütz (ed.), *Thirty Years of Linguistic Evolution*, 243–256. Amsterdam: Benjamins.

Levinson, Stephen C. (1994). Vision, shape, and linguistic description: Tzeltal body-part terminology and object description. *Linguistics* 32: 791–855.

Levinson, Stephen C. & Penelope Brown (1994). Immanuel Kant among the Tenejapans: Anthropology as empirical philosophy. *Ethos* 22: 3–41.

Lewis, Gilbert (1974). Gnau anatomy and vocabulary for illness. *Oceania* 45: 50–78.

Lyons, John (1967). A note on possessive, existential and locative sentences. *Foundations of Language* 3: 390–396.

Malotki, Ekkehart (1983). *Hopi Time: A Linguistic Analysis of the Temporal Concepts in the Hopi Language*. Berlin: Mouton.

Matthews, Stephen & Virginia Yip (1994). *Cantonese: A Comprehensive Grammar*. New York: Routledge.

McConvell, Patrick (1991). Cultural domain separation: Two-way street or blind alley? Stephen Harris and the NeoWhorfians on Aboriginal education. *Australian Aboriginal Studies* 1: 13–24.

Mel'čuk, Igor A. (1988). Semantic description of lexical units in an explanatory combinatorial dictionary: Basic principles and heuristic criteria. *International Journal of Lexicography* 1: 165–188.

— (1989). Semantic primitives from the viewpoint of the Meaning-Text linguistic theory. *Quaderni di Semantica* 10: 65–102.

Mondry, Henrietta & John R. Taylor (1992). On lying in Russian. *Language and Communication* 12: 133–143.

Mosel, Ulrike (1984). *Tolai Syntax and its Historical Development*. (Pacific Linguistics, B-92.) Canberra: Australian National University.
— (1987). Subject in Samoan. In Donald C. Laycock & Werner Winter (eds.), *A World of Language: Papers Presented to Professor S. A. Wurm on his 65th Birthday* (Pacific Linguistics, C-100), 455–479. Canberra: Australian National University.
— (1994). Samoan. In Goddard & Wierzbicka (eds.) 1994, 331–360.
Myhill, John (1996). Is BAD a semantic primitive? Evidence from Biblical Hebrew. *Lexicology* 2: 99–126.
Newman, John (1996). *Give: A Cognitive Linguistic Study*. Berlin: Mouton de Gruyter.
— (1997). Preface. In Newman (ed.) 1997, vii–xv.
Newman, John (ed.) (1997). *The Linguistics of Giving*. Amsterdam: Benjamins.
Ngakulmungan Kangka Leman (1997). *Lardil Dictionary*. Gununa, Queensland: Mornington Shire Council.
Nida, Eugene A. (1947). *Bible Translating*. New York: American Bible Society.
Nordlinger, Rachel (1998). *A Grammar of Wambaya, Northern Territory (Australia)*. (Pacific Linguistics, C-140.) Canberra: Australian National University.
O'Connor, Mary Catherine (1992). *Topics in Northern Pomo Grammar*. New York: Garland.
Onishi, Masayuki (1994). Semantic primitives in Japanese. In Goddard & Wierzbicka (eds.) 1994, 361–386.
— (1997). The grammar of mental predicates in Japanese. *Language Sciences* 19: 219–233.
Parsons, Terence (1994). *Events in the Structure of English: A Study in Subatomic Semantics*. Cambridge, Mass.: MIT Press.
Pawley, Andrew (1993). A language which defies description by ordinary means. In William A. Foley (ed.), *The Role of Theory in Language Description*, 88–129. Berlin: Mouton de Gruyter.
— (1994). Kalam exponents of lexical and semantic primitives. In Goddard & Wierzbicka (eds.) 1994, 387–422.
— (1997). The Trans New Guinea Phylum hypothesis: A reassessment. Unpublished manuscript.
Reilly, Judy Snitzer (1986). The acquisition of temporals and conditionals. In Elizabeth C. Traugott (ed.), *On Conditionals*, 309–331. Cambridge: Cambridge University Press.
Rhodes, Richard A. (1985). *Eastern Ojibwa–Chippewa–Ottawa Dictionary*. Berlin: Mouton.
Rice, Sally (1997). Giving and taking in Chipewyan: The semantics of THING-marking classificatory verbs. In Newman (ed.) 1997, 97–134.
Richards, Eirlys & Joyce Hudson (1990). *Walmajarri–English Dictionary*. Darwin: Summer Institute of Linguistics.
Roberts, John (1987). *Amele*. London: Croom Helm.
— (1997). GIVE in Amele. In Newman (ed.) 1997, 1–34.
Russell, James A. (1991). Culture and the categorization of emotions. *Psychological Bulletin* 110: 426–450.
Sasse, Hans-Jürgen (1987). The thetic/categorical distinction revisited. *Linguistics* 25: 511–580.
— (1996). Cayuga: A grammatical sketch. Unpublished manuscript, Australian Linguistic Institute, Australian National University.
Scheffler, Harold W. (1978). *Australian Kin Classifications*. Cambridge: Cambridge University Press.
Shibatani, Masayoshi (1990). *The Languages of Japan*. Cambridge: Cambridge University Press.
Sinha, Chris (1995). Introduction [to special issue: *Spatial Language and Cognition 1*]. *Cognitive Linguistics* 6: 7–9.
Suzuki, Takao (1978). *Japanese and the Japanese: Words in Culture*. Tokyo: Kodansha International.
Swadesh, Morris (1972 [1960]). What is glottochronology? In Morris Swadesh (ed. by Joel Sherzer), *The Origin and Diversification of Languages*, 271–284. London: Routledge & Kegan Paul.

Talmy, Leonard (1985). Lexicalization patterns: Semantic structure in lexical forms. In Timothy Shopen (ed.), *Language Typology and Syntactic Description*, Volume 3: *Grammatical Categories and the Lexicon*, 57–149. Cambridge: Cambridge University Press.

Timberlake, Alan (1993). Russian. In Bernard Comrie & Greville G. Corbett (ed.), *The Slavonic Languages*, 827–886. London: Routledge.

Tong, Malindy, Michael Yell, & Cliff Goddard (1997). Semantic primitives of time and space in Hong Kong Cantonese. *Language Sciences* 19: 245–261.

Verhaar, John W. M. (ed.) (1967–73). *The Verb 'Be' and its Synonyms: Philosophical and Grammatical Studies*. 6 volumes. Dordrecht: Reidel.

Whorf, Benjamin Lee (1956). *Language, Thought and Reality: Selected Writings of Benjamin Lee Whorf*. Cambridge, Mass.: MIT Press.

Wierzbicka, Anna (1972). *Semantic Primitives*. Frankfurt/Main: Athenäum.

— (1976). In defense of YOU and ME. In Wolfgang Girke & Helmut Jachnow (eds.), *Theoretische Linguistik in Osteuropa*, 1–21. Tübingen: Niemeyer.

— (1980). *Lingua Mentalis: The Semantics of Natural Language*. Sydney: Academic Press.

— (1986). Human emotions: Universal or culture-specific? *American Anthropologist* 88: 584–594.

— (1992). *Semantics, Culture, and Cognition*. Oxford: Oxford University Press.

— (1994). Semantic primitives across languages: A critical review. In Goddard & Wierzbicka (eds.) 1994, 445–500.

— (1996). *Semantics: Primes and Universals*. Oxford: Oxford University Press.

— (1998). Anchoring linguistic typology in universal semantic primes. *Linguistic Typology* 2: 141–194.

— (1999). Lexical prototypes as a universal basis for cross-linguistic identification of "parts of speech". In Petra M. Vogel & Bernard Comrie (eds.), *Approaches to the Typology of Parts of Speech*, 285–317. Berlin: Mouton de Gruyter.

— (in press). The syntax of universal semantic primes in Polish. In Goddard & Wierzbicka (eds.) in press.

Wilkins, David P. (1996a). Natural tendencies of semantic change and the search for cognates. In Mark Durie & Malcolm Ross (eds.), *The Comparative Method Reviewed*, 264–304. New York: Oxford University Press.

— (1996b). Does 'tyerrtye' mean 'person'? Issues in the exploration of meaning, use and word-relations. Paper presented at Max Planck Institute for Psycholinguistics, Nijmegen.

Wilkins, David P. & Deborah Hill (1995). When "go" means "come": Questioning the basicness of basic motion verbs. *Cognitive Linguistics* 6: 209–259.

Witkowski, Stanley R. & Cecil H. Brown (1978). Lexical universals. *Annual Review of Anthropology* 7: 427–451.