Studies in Language 13-1. 51-103 (1989).

THE CREATION OF TENSE AND ASPECT SYSTEMS IN THE LANGUAGES OF THE WORLD

JOAN L. BYBEE & ÖSTEN DAHL
University of New Mexico *University of Stockholm*

Introduction

A characteristic and universal property of natural languages is the use of grammatical morphemes — morphemes which belong to closed classes and exhibit grammatically regular distributional properties — alongside lexical ones. Grammatical morphemes perform a large share of the work of grammar, for perhaps more than distinctive position (word order), grammatical morphology is the major signal of grammatical and discourse structure, as well as temporal and aspectual relations.

The formal properties of grammatical morphemes have been well-studied and a rich terminology has been developed to deal with differences of expression properties. Thus we have prefixes, suffixes, infixes, prepositions, postpositions, clitics, auxiliaries, reduplication, stem change, ablaut and so on. A major distinction in grammatical morphemes may be made between **bound** and **periphrastic** expression, and these have been traditionally treated under morphology and syntax respectively. The fact that grammatical morphemes are regarded as belonging to two different domains of grammar makes it difficult to find terms that are generally applicable. 'Morpheme' is a term that is difficult to apply to morphological processes such as ablaut and stem change, for example. 'Inflectional category' is clearly too narrow and 'grammatical category' too wide. To solve the terminological problem, we will here use the neologism 'gram' (a shortening of 'grammatical morpheme') to apply to the whole class of items enumerated above.

Another terminological problem arises regarding the content of grams. Many terms used in grammar may be understood as ambiguous between referring to 'notional' ('semantic') or 'grammatical' categories. Thus 'pro-

gressive' may either denote a certain meaning or context of use (which presumably can occur in any language) or a category, like the English Progressive, which has a certain meaning and a certain expression.[1] Our 'grams' are entities of the second kind.

Notional and grammatical categories are not always kept apart in the literature. In this paper, we shall try to keep the distinction clear, and we shall furthermore deviate from the not uncommon identification of the distinction 'notional vs. grammatical' with the distinction 'universal vs. language-specific'. In fact, our main thesis is that the meanings of grams are cross-linguistically similar, making it possible to postulate a small set of cross-linguistic **gram-types**, identifiable by their semantic foci and associated with typical means of expression. The basis for this claim, as well as for the others that we shall make, will be two independent studies of tense and aspect (Dahl 1985 and Bybee 1985). The present paper is an attempt to integrate the results of these studies in order to come closer to a general theory of grams with tense and aspect as a special case.

1. Comparing the meaning of grams across languages

The meaning of grams is characterized by an abstractness and relational quality which makes it notoriously difficult to pin down in a single language, and even more difficult to compare across languages. The idea that the meaning of grams may be similar across languages has often been denied in American linguistics during the current century. Reacting against the view that languages should be judged on a Latin-based scale of adequacy, the Boasian tradition, as continued by Whorf and Sapir, denied semantic universals and treated each language as expressing meanings reflecting, and perhaps even molding, a culture-specific world-view. Thus, not only were similarities across languages denied, the investigator was considered to be better off forgetting all he or she knew about other languages when embarking on the analysis of a new one. A similar result obtained in European structuralism where grams were taken to form sets of semantic oppositions, which, like phonemes, were necessarily language-specific. Subsequent developments in American linguistics took us even further from universals of grammatical meaning, as Bloomfield began a tradition which explicitly denied access to meaning at all. From a methodological principle this was elevated to a theoretical principle by Chomsky who asserted not only that grammatical description should be autonomous from semantic,

but also that there is not any interesting semantic difference between grammatical and lexical morphemes (Chomsky 1957:104-105).

However, a trend in the opposite direction has emerged in the last decade. Especially in the area of tense and aspect, which are the focus of the current paper, the evidence that there may be significant generalizations to make cross-linguistically has been growing. The appearance of books entitled *Aspect* (Comrie 1976) and *Tense* (Comrie 1985), which draw on data from multiple languages, and conferences on "tense and aspect" with resulting volumes such as that edited by Tedeschi and Zaenen (1981) or Hopper (1982), in which scholars discussing tense and aspect in diverse languages seem to feel as though they are talking about the same thing, suggest that there may be some way of arriving at a cross-linguistic understanding of grammatical meaning, or more ambitiously, a universally valid theory of grammatical meaning.

The impetus for the current paper is the fact that the two authors independently and using different methodologies undertook extensive cross-linguistic comparison of verbal morphology and arrived at very similar results concerning the meanings and expression properties of tense and aspect grams in the languages investigated. We found that certain meanings tend to have grammatical expression so that if the domain of tense and aspect were to be compared to the color spectrum, we would be able to identify certain areas, comparable to the focal colors of Berlin and Kay (1969), that are commonly expressed by grams in the languages of the world. In addition, we each independently discovered that certain correlations between meaning and mode of expression exist for grams viewed cross-linguistically, and this correlation suggests a theory that helps us understand both the similarities and the differences among grams and their meanings across languages. Before saying more about the results of these two studies, we will present a sketch of the two methodologies.

2. The two studies

The study reported on in Bybee (1985) was designed to test certain hypotheses concerning the relation between the meaning of inflectional grams and the degree of fusion they exhibit with a lexical stem. To test these hypotheses on a world-wide sample of languages, Bybee used a sample drawn up by Revere Perkins (1980), which contained fifty languages that were randomly selected while controlling for genetic and areal bias.

Information on these languages was available only through published material, usually reference grammars. The test of the hypotheses required identifying verbal inflection as belonging to one of the super-categories of valence, voice, aspect, tense, mood or agreement. Despite the fact that some descriptions were brief and offered only a few examples, it was nonetheless possible to make this assignment with some degree of confidence in most cases. In addition, an attempt was made to identify meanings expressed by particular inflections as belonging to categories traditionally defined in the literature, such as **perfective**, **imperfective**, **past** and **future**.

At this level, the study was primarily exploratory; it was an attempt to see to what extent there was cross-linguistic comparability of grammatical meaning. In order to categorize the meanings of the verbal inflections, it was not possible to rely on the labels that authors assigned to the inflections, for terminology varies widely, and the confusion of aspect with tense is commonly encountered. Rather it was necessary to examine the author's description of the appropriate contexts of use of an inflection, as well as the examples, in order to decide how to categorize it. The list of possible categories was not closed, so that an inflection which did not fit under any existing labels could also be accomodated.

In Dahl's project, data were collected through a questionnaire containing about 150 sentences with indications of contexts, chosen in such a way as to give as good a sample of the tense-mood-aspect field as possible. The questionnaire was translated into 64 languages by native informants. Interference from English was minimized by giving the verbs in the questionnaire in the base form and letting the informants choose the right categories in their own languages on the basis of the contextual indications given. The translations were analyzed, i.e. labels were assigned to each predicate indicating the tense or aspect category chosen. The results of the analysis were fed into a computer and processed in a data base system whereby inter- and intralinguistic comparisons of the distribution of categories was facilitated.

The sample of languages used in Dahl's investigation was mainly a convenience one with the bias towards European languages that is inevitable if one does not use a strict sampling method. Still, the sample has reasonably good coverage of the world's languages: all major continents are represented by at least two items. A smaller sample of 18 languages which were chosen more carefully was used to check the validity of any statistical generalizations.

Despite differences between the studies in source of information and scope, they both found that a large majority (between 70% and 80%) of the grams marking tense or aspect notions in the languages studied belongs to one of the following six gram-types (characterized only roughly here, but discussed in more detail below):

a. **perfective**, indicating that a situation is viewed as bounded;

b. **imperfective**, indicating that the situation is viewed as not bounded;

c. **progressive**, (called **continuous** in Bybee's study) indicating the situation is in progress at reference time;

d. **future**, indicating that the speaker predicts a situation will occur subsequent to the speech event;

e. **past**, indicating that the situation occurred before the speech event;

f. **perfect**, (called **anterior** in Bybee's study) indicating that a situation is being described as relevant at the moment of speech or another point of reference.

The **present** in Bybee's survey was unmarked in the majority of cases. Dahl does not postulate present as a gram-type: present tenses were in general treated as 'default' members of categories. The majority of these had zero marking.

In addition to these six major gram-types, combinations of them were found. For instance, the perfect and progressive are often combinable with a past tense to give a past perfect (or pluperfect) and a past progressive respectively. To what extent these should be regarded as separate gram-types is partly an open question (for a discussion, see Dahl 1985: 67). Furthermore, both surveys turned up other gram-types, which included an habitual in several languages, remoteness distinctions mainly restricted to Bantu languages, and a number of unique grams, whose occurrence was restricted to a single family.[2] Aspectual grams such as iterative and inceptive were noted in Bybee's study, but classified as derivational since such markers are usually restricted to verbs of certain semantic types. Thus the six gram-types listed above are far and away the most common and the most widespread of grams marking notions of tense and aspect.

3. Form/meaning correlations

Both of our studies show that it is possible to make substantive generalizations concerning the semantic content of grams of tense and aspect. This means that a form / meaning correlation exists in the languages of the world, such that the meanings expressed by grammatical morphemes are not only distinct from those expressed by lexical morphemes, but are also universally characterizable. In addition, however, both studies uncovered a correlation between meaning and form even among grammatical morphemes.

Although Dahl's study focuses on the uses of tense and aspect grams, certain expression properties of the grams were noted, in particular, whether they are periphrastic, bound or zero-marked expressions. When expression is compared with meaning, a striking correlation becomes evident: certain categories usually have periphrastic expression while others usually have bound expression. In particular, perfect, and progressive usually have periphrastic expression, while past, and perfective and imperfective usually have bound expression. The future is fairly evenly split between periphrastic and bound expression, as shown in Table 1.

Table 1. Expression of major gram-types in Dahl (1985)

Periphrastic		Bound	
perfect (16/18)	88%	past (33/45)	73%
		perfective (17/20)	85%
progressive (18/19)	95%	imperfective (7/7)	100%
future (27/50)	54%	future (23/50)	46%

We propose that an explanation for this correlation must take into account the manner in which grams develop over time: grams develop out of lexical material by a gradual generalization of meaning which is paralleled by a gradual reduction in form and fusion with the head (in this case the verb).[3] Perfect and progressive are less grammaticized, less general meanings, and thus show less grammaticization of form. Past, perfective and imperfective are more abstract and general grammatical meanings, and thus show more grammaticization of form. Furthermore, an actual diachronic relation can be demonstrated between pairs of grams: a perfect tends to develop into a past or perfective (as in Romance languages, Harris 1982), and a progressive tends to develop into an imperfective (as in Celtic and Yoruba [Comrie 1976] and Kru languages [Marchese 1986]).

Further evidence of a form / meaning correlation due to evolution is given in Bybee 1985. While this study included only inflections, detailed information about their mode of expression was recorded, including information about the order of inflections with respect to the stem, and the degree of fusion with the stem as measured by the presence of stem changes conditioned by the inflection. A comparison of the semantically more specific gram-types of habitual and continuous, and the more general perfective and imperfective, revealed that habitual and continuous grams show a strong tendency to be farther from the stem, and to condition stem changes much less often than perfective and imperfective inflections.[4] We take these facts to indicate that perfective and imperfective have undergone a longer course of development than habitual and continuous grams.

The form/meaning correlation suggests a universal theory of tense and aspect which includes a diachronic dimension. This theory proposes that the paths along which grams develop may be the same or similar across languages, and that the differences among the meanings expressed by tense and aspect grams across languages correspond to the location the particular gram occupies along one of these universal paths at a particular time.

The major tense and aspect categories identified in the Dahl (1985) and Bybee (1985) studies point to three major paths of development:

a. expressions with a copula or possession verb plus a past participle, or verbs meaning 'finish', 'come from' or 'throw away', develop into grams marking anterior or perfect, which in turn develop into perfectives or pasts;

b. expressions with a copula, locative or movement verb develop into progressives which in turn develop into imperfectives;

c. expressions with a verb meaning 'desire', 'movement towards a goal' or 'obligation' develop into grams expressing intention and future.

Further evidence for the universal semantic basis of grammaticization is the fact that the lexical material which evolves into grammatical material expressing tense and aspect meanings also appears to be the same or similar across languages. The following examples document this fact by showing that each of the major lexical sources for the three major paths can be found to occur in at least three unrelated languages. (See Table 2.)

Table 2. Lexical sources for tense and aspect grams

Lexical source	Grammatical category	Languages/Examples
desire	› intention › future	English: *willan* 'want' › *will* 'future' Swahili: *taka* 'want' › *ta* 'future' Mandarin: *yào* 'want' › *yào* 'future'
movement towards goal	› intention › future	English: *be going to* Ewe: *vá* 'come' › *á-* 'future' Duala: **ɛndɛ* 'go' › *-ɛndɛ* 'future'
have/be + inf	› obligation › future	Latin: inf + *habeo* › Spanish *-ré* future Vata (Kru): *ka* 'has' › *ká* 'future' English: *be to* + verb
have/be + Past Participle	› perfect › past or perfective	English: *have done* French: passé composé Finnish: copula + past part.
finish	› perfect › past or perfective	Mandarin: *liao* 'finish' › *le* perfective Ewe (Dahome): *kɔ* 'finish' › *kɔ* past tense Spanish: *acabar* 'fo finish' › "to have just"
throw away	› perfect	Palaung: *pět* 'throw away, finish' › perfect Korean: *pelita* 'to throw away' › perfect Fore: *kai* 'cast aside' › perfect
movement from source	› perfect › past or perfective	French: *venir de* 'come from' › 'to have just' Teso (E. Nilotic): *-bu/-potu* 'come' › past Somali (Jiddu): *-ooku* 'come' › past Palaung: *yu* 'to rise up, to come from' › past
locational/ postural verb + verb	› progressive › imperfective	Iris: copula + *ag* 'at' + verbal noun Cocama: *yuti* 'be located' Diola-Fogny: verb + copula + locative prep Ngambay-Mundu: *ísi* 'sit' + verb *ár* 'stand' + verb
movement	› progressive	Spanish: *andar, ir, venir* + present part. Tarahumara: verb + *eyéna* 'go' Tatar: gerund + *kil-* 'come'; + *bar-* 'walk' Turkish: 'to go, to walk' › *-yor* present

Since we regard these diachronic paths as essential for our understanding of tense and aspect as a cross-linguistic as well as a language-particular phenomenon, this paper is concerned primarily with describing the three paths of development as established by data from different languages. It is also necessary, however, to explicate our view of the grammaticization process, which calls for a discussion of the properties that characterize grams, a matter to which we now turn.

4. Formal and semantic properties of grams

The development of grams out of lexical material is a gradual process, which means that in any particular language at any particular time we will find grams in various stages of development. The theory proposed here views inflectional material as the most highly grammaticized, which means that inflection exhibits the largest number of grammatical properties. Other grams are considered more or less grammaticized depending upon the number of properties they share with inflection. We do not exclude derivational morphemes from consideration entirely, since they are also grammatical morphemes. In fact, in Section 7, we discuss derivational perfective grams. However, we have chosen to view inflectional grams as exhibiting the highest degree of grammaticization, since derivational grams tend to show more of the properties of lexical morphemes (i.e. having idiosyncracies of meaning, restricted distribution, etc.).

4.1 *Closed classes*

The defining property of grams that distinguishes them from lexical morphemes — membership in a closed class — is in some ways a gradient notion itself. Classhood is determined by formal properties, such as positional constraints, and these may be more or less rigid. Furthermore, the size of the class may vary, and in some cases it may be difficult to determine if a class is indeed closed or not. For instance, in some languages with noun incorporation, a fairly large, but apparently closed class of semantically determined nouns may be incorporated (e.g. Pawnee, Parks 1976). Whether the class is closed or not may be difficult to determine, since adding new nouns in the appropriate semantic field (such as body parts) may not occur frequently. It is sometimes difficult to determine whether auxiliary constructions constitute closed or open classes. For instance, is the

class that includes *have to*, *want to*, *need to*. *try to*, *start to*, etc. a closed or open class? In this case, as well as in the case of incorporated nouns, the closer the classes are to lexical classes, the more difficult it is to determine if they are open or closed.

On the other hand, many closed classes consist of a single member, or a single overt member contrasting with zero. The English Perfect gram, *have* + Past Participle and the Progressive, *be* + *ing*, are each the sole occupants of their classes.

Since we are arguing that lexical morphemes can become grammatical, it would seem to follow that new closed classes items may be added to a language. While this is true, it is also the case that new grams are rarely added to existing closed classes, rather, as they grammaticize, they create new closed classes. Thus the original modal auxiliaries in English, *may*, *can*, *must*, *will*, *shall*, etc. constitute a closed class. Some of the properties of this class are that its members appear before the verb in questions, the negative marker occurs after each one, they take the main verb without *to*, and so on. Newer auxiliary constructions in English, such as *be going to* and *have to* do not share these properties, and cannot, because these properties are specific to the items with which they are associated and the period in which they developed. For example, the older modals do not take the infinitival *to* because at the period in which they began to grammaticize, the infinitive marker was a suffix, *-an*. The newer auxiliaries take *to* because they have been formed in a period in which *to* is the infinitive marker. Grams fossilize the particular syntactic and morphological properties operative at the time the construction from which they develop is first formed. Just as a new word does not undergo old unproductive sound changes, a new gram does not acquire old unproductive grammatical properties.

It should also be remembered that not all members of a closed class are at the same stage in the grammaticization process, rather the individual members may have differing properties, especially to the extent that the properties are determined by the meaning of the gram. For example, some of the English modal auxiliaries contract with the negative gram while others do not. Compare, *can't*, *won't*, *shouldn't* with *may not*.

Finally, it is not true that grams that occur in the same closed class necessarily form a semantically coherent set. In Bybee 1986 it is shown that the inflections in the 50-language sample fail to show a correlation between membership in the semantic categories of aspect, tense and mood and membership in positional classes. Rather the situation as seen in English is

typical: each member of a semantic category may belong to a different class, as the tenses past (*-ed*), future (*will*) and perfect (*have* + past participle) do, or a single class may express meanings from different semantic categories, as the modal auxiliaries, which express tense (*will* and *shall*), deontic modality (some readings of *must* and *should*) or epistemic modality (*may* and *might*)). This lack of correlation between structural and semantic classes is predicted by grammaticization theory as we are developing it here: if each gram follows a path of development according to its original meaning, then it develops independently of other grams. It belongs to a structural class if other grams from structurally similar sources (such as auxiliary verbs) undergo grammaticization at approximately the same period of time. Its membership in a structural class, then, is not determined solely by its meaning, but at least in part by chronological coincidence.

4.2 *Loss of lexical meaning and fixed position*

Classhood is determined by distributional or positional properties. Grams have a fixed, or grammatically determined, position in the phrase or clause. It is not necessary that a gram have only one possible position, it is only necessary that the position or positions be determined by grammatical criteria. The English auxiliaries have two possible positions, but these are grammatically defined: they occur before the subject only under specific conditions, in questions, after a preposed negative, and so on. The possibility of more than one position may indicate a lesser degree of grammaticization.

A fixed position for tense and aspect grams is usually a position fixed with respect to the verb, or occasionally the verb phrase or the whole clause. In addition the order of grams is usually fixed with respect to one another. To the extent that aspect, tense, mood and agreement are grammaticized, they will have a single, fixed position with respect to one another.

The rigidity of the positioning of a gram corresponds to the nature of the semantic relations a gram is capable of entering into. Perhaps the best way to characterize the semantic changes that take place in grammaticization is to say that specific components of lexical meaning are gradually lost (Givón 1973). For example, the English verb *willan* which formerly meant "to want" has lost (in most contexts) that specific lexical meaning of "an agent desires". The loss of specific lexical meaning has a number of conse-

quences for developing grams, which we will discuss in two groups. Under the heading of 'loss of semantic autonomy' we will discuss in this section the changes in the types of semantic relations grams may enter into. In the next section under the heading of 'generalization of meaning' we will discuss the properties derivable from the wider applicability of meaning which lacks specific lexical content.

The fixing of a gram's position reflects the loss of semantic autonomy in the following way: grams differ from lexical morphemes in that the position of grams is not manipulable for semantic or pragmatic purposes. Rather, the possibilities of combining grams with each other or with lexical morphemes is often heavily restrained.

To illustrate this fact, let us compare the tenses of English with the 'tenses' of the formal languages employed in so-called tense logic. In the latter, tenses are operators with propositional scope, e.g. 'it was the case that p' or 'it will be the case that p'. These operators may be freely combined with each other and also iterated, giving rise to arcane constructions like 'it was the case that it will be the case that it will be the case that John loves Mary'. If we change the order of two operators, or increase the number of occurrences of an operator, the semantic value of the expression will change. The morphologically expressed tenses in English on the other hand cannot by themselves be used in this way: whenever an inflectional gram may occur it occurs once, or not at all, and in a fixed position. Periphrastic and derivational grams, on the other hand, are often freer in this regard, but here, too, the limitations are significant, and the loss of freedom with regard to combination and iteration may be taken as one of the symptoms of grammaticization. We may, for instance, compare the English modals with their etymological cognates in other Germanic languages such as Swedish. A sequence such as *kan skola kunna* 'can have to be able to', which is at least in principle possible in Swedish will be completely ungrammatical if translated with the same morphemes into English: **can shall can*.[5] Another example is negation, the iteration of which in simplex sentences is ungrammatical in many languages, in spite of the popularity of double negation among logicians.

Another semantic property that grams gradually lose as they lose their lexical content is the ability to be modified independently (Spang-Hanssen 1983, Davidsen-Nielsen 1986). Modifiers such as adverbs in the verb phrase modify the lexical or main verb and not the inflections. A marginal case is seen with the least developed of the modal auxiliaries in English, *can*,

which may be modified in its ability sense:

(1) He can easily swim a mile.

However, the more developed grams do not take modification. Compare the following, in which the adverb can only be interpreted as modifying the main verb:

(2) He will easily swim a mile.
(3) He is easily swimming the mile.

4.3 *Generality of meaning*

Whereas lexical meaning is specific and referential, grammatical meaning is highly general and relational in quality, serving to relate parts of clauses or parts of discourses to one another. Both Boas and Sapir postulated a difference between material and relational concepts in an attempt to characterize the difference between lexical and grammatical meaning. While the validity of this distinction has been challenged (e.g. by Weinreich 1963:169), certain clear semantically-based diagnostics for grammatical meaning can be applied to making this distinction. One of these we have just mentioned in connection with the rigidity of positioning of grams — the inability of grams to be permuted or iterated — is a semantically-based restriction.

The other important correlate of the abstractness and generality of grammatical meaning is the absence of lexical and contextual restrictions on the occurrence of highly developed grams. Lexical verbs and auxiliary verbs in early stages of development often are restricted to sentences with certain types of subjects, and in the case of auxiliary verbs, certain types of main verbs. For instance, the verb *want* occurs most felicitously with an animate subject, and its use with an inanimate subject must be viewed as metaphorical:[6]

(4) It looks like it wants to rain.
(5) This door doesn't want to open.

However, the auxiliary *will*, which previously had the lexical sense of 'to want' can now be used with subjects of any sort. This difference in appropriate contexts is a direct reflection of a difference in semantic content. *Want* expresses an internal state or drive that is only possible for animate beings, while *will* has lost this specific element of meaning in its most com-

mon uses, which makes it applicable to subjects of all sorts. (See Section 8 for a discussion of the development of *will*.) In fact, highly developed grams do not affect the lexical restrictions of the verbs which they modify (Spang-Hanssen 1983).

In addition to distinguishing lexical morphemes from grammatical ones, generality of distribution also constitutes one of the main factors distinguishing inflectional from derivational morphology (Bybee 1985). This fact is of some relevance to our present discussion since aspectual notions may also be expressed in derivational morphology (as distinct from notions of tense, which are not expressed derivationally, again see Bybee 1985). For instance, morphemes signalling iteration tend to be restricted to verbs that are inherently telic or punctual and are thus usually considered derivational. Similarly, morphemes signalling inception usually apply only to stative verbs. On the other hand, inflectional grams such as perfective, past and imperfective tend to be applicable to all the verbs of a language.

A direct consequence of the loss of contextual co-occurrence restrictions is the rapid increase in token frequency which accompanies grammaticization. A construction or morpheme which appears only in clauses with selected types of subjects and verbs increases its frequency greatly as it begins to appear in clauses with any type of subject or verb. High text frequency moreover seems to be a factor in the development of other formal properties of grams as illustrated in the next sections.

4.4 *Obligatoriness and redundancy*

One of the defining properties of inflectional grams is their membership in obligatory sets. Here obligatory means that the presence of one member of the set is required by the grammatical context. The absence of an overt marker in such a case is meaningful and consitutes zero expression. The notion of obligatoriness is most useful as a diagnostic for distinguishing inflectional morphology from derivational. A word formed by derivational morphology belongs to a major syntactic category, such as noun, verb or adjective and may be used in a clause anywhere one of these categories is appropriate. For instance, the adjective *unhappy* has the same distribution as *happy*, the derived noun *electricity* may occur in the same grammatical contexts as the basic noun *water*. The distribution of an inflected word, however, is heavily constrained by the grammatical context, being sensitive to agreement, sequence of tenses, subordination, and so on (Malkiel 1978).

It is much more difficult, and perhaps unnecessary, to apply the notion of obligatoriness to periphrastic grams. As periphrastic grams develop they are gradually becoming obligatory, but there is no one point which can be singled out as the point at which a gram becomes obligatory. For instance, to ask the question of whether or not the English modal auxiliaries are obligatory would mean asking whether the lack of an auxiliary signals a particular mood or modality, such as indicative or assertive. Such a question is very difficult to answer. However, there is some interest in investigating the process by which a set of grams comes to be obligatory, since this often entails the creation of an unmarked category — a meaningful unit which has no overt marker.

The notion of obligatoriness also has what might be called a semantic side: the notional domain encoded by a set of obligatory grams must be touched on whenever the appropriate grammatical context arises. Thus in languages with obligatory evidential grams, some indication of the source of the information must be given in every sentence. In languages without such obligatory evidentials, the speaker may chose when to supply the added indication of the source of the information.

As the meaning of a gram continues to generalize, grow in frequency and become obligatory, its occurrence in certain contexts may be redundant. That is, it occurs with other indicators of meaning that make the small contribution of the gram strictly speaking unnecessary. For instance, in English, if a narrative is framed in the past tense, then all the verbs in a sequence must bear the past tense marker, even though its appearance on all but the first verb is redundant. In some languages the occurrence of a tense marker is optional in the sense that it appears on the first verb in a sequence, but may be omitted on subsequent verbs (e.g. Tongan, Churchward 1953).

4.5 *Affixation*

We have been arguing that the formal properties of grams are closely linked to their semantic properties. Whether a gram is an affix or not would seem to have less to do with its meaning and more to do with the morphological type of the language and its phonological processes of fusion. However, the fact that our data show a strong correlation between the meaning of a gram and its expression as an inflection or periphrasis indicates that affixation is not simply a formal process, but depends to some

extent upon semantic factors. That is, the phonological reduction necessary for affixation moves hand in hand with the reduction of semantic content in grammaticization.

Affixation is not a discrete event with clear before (unbound) and after (bound) stages, but rather a gradual process that involves several factors. In the Dahl (1985) and Bybee (1985) studies all of these factors could not be evaluated independently for lack of sufficient data, so the most salient criterion was chosen to determine affixhood: whether the gram was written bound by the analyst or author of the grammar. It is important to understand, however, what the prerequisites are for a gram to be considered an affix and written bound.

One requirement is that the gram appear in a fixed position — movable and permutable elements are not considered affixes and are hardly ever written bound. Another requirement is that no open class items may intervene between the gram and its head, i.e. the noun or verb stem it modifies. For instance, the English articles, *the* and *a/an*, although they appear in a fixed position, are not considered affixes because an adjective may come between them and the noun. However, a gram could meet both of these requirements and still be written separately. For instance, the Mandarin perfective marker *le* which occurs directly after the verb, meets the criteria for affixhood and yet is often written separately. As long as a gram is not phonologically fused with the stem, it does not have to be considered an affix.

Once the gram loses syllabicity or assimilates to or is assimilated to by the stem, it becomes very difficult **not** to write it as an affix. Processes of phonological reduction and fusion mirror the loss of grammatical and semantic autonomy in grammaticization. For instance, loss of independent stress or tone goes along with the fixing of grammatical position and semantic scope, as does extreme reduction in size and phonological fusion with the stem. Thus all the factors that lead to affixation are directly connected with the process of semantic generalization in grammaticization.

Fleischman (1982) suggests that affixation may correlate with a particular stage in the semantic development of grams. In dealing with the Romance futures, she claims that the development of temporal or tense functions for future grams tends to occur at approximately the same time as the agglutination with the verb stem. Such a correlation would imply a definite connection between semantic and formal developments in grammaticization. Of course, the data we have mentioned above shows a clear correlation between past, perfective and imperfective meaning on the one hand and affixation on the other. (Also our data on future grams, to be discussed below, show some correlations of affixation with function, however it is not clear whether this is precisely the relation that Fleischman predicted.)

What is behind this correlation between meaning and affixation? We refer again to the properties of highly grammaticized elements that we have mentioned before. Inflectional affixation tends to correlate with the loss of semantic autonomy, and with complete lexical and contextual generality, both of which occur during the stages of semantic generalization. We do not, however, think that it is necessarily the case, as Fleischman argues for Romance, that agglutination correlates with tense meanings rather than modal or aspectual ones. We propose instead that for each path along which tense and aspect grams develop, a semantic stage is reached which necessitates the expression properties that lead to affixation. We turn now to a discussion of these paths of development.

5. Perfects

The perfect is a wide-spread gram-type among the languages of the world, and in our material it occurs in 25 to 35 per cent of the languages. Semantically, the most important characteristic of perfects is that the situation described in the sentence is viewed from the perspective of — or described as being relevant at — a later point in time, most typically the point of speech. In the theory of Reichenbach 1947, the 'point of the event' is said to precede 'the point of reference' in the perfect, in contradistinction to the simple past in English where these points are said to coincide.

Normally, perfects are used both in 'resultative' (or 'stative') cases, i.e. those that are usually described in terms of 'the present result of a previous event', such as *John has gone to Paris (and is there now)*, and in a range of other cases, such as those sometimes called 'experiential', such as *Have you ever been to Paris?*. The relative semantic uniformity of perfects cross-linguistically comes forth clearly in the analysis of Dahl's questionnaire material presented in Dahl (1985, Chapter 5.).

From the point of view of mode of expression, perfects are predominantly — maybe in four cases out of five — expressed periphrastically. We may note at least four common types of such periphrastic constructions, which also correspond to common diachronic sources for perfect:

i. copula + past participle (or similar form) of the main verb (ex.: Hindi, Bulgarian, Tamil);

ii. constructions based on original possessive constructions, e.g. auxiliary 'have' + past participle of the main verb (ex.: most Germanic and Romance languages, North Russian dialects);[7]

iii. main verb + particle with an original meaning 'already' (ex.: the Kwa languages Yoruba and Isekiri);

iv. constructions involving auxiliaries historically derived from verbs meaning 'finish', or less frequently other lexical verbs, such as 'throw away' or 'come from' (ex.: 'finish' Sango (Samarin 1967)), Ewe (Heine and Reh 1984), 'throw away, cast aside' Fore (Scott 1978)).

We shall have relatively little to say about the last two of these types, except that in the two languages mentioned under (iii) the particles used seem to retain their original uses, i.e. they are also used to translate English sentences containing the word *already*. It may be noted that in languages without a grammatical perfect (such as Russian), morphemes meaning 'already' may be used more extensively than in English to make up for the lack of a perfect as it were. In some languages in Dahl's material (e.g. Karaboro, a Gur language), similar particles seem to be on their way to being grammaticized as perfects.

5.1 *Resultative to perfect*

The first two types of perfects really represent the same kind of development, viz. that from what we shall call resultative constructions to perfects. An example of a resultative is the construction 'Copula + Past Participle' in some Germanic languages, such as Swedish, as in (6).

(6) *Han är bortrest*
'He is away-gone' (lit.)

A thorough survey of resultative constructions is given in Nedyalkov et al. 1983, who point to a number of characteristics that differentiate resultatives from perfects, such as the following:

i. Resultatives are restricted in their meaning by having only the reading 'the direct result of such-and-such an event prevails' where the nature of the result is directly defined by the meaning of the verb. Perfects, on the other hand, typically do not imply the presence of a direct result: they can be used both in cases where no such result can be defined at all (e.g. with statives and 'activities') or where a former result has been cancelled at the point of reference (as in *Poland has been divided by her neighbors several times*).

ii. Whereas perfects can in general be formed from any verbs, resultatives are commonly lexically restricted: it follows from what has already been said about their meaning that they can only be formed from verbs whose interpretation involves some type of change (basically 'telic' verbs), but in many cases, resultatives are further restricted within this group in sometimes idiosyncratic ways. For example in Nivkh, transitive verbs whose objects are typically animate do not usually form resultatives (Nedyalkov *et al.* 1983:86).

iii. Whereas perfects usually do not have any effect on the valence or voice of the verb, resultatives are commonly valence-changing and/or part of the voice system: for example, the subject in a resultative construction often corresponds to the direct object of a non-resultative sentence.[8]

iv. Perfects and resultatives differ in how they can be combined with various kinds of temporal qualifications. Perfects do not readily combine with adverbs like 'still', whereas resultatives do very easily. Nedyalkov *et al.* (1983:12) provide the following examples from Armenian (similar examples are found e.g. in Swedish, see Dahl 1985:134):

(7) *na (*der) ənkel e* (perfect construction - ungrammatical with 'still')
he still fallen is
'he has (*still) fallen'
na der ənkats e (resultative construction - 'still' possible)
he still fallen is
'he is still fallen'

The differences in cooccurrence restrictions are important, since they show that the contexts of use of resultatives are not just a subset of the contexts of use of perfects: if that were the case, a perfect would be possible wherever a resultative can be used. The lack of a subset relation means that the development from resultatives to perfects involves a shift in meaning

rather than an extension of the meaning of resultative. The shift represents a change in emphasis: whereas resultatives focus on the state which is the result of a previous event, perfects focus on the event itself which leads to the extension to non-resultative cases. However, it is diffult to document in a convincing way that the difference in cooccurrence restrictions that we have been discussing corresponds to actual changes in the development of grams. Also, it appears that some perfect categories retain some of the properties of their resultative source; Nedyalkov *et al.* (1983:29) quote Lithuanian as a case in point.

As predicted by the general theory of grammaticization outlined in Section 4, the semantic change connected with the transition from resultative to perfect is accompanied by changes in grammatical properties. One is that of lexical generalization: as mentioned already, resultatives are generally lexically restricted, whereas perfects are not. Another is the tendency to adapt the construction to suit the general pattern for auxiliary constructions, in that the passage from resultative to perfect is accompanied by the disappearance of agreement between a participle and the subject or object in the sentence. One way of interpreting this is to say that the participle becomes part of the verb group rather than a modifier of subject or object.

These grammatical developments may be illustrated with examples from some West European languages. Most contemporary Germanic languages have a perfect, although in some cases the original perfect has undergone further development (see below). There are two historical sources: one is the construction copula + Past Participle, the other is a transitive 'possessive' construction with resultative meaning, a literal transfer of which into Modern English would yield something like *I have two books written*, where the participle *written* agreed with the direct object. A completely analogous construction is alive and well in Czech and was used several times in the Czech questionnaire, e.g.:

(8) *Máš výčištěné zuby?*
have-you cleaned (pl) teeth
'Have you brushed your teeth?'

Although these two constructions seem to have been available in all dialects, the ways in which they were employed in the formation of the perfect differ from language to language. In some Germanic languages, e.g. German, Dutch, and many dialects of Western Scandinavia, the perfect was formed from both the sources mentioned. The copula construction had an

active interpretation only for a limited set of intransitive verbs (mainly verbs of change): for transitive verbs its subject was interpreted as the 'deep' object (as is the normal case in passive constructions). Consequently, the perfect came to be formed with the copula only with the former group of verbs (e.g. German *ist gefallen* 'has falled'); in all other cases, the transitive construction was used as the basis for the perfect.

Dal (1952, 128) gives the following account of the development of the transitive resultative into perfect in Middle High German: the process started with constructions in which the verb 'have' still had its original meaning and the participle denoted a state of the object, e.g.:

(9) a. *then tōd habēt funtan thiu hella*
the death has found the hell
'hell has found death'

b. *sie eigun mir ginomanan lioban druhtīn mīnan*
they have me taken dear lord my
'they have taken my dear lord from me'

The construction was extended to all transitive verbs, e.g.

(10) *thaz eigut ir gihōrit*
that have you heard
'you have heard that'

The extension to intransitive verbs started out with verbs with a 'that'-complement and 'transitive verbs in absolute use':

(11) *sō wir eigun nū gisprochan*
so we have now spoken
'so we have now spoken'

Finally, intransitive verbs proper (except those that took 'be' as an auxiliary) were pulled in:

(12) a. *er habēt sīn ein luzzel ergezen*
he has himself a little enjoyed
'he has enjoyed himself a little'

b. *nu habēnt sie dir ubelo gedanchōt*
now have they you ill thanked
'now they have thanked you badly'

c. *habe ih keweinōt*
have I wept
'I have wept'

A similar story is told about the development of the 'habeo factum' construction in Latin and Romance (Harris 1982:47), which in Classical Latin still was lexically restricted (to begin with, to verbs relating to possession and then to verbs like 'learn', 'discover', 'persuade', 'compel') and involved agreement between the participle and the object, as in this example from Plautus:

(13) *multa bona bene parta habemus*
much goods well obtained we-have
'we have obtained many good things'

In the 6th century there were already examples of transitive verbs without agreement:

(14) *haec omnia probatum habemus*
this all tried we-have
'we have tried all this'

Eventually, the construction was extended to intransitives, again, with the exception of those which formed their perfect with 'be' in some daughter languages (e.g. French) and to all intransitives in others (e.g. Spanish).

In English and many varieties of 'Peninsular Scandinavian', including Standard Swedish, the only source of the modern perfect is the transitive construction, which has undergone several changes: (i) generalization to all verbs, (ii) loss of agreement between participle and object, and (iii) a change in word order which placed the participle closer to the auxiliary. In Standard Swedish, one can even see the appearance of a new non-finite verb form, the so-called 'supine', which is used instead of the past participle in the perfect.[9] The fate of the old resultative copula construction varies within the group: at least in Swedish and Norwegian, it still retains its old use, whereas it is virtually gone in English (with the possible exception of a few cases like the use of *gone* illustrated by this sentence).

Notice that what has happened in German, for example, is a virtual merger of two different constructions: except for the rather marginal resultative uses of *sein* + Past Participle, this construction is functionally equivalent to the Perfect with *haben* 'have'. This convergence of function and the fact that functionally equivalent categories in other languages may arise from other combinations of historical sources strongly supports the validity of cross-linguistic gram-types, such as perfect.

5.2 *Further developments*

The development of perfect does not stop at this point. Three directions for further change are frequently attested (and a few others may be possible). One involves the development of evidential function in perfects, the other the development into a past or perfective marker, and the third the use of perfect categories to express remoteness distinctions. We shall here concentrate on the first two developments, since they are of most direct relevance to the topic of this paper.

The first of these involves the use of perfect for evidential functions such as signalling that the assertion is based on inference (including inference from results), or a first or second hand report. The development of evidential meaning is particularly common in a geographical region comprising the Balkans and adjacent parts of the Middle East, but it is also attested in at least one Tibeto-Burman language, Newari (Genetti 1985), where perfects with evidential functions derive from the lexical verbs 'keep' (cf. 'have') and 'finish'.

In some languages (e.g. Persian, Georgian, Azerbaijani from Dahl's sample, Macedonian (Friedman 1986) and Newari (Genetti 1985)), the perfect retains its original uses in addition to the newly acquired ones. In Turkish and Kurdish, however, the gram loses its original perfect uses and functions primarily as an evidential. A peculiar situation obtains in Bulgarian, where according to standard grammars (Stojanov 1964, Maslov 1981), there exists a 'reportive evidence mood' with three distinct sets of forms for tense and aspect. The one used for perfective past contexts differs from the ordinary Perfect only by the absence of a copula in the third person. On this analysis, the Perfect of Old Slavonic would, in modern Bulgarian, have split up into two categories, albeit still formally distinct only in some contexts. It is somewhat hard to judge to what extent the forms with and without copula are differentiated in actual usage. The data presented in Roth (1979) indicate a considerable overlap between them, while Friedman (1986) argues that there is no semantic difference.

The evidential uses of perfects develop because the perfect is used to describe past actions or events with present results. If the focus of the meaning is on the idea that the present results are connected to and perhaps

attest to past actions or events, then the notion of an action known by its results can be extended to actions known by other indirect means, such as by inference (from reasoning in addition to inference from results) and by reports from other parties. (Discussions of these diachronic developments can be found in Aksu-Koç and Slobin 1987; Friedman 1987; Willett 1988.)

The most well-known path of development, however, is that of perfect developing into either a general past or into a perfective (past) category. Examples of these two possibilities are the Southern German dialects and spoken French, respectively. It appears that the choice between them is conditioned by the presence or absence of a separate 'Imperfective Past' in the language: in French, due to the existence of the morphologically expressed '*Imparfait*' the '*Passé Composé*' has not taken over the whole Past area. (A possible counterexample to this generalization is Dutch, where, in spite of the lack of an 'Imperfective Past' there is at this stage a tendency to restrict the Perfect to perfective contexts (de Vuyst 1985)).

Semantically, the passage from perfect to perfective or past can be said to be a generalization: the contexts-of-use of the perfect are much more limited than that of the other grams. However, it would be an over-simplification to say that this change is simply a weakening of the original meaning: the perfect is sometimes used with non-past reference (e.g. in a German sentence such as *Morgen bin ich schon abgefahren*, 'Tomorrow I will already have gone' in which the other categories are not always possible (cf. the impossibility of I **Morgen fuhr ich schon ab.* 'Tomorrow I already went'). Thus in order for a perfect to become a past the point of reference must be restricted to the moment of speech, and the part of its meaning that specifies that the past event is especially relevant to the current moment must be lost.

Whereas perfect, as we mentioned earlier, tends to be expressed periphrastically, both past and perfective are expressed by bound morphology in the majority of cases.[10] If it is the case that perfects often develop into past or perfective, these differences seem strange if we do not at the same time assume that the continued grammaticization of meaning is accompanied by a continued grammaticization of form. In the most typical cases the periphrastic marker of perfect gradually becomes affixed to the main verb stem, as the meaning gradually changes from perfect to past or perfective. The end results of this process can be seen in Swahili, where the copula *li* has become a past tense marker (Heine and Reh 1984: 130):

(15) *a - li - kwenda*
he - past - go
'He went'

Similary, in the Dahome dialect of Ewe, the verb 'be finished' is prefixed to the main verb and has simple past meaning (Heine and Reh 1984: 130):

(16) *m - kɔ - sa*
I - finish - sell
'I sold'

Another interesting case is that of Mandarin Chinese in which the perfect and the perfective are expressed by homonymous grams *le*. With perfect functions, *le* occurs sentence-finally, but when it has perfective functions, it occurs immediately after the main verb with nothing allowed to intervene, which means that it is basically affix-like in nature. Both of these instances of *le* probably derive from the erstwhile main verb *liao* 'to finish'.

Another way in which a periphrastic construction may become non-periphrastic is through the deletion of the auxiliary. At least one fairly clear case is cited in the literature, viz. that of the Slavic languages, in many of which an old periphrastic perfect, consisting of a copula plus a past participle of the verb, has developed into a general past by replacing the old Aorist and Imperfect forms of the verb. This process was accompanied by an approximately simultaneous loss of the copula, globally (in the East Slavic languages) or just in the 3rd person (in e.g. Czech), effectively turning the periphrastic construction into a morphological one. In the languages where the Perfect has retained or until recently retained its old function (Bulgarian and Sorbian [Wendish]), the copula has also been retained. These facts suggest a close connection between the shift in meaning and the change from periphrastic expression of the old Perfect. However, since there was also in some languages, e.g. Russian, a more general process of copula loss, extending to all copula constructions in the present tense, it may legitimately be asked what kind of connections there are between the different processes.

The very nature of copula deletion as a diachronic process is also of interest. Both the restructuring of the tense-aspect system — as manifested in the disappearance of the Aorist and Imperfect — and the disappearance of the copula were (or rather are, since both of them seem to be still going on in some areas) gradual processes, taking several centuries to be completed. At least for Russian (to judge from works such as L'Hermitte 1978),

they span more or less the same period in time (from the 11th century onwards). The detailed data presented in L'Hermitte (1978) show that even if it is not completely obvious how the process started, the disappearance of the copula went much faster when it was used as an auxiliary — that is, mainly in the Perfect — than e.g. before adjectives and nouns. It is thus not excluded that the semantic change in the old Perfect construction triggered the copula deletion rather than the other way round.

The work of Labov and his collaborators on Black English shows that in that language, copula deletion is a complex phenomenon governed by a number of phonological, lexical and syntactic factors. In fact, the picture is somewhat similar to that of Old Russian: in both languages, the copula is omitted more often before verbs than before nouns and adjectives. There is nothing to contradict the assumption that copula deletion might be favored also by semantic factors pertaining to the use of tense and aspect categories. If a perfect is extended in its use to 'ordinary past tense contexts', a situation might arise where the copula disappears in these very contexts, thus in effect creating a formal differentiation between a periphrastic perfect and a morphologically expressed past differing only in the presence of a copula in the former.

Such a synchronic situation is indeed attested in a number of languages. Comrie (1976, 107) mentions Hindi, Urdu, and Punjabi as examples. Outside Indo-European, Oromo/Galla (Cushitic) exhibits a similar system (Moreno 1964): the perfect is formed by adding the auxiliary *gira* 'exist' to the Perfective form of the verb, although this is obscured by the fact that the auxiliary is normally contracted with the verb stem. From Dahl's material it appears that Tigrinya (Semitic) may also be a case in point.

There is an interesting parallel from another diachronic path involving perfect. According to the standard analysis of Bulgarian, as we mentioned above, a special 'reportive evidence' category has developed alongside the perfect, the only formal distinction between these two being in the third person, where the 'reported' forms lack the copula used in the perfect. Even if the standard analysis makes the situation look more clear-cut than it is, the case still appears to be another example of a situation where a gram that has undergone more development differs from a less developed gram only by the absence of a copula. We know of no example of the inverse situation and indeed, it is tempting to predict that it will never occur.

We have examined in this section the development of perfect grams, concentrating our attention on those which develop from auxiliary plus past participle constructions. The semantic path followed by such constructions leads from resultative to perfect to perfective or past. A common element appears in this long path of development: whereas the resultative views a past event in terms of its prevailing results, the perfect de-emphasizes the perspective of the present moment, by focusing more on the past event, requiring only that that event have some relevance to the present moment, and not that it produce some current state. The change to past or perfective is in the same direction: the sense of relevance to the current moment disappears altogether. For lack of evidence we have not examined in detail the semantic development of pasts and perfectives from perfects whose lexical sources are active verbs, such as 'finish'. It would be of considerable interest to compare such a developmental path to the one discussed here.

6. Progressives

The gram-type 'progressive' occurs in approximately one-third of the languages of Dahl's sample, making it comparable in frequency to the perfect. In most languages, the progressive occurs in combination with the present, past, and less frequently the future tense. Like the perfect, the progressive shows a very strong tendency to have periphrastic rather than inflectional expression. As we mentioned before, this indicates that progressive grams are relatively young grammaticizations, and this is supported further by the fact that their lexical sources are often transparent. In this section we will discuss the sources of progressives and the general path of development that they take. We will present evidence that progressives develop into imperfectives.

By far the most common source of progressive grams are locative expressions paraphraseable as 'to be located in or at an activity' (Blansitt 1975; Traugott 1978). Blansitt divides sources for progressives into copula and non-copula sources, ignoring the fact that some copulas incorporate location and some do not, as well as the fact that some copula constructions are accompanied by locative adpositions while some are not. We propose that it is the semantics and not the form of the sources that determine their subsequent development, and thus classify the constructions according to the semantic elements that compose them.

Explicitly locative phrases seem to be the most common sources. These

usually take the form of a copula plus a locative adposition and a nominalized form of the verb: e.g. Irish, copula + *ag* 'at' + Verbal Noun.

(17) *Tá sé ag dúnadh an dorais.*
He is at shutting the door
'He is shutting the door (GEN)'

Also in this category are expressions such as the French *être en train de* which consists of the copula plus the phrase *en train de* which is originally a locative expression. Another example is the Kru language Godié, which uses a verb *kù* meaning 'be at' (Marchese 1986:63).

(18) *ɔ̃ kù ɓlι -dʌ*
he be-at sing-place
'He is singing'

In the last example, there are two locative elements, a verb that expresses location, and the nominalizing suffix that means "place".

Another way that locative meaning may enter into a progressive construction is through the use of a verb meaning 'be in, be at or be located' plus a main verb form, as in the following Cocama example, which uses the verb *yuti* 'be located' (Faust 1972:55).

(19) *Iquiaca ta camata yuti.*
Here I work be
'I am working here'

Postural verbs, such as 'sit' and 'stand', and durative verbs, such as 'stay' or 'live', are also often mentioned as lexical sources for progressive (Blansitt 1975; Traugott 1978), but such constructions are not really distinct semantically from the locative constructions, since postural verbs and durational verbs themselves involve a notion of location, and furthermore often serve as the lexical sources for copulas and locational verbs. For instance, Spanish *estar* comes from Latin *stare* "to stand". Such verbs may enter into progressive constructions before their lexical semantics are entirely lost, creating contrasts such as that found in Ngambay-Moundou between a progressive using *ísi* 'sit' and *ár* 'stand' (Vandame 1963:93)

(20) *m-ísi m-úsa da*
I-sit I-eat meat
'I am eating meat'

m-ár m-úsa da
I-stand I-eat meat
'I am eating meat'

Less commonly, motion verbs are found as progressive auxiliaries, for instance in the Mouroum dialect of Ngambay-Moundou, and in Spanish:

(21) *Anda buscando su reloj.*
'He's (going around) looking for his watch'

Tiwi (an Australian language) has a 'moving' aspect formed with a suffix *-ami* (perhaps cognate with the verb 'to go' *-mi*) (Osborne 1974):

(22) *a -untiŋ -apu-kami*
he-durative-eat -moving
'He's eating moving about'

In Turkish a verb meaning 'to go, walk' supplied the suffix *-yor*, which is today used as an aspectual morpheme of progressive or imperfective (see below) without any sense of movement (Lewis 1967). These motion verb constructions may be considered a special type of locative progressive, since they do in effect specify that the agent is located in the activity. The difference is in the specification of a moving rather than a static location.

Since copulas derive from postural, durational or locational verbs, or locational adpositions (see Li and Thompson 1979), and since nominalization markers often have locative sources, it is possible that locative meaning contributes to most if not all progressive constructions. In cases where the progressive construction involves unidentified elements, it is reasonable to surmise that these might have had locative meaning originally. Even the English Progressive, which has no overt locative element can be plausibly traced to the construction exemplified by *he is a-working*, where *a-* is from the preposition *on*, which was deleted as the construction became more common. We have not found a clear example of a progressive construction formed with a non-locative copula and a main verb with no other elements involved.

One other source of progressives, which is not very common, but definitely attested, is a periphrasis meaning 'to keep on, to continue' an activity. One example is the Swedish phrase *hålla på att* + infinitive ('keep on'), which emerges as a good example of a progressive in Dahl's questionnaire. This source is related to the durational source ('stay'), and, it should be noted, involves a locative element *på*, 'on'. If the English phrase 'to keep

on' + present participle is any indication, in its early stages it implies an active continuation of an activity beyond normal expectation or despite certain obstacles.

The most common use of the progressive construction in Dahl's questionaire is to signal an activity that is ongoing at speech time and many progressives are restricted to ongoing activities, e.g. the Dutch progressive (Donaldson 1981:165):

(23) *Ik ben druk aan 't koken.*
I am busy at the cooking
'I am busy cooking'

However, careful studies of constructions such as the Progressive in English reveal that more than a sense of on-going activity is conveyed by this construction. Hatcher (1951) argues that the Present tense of the English Progressive (in the use which appears to be the same as the one that is prototypical in Dahl's questionnaire) is favored with verbs signalling overt activity:

(24) She is washing the dishes, combing her hair, chewing gum.
It's boiling over, its spilling.

while the Simple Present is favored where no overt activity is displayed:

(25) It stings. It tickles. Does this light bother you?
I remember her. I understand. I love your hat!

Internal processes may occur in the Progressive, according to Hatcher, if they are seen as developing by degrees:

(26) I'm getting hot. One of my headaches is coming on.
He is progressing, improving, getting worse.

Another way of looking at this, according to Hatcher, is that the Progressive is favored to describe situations in which the subject is involved in the activity, either by being affected by it, by being engrossed in it or by accomplishing something by it. Chafe (1970:175) argues that the Progressive in English has an implicit limit on the temporal dimension of the activity. He describes the meaning of the Progressive as indicating "that the event ... (is) spread out over a certain period of time." He continues "This period, furthermore, is not unlimited in extent but is understood to be subject to eventual termination."

We do not feel that it is necessary to decide which of these is the 'correct' characterization of the English Progressive, nor do we feel that it is necessary to limit the Progressive to aspectual meaning. The richness of nuance that Hatcher and others have discovered can be recognized as a part of the meaning of the progressive that it inherits from its original semantics, or implications therefrom.

Let us consider what the initial uses of a progressive must be like, taking the locative source as the paradigm case. The durativity or sense of "ongoing" activity comes from the stative sense of "be located at", but the construction **implies** much more than that. In order to understand the implications that develop from this meaning, we must first consider in what cases such a locution would most appropriately be used. Since in any language in which a progressive is developing there already exists a means of making simple predications, the progressive would develop to signal some extra or special meaning over and above what the default predication conveys. It would thus be useful to say that a subject X is located in or at an activity (i) if that activity has a concrete, physical location, (ii) if X is mobile (and capable of being located elsewhere) and (iii) if X is perhaps even a volitional agent who may at times be involved in other activities. The prototypical use of such a construction, then, would emphasize the subject's ongoing involvement in an activity much as Hatcher describes, while the temporal limits that Chafe notices are derivable from the fact that an activity requires a steady input of energy to be maintained and is thus inherently limited in the way that a state is not. In this view the non-occurrence of progressive with stative verbs would not be due to an absolute prohibition, but merely reflects the inappropriateness under most conditions of claiming that a state takes place in a physical location and the subject is actively involved in maintaining a state, which is typically involuntary, such as being tall. On the other hand, developing states, such as 'improving, getting hot, growing tall' may be viewed as activities that involve the subject.

Much has been made of the use of locative expressions for temporal notions (Anderson 1973, Traugott 1978), but in this case no great conceptual leap is needed for the locative construction to take on temporal value. To be located spatially in an activity is to also be located temporally in an activity, so that from the beginning the meaning of such constructions has temporal implications. Gradually the locative meaning weakens while the temporal implications stabilize, giving rise eventually to the aspectual meaning of progressive.

Of course, this analysis is based on descriptions of English and may not

be generalizable to other languages. The information gleaned from Dahl's questionnaire cannot supply the full range of nuances available in a language-specific study. However, since the distribution of progressive in the questionnaire is very similar across the twenty-six languages displaying the gram-type, we can hypothesize that progressives are similar in their semantics across languages. Moreoever, the questionnaire does show the absence of progressive grams with verbs expressing intellectual or emotive states, such as *think*, *believe*, *hope*, *feel* and *doubt*, even when the context specifies 'right now' (questionnaire examples (115) - (124)). The progressive is also absent in the questionnaire data, as in English, with performatives such as *I promise* (see Hatcher 1951:267), suggesting considerable cross-language comparability in this gram-type.

While the most restricted and more prototypcial use of the progressive requires that the activity actually be in progress at event time, it has often been pointed out (Chafe 1970, Comrie 1976) that the English Progressive is more general in its use, since it may be used to describe activities that are not actually in progress at the reference time, but that are characteristic of a certain time frame which includes the reference time:

(27) He is studying Chaucer.

The limits on the time period are evident here if one compares the simple Present, *He studies Chaucer*, which may characterize an entire career. The extended usage in English is also compatible with temporal phrases indicating repeated or habitual activities, such as:

(28) He is working on his book everyday.

This indicates an habitual activity within a certain limited time frame, but not necessarily literally in progress at reference time. Such usage, then, represents a generalization of the earlier progressive meaning.

Comparative evidence suggests that in Igbo, Yoruba, Scots Gaelic (Comrie 1976:99-101) and Kuwaa (a Kru language, Marchese 1986) a progressive construction (in all cases derived from a locative construction) has evolved into a general imperfective. The Turkish suffix *-yor* which comes from a verb meaning 'to go, to walk' shows the distribution of an imperfective in Dahl's questionnaire, and has presumably developed from a progressive. In order for a progressive to become an imperfective it must lose the special senses associated with its original semantics, such as specifying a limit on the period of time in which an activity is ongoing, and implying involvement of the subject, and come to signal a situation that is simply

ongoing at reference time, which may encompass situations that are repeated, habitual or continuous. We may speculate that the English Progressive represents an intermediate stage in this development, in which the construction has become compatible with habitual adverbs, and thus can be used to describe habitual situations within a limited time frame. Then the only additional step needed is the loss of the specification of a limited time frame.

7. The perfective / imperfective distinction

Our data show that the two most common distinctions expressed consistently through bound morphology are the past / non-past distinction and the perfective / imperfective distinction. These two distinctions are discussed together in this section because they interact in specific ways, making it possible to characterize a typical tense-aspect system.

Theoretically, a combination of three tenses (present, past, and future) and two aspects (perfective and imperfective) would yield at least six combinations. This possibility, however, is practically never realized in a single language. In fact, it seldom happens that more than four combinations are distinguished, due to a universal tendency for there to be a coupling between perfectivity and the past, in most cases resulting in a system where a gram with the meaning 'perfective past' is opposed to everything else, as shown in (29).

(29) perfective / imperfective

past / non-past

Judging from our material, this kind of opposition seems to occur in about every second language in the world. Thus, in Dahl's 18 language sample (see Introduction), there are the following 9 clear cases: Alawa, Bandjalang, Mandarin Chinese, Azerbaijani, French, Kurdish, Beja, Kikuyu, Seneca. (A problematic case, Georgian, will be discussed below.) In most of these — clear exceptions are Chinese and Seneca — there is also a grammatical opposition between past and non-past in the imperfective aspect. Systems containing these two oppositions will be referred to here as **tripartite** tense-aspect systems emphasizing the trilateral nature of the basic inflectional categories, even though the language might also have other tense or aspect grams such as perfect or progressive.

We shall take the above observations about tripartite tense-aspect sys-

tems as a point of departure for a discussion of the concept of perfectivity. It is important here to distinguish between notional perfectivity and the gram 'perfective aspect'. The view of the former that we shall argue for is that it is a set of related concepts rather than one single notion. It could thus be called a 'family concept' although 'family of concepts' is perhaps more adequate, as we would argue that the members of the family share a common focus — that is, the prototypical cases are the same. In Dahl 1985:78 the prototypical perfective was described in the following way:

> A perfective verb will typically denote a single event, seen as an unanalyzed whole, with a well-defined result or end-state, located in the past. More often than not, the event will be punctual, or at least, will be seen as a single transition from one state to its opposite, the duration of which can be disregarded.

It is argued in Dahl (1985) that among the properties that characterize a prototype, some may be more essential to the category in question than others. In the perfective, the aspectual properties could thus be seen as 'dominant' relative to the temporal properties: both kinds of properties characterize the prototypical instances, but there is considerably more variation both intra- and inter-linguistically as to how the 'past time reference only' restriction is manifested. It thus happens fairly frequently that perfective categories may have non-past reference in non-indicative moods or (which is often the same thing) certain non-assertive contexts, such as conditional clauses. Well-known examples are the Aorist in Classical Greek and the Perfective in Arabic.

Furthermore, it will be argued that notional perfectivity may be manifested in different ways in the grammar, and that a certain kind of manifestation favors a certain more precise interpretation of it. In other words, different members of the family will show up in different parts of the grammar.

The skewing of tense-aspect systems with respect to the relation between aspect and time reference is also parallelled in morphology. Aspect is in general closer to the derivational end of the derivational-inflectional scale than tense (Bybee 1985). This is true also of the perfective / imperfective distinction, although maybe to a lesser extent than for some other aspectual categories, such as inceptive or iterative. Thus, we find in many Indo-European languages, e.g. Greek, that the distinction between Aorist on the one hand and Present and Imperfect on the other tends to be marked irregularly, often by modifications of the verb stem, whereas the

distinction between Present and Imperfect is regularly marked by inflectional suffixes. Similar systems are found in languages as diverse as Nahuatl and Burushaski (from Bybee's sample). This suggests that if perfective / imperfective and past / non-past are hierarchized, it is the aspectual distinction that is more basic, which means that the members of tripartite tense-aspect systems should be described as 'perfective' in contrast with 'past imperfective' and 'non-past imperfective' rather than as 'present' in contrast with 'past perfective' and 'past imperfective'. Such an analysis also makes the systems of Chinese and the other languages with only a perfective / imperfective distinction seem more similar to the tripartite system in that they differ only in not having a past / non-past distinction in the imperfective aspect.

Historically, there are a number of ways in which systems of the kind described may arise. One well-attested process is the one discussed in Section 5, whereby constructions with an original perfect function develop into perfectives. Another possible development, discussed in section 6, is that of original progressives extending to non-progressive, imperfective uses. Of course, in many cases the historical origins of perfective / imperfective distinctions are obscure, but it is only reasonable to hypothesize that they are of attested types, such as the ones discussed here.

In spite of the pervasiveness of tripartite tense-aspect systems, there are some notable exceptions to the generalizations made above. In particular, the tense-aspect systems found in a number of Slavic languages, such as Russian, do not conform to them. Since the perfective / imperfective opposition in Russian is often seen as a paradigm example of an aspectual category, this may seem somewhat astonishing. It is possible, however, that it is the rather special features of the Slavic aspectual systems that attracted the attention of grammarians and which, paradoxically, have given them the status of first-cited examples in all textbooks. What we shall argue here is that the Slavic aspectual systems differ from the tripartite systems in their origins, their semantics, their means of expression and their relation to other parts of the system of verbal grams such as tense. Moreover, we argue that these differences are related to one another.

Probably most languages have counterparts to particles like the English *out*, *up*, *apart* etc., although they vary in their manifestations, sometimes being prefixes or suffixes on verbs rather than free morphemes. This makes it somewhat difficult to find a good term for them: since they have the effect of making the process denoted by the verb 'bounded' ('telic') we

shall call them 'bounders'. Adding a bounder to a verb often has effects both on its syntactic valency and its aspectual potential or Aktionsart. Thus, *eat up* in English differs from the simple *eat* both by being more clearly transitive and by implying a definite limit or end-state of the process (the total consumption of the object). In a number of languages from different parts of the world, one can see a tendency for bounders to become grammaticized as aspectual markers. Relatively well-known examples are, in addition to the Slavic languages, Latvian and Lithuanian (Indo-European: Baltic), Hungarian (Finno-Ugric) and Georgian (Kartvelian) (For a discussion of aspect in these languages, see Comrie 1976). Some less often quoted parallels in other parts of the world are Margi (Chadic: Hoffmann 1963; Dahl 1985) and some Micronesian languages such as Kusaeian (Lee 1975) and Mokilese (Harrison 1976; Chung and Timberlake 1985). In all these, verbs with bounders are interpreted as being in some sense 'perfective'. Although the limited amount of information available on some of these languages makes it difficult to make generalizations, there seems to be considerable variation both in the exact functions of the 'perfective' aspect so obtained and in the degree of grammaticization of the perfectivizing process. In most cases only certain groups of verbs can take bounders. In addition, the choice of bounder for a particular verb is usually unpredictable or at least heavily dependent on the meaning of the verb. In other words, perfectivization by bounders usually has a marginal grammatical status in the languages where it can be found. In spite of the parallels that can be found in other languages, Slavic languages seem to have gone further than other languages towards generalizing the applicability of bounder perfectivization, and making it an essential part of the aspect system.

Moreover, the other languages that use bounders for perfectivization do not seem to have the additional feature exhibited in most Slavic languages of a derivational imperfectivization process by which secondary imperfective verbs are formed from perfective ones, as when an imperfective verb *perepisyvat'* is formed from the Russian *pere-pisat'* 're-write' (with the perfectivizing prefix *pere-* 're') by the suffix *-va-*. The existence of this kind of process, which is basically complementary to perfectivization by bounder prefixes, is one of the reasons why a much larger set of the verbs in e.g. Russian can be assigned to 'aspectual pairs' than seems to be the case in the other languages we have mentioned. It should be noted, however, that even in Russian, a relatively large group of verbs, including some productive formations (so-called 'biaspectual verbs' are aspectually neutral

and thus remain outside the aspectual system. This property makes the Slavic aspectual systems much more derivational in their character than the tripartite systems which are typically inflectional.

A further parameter of variation is the extent to which bounder perfectivization is integrated with the rest of the tense-aspect system of the language in question. In at least two cases, Georgian and Bulgarian, the bounder system coexists with a tripartite tense-aspect system of the kind described above. In the former language, the presence of bounders is largely predictable from the choice of inflectional tense-aspect categories, in such a way that e.g. the Aorist occurs with, and the Imperfect, without bounders. Although there is also a rather high correlation between the two systems in Bulgarian, one can find a sufficient number of disharmonic choices, such as Imperfective Aorists, for this language to be of primary interest when comparing the semantics of the two types of perfective / imperfective oppositions.

One clear way in which a tense-aspect system like that of Russian differs from the tripartite system is in the relation between aspect on the one hand and tense and time reference on the other. To start with, we may note that morphologically the opposition between perfective and imperfective aspect is almost wholly independent of the category of tense in Russian: verbs of both aspects have both past and non-past (present) forms — in contradistinction with the more common system, where only the imperfective aspect evinces tense distinctions. The independence of morphological tense and aspect is somewhat obscured by two facts: (i) although there is no morphological future / non-future distinction, there is a periphrastic future construction, which can be formed only from imperfective verbs, (ii) similar to the tripartite system, there is a restriction on the time-referential potential of the perfective aspect, although it has to be formulated as 'non-present reference only' rather than 'past reference only'. That is, non-past forms of perfective verbs are either used with non-specific reference or else to refer to the future. (This is subject to some variation among the Slavic languages: in Bulgarian, perfective presents seem to be used exclusively with non-specific time reference.)

Among the various accounts of the semantics of the perfective / imperfective distinction found in the literature, one may distinguish two important trends: one which considers the crucial property that characterizes the perfective aspect to be that it involves a 'total view of the situation', and another which emphasizes the connection between perfectivity and the pre-

sence of a limit or end-state for the process. A comparison of the use of aspectual categories in different languages suggests that it is not the case that one of these views is right and the other is wrong; rather, they are not equally adequate for all languages. To demonstrate this, we shall look at the cases where the two views of perfectivity make different predictions about aspectual choice.

One may talk of a process without indicating any limit to it and without taking the 'internal' perspective supposedly characteristic of the imperfective aspect. From Dahl's questionnaire, we may cite the following as an example of this:

(30) (What did your brother do after dinner yesterday?)
He wrote letters.

This example was translated using the imperfective aspect in only 25 per cent of the languages that have a perfective / imperfective distinction in Dahl's sample. What is interesting about this minority group is that all the Slavic languages in the material are included in it. Even more strikingly, Bulgarian, which is the language that has the most well-developed 'double' perfectivity distinction, makes a seemingly contradictory choice here: Bulgarian informants consistently choose the 'Imperfective Aorist' for those cases. This strongly suggests that there is a systematic semantic difference between the two kinds of perfectivity distinctions. Further data from Dahl's material support this supposition. In sentences containing adverbials which answer the question 'For how long?', about two thirds of the tripartite aspectual systems choose the perfective aspect, whereas Russian, Polish, and Czech use the Imperfective and Bulgarian again the Imperfective Aorist. A similar tendency, although a less consistent one, can be seen in sentences containing manner adverbials such as 'slowly'.

In our view, these facts taken together strongly support the view that the tripartite aspectual systems and the 'Slavic type systems' differ with respect to the weight they give to different components of prototypical perfectivity.

The differences between the two kinds of aspect systems that we have been discussing may be looked at both from the synchronic and the diachronic point of view. From the synchronic point of view, it is noteworthy that the non-inflectional character of the Slavic-type aspectual systems, which distinguishes them from the tripartite system, is paralleled by clear differences in the semantics of the categories involved. From the diachronic

point of view the Slavic systems represent a radically different path of development than the tripartite ones, and this development leads to a different semantic character. On the other hand, the similarity between the perfective meaning evolved historically from bounders and that which evolves from periphrastic constructions (i. e. perfects) is strong evidence for the validity of universal gram-types for perfective aspect.

At this point, a third kind of morphological system should be mentioned, viz. the case-marking systems of Fenno-Ugric languages. Finnish is an example in that the case used to mark prototypical direct objects — the accusative — entails perfectivity. It seems from our material that to the extent that these factors can be identified with those underlying the distinction between perfective and imperfective aspect, they seem to be more like the 'Slavic' than the tripartite distinction. For instance, the use of the 'perfective' member of the Finnish opposition, the accusative, has a 'non-present time only' restriction on its interpretation, just like the perfective aspect in Russian.

Finnish is an example of how the transitivity system of a language involves notional aspect in that the prototypical direct object case marking — the accusative — entails perfectivity. Hopper and Thompson (1980) give various other examples of the tendency for transitivity and (notional) perfectivity to go together. It should be noted that the Slavic aspect systems are also clearly linked to transitivity: purported imperfective / perfective pairs often differ in their propensity to take direct objects (e.g. Russian *napisat'* 'write (pf.)' is almost exclusively used as a transitive verb, whereas *pisat'* frequently occurs without an object), and there is considerable interaction between the aspect and the case marking systems in a way strongly reminiscent of Finnish (Dahl and Karlsson 1976). When evaluating Hopper and Thompson's claim about the link between transitivity, aspect, and concepts such as 'backgrounding', it should be emphasized that their term 'Aspect' stands for a notional category rather than a grammatical one. There is much less evidence to suggest that inflectional tripartite aspect is very closely linked to the transitivity system.

Summing up, the variation in the language-specific manifestations of the perfective / imperfective distinction shows a clear correlation of the semantics of the categories, their means of expression (as inflectional or derivational), and their historical source.

8. Futures

Future grams are extremely common both in Bybee's reference grammar survey and Dahl's questionnaire. They are unique among the major tense and aspect categories in that they do not favor either periphrastic or morphological expression, but are evenly split in Dahl's sample between the two. In this section we will make use of this fact to test our hypotheses about the correlation of meaning and means of expression, but first we will describe the evolution of future grams and their distribution in the questionnaire.

The major lexical sources for future grams, which have been well-documented across numerous examples, are the following three (see Ultan 1978; and Bybee and Pagliuca 1987):

i. An auxiliary verb with the original meaning of 'want' or 'desire', or less commonly a derivational desiderative morpheme, which in turn has as its source a main verb meaning 'want' or 'desire'. Examples may be found in English, Serbo-Croatian, Swahili and Mandarin, to name but a few.

ii. A construction meaning 'movement towards a goal' (such as English *be going to*), which contains a movement verb in a progressive or imperfective aspect, and an allative component either explicit or incorporated in the verb. Less commonly, a derivational andative construction (whose source is also a verb meaning 'movement towards a goal') may develop into a future gram. Examples may be found in Hausa, Logbara, Haitian Creole, Isthmus Zapotec and many more.

iii. A verb meaning 'to owe' or 'to be obliged', or more commonly a construction with a copula or possession verb, and a non- finite main verb, such as English *to have to* or *to be to*. Examples may be found in the Western Romance languages, the Eastern Kru languages, Korean and Ecuadorian Quechua.[11]

English offers a unique opportunity to study the diachronic development of future grams, since it has developed one from each of these sources during the documented period. *Shall* has developed from a main verb meaning 'to owe', *will* from a main verb meaning 'to want', and the source of *be going to* is still transparent. With their original semantics, each of these constructions has very specific meanings, which require the presence of an animate and volitional agent and an active verb. The data from Old

and Middle English examined by Bybee and Pagliuca (1987) show that the generalization of the meaning of these grams takes place through their use to signal intention, especially of a first person subject. Such uses are attested for *shall* from the Old English period:

(31) *Ic thæm godan* ***sceal****, for his mod-thræce, madmas beodan.*
'I shall offer the good (man) treasures for his daring.'
(*Beowulf*, l. 384)

Will is infrequent in Old English, but increases in frequency in Middle English, both in its use for 'want' and in the expression of intention:

(32) *Now* ***wyl*** *I of hor servise say yow no more ...*
'Now I will tell you no more of their service.'
(*Sir Gawain and the Green Knight*, l. 130)

These intention uses probably retain much of the original lexical meaning of the auxiliary, with *shall* conveying the sense of 'I intend to because of externally imposed obligation or necessity' and *will* conveying the sense of 'I intend to because of internal desire'. Consider the following sentence which contains both auxiliaries:

(33) *And I* ***schal*** *erly rise, on hunting* ***wyl*** *I wende.*
'And I gotta get up early, (because) I wanna go hunting.'
(*Sir Gawain*, ll. 1101-2)

A possible interpretation of the positions of *schal* and *wyl* in this sentence appeals to the fact that hunting is something the speaker **wants** to do, but it entails the **necessity** of getting up early. That is, both auxiliaries express intention, but one is intention born of desire and the other of necessity.

The *be going to* construction is another example (along with the perfect formed from *have* or *be* plus a past participle) of a gram that is made up of what are originally several different morphemes. In fact, it is typical of *go*-futures that they comprise a semantic element indicating 'movement', one indicating 'towards a goal' put into an imperfective or progressive aspect, yielding a literal meaning of 'an agent is on a path towards a goal'. In Bybee and Pagliuca (1987) it is argued that the meaning of this gram in Modern English is still very close to its original meaning, though the restriction that an animate agent and a change in physical location be involved has been lost. This understanding of the meaning of *be going to* as closely tied to its source meaning helps explain the differences between the following pairs:

(34) The ladder is going to fall
The ladder will fall.

There's going to be a storm.
There will be a storm.

Mary's going to have a baby.
Mary will have a baby.

In each case where *be going to* is used, the entity involved is interpreted as already on the path leading to the goal expressed by the main verb — the ladder in an unstable position, the storm clouds gathering, Mary already pregnant — whereas the statements with *will* make predictions for some unspecified future with no implication of connection to present states.

Be going to is used in Modern English to state intentions, particularly with a first person subject, as in the following example from Coates 1983:139:

(35) "Listen, my dear, I asked you to marry me, didn't I? And I'm **going to** do my very best to make you happy."

Thus grams from all three lexical sources converge at least partially in their early development by being used to state intentions. The intention uses of futures persist, even after the more abstract prediction use has developed, as shown for English by the examples cited in Coates (1983) of intention uses for *will* and *shall*, and as shown by the fact that among the questionnaire sentences using the highest number of future grams are those expressing intention, for instance the second verb in the following sentence:

(36) [Said by a young man]
When I **grow** old, I **buy** a big house.

With a third person human subject, a future gram such as *be going to* may be interpreted as expressing the intentions of the subject, or a prediction by the speaker. The second verb in the following sentence tied with the previous one for the highest number of future grams in the questionnaire:

(37) [The boy is expecting a sum of money.]
When the boy **get** the money, he **buy** a present for the girl.

It is perhaps via such sentences that the prediction sense of futures develops, and makes possible the use of future grams in sentences with inanimate subjects, such as the following questionnaire example, which also received a high number of future grams.

(38) [It's no use trying to swim when we get to the lake tomorrow]
The water **be cold** (then).

A sense of prediction by the speaker is the one element most commonly associated with grams labeled futures, since it is present whether or not intention is present.

It is important to note that the modal uses or nuances of future grams, such as intention, volition or obligation, in these cases are due to the historical sources of future grams. Bybee and Pagliuca (1987) argue in more detail that the differences observable in the Modern English uses of *shall*, *will*, and *be going to* are traceable to their distinct historical origins. For instance, the use of *shall* in first person questions, where *will* is not allowed (except in Irish English) is derivable from the original obligation sense of *shall*, which asks for external motivation for actions. *Will* in such questions gives the odd implication that I do not know my own intentions and desires:

(39) Shall I pick you up at seven?
Will I pick you up at seven?

Moreover, certain of the more lexical senses of both *will* and *shall* are retained in specific contexts: *shall* is used in formal or legal language to indicate obligation (Coates 1983) and *will* has its volitional sense in *if*-clauses:

(40) If you will come by at seven, I'll give you the book.

Thus despite the remarkable convergence of future grams from different sources to the expression of intention and prediction, retention of lexical meaning creates subtle nuances among futures. Such retentions make grammatical meaning difficult to characterize and compare across languages.

Two other commonly occurring modal uses of futures develop out of the prediction sense. One of these is the imperative use of futures (as in *You **will** go to bed.*) in which a prediction is used in an indirect speech act with the force of an imperative. The second is often referred to as expressing probability and is exemplified by statements such as *That will be Todd* uttered upon hearing the phone ring. This use is also a prediction, though it is a prediction about a state of affairs in present time. It is typically used when the verification of the present prediction lies in the future.

The hypothesis that in grammaticization, semantic development and phonological development run in parallel predicts that the meaning (and consequently the uses) of inflectional futures will be different from that of

periphrastic futures, and further that the meaning of inflectional futures will show that they have undergone more development. It is not immediately obvious from the data we have examined so far that futures bear out these predictions. As mentioned above, we find that both inflectional and periphrastic futures occur in environments which require a signal of intention, future prediction or a combination of these two senses. In these cases, the ratio of inflection to periphrastic grams is approximately equal. In addition, however, the inflectional futures occur in another important context from which periphrastic futures are for the most part excluded. These contexts are precisely the ones in which we would expect an older gram to occur, for they are contexts in which the reference to future time by a verbal gram is redundant. In subordinate clauses introduced by *whatever*, *if*, *even if* or *when* there were on the average eight inflectional futures but no periphrastic futures (out of a total of 47 future grams). The *when* clause in (36) had the highest number of inflectional futures of such subordinate clauses, but this was only 10 out of the 47. Notice that such clauses represent pure future time reference, without a sense of prediction. The fact that such clauses take a future gram in less than one-fourth of Dahl's sample, compared to clauses that signal prediction or intention, which take a future in 95% to 100% of the sample, shows clearly that simple future time reference is not the central meaning of future grams.

The only other redundant context for future in the questionnaire is a complement clause following the main verb *promise*;

(41) [Uttered as a promise]
I promise to **come** to you tomorrow.

The second verb in this sentence took fourteen inflectional futures as against nine periphrastic ones. The predominance of inflectional futures in both of these types of redundant contexts seems to support the hypothesis that inflectional futures have undergone more semantic reduction, although this is a question that clearly requires further investigation.

In the case of futures, then, where we have been able to examine grammaticization from three different lexical sources, we find that retentions of lexical meaning yield grams with slightly different nuances and ranges of use. However, the data also show that futures from distinct sources reach a remarkable convergence in the areas of intention and prediction, since the futures in Dahl's questionnaire are from all the different sources and yet show considerable overlap in function.

9. Conclusions

Our data show that the most common type of inflectional tense / aspect system is what we have called the tripartite system, in which the perfective / imperfective distinction is supplemented with a past / non-past distinction in the imperfective. It is also common for either of these distinctions to occur without the other; that is, for a language to have the perfective / imperfective distinction without having past (e.g. in Arabic), or for a language to have a past without the perfective / imperfective distinction (e.g. English). The two distinctions appear to be reasonably independent of one another.

The ubiquity of perfective and / or past inflection is shown by the fact that all of the languages in Dahl's sample that have inflectional tense or aspect morphology have either a perfective or a past or both, with the possible exception of Greenlandic Eskimo. It will be remembered that both past and perfective / imperfective are almost always expressed by bound or affixal morphology. The only languages that do not have one of these gram-types are languages in which **all** tense / aspect notions are expressed by periphrasis.[12] In Bybee's material a similar strong trend is seen: all the languages that have inflectional tense or aspect have a past or perfective with the exception of Diegueño. Considering that we have claimed that past and perfective tend to arise from the same historical sources (perfects), and that the prototypical use of both involves past time reference, the following universal is actually stronger than the disjunction 'past or perfective' makes it appear:

All languages with inflectional tense or aspect have grammatical expression for past or perfective or both.

It should also be remembered that there are a few cases of languages with no inflection that have a perfective (e.g. Mandarin), and languages which have inflection, but a periphrastic perfective (e.g. French).

In addition to these basic inflectional distinctions, a language may also have one or more of the gram-types perfect, progressive or future. In our data we find no indications of any implicational relations among these gram-types nor among the ones participating in the tripartite system. Each gram-type appears to be able to develop independently of the others (Bybee 1986).

Thus we have argued that behind the commonly occurring tense and aspect gram-types lie three major paths of historical development:

a. perfects from various sources becoming past or perfective
b. progressives becoming imperfective
c. futures evolving from various sources

It is important to stress that not only are the mechanisms by which grammaticization is implemented the same across languages, but also the actual semantic material that is molded by this process appears to be very similar across languages. In support of this claim, we have argued that there are a small number of possible lexical sources for each gram-type, and these are documented in unrelated languages. As each of these lexical sources enters a grammaticization path it begins a process of change that leads to one of the major tense / aspect gram-types. This process of change involves a partial convergence for grams from different sources (such as futures from 'want', 'owe' or 'be going to') as they evolve toward one of the major gram-types. If we assume that language change takes place as language is used, the fact that such a small number of paths and gram-types in the tense / aspect domain may be identified for a large number of languages points strongly to a small set of highly generalized discourse or pragmatic functions served by tense and aspect grams.

On the other hand, the observed diversity in tense and aspect in the languages of the world is due to the particular properties of the grammaticization process, such as the following:

i. the independence of the development of each gram-type allows considerable variation as to which of the gram-types are interacting in a language at any given time;
ii. moreover, at any one stage a language may have grams that are close to one another semantically; e.g. it is not uncommon for a language to have more than one gram expressing the notions associated with future;
iii. languages may vary with respect to the possible combinations of tense and aspect grams; that is, future perfect or past progressive may or may not be possible combinations;
iv. languages may vary in the extent to which other less common gram-types such as habitual, evidential, or degrees of remoteness interact with the major grams;
v. in the case of aspect, languages may have derivationally expressed meanings, as in the Slavic languages, which interact with the inflectional tenses;
vi. differences may be found among grams compared cross- linguistically, (a) according to the original lexical source of the gram, since certain nuances of meaning may be retained for long periods of time, and (b) according to where the gram stands on its particular path of grammaticization.

We propose, then, that further study of tense and aspect could profit from taking into account the universal paths of development as sources of similarity among grams of different languages, while at the same time using the particular properties of such development to understand the diversity among grams across languages.

The supercategories mentioned in the title and referred to throughout the paper — tense and aspect — are valid and useful as notional domains that are often grammaticized in language. However, their validity and usefulness as grammatical domains is undermined by the facts presented here.[13] It is not the case that tense and aspect each present a domain that languages divide into distinctive members in idiosyncratic ways, it is rather the case that there are a few major gram-types, each representing a section or range of one of a smaller number of frequently occurring paths of development. A better understanding of a gram 'perfective' is not as a member of a supercategory of aspect, but as an instantiation of a range on a path of development, comparable to other perfectives from similar sources and at similar stages of development. This means further that we do not have to concern ourselves with defining 'tense' or 'aspect' or the more recalcitrant 'mood' as overarching categories, nor with deciding whether perfect is a tense or an aspect, or whether future is a tense or a mood. Rather the relevant entity for the study of grammatical meaning is the individual gram, which must be viewed as having inherent semantic substance reflecting the history of its development as much as the place it occupies in a synchronic system.

Authors' addresses:

Joan L. Bybee
Department of Linguistics
Humanities Bldg 526
University of New Mexico
Albuquerque, NM 87131
U.S.A.

Östen Dahl
Institute of Linguistics
University of Stockholm
S-106 91 STOCKHOLM
Sweden

ACKNOWLEDGEMENT

* We are grateful to Suzanne Fleischman for commenting on an earlier version of this paper, and to input from Revere Perkins and William Pagliuca. Bybee acknowledges the support of NSF grant BNS 8318262.

NOTES

1. Language-specific grams are designated by their proper names, which are therefore capitalized, e.g. English Progressive.
2. As noted in Dahl (1984) and Dahl (1985: Chapter 4), it is not uncommon for remoteness distinctions, such as the distinction between 'near past' and 'remote past', to be of importance in tense-aspect systems. Less often, however, does one find independent grams whose main function is to signal such distinctions; the most common case is for remoteness concepts to be secondary meanings of other tense- aspect grams. Thus (present) perfects sometimes develop a secondary meaning of 'today-past' whereas pluperfects are sometimes used as remote pasts.
3. The process of grammatic(al)ization has been abundantly documented, although its theoretical study has just begun. See Givón (1971; 1979), Traugott (1982), Lehmann (1982), Heine and Reh (1984), Bybee and Pagliuca (1985; 1987).
4. Bybee and Pagliuca (1985) report a similar correlation between modals with more specific (agent-oriented) meaning, which tend to have periphrastic expression, and modals with more general epistemic meaning, which may have bound expression.
5. It should be pointed out that some modals in Swedish (e.g. *måste* 'must') are also restricted in their combinatory possibilities due to a lack of forms like the infinitive.
6. The use of *want* in sentences such as *This hinge wants oil* is a very old lexical use surviving from the period in which *want* meant 'need'.
7. What we here call 'Possessive constructions' are those that semantically correspond to the main verb use of 'have' in English. Expressing 'having' by a transitive verb appears to be an areal trait of Western and Southern Europe. It may or may not be an accident that some kind of possessive-based perfect shows up in many if not most languages which have this trait. The clearest example of a possessive-based perfect which does not involve a transitive verb is that found in some Russian dialects (cf. examples such as *U syna zenenos'* 'The son has married', lit, 'at the son is married').
8. There are some borderline cases, viz. the so-called split ergativity systems (e.g. Kurdish), where the choice of tense-aspect category influences the case-marking of subject and object, commonly in such a way that perfect (or grams historically derived from it) follows an 'ergative' principle while e.g. the present tense has 'accusative' case-marking.
9. This form, which differs from the neuter past participle only in strong verbs, has a somewhat obscure origin. It is sometimes claimed to be an invention by normative grammarians.
10. In the case of perfect expressed by a present tense auxiliary plus a past participle, the two parts of the construction could be argued to transparently reflect the meaning: the participle refers to the 'point of the event' and the auxiliary to the 'point of reference' (cf. Comrie 1976:107). In the past or perfective, where those two points are not separated, such a bipartite construction would be less well-motivated.

11. Other less common sources that are probably related to (iii) are exemplified by the *werden* 'to become' plus infinitive future in German and the Russian future formed from the Future of the copula plus the main verb.
12. The fact that the only languages that lack past or perfective in Dahl's sample have no tense or aspect inflection and many have no inflection at all suggests that languages of different morphological typology may actually express different concepts as Sapir (1921:136-146) argued. Incidentally, all of these languages do have a gram for perfect. We leave for later research the question of why some languages have no inflection.
13. This point has also been put forth independently and for different reasons in Dahl (1985:21-22) and in Bybee (1985: Chapter 9).

REFERENCES

Aksu-Koç, A.; and Slobin, D. 1987. "A psychological account of the development and use of evidentials in Turkish". In: Chafe and Nichols (eds) 1987: 159-167.

Anderson, J.M. 1973. *An essay concerning aspect*. The Hague: Mouton.

Anon. (ed.). 1971. *Papers from the Seventh Regional Meeting, Chicago Linguistic Society, April 16-18, 1971*. Chicago: Chicago Linguistic Society.

Berlin, B.; and Kay, P. 1969. *Basic color terms: their universality and evolution*. Berkeley: University of California Press.

Blansitt, B.L. 1975. "Progressive aspect". *Working Papers on Language Universals* 18:1-34. *Explanations in language universals*.

Butterworth, B.; Comrie, B.; and Dahl, Ö (eds). 1984. *Explanations of language universals*. The Hague: Mouton.

Bybee, J.L. 1985. *Morphology: A study of the relation between meaning and form*. Amsterdam: John Benjamins.

Bybee, J.L. 1986. "On the nature of grammatical categories: A diachronic perspective". In: Choi (ed.) 1986: 17-34.

Bybee, J.L.; and Pagliuca, W. 1985. "Cross-linguistic comparison and evolution of grammatical meaning". In: Fisiak (ed.) 1985:59-83.

Bybee, J.L.; and Pagliuca, W. 1987. "The evolution of future meaning". In: Ramat, Carruba, and Bernini (eds) 1987:109-122.

Chafe, W. 1970. *Meaning and the structure of language*. Chicago: The University of Chicago Press.

Chafe, W.; and Nichols, J. (eds). 1987. *Evidentiality: The coding of epistemology in language*. Norwood, N.J.: Ablex.

Choi, S. (ed.). 1986. *Proceedings of the Second Eastern States Conference on Linguistics*.
Chomsky, N. 1957. *Syntactic structures*. The Hague: Mouton.
Chung, S.; and Timberlake, A. 1985. "Tense, aspect and mood". In: Shopen (ed.) 1985: 202-258.
Churchward, C.M. 1953. *Tongan grammar*. London: Oxford University Press.
Coates, J. 1983. *The semantics of modal auxiliaries*. London: Croom Helm.
Comrie, B. 1976. *Aspect*. Cambridge: Cambridge University Press.
Comrie, B. 1985. *Tense*. Cambridge: Cambridge University Press.
Dahl, Ö. 1984. "Temporal distance: remoteness distinctions in tense-aspect systems". In: Butterworth, Comrie, and Dahl (eds) 1984: 105-122.
Dahl, Ö. 1985. *Tense and aspect systems*. Oxford: Basil Blackwell.
Dahl, Ö.; and Karlsson, F. 1976. "Verbal aspects and object marking. A comparison between Finnish and Russian". *International Review of Slavic Linguistics* 1:1-30.
Dal, I. 1952. *Kurze deutsche Syntax*. Tübingen: Niemeyer.
Davidsen-Nielsen, N. Forthcoming. *Tense and mood in English and Danish*.
Donaldson, B.C. 1981. *Dutch reference grammar*. The Hague: Martinus Nijhoff.
Faust, N. 1972. *Gramática cocama: Lecciones para el aprendizaje del idioma cocama*. Lima: Instituto lingüístico del Verano.
Fisiak, J. (ed.). *Historical semantics, historical word- formation*. The Hague: Mouton.
Fleischman, S. 1982. *The future in thought and language*. Cambridge: Cambridge University Press.
Friedman, V. 1987. "Evidentiality in the Balkans: Bulgarian, Macedonian and Albanian". In: Chafe and Nichols (eds) 1987: 168-187.
Genetti, C. 1985. "The grammaticalization of the Newari verb *tol*". *Linguistics in the Tibeto-Burman Area* 9, 1.
Givón, T. 1971. "Historical syntax and synchronic morphology: An archeaologist's field trip". In: Anon. (ed.) 1971:394-415.
Givón, T. 1973. "The time-axis phenomenon" *Language* 49:890-925.
Greenberg, Joseph H. (ed.). 1963. *Universals of language*. Cambridge, Mass.: M.I.T. Press.
Greenberg, Joseph H. (ed.). 1978. *Universals of human language. Volume 3. Word structure*. Stanford: Stanford University Press.

Harris, M. 1982. "The 'Past simple' and the ''Present perfect' in Romance". In: Vincent and Harris (eds) 1982:42-70.
Harrison, S.P. 1976. *Mokilese reference grammar*. Honolulu: University of Hawaii Press.
Hatcher, A.G. 1951. "The use of the progressive in English: A new approach". *Language* 27:254-280.
Heine, B.; and Reh, M. 1984. *Grammaticalization and reanalysis in African languages*. Hamburg: Helmut Buske Verlag.
Herslund, M; Mordrupk, O.; and Sorensen, F. (eds). 1983. *Analyses grammaticales du français*. Copenhagen: Akademisk Forlag.
Hoffmann, C. 1963. *A grammar of the Margi language*. Oxford: Oxford University Press.
Hopper, P. J. (ed.) 1982. *Tense-aspect: Between semantics and pragmatics*. Amsterdam: John Benjamins.
Hopper, P.J.; and Thompson, S.A. "Transitivity in grammar and discourse". *Language* 56:251-299.
Labov, W. 1972. *Language in the inner city*. Philadelphia: University of Pennsylvania Press.
Lee Kee-dong. 1975. *Kusaiean reference grammar*. Honolulu: University of Hawaii Press.
Lehmann, C. 1982. "Thoughts on grammaticalization". *Arbeiten des Kölner Universalienprojekts*, 48.
Lehmann, W.; and Malkiel, Y. (eds). 1982. *Perspectives on historical linguistics*. Amsterdam: John Benjamins.
Lewis, G.L. 1967. *Turkish grammar*. Oxford: Clarendon Press.
L'Hermitte, R. 1978. *La phrase nominale en russe*. Paris: Publications de la Sorbonne.
Li, C.N. (ed.) 1979. *Mechanisms of syntactic change*. Austin: University of Texas Press.
Li, C.N.; and Thompson, S.A.. 1979. "A mechanism for the development of copula morphemes". In: Li, C.N. (ed.) 1979:419-444.
Malkiel, Y. 1978. "Derivational categories". In: Greenberg (ed.) 1978:125-149.
Mańczak, W. 1978. "Irregular sound change due to frequency in German". In: Fisiak (ed.) 1978:309-319.
Marchese, L. 1986. *Tense / aspect and the development of auxiliaries in Kru languages*. Arlington: The Summer Institute of Linguistics.

Maslov, Yu. S. 1981. *Grammatika bolgarskogo jazyka*. Moskva: Vischaya shkola.

Moreno, M.M. 1964. *Grammatica teorico-pratica della lingua galla con esecizi*. Farnborough, Hants.: Gregg Press.

Nedyalkov, V.P.; *et al.* 1983. *Tipologia resul'tativnych konstruktsij*. Moskva: Nauka.

Osborne, C.R. 1974. *The Tiwi language*. Canberra: Australian Institute of Aboriginal Studies.

Parks, D. 1976. *A grammar of Pawnee*. New York: Garland Publishing [*Garland Studies in American Indian Linguistics*].

Perkins, R.D. 1980. *The evolution of culture and grammar* [SUNY at Buffalo dissertation].

Ramat, A.G.; Carruba, O.; and Bernini, G. (eds). 1987. *Papers from the VIIth International Congress on Historical Linguistics*. Amsterdam: John Benjamins.

Reichenbach, H. 1947. *Elements of symbolic logic*. New York: Macmillan.

Roth, J. 1979. *Die indirekten Erlebnisformen im Bulgarischen. Eine Untersuchung zu ihrem Gebrauch in der Umgangssprache*. Munich: Verlag Otto Sagner.

Samarin, W. 1967. *A grammar of Sango*. The Hague: Mouton.

Sapir, E. 1921. *Language*. New York: Harcourt, Brace and World.

Scott, G. 1978. *The Fore language of Papua New Guinea*. Canberra: Linguistic Circle of Canberra [*Pacific Linguistics*, B-47].

Shopen, T. (ed.). 1985. *Language typology and syntactic description. Volume III. Grammatical categories and the lexicon*. Cambridge: Cambridge University Press.

Spang-Hanssen, E. 1983. "La notion de verbe auxiliaire". In: Herslund, Mordrupk and Sorensen (ed.) 1983: 5-37.

Stojanov, S. 1964. *Gramatika na bolgarskija khizhoven ezik*. Sofia: Nauka i iskustvo.

Tedeschi, P.; and Zaenen A. (eds). 1981. *Tense and aspect*. New York: Academic Press. [*Syntax and Semantics* 14].

Traugott, E.C. 1978. "On the expression of spatio-temporal relations in language". In: Greenberg (ed.) 1978:369-400.

Traugott, E.C. 1982. "From propositional to textual and expressive meanings: Some semantic-pragmatic aspects of generalization". In: Lehmann and Malkiel (ed.) 1982:245-271.

Ultan, R. 1978. "The nature of future tenses". In: Greenberg (ed.) 1978:83-123.

Vandame, R.P.C. 1963. *Le Ngambay-Moundou: Phonologie, grammaire et textes*. Dakar: Ifan [*Mémoires de l'Institut Français d'Afrique Noire* 69].

Vincent, N.; and Harris, M. (eds). 1982. *Studies in the Romance verb*. London: Croom Helm.

Vuyst, J. de. 1985. "The Present Perfect in Dutch and English". *Journal of Semantics* 4:137-163.

Weinreich, U. 1963. "On the semantic structure of language". In: Greenberg (ed.) 1963: 142-216.

Willett, T. 1988. "A cross-linguistic survey of the grammaticalization of evidentiality". *Studies in Language* 12:51-97.