

Heterosemy and the grammar-lexicon trade-off

N. J. Enfield

1. Introduction

A fundamental task for the grammarian is to determine what will be listed in the lexicon, and what will be derived from this list by productive rules specified in the grammar. The balance of information between grammar and lexicon becomes – in principle – an economical trade-off. While a grammarian can theoretically make as many detailed distinctions as he or she likes, the description should capture useful generalisations in an accessible way. Ideally this will reflect natural relationships at higher levels within the linguistic system being described. The need to balance parsimony, accuracy and comprehensiveness, readability, and intuitive faithfulness to the genius of a language is what makes grammar-writing as much an art as it is a technical craft. This tension is exemplified by the problem to be discussed in this chapter, namely: How should the grammarian treat a morpheme which apparently has membership in more than one grammatical class?

2. Polysemy and heterosemy

If a morpheme is polysemous it has more than one meaning, where there is some significant overlap in semantic content between the meanings. An example is the English verb *fly*, meaning either ‘propel oneself through air’ or ‘go very fast’. A special case of polysemy is termed *heterosemy* (Lichtenberk 1991, after Persson 1988), where the different but related meanings of a given morpheme are associated with distinct grammatical contexts. Examples from English are *father*, *walk*, *stone*, and *talk*, each of which have related noun and verb meanings. It is useful to distinguish between different logical possibilities for heterosemy with reference to the *open* versus *closed* form class distinction. When two grammatical classes

associated with a heterosemous word's multiple meanings are both open classes, this may be termed *open-class heterosemy*. The English verb *fly* is an example. When two grammatical classes associated with a word's multiple meanings are both closed classes, this may be termed *closed-class heterosemy*. An example is the English word *to*, which (among other things) marks a non-finite verb in a subordinate clause (*I want to go*), or a goal of motion in a prepositional phrase (*I went to London*). When the different grammatical classes associated with a word's multiple meanings are each of different status – one open, one closed – we may call this *cross-class heterosemy*. An example is English *have*, which occurs as a regular verb (*I have two bicycles*) and as an auxiliary verb (*I have sold my car*).

That lexical items can enter into alternations across word classes with no overt marking raises a fundamental question for the organisation of a descriptive grammar. Are these relationships derived by grammatical rule? We consider the issue with reference to examples from Lao (a Southwestern Tai language of Laos, Thailand, Cambodia) and other isolating languages of mainland Southeast Asia.

3. Open-class heterosemy: Noun-verb alternation in Lao culinary terms

There is a set of items in Lao each of which functions in one grammatical environment as a transitive verb meaning 'to prepare a certain food or ingredient in a certain way', and in another grammatical environment as a nominal denoting the dish so prepared. Compare two uses of the word *kèng3* in the following:¹

- (1) *kuu3 siø kèng3 paa3 too3 nii4*
 1SG IRR fish CLF DEM.GEN
 'I'm going to (make) soup (with) this fish.'
- (2) *qan3 nii4 mèn1 kèng3 paa3*
 CLF DEM.GEN COP fish
 'This is fish soup.'

Lao noun phrases are head-initial. Thus, in the noun phrase *kèng3 paa3* [soup fish] in (2), *kèng3* 'soup' is the head, modified by *paa3* 'fish'. (It is also possible to use *kèng3* 'soup; make soup' as a modifier of *paa3* 'fish'; this would give *paa3 kèng3* [fish soup] 'fish for (making) soup' or 'fish [which has been] made into soup'.)

Here is a list of words which enter into the same alternation:

Table 1. Some items entering into a noun-verb alternation, where modified noun refers to ‘dish prepared with ingredient x in a certain way’, and transitive verb refers to ‘preparing a dish with ingredient x in that way’.

Term	Occurs as head noun (modified by ingredient name)	Occurs as transitive verb
<i>khua5</i> ‘fry’	yes	yes
<i>mòk1</i> ‘bake in leaves’	yes	yes
<i>qòq2</i> ‘make stew of’	yes	yes
<i>ñam2</i> ‘make salad of’	yes	yes
<i>kèèng3</i> ‘make soup of’	yes	yes
<i>laap4</i> ‘make larb of’	yes	yes
<i>cèèw1</i> ‘make chutney of’	yes	yes
<i>tam1</i> ‘pound; make salad of’	yes	yes
<i>tôm4</i> ‘boil’	yes	yes
<i>nùng1</i> ‘steam’	yes	yes
<i>qôp2</i> ‘broil/bake’	yes	yes
<i>cùùn3</i> ‘shallow fry’	yes	yes
<i>thòòt4</i> ‘deep fry’	yes	yes
<i>piing4</i> ‘grill’	yes	yes
<i>luak4</i> ‘parboil’	yes	yes

A grammarian may want to avoid listing these words separately with their noun and verb meanings, but instead derive one from the other with a generalising rule. One could list *kèèng3* as a noun meaning ‘soup’, and state that its use in a transitive verb frame derives a verb with the meaning ‘to make soup from (something)’. To pursue this analysis for the whole set, one’s statement of derivation would need to be valid for each alternating item. A problem is that, according to informants, some of the items in Table 1 refer ‘basically’ to a kind of dish (i.e. are ‘basically nouns’), while others refer ‘basically’ to a manner of preparation (i.e. are ‘basically verbs’). In other words, it is not clear that all items would have the same direction of derivation. Lao consultants regard *tôm4*, for example, as basically a verb meaning ‘to boil something’, and its meaning in a noun phrase like *tôm4 paa3* [boil fish] ‘a dish made by boiling fish’ is regarded as an extended sense. This division among items in Table 1 has grammatical consequences. While all of the items listed in Table 1 can appear as nominal heads taking a modifier which specifies the ingredient from which the dish was made, only half can appear as stand-alone nominals referring to a

certain dish whose main ingredient is unspecified. Compare *tôm4* ‘boil’ and *kèng3* ‘soup’ in the following examples:

- (3) *kuu3 siø kin3 tôm4 *(paa3)*
 1SG IRR eat boil fish
 ‘I’m going to eat (a dish of) boiled *(fish).’
- (4) *kuu3 siø kin3 kèng3 (paa3)*
 1SG IRR eat soup fish
 ‘I’m going to eat (fish) soup.’

Table 2 shows that about half these terms cannot be stand-alone nominal heads:

Table 2. Distinctions between items in Table 1, by occurrence as unmodified nouns.

	Term	Occurs as free noun (e.g. as object of <i>kin3</i> ‘eat’)
‘basically nouns’	<i>khua5</i> ‘fry’	yes
	<i>môk1</i> ‘bake in leaves’	yes
	<i>qòq2</i> ‘make stew of’	yes
	<i>ñam2</i> ‘make salad of’	yes
	<i>kèng3</i> ‘make soup of’	yes
	<i>laap4</i> ‘make <i>larb</i> of’	yes
	<i>cèw1</i> ‘make chutney of’	yes
‘basically verbs’	<i>tam1</i> ‘pound; make salad of’	no
	<i>tôm4</i> ‘boil’	no
	<i>nùng1</i> ‘steam’	no
	<i>qôp2</i> ‘broil/bake’	no
	<i>cùùn3</i> ‘shallow fry’	no
	<i>thòt4</i> ‘deep fry’	no
	<i>püing4</i> ‘grill’	no
<i>luak4</i> ‘parboil’	no	

This issue is treated by Dixon (1991: 51) in a discussion of noun-verb alternations in English (cf. *walk*, basically a verb, versus *stone*, basically a noun), for which two possible analyses are raised. A first possibility is to allow that some nouns (e.g. *stone*) can be heads of verb phrases, and some verbs (e.g. *walk*) can be heads of noun phrases. “The other approach”, writes Dixon, “is to say that the head of a VP can only be a verb, etc., and that English has derivational processes, with zero marking, that derive verbs from some nouns (verb *stone* from noun *stone*), and so on” (Dixon 1991: 51). The challenge in a grammatical description is to achieve a se-

mantic mapping between the noun and verb senses which is consistent across an entire class of alternating items. Dixon (1991: 52) suggests that some generalisations are possible. For example, ‘many nouns referring to implements can also be used as verbs, e.g. *spear, knife, saw, hammer, whip, nail, screw*’. But this observation does not yet constitute the kind of grammatical generalisation which would genuinely reduce the burden on the lexicon. The semantic details of each alternation require further specification. Consider *knife* and *saw*. There are differences between the grammatical and semantic properties of these verbs in relation to their respective nominal counterparts. The verb *knife* has a more specific meaning than simply ‘use a knife (to cut something)’ (cf. *John knifed Bill* vs. **John knifed the onions*). The same issues arise with respect to the data in Table 1 above. It may be possible to make a generalisation to the effect that there is a set of words which function both as verbs and as nouns, such that the verb meaning refers to the action(s) involved in preparing a certain dish, and the noun meaning refers to the dish itself. However, as shown above, one does not have to look far before finding both formal and semantic differences among alternations in the list of terms.²

The descriptive problem of open-class heterosemy has spawned volumes of research (cf. Apresjan 1974, Jackendoff 1975, Clark and Clark 1979, Voorhoeve 1981, Dixon 1991, Levin 1993, Levin and Rappaport Hovav 1995, Vonen 2000, Croft 2001, inter alia). The point to be taken from the present discussion is that what appear to be regular heterosemous alternations can quickly turn out not to involve a homogeneous set of items, and the alternations observed are neither regular nor productive enough to be captured in the grammar in the strict sense (i.e. such that knowing the rule means not having to consult the lexicon in order to interpret or produce the data).³ To capture the facts, the descriptive linguist is forced to list these items and alternations in the lexicon, with specifications of the distinct semantic and grammatical properties of each case.

4. ‘Derivation’: The parallel with morphology

Issues arising in above discussion of a heavily isolating language are similar to those arising in the study of derivational morphology. Dixon (1991: 51) invokes derivational morphology in arguing for a ‘zero derivation’ account of noun-verb alternations in English. He draws the following analogy:

talk (v.) *talk* (n.)
converse (v.) *conversation* (n.)

The analogy suggests that heterosemy is as much a grammatical process as the explicit morphological marking of functionally equivalent alternations. While this may encourage the grammarian to look for rules akin to morphological derivations, an important point in the present context is that derivational morphology itself often does *not* allow for statements in the grammar which genuinely reduce the burden on the lexicon. For example, the English inchoative suffix *-en*, which derives verbs from adjectives, often signals the presence of a meaning more specific than simply ‘(cause to) become (more) ADJ’. *Blacken* does not mean ‘cause (sth.) to become black’, hence the oddity of saying *I blackened my house* with reference to having painted it with black paint. With reference to the derivational morphology of Khmer (Eastern Mon-Khmer, Cambodia), Huffman (1970: 311) writes:

... it is one of the clichés about Cambodian [Khmer] that it has a complex system of prefixes and infixes. This statement is misleading if it leads to the conclusion that Cambodian speakers “use” affixation as a *derivational process* in speaking. It is a fact that the Cambodian lexicon contains a large number of words (*derivatives*) which are *related to* other words (*bases*) by various prefixes and infixes, but these affixes are not productive in the modern language.

Nevertheless, Huffman usefully describes a list of derivational prefixes and infixes in the language (1970: 312–316). Similarly, Macdonald and Darjowidjojo (1967: Chapter 2) list a range of morphological derivations in Indonesian, many of which are semantically idiosyncratic and cannot be accounted for by general rules. It is normal, and indeed useful, for grammarians to list and analyse derivational morphemes, even when precise semantic generalisations are not forthcoming, and when there is no genuine reduction in the number of necessary lexical entries. As Anderson (1985: 163) points out, a derivational morphological relationship is “quite properly part of the lexicon of the language”, not “a property of the grammar itself”.

5. Cross-class heterosemy: ‘Verb/preposition alternations’ in Lao

Consider the words in boldface in the following Lao examples:

- (5) a. *man2* ***nam2*** *kuu3*
 3SG 1SG
 ‘S/he followed me.’
- b. *man2* *paj3* *talaat5* ***nam2*** *kuu3*
 3SG go market 1SG
 ‘S/he went to the market with me.’
- (6) a. *man2* ***khùn5*** *phuu2*
 3SG mountain
 ‘S/he ascended the mountain.’
- b. *man2* *khap1* *lot1* ***khùn5*** *phuu2*
 3SG drive vehicle mountain
 ‘S/he drove a vehicle up the mountain.’

The words *nam2* and *khùn5* appear to function like verbs in one slot (in the (a) examples), and like prepositions in another slot (in the (b) examples). This is suggested not only by an apparent difference in meaning, but also by different possibilities in grammatical behaviour of the words in question. Only in the (a) examples can the words in boldface be immediately preceded by verbal aspect-modality markers (such as *siø* ‘irrealis’, *bòø* ‘not’, *kamlang2* ‘progressive’, *daj4* ‘achievement’, and the experiential marker *kheej2* ‘to have ever v-ed’). In the (a) examples in (5) and (6), the word in boldface is the core predicate of the clause, while in the (b) examples it is a secondary or oblique element. Table 3 (next page) shows 12 items which enter into this alternation.

How does the grammarian best capture the content of Table 3? One approach – which we may call the ‘lexical specification’ approach – would be to say that each of the items in Column b of Table 3 has two distinct senses, represented by the cells in Columns c and d, respectively. Thus, for example, we would find in the lexicon two entries for *khùn5*, one a verb meaning ‘ascend’, another a preposition meaning ‘up’. Similarly, for *nam2*, we would find one entry meaning ‘follow’ (a verb), and another meaning ‘with’ (a preposition). This solution puts the entire burden on the lexicon, meaning that the 12 cells in Column b of Table 3 correspond to 24 distinct dictionary entries (the combined total of cells in Columns c and d). This solution, however, would fail to make explicit a link between senses listed in Columns c and d, and would thus fail to capture the intuition that there are systematic relationships between whole classes of such pairs in the lexicon. The notions ‘ascend’ and ‘up’ are not only closely related semantically, but their relationship to each other is analogous to that between ‘de-

scend' and 'down', 'enter' and 'into', and so on. It would seem preferable to capture this set of relationships, if possible, with a general statement in the grammar. Let us consider two possible solutions, a zero derivation approach and a precategoriality approach.

Table 3. Some items entering into a 'verb-preposition alternation' in Lao.

a. semantic class	b. item	c. English translation in core verb slot	d. English translation in non-core slot
<i>Directional</i>	<i>khùn5</i>	ascend	up
	<i>long2</i>	descend	down
	<i>khaw5</i>	enter	into
	<i>qòðk5</i>	exit	out of
	<i>khaam5</i>	cross over	across
	<i>lòðt4</i>	cross under	across.under
	<i>taam3</i>	follow	along
	<i>phaan1</i>	pass	past
	<i>liap4</i>	move along edge of	along edge of
		<i>qòðm4</i>	move around (sth.)
<i>Comitative</i>	<i>nam2</i>	follow	with
<i>Benefactive</i>	<i>haj5</i>	give	for

According to the zero derivation approach, one of a morpheme's two senses is basic, and the morpheme is entered in the lexicon with this basic sense. Then, a 'zero' morpheme is used to derive the second sense (in this case realised in a different grammatical context). Thus, for example, we would find in the lexicon just one entry for *khùn5* (a verb meaning 'ascend'), and from this verb a preposition meaning 'up' would be 'zero-derived'. Clark and Prasithrathsint (1985; cf. Clark 1989) propose this solution for dealing with data such as that in Table 3, namely verb-preposition alternations in isolating languages of mainland Southeast Asia (including Khmer, Hmong, Thai, and Vietnamese; cf. Lefebvre 2001: 126 for discussion of the same suggestion for typologically similar Creole languages). Similarly, Matthews and Yip (1994: 55) posit 'conversion' in their description of Cantonese. One attraction of this solution for the grammarian is that it would dramatically reduce the number of lexical entries required for dealing with data such as in Table 3, by supplying one derivational rule to be specified in the grammar, representing a productive relationship between

the putatively related senses, and generalising across a class of lexical items. This is easier said than done. As stated already for the examples of open-class heterosemy, such a solution can only be justified if the semantics of the derivational rule are explicitly stated, and the rule can be shown to be consistently productive across the class of items to which the rule is applied.

What kind of derivational rule could work for the data in Table 3? Harrison (1992), writing on Thai (closely related to Lao, and not distinct from it in regard to the present topic), argues that the distinction between *khùn5* as a verb ‘ascend’ and *khùn5* as a preposition ‘up’ is one of presence or absence, respectively, of a ‘temporal profile’ (i.e. whereby ‘its evolution through conceived time is scanned in sequential fashion’; Langacker 1987: 244). In all other respects, the ‘verb’ and ‘preposition’ meanings are identical. The claim, then, could be that the semantic specification of the zero derivation rule is ‘suppress temporal profiling’, with the result that as a verb, *khùn5* predicates a process, while as a preposition it does not. What remains constant across the two meanings of *khùn5* is ‘path of motion’ information. Thus, in the case of *khùn5*, the lexical entry would mean ‘move up’, and the derivational rule would suppress the element of extension through time, thereby suppressing the element of movement, and deriving a preposition meaning ‘up’. Schematically, then, if a verb means ‘to move along path *x*’, a preposition derived from it would mean ‘along path *x*’. This generalisation seems possible for the set of ten items marked as ‘directional’ in Table 3. But it would not work for the benefactive *haj5*, nor the comitative *nam2*, since neither encode spatial (path) information, nor do the differences between the verb and preposition uses of these two items concern mere presence or absence of a ‘temporal profile’. For benefactive *haj5* and comitative *nam2*, the semantic relationship between the verb and preposition meanings is not captured by a general rule. Unless one wants to propose a rule which applied to a single-member class, the two meanings will have to be separately listed in the lexicon.

A zero derivation account for the data in Table 3 could look like this:

Table 4. A zero derivation approach to the data in Table 3.

a. sem'ic class	b. item	specified in the lexicon		derived by rule	
		c. verbs	d. prepositions		
Directional	<i>khùn5</i>	ascend	-	up	
	<i>long2</i>	descend	-	down	
	<i>khaw5</i>	enter	-	into	
	<i>qòòk5</i>	exit	-	out of	
	<i>khaam5</i>	cross over	-	across	
	<i>lòòt4</i>	cross under	-	across.under	
	<i>taam3</i>	follow	-	along	
	<i>phaan1</i>	pass	-	past	
	<i>liap4</i>	go along edge of	-	along edge of	
		<i>qòòm4</i>	go around	-	around
Comitative	<i>nam2</i>	follow	with	-	
Benefactive	<i>haj5</i>	give	for	-	

Lexical entries: 14

Grammatical rules: 1 [i.e. in peripheral ('preposition') slot, 'Suppress temporal profiling']

A second proposal for reducing the burden on the lexicon and capturing a generalisation in the grammar would be to view the relevant lexical entries as 'precatatorial'. According to this approach, the item to be stored in the lexicon is given no inherent specification of grammatical class membership, and whether the form emerges as, say, a verb or a preposition depends on the grammatical context in which it appears (cf. Broschart 1997, Lefebvre 2001, inter alia). Taking such an approach, we would list in the Lao lexicon an entry for *khùn5* which has no specification for word class identity. The verb and preposition functions ('ascend' and 'up') would be concrete realisations of this precatatorial item triggered by different syntactic contexts.

Table 5. A precatatorial approach to the data in Table 3.

a. sem'ic class	b. item	specified in lexicon			derived by rule, in different slots	
		c. (no word class)	d. verbs	e. prepositions		
Directional	<i>khùn5</i>	ascend/up	ascend	up		
	<i>long2</i>	descend/down	descend	down		
	<i>khaw5</i>	enter/into	enter	into		
	<i>qòòk5</i>	exit/out of	exit	out of		
	<i>khaam5</i>	cross over/across	cross over	across		
	<i>lòòt4</i>	cross under/ across.under	cross under	across.under		
	<i>taam3</i>	follow/along	follow	along		
	<i>phaan1</i>	pass/past	pass	past		
	<i>liap4</i>	(go) along edge of	go along edge of	along edge of		
	<i>qòòm4</i>	(go) around	go around	around		
Comitative	<i>nam2</i>	follow	follow	with		
Benefactive	<i>haj5</i>	give/for	give	for		

Lexical entries: 12

Grammatical rules: 4

- i. for deriving verb forms from the underlying abstract lexical entries of the whole set;
- ii. for deriving preposition forms from the underlying lexical entries of the directional set;
- iii. for deriving preposition form from underlying lexical entry of benefactive *haj5*;
- iv. for deriving preposition form from underlying lexical entry of comitative *nam2*.

This account differs from the zero derivation account in positing an abstract underlying meaning for each item (Column c. of Table 5). Furthermore, one needs to state *two* semantic mapping rules, one stating the function by which the verb meaning is derived from the precatatorial entry, and the other stating the function by which the preposition meaning is derived from the precatatorial entry. The main problem is the question of just what the precatatorial item itself means, if it is to be genuinely abstract with respect to the surface realisations, and if the meanings of the output forms are to be proper functions of the input form plus the rule. If the ‘abstract’ underlying meaning of the precatatorial element were specified as a disjunction (e.g. ‘ascend/up’ for *khùn5* or ‘give/for’ for *haj5*; cf. Lefebvre 2001: 132), then

the solution has achieved nothing, since in fact *two* lexical items have been stored in the lexicon (both ‘give’ and ‘for’ being smuggled into a single entry), and the ‘mapping rule’ has become superfluous. That is, the rule wouldn’t *derive* the surface meaning, but would simply specify which disjunct applied. Further, if the ‘abstract’ underlying meaning of the precategorial element were identical to one of the surface realisations (e.g. ‘ascend’ for *khùn5*), such that the mapping onto one syntactic frame required no semantic change in the derivation, then this would be indistinguishable from the zero derivation approach.

Both the zero derivation and precategoriality proposals would significantly reduce the number of entries required in the lexicon (cf. Table 3), but both require care in accounting consistently for the semantic relationship between input and output forms. For example, in the zero derivation approach, it is not sufficient to point out *that* there is a semantic relationship between the source (verb) and output (preposition) forms. The hypothesis of derivation must be supported by an explicit statement of *just what* the semantic relationship between input and output forms is (as for pairs like *knife* (v.) vs. *knife* (n.), and *saw* (v.) vs. *saw* (n.), above). A generalisation seems possible for the set of directionals in Table 3, but it does not cover the remaining two items, which belong to semantic categories of their own. One can only justify a derivational approach to heterosemy (i.e. where only one form is listed, and further forms are grammatically derived) if one can explicitly state the *semantic* conditions of the derivation (specifying input, function, and output), and show that they hold consistently across all members of the relevant class. ‘Semantic regularities’ (Jackendoff 1975) are necessary for generalisations in lexical derivation. The challenge of getting the semantics right is even greater for a precategoriality account, since the precategorial form needs to be characterised semantically such that not one but two source-to-target semantic mappings apply consistently across a class of items.⁴

6. Previously offered solutions

Some scholars of Southeast Asian languages have assumed heterosemy for data such as in Table 3, and have argued for an ‘economising’ solution by grammatical rule. As already mentioned, Clark and Prasithratsint (1985) propose a zero derivation account for directional verbs in prepositional functions, among other cases of cross-class heterosemy, in a number of

mainland Southeast Asian languages. Descriptive grammars of these languages do not offer this kind of solution, but do depict the ‘verb’ and ‘preposition’ roles as distinct.

Huffman (1970), in his description of Khmer (Mon-Khmer), lists a class of ‘directional verbs’ which function as directional prepositions and particles, but he does not explicitly reflect on their status as cross-class heterosemous. He writes: ‘Directional verbs occur after primary verbs which are non-specific as to direction or goal, and specify the direction or general orientation of the action initiated by the verb’ (Huffman 1970: 138). This passage presupposes that directional verbs in this sense are not ‘primary verbs’. Although the term ‘directional verb’ clearly points to the identity of these expressions as a type of verb, Huffman does not discuss the fact that they also function as primary or open class verbs in other contexts.

Li and Thompson (1981) describe a Mandarin (Sinitic) category analogous to Lao directionals in oblique function, calling them *coverbs*, and stating that they are ‘essentially prepositions’ (Li and Thompson 1981: 360). They acknowledge the possibility that one may therefore simply want to call them ‘prepositions’, but reject this on grounds that the class “contains words that are partly like verbs and partly like prepositions” (Li and Thompson 1981: 360). The status of these words is ‘mixed’, they write, because “most of these present-day coverbs used to be verbs at earlier stages of the language, and many of them still have characteristics of verbs and can be used as verbs that have similar meanings” (Li and Thompson 1981: 360). Nevertheless, they list the forms in a distinct section of the grammar describing their role as prepositions, and make no attempt to economise by deriving them from verbs by rule. For Cantonese, Matthews and Yip (1994: 60ff) write of ‘coverbs’, discussing their problematic status as ‘prepositions’, and concluding that ‘as far as prepositions exist in Cantonese, they are a subclass of verbs which may be used as prepositions’ (Matthews and Yip 1994: 62).

A final example (from a somewhat less isolating language) is the class of ‘directional suffixes’ in Manam (Oceanic) (Lichtenberk 1983), which are “added to verbs to indicate the direction of the action ... denoted by the verb” (Lichtenberk 1983: 580). Lichtenberk notes that these suffixes “are homonymous with (and so presumably derived from) the corresponding directional verbs” (Lichtenberk 1983: 580; cf. Lichtenberk 1991).

7. A monosemy analysis

An alternative approach to the directionals in Table 3 is to ask whether they present a case of heterosemy at all. Consider example (6), repeated here:

- (6) a. *man2 khùn5 phuu2*
 3SG mountain
 ‘S/he ascended the mountain.’
- b. *man2 khap2 lot1 khùn5 phuu2*
 3SG drive vehicle mountain
 ‘S/he drove a vehicle up the mountain.’

While *khùn5* is translated in (6a) by English *ascend* (or *go/come up*) and in (6b) by English *up*, one may ask whether this really reflects an underlying distinction in Lao. To posit heterosemy, one must be sure that the item concerned is both polysemous and a member of two different word classes. But it seems possible to say that *khùn5* in (6a) and (6b) has a consistent meaning ‘move along path upwards’, where any apparent difference between ‘verb’ and ‘prepositional’ uses emerges from the nature of the constructional slot in which it appears. If we assume that there is a ‘main verb’ slot, immediately after the subject, then this is the slot in which *khùn5* occurs in (6a) (being the only verb in the clause), and it takes as its direct complement a nominal denoting the ground referent of the path-oriented motion event. The role of *khùn5* in (6a) is to head a clause with the meaning that a certain entity (referred to by the subject noun phrase) moves up with respect to another entity (referred to by the direct complement noun phrase). Accordingly, *khùn5* in (6a) may take regular verb aspect-modality marking, such as the irrealis marker *siø* in the following example:

- (7) *man2 siø khùn5 phuu2*
 3SG IRR mountain
 ‘S/he will ascend the mountain.’

In (6b), however, *khùn5* is not in the same post-subject main verb position. In main verb position is the verb *khap2* ‘drive’, denoting cause and manner of motion, along with a complement noun phrase *lot1* ‘vehicle’ referring to the means of motion. This is followed by *khùn5* and its ground-denoting nominal complement *phuu2* ‘mountain’. Placement of aspect-modality marking in the structure of (6b) must be in the slot immediately before the main verb *khap2* ‘drive’, as follows:

- (8) *man2* *siø* *khap2* *lot1* *khùn5* *phuu2*
 3SG IRR drive vehicle mountain
 ‘S/he will drive a vehicle up the mountain.’

If an aspect-modality marker such as *siø* were placed immediately before *khùn5* in (6a), *khùn5* would be treated as a clausal head. The sentence would then comprise of two separate clauses, whereby the manner of motion expressed in the first verb and the path of motion expressed in the second verb are not understood as features of a single event (thus, with no entailment that the means of ascending the mountain were by vehicle):

- (9) *man2* *khap2* *lot1* *siø* ***khùn5*** *phuu2*
 3SG drive vehicle IRR mountain
 ‘S/he drove (or: is driving) a vehicle, (s/he was/is) going to go up the mountain.’

The inability of *khùn5* in its oblique role in (6b) to take direct aspect-modality marking could be taken as an index of its non main-verb status, supporting a heterosemy analysis. (This is the sense in which Lord 1993 refers to grammaticalising serial verbs as ‘defective’.) But perhaps we could instead interpret this constraint as a property of the *construction*. (6b) would then exemplify a construction which provides two slots for two types of verb. The first slot is open to verbs from a large (probably open) class, denoting any kind of activity which can be understood as a cause of motion. The second slot is open to verbs from a closed class (10 members), denoting motion along an abstract path (‘up’, ‘down’, ‘across’, etc.). The resulting construction is a complex predicate in which the order of the two verbal components is fixed, and where the two verbs do not represent two events but complementary specifications of the same event.

It is notable that none of the grammarians mentioned in the previous section, faced with the same puzzle posed by the Lao data in Table 3, opt for a monosemy analysis (putting the ‘verb/preposition’ alternation down to a difference in constructional environment), nor do they try stating a single rule which would derive the directional ‘prepositions’ from the class of directional verbs (even though this could have been done in each case). While the theorising, tinkering, ‘structuralist grammarian’ portrayed by Pawley (1996) would certainly have wielded Ockham’s Razor and opted for economy in one way or another, these non-mythical descriptive grammarians have approached the task more pragmatically.

8. Cross-class heterosemy in the Lao verbal complex: Extension in diachrony

Many grammatical markers of the verb in Lao are cross-class heterosemous. However, the synchronic semantic relationships between open and closed class senses are highly idiosyncratic. For this reason, there are few if any possibilities for lexicon-reducing grammatical generalisations.

The template in Figure 1 shows the Lao verbal complex, with slots for aspect/modality marking on the verb:

ASP/MOD_A – IRR – NEG – ASP/MOD_B – *daj4* – DIR.PCL – [VERB (OBJ)] – ASP/MOD_C

Figure 1. Elements of the Lao verb complex, in order.

Most morphemes marking grammatical meanings on the verb – specifically, those appearing in the slots marked ‘ASP/MOD’ (aspect-modality), ‘*daj4*’ (achievement marker), and ‘DIR.PCL’ (directional particles) – also function as main verbs or nouns. The following examples feature the item *lèèw4* (which functions either as a main verb meaning ‘finish’ or as a post-verbal perfective marker) and the item *daj4* (which functions either as a main verb ‘acquire’ or as a preverbal modal marker of achievement):

- (10) a. *khòj5 lèèw4 viak4 nii4*
 1SG work DEM.GEN
 ‘I have finished this work.’
 b. *khòj5 hèn3 man2 lèèw4*
 1SG see 3SG
 ‘I already see him/her.’
- (11) a. *khòj5 daj4 ngen2 laaj3*
 1SG money much
 ‘I got a lot of money.’
 b. *khòj5 daj4 hèn3 man2*
 1SG see 3SG
 ‘I did see him/her.’

Here is a list of some of the cross-class heterosemous items which appear in the template depicted in Figure 1:

Table 6. Some cross-class heterosemous terms found in the Lao verb complex.

a. item	b. meaning in open class	c. meaning in closed class (and position in 'Lao verb complex' template)
<i>daj4</i>	v. obtain	i. can (ASP/MOD _C) ii. ACHIEVEMENT (<i>daj4</i>)
<i>lèw4</i>	v. finish	PERFECTIVE (ASP/MOD _C)
<i>kamlang2</i>	<i>n.</i> energy	PROGRESSIVE (ASP/MOD _A)
<i>kheej2</i>	v. accustomed	EXPERIENTIAL (ASP/MOD _B)
<i>khanaat5</i>	<i>n.</i> extent	INTENSIFIER (ASP/MOD _C)
<i>ñang2</i>	v. remain	not yet (ASP/MOD _A , with negation)
<i>thua1</i>	<i>n.</i> occasion, time	not yet (ASP/MOD _C , with negation)
<i>jaak5</i>	v. want	somewhat (ASP/MOD _B)
<i>mak1</i>	v. like	tend to (ASP/MOD _B)

A grammarian might hope to economise by removing such alternations from the lexicon and capturing them in grammatical rules. As stated earlier, to qualify as a derivational statement which applies to a class with more than a single member, the required rule would need to specify a function with consistent input and output, both semantic and morphosyntactic, across a class of items. However, the semantic idiosyncrasies of the alternations in Table 6 prevent any such rule from being found. As closed-class items, Lao words such as *daj4* (as preverbal modal) and *lèw4* (as postverbal aspectual) must be individually treated in the relevant section of the grammar (e.g. in a section on grammatical marking of modality and aspect in the verbal complex). But as open class items (i.e. as verbs meaning 'acquire' and 'finish', respectively), they deserve no special treatment.

Ultimately, however, the grammarian would be taking too far his economising mission were he or she not to draw attention to these relationships, despite the inability of rules to capture them. Some grammarians include dedicated sections on cross-class heterosemous items, enumerating their various open and closed class functions. Thompson's Vietnamese grammar includes a long appendix entitled 'Glossary of difficult forms', in which many entries are cross-class heterosemous (Thompson 1965: 336ff). Huffman's description of Khmer dedicates a section to *baan*, a cross-class heterosemous item which occurs 'in three different syntactic positions and with three different functions' (Huffman 1970: 74; cf. Haiman 1999 and Enfield 2001, 2003). Smyth's sketch grammar of Thai includes an appendix dedicated to three cross-class heterosemous items *hây* ('give', 'for',

‘cause’, etc.), *dây* (‘obtain’, ‘can’, ‘perfective’, etc.), and *pen* (‘be’, ‘be able’, ‘as’, etc.; Smyth 2002: 218).

While there is no productive synchronic derivational relationship attributable to most cross-class heterosemy, it is usually obvious that there is a derivational relationship on the diachronic dimension (Durie 1988; Lichtenberk 1991; Lord 1993; Diller 2001; Enfield 2003: 125ff; note, however, Lefebvre’s 2001: 132ff warning that synchronic multifunctionality is not necessarily due to language-internal processes of grammaticalization). The source of cross-class heterosemy in isolating languages is often verb serialisation, which facilitates the emergence of new meanings for open-class words in closed class positions. Lord (1993) focuses on serial verb constructions ‘in which one of the verbs is defective in some respect, such as phonological assimilation, failure to take the usual verb inflections or negation affixes, or showing unexpected syntactic properties’ (Lord 1993: 3). This formal ‘defectiveness’ in a restricted grammatical position is an index of semantic change, revealing a new status of the erstwhile open-class item as a closed-class grammatical marker.

The process of grammaticalisation by definition involves the historical passage of a morpheme from open- to closed-class status (Traugott and Heine 1991, Hopper and Traugott 1993, *inter alia*), as diagrammed here:

$$(12) [a_{\text{OPEN}}]_{t1} \rightarrow [b_{\text{CLOSED}}]_{t2}$$

Research on historical semantics has established that a simple ‘*a* becomes *b*’ scenario is too simple, and that a medial stage of polysemy must be recognised in the unfolding of meaning change (Sweetser 1990, Evans 1992, 1997, 2003, Wilkins 1996, Enfield 2003). That is, if a word meaning ‘*a*’ comes to mean ‘*b*’, there must be an intervening stage in which the word is ambiguous, with possible meanings ‘*a*’ and ‘*b*’. This is illustrated in stage ‘t2’ in the following diagram:

$$(13) [a_{\text{OPEN}}]_{t1} \rightarrow [a_{\text{OPEN}}, b_{\text{CLOSED}}]_{t2} \rightarrow [b_{\text{CLOSED}}]_{t3}$$

Cross-class heterosemy arises precisely in stage ‘t2’ in (13), i.e. where the ‘*a*’ and ‘*b*’ meanings are associated with open and closed class positions, respectively. The importance of cross-class heterosemy to the typological study of semantic change and grammaticalisation is reason enough for the grammarian to note it (and elaborate on it, to the extent possible) in a grammatical description.

9. Summary and conclusion

Pawley (1996) depicts the ‘structuralist grammarian’ as an economy hound, bent on looking for grammatical generalisations by which he or she can minimise information to be included in the lexicon. To the extent that it really is the grammarian’s job to do so, balancing the respective burdens of lexicon and grammar must be done with simultaneous consideration of formal and semantic criteria. Treatments of the issue to date have failed to deal adequately with semantics. If a derivational rule is to account for regular patterns of heterosemy, then the semantic relationships between putative source and target structures must be explicitly stated and shown to be productive. Only when one can identify semantic relationships which are consistent for a whole set of alternations can one state a rule which allows predictable output across a class of lexical entries, thus allowing a statement in the grammar to genuinely reduce the size of the lexicon. Typically, in a description of derivational morphology, the grammarian’s observation of apparently consistent relationships across lexical items fails to achieve economy, in that it does not obviate the need to list separate lexical entries for each derived item in the pattern. Apparently regular alternations arising from heterosemy are no different, and relationships of cross-class heterosemy in particular tend to be semantically idiosyncratic enough to resist even the most abstract generalisation. It is clear, however, that failure to achieve ‘economy’ does not detract from the utility of discussing general patterns observed in the lexicon of a language. Such perceived sets of relationships, particularly given their common diachronic significance, are of intrinsic interest in a grammatical description.

Notes

1. Here, and in other examples below, some glosses are omitted intentionally so as not to pre-judge the unity or separateness of meaning of a given item in different contexts. Transliteration of Lao used here follows IPA convention except for velar nasal /*ŋ*/, glottal stop /*q*/, mid-high front vowel /*ê*/, low front vowel /*è*/, central vowel (schwa) /*e*/, low central vowel /*a*/, high back unrounded vowel /*ù*/, low back vowel /*ò*/. Tones are represented by numbers following each syllable (\emptyset =unstressed/atonal, 1=/33/, 2=/35/, 3=/13/, 4=/53/, 5=/31/). Abbreviations in interlinear glosses are: 1=1st person, 2=2nd person, 3=3rd person, CLF=classifier, COP=copula, DEM=demonstrative, GEN=(semantically) general, IRR=irrealis, SG=singular.

2. When the meaning of the derived term cannot be computed or uniquely predicted by the meaning of the constituent parts (e.g. from ‘under’ and ‘arm’ in *underarm* or ‘knife’ and ‘USED AS VERB’ in *He knifed me*), the derived form has an indexical relationship to the source form. The constituent morphemes do not constitute the whole derived meaning, they constitute only part of it, indexing the lexically specified whole meaning. The Peircean distinction between symbolic and indexical meaning is useful in distinguishing between inflection (where the symbolic values of the constituent morphemes and the construction directly encode the whole meaning) and derivation (where, usually, the symbolic values of the constituent morphemes point to, but do not by themselves constitute, the lexically specified whole meaning).
3. Clark and Clark (1979) argue for pragmatic inference as a way of accounting for effects which arise ‘when nouns surface as verbs’, citing examples like *He porched the newspaper*. This is undoubtedly an important basis for speakers’ production and interpretation of novel uses, and is a promising avenue for explaining historical extensions across word classes, but once an extension has become lexicalised (e.g. the verb *knife*), the meaning becomes context-independent, and comprehension by pragmatic inference is no longer a viable account. Conventional meanings will pre-empt conceivable pragmatic inferences. Other authors who have discussed possible rules for derivation (cf. Voorhoeve 1981, Vonen 2000) fail to adequately specify the *semantic* output of the said rule. The same can even be said for those who have argued that semantics is the crucial issue (Ruhl 1989; cf. Lakoff 1987).
4. The two approaches have different implications if meant as hypotheses about the grammar as a mental object. Thus, the zero derivation analysis would entail that one of the two forms was conceptually basic and/or prior with respect to the other, while the precategorial approach would not. In the absence of any independent or principled reason to regard one of the two meanings as more basic, this would perhaps be an attraction of the latter approach.

Acknowledgements

I would like to thank Felix Ameka, Louise Baird, Melissa Bowerman, Alan Dench, Nick Evans, Bhuvana Narasimhan, and Rachel Selbach for providing helpful comments and discussion on earlier drafts of this chapter. I also gratefully acknowledge the support of the Max Planck Society.

References

- Anderson, Stephen R.
1985 Inflectional morphology. In *Language Typology and Syntactic Description (Volume III Grammatical Categories and the Lexicon)*, Timothy Shopen (ed.), 150–201. Cambridge: Cambridge University Press.
- Apresjan, Juri D.
1974 Regular polysemy. *Linguistics* 142: 5–32.
- Broschart, Jürgen
1997 Why Tongan does it differently: categorial distinctions in a language without nouns and verbs. *Linguistic Typology* 1 (1): 123–165.
- Clark, Eve V. and Clark, Herbert H.
1979 When nouns surface as verbs. *Language* 55 (4): 767–811
- Clark, Marybeth
1989 Hmong and areal Southeast Asia. In *Southeast Asian Syntax*, David Bradley (ed.), 175–230. (A-77) Canberra: PL.
- Clark, Marybeth, and Prasithratsint, Amara
1985 Synchronic lexical derivation in Southeast Asian languages. In *Southeast Asian Linguistic Studies Presented to Andre-G. Haudricourt*, Suriya Ratanakul, David Thomas, and Suwilai Premsrirat, (eds.) 34–81. Bangkok: Mahidol University.
- Croft, William
2001 *Radical Construction Grammar: Syntactic Theory in Typological Perspective*. Oxford: Oxford University Press.
- Diller, Anthony
2001 Grammaticalization and Tai syntactic change. In *Essays in Tai Linguistics*, M. R. Kalaya Tingsabath and Arthur S. Abraham (eds.), 139–175. Bangkok: Chulalongkorn University Press.
- Dixon, R. M. W.
1991 *A New Approach to English Grammar, on Semantic Principles*. Oxford: Clarendon Press.
- Durie, Mark
1988 Verb serialization and ‘verbal-prepositions’ in Oceanic languages. *Oceanic Linguistics* 27 (1/2): 1–23.
- Enfield, N. J.
2001. Remarks on John Haiman, 1999. ‘Auxiliation in Khmer: the case of baan’, *Studies in Language* 23 (1): 149–172. *Studies in Language* 25 (1): 115–124.

- Enfield, N. J.
2003 *Linguistic epidemiology: Semantics and Grammar of Language Contact in Mainland Southeast Asia*. London: RoutledgeCurzon.
- Evans, Nicholas D.
1992 Multiple semiotic systems, hyperpolysemy, and the reconstruction of semantic change in Australian languages. In *Diachrony within Synchrony: Language History and Cognition*, Günter Kellerman and Michael D. Morrissey (eds.), 475–508. Frankfurt am Main: Verlag Peter Lang.
- Evans, Nicholas D.
1997 Sign metonymies and the problem of flora-fauna polysemy in Australian linguistics. In *Boundary Rider: Essays in Honour of Geoffrey O'Grady*, Darrell Tryon and Michael Walsh (eds.), 133–153. (C-136) Canberra: PL.
- Evans, Nicholas D.
2003 Context, culture, and structuration in the languages of Australia. *Annual Review of Anthropology* 32: 13–40.
- Harrison, Colin
1992 The grammar of directional serial verb constructions in Thai. Honours dissertation, Department of Linguistics, Australian National University.
- Hopper, Paul J. and Elizabeth Closs Traugott
1993 *Grammaticalization*. Cambridge: Cambridge University Press.
- Huffman, Franklin E.
1970 *Modern Spoken Cambodian*. Ithaca, NY: Cornell University, Southeast Asia Program.
- Jackendoff, Ray
1975 Morphological and semantic regularities in the lexicon. *Language* 51 (3): 639–671.
- Lakoff, George
1987 *Women, Fire, and Dangerous Things: What Categories Reveal About the Mind*. Chicago: University of Chicago Press.
- Langacker, Ronald W.
1987 *Foundations of Cognitive Grammar: Volume I, Theoretical Prerequisites*. Stanford: Stanford University Press.
- Lefebvre, Claire
2001 Multifunctionality and the concept of lexical entry. *Journal of Pidgin and Creole Languages*. 16 (1): 107–145.
- Levin, Beth, and Malka Rappaport Hovav
1995 *Unaccusativity: at the Syntax-Lexical Semantics Interface*. Cambridge, Mass./London: MIT Press.

- Li, Charles and Sandra Thompson
1981 *Mandarin Chinese: A Functional Reference Grammar*. Berkeley: University of California Press.
- Lichtenberk, Frantisek
1983 *A Grammar of Manam*. Hawaii: University of Hawaii Press.
- Lichtenberk, Frantisek
1991 Semantic change and heterosemy in grammaticalization. *Language* 67 (3): 475–509.
- Lord, Carol
1993 *Historical Change in Serial Verb Constructions*. Amsterdam/Philadelphia: John Benjamins.
- Macdonald, R. Ross, and Soenjono Darjowidjojo
1967 *A Student's Reference Grammar of Modern Formal Indonesian*. Washington, DC: Georgetown University Press.
- Matthews, Stephen, and Virginia Yip
1994. *Cantonese: A Comprehensive Grammar*. London: Routledge.
- Pawley, Andrew
1996. Grammarian's lexicon, lexicographer's lexicon: worlds apart. *KHVAA Konferenser*, 36: 189-211.
- Persson, Gunnar
1988 Homonymy, polysemy, and heterosemy: three types of lexical ambiguity in English. Paper read to Third International Symposium on Lexicography, University of Copenhagen: May 1–16, 1986.
- Ruhl, Charles
1989 *On Monosemy*. Albany: State University of New York Press.
- Smyth, David
2002 *Thai: An Essential Grammar*. London and New York: Routledge.
- Sweetser, Eve
1990 *From Etymology to Pragmatics: Metaphorical and Cultural Aspects of Semantic Structure*. Cambridge: Cambridge University Press.
- Thompson, Laurence C.
1987[1965]. *A Vietnamese Grammar*. (Second edition.) University of Hawaii Press.
- Traugott, Elizabeth Closs and Bernd Heine (eds.)
1991 *Approaches to Grammaticalization*. Amsterdam: John Benjamins.
- Vonen, Arnfinn Muruvik
2000 Polynesian multifunctionality and the ambitions of linguistic description. In *Approaches to the Typology of Word Classes*,

Petra M. Vogel and Bernard Comrie (eds.), 479–487. Berlin: Mouton de Gruyter.

Voorhoeve, Jan

1981 Multifunctionality as a derivational problem. In *Generative Studies on Creole Languages*, Pieter Muysken (ed.). Dordrecht: Foris, 25–34.

Wilkins, David

1996 Natural tendencies of semantic change and the search for cognates. In *The Comparative Method Reviewed: Regularity and Irregularity in Language Change*, Mark Durie and Malcolm Ross (eds.), 264–304. New York: Oxford University Press.