Back to the Future

Joan L. Bybee, William Pagliuca, and Revere D. Perkins
The University of New Mexico, The University of Wisconsin, and New York, New York

1. INTRODUCTION

In earlier works, the semantic development of future markers in certain well-known languages has been discussed — Fleischman (1982a) studies the development of futures in Romance languages, and in a much briefer study, Bybee and Pagliuca (1987) offer an analysis of the development of three futures in English.* Parallels in other unrelated languages are easy to find: Futures developing from the same lexical sources as in English and Romance languages are common (Ultan, 1978; Bybee and Pagliuca, 1987). Dahl's (1985) cross-linguistic study of grammatical meaning identified a universal prototype for future, claiming that the meaning signalled by future morphemes is similar cross-linguistically. In this paper, we go back to the future to test some hypotheses about grammaticization on a representative sample of the world's languages. In particular, we test the following hypotheses:

(1) That futures in all languages develop from a small set of lexical sources and that all future morphemes from a given source go through similar stages of development;

(2) That the semantic change in grammaticization is accompanied by formal reduction, whereby the morpheme loses its independence and may fuse to contiguous material.

The material upon which these hypotheses can be tested is a database constructed by the GRAMCATS project, which is an exhaustive survey of morphology associated with the verb in 75 languages. The languages were randomly selected from maximally unrelated genetic groups.[1] Our access to the languages is largely through reference grammars and partial descriptions.

We coded information about all of the closed class items that were fixed in position with respect to the verb.[2] The following types of information were included:

(1) *Meaning*: The separate senses conveyed by the gram are recorded as separate *uses*, and each use is characterized by one or more standardized meaning labels, called *meaning components*.

(2) *Form*: The formal properties of grams are recorded, as, for example, their phonetic shape, their position, whether they are bound, the existence of allomorphs, the type of conditioning for allomorphs, and so on.

(3) *Related strings*: Here we record whatever information is available about the historical relations of the gram, i.e. what other forms in the language it is homophonous with or similar to, whether they are grammatical or lexical, or any etymological information the author supplies.

Our general hypothesis is that there are only a small number of major grammatical categories or gram-types (such as 'past', 'future', or 'perfective') in the languages of the world, and that each of these develops historically via a small number of separate paths. Thus it has been found that futures — grams that have 'future' as one of their uses (hereafter *future grams*) — may develop from auxiliary constructions with the meanings of 'desire', 'obligation' or 'movement towards a goal' (Ultan, 1978; Bybee and Pagliuca, 1987). By postulating independent and unidirectional paths of development, we make predictions about the expression of grammatical meaning in languages. In particular, we predict that certain combinations of meanings will commonly be expressed by a single form, and that certain meanings are characteristic of early stages of development, while others are characteristic of later stages.

In the GRAMCATS sample to date, we have 129 grams which have 'future' as one of their meaning components, and these are distributed in 63 languages.[3] Analyzing the meanings and related strings of these grams, we come up with three major types of futures, each of which is represented by at least two separate paths of development, and a fourth, less common type:

(1) *Aspectual forms*: These are grams whose main function is to mark imperfective, and, in the rare case, perfective, but which also can be used for statements about future time.

(2) *Agent-oriented modalities*: Verbs or constructions with agent-oriented meaning, such as 'desire' or 'obligation' and, less commonly, 'ability', may develop into grams with future as one of their uses.

(3) *Movement towards*: Verbs or constructions that signal movement towards a goal may develop into future grams, with some differences discernible between those using 'come' and those using 'go'.

(4) *Temporal adverbs*: Adverbs indicating a time after the moment of speech or a reference time, e.g. *soon*, *then*.

We arrive at this classification into types by comparing the other meanings expressed by individual future grams, and by comparing their related strings. A detailed consideration of each of these types begins in section 3, but first we will make some general statements about the term 'future', what it stands for in actual language and how it is used in reference grammars.

2. WHAT IS FUTURE?

The future does not represent 'future time reference' in the same way that the past represents 'past time reference', either in the use of the term in reference grammars, or in the actual usage of future grams in languages. It appears that the prototypical use of future grams, as well as what authors of reference grammars mean when they use the term 'future', is to signal that *an assertion about future time* is being made, or in other words, to signal a *prediction* (Bybee and Pagliuca, 1987; cf. Coates, 1983; Dahl, 1985). Because a prediction is a type of assertion, future grams often do not occur in subordinate clauses, even where future time reference is clearly intended, as the English examples in (1) and (2) show:

(1) When I grow up, I want to be a pilot
*When I will grow up, I want to be a pilot
*When I'm gonna grow up, I want to be a pilot

(2) If he asks for it, I'll give it to him
If he asked for it, I would give it to him
*If he will ask, I'll give it to him

Note that the last sentence is acceptable only with the more agent-oriented reading of *will*, that of 'willingness'. Note also that an *if*-clause with a Past

Tense verb refers to a time that includes the future, albeit future hypothetical, but it still does not use the future gram.[4] This situation is typical of future grams (cf. Dahl, 1985): They only rarely occur in *if* and *when* clauses that refer to future time.[5] By contrast, a typical past tense does occur to signal past time in non-assertive clauses, such as *when*-clauses.

While prediction is the prototypical, and in our study, the defining, use for future grams, more than half of the future grams in our database have other uses as well. These other uses provide us with the clues to the semantic development of futures, telling us where the grams come from and how far they have advanced in their development. In the following, we examine the semantic development of future grams as indicated by these other uses, structuring our discussion in terms of the three most common types. Our goal will be to develop hypotheses about the semantic development of each of these types that can be tested on our cross-linguistic database. We begin with aspectual futures and follow with discussions of modality-derived and movement-derived futures.

3. ASPECTUAL GRAMS USED FOR FUTURE

The aspectual grams used for future in the database were identified by finding all the grams that had future as one of their meaning components, and in addition one of the following aspect or tense meaning components. The numbers indicate the number of co-occurrences with future: (See Appendix 1 for definitions of these meaning components.)

present	9	imperfective	7
continuous	9	perfective	3
habitual	7	past	3

We were looking here for cases of general perfective or imperfective grams that could be used as futures, as well as cases of presents with future uses, such as the English *'I go to Chicago tomorrow'*.

The prediction that this is a distinct (although not necessarily homogeneous) type from the others is borne out by the distribution of meaning components in these grams. These tense and aspect meanings tend to co-occur with one another in the same set of forms: The 44 aspect or tense meaning components co-occurring with future are clustered in only 24 forms. In fact, these meaning components tend to co-occur to the exclusion of other meaning components commonly found with futures which we will discuss

below, i.e. epistemic modalities, such as possibility and probability, and the agent-oriented modalities of desire, obligation and ability.

The aspectual grams that signal future naturally divide into two groups — those expressing imperfective meanings (such as present, continuous, habitual and imperfective) and those expressing perfective meanings (such as anterior and perfective). The former group comprises twenty grams in our database, while the latter contains only four.

For both groups, we suspect that the future reading is only obtained when the context warrants it; that is, the semantics of the gram does not explicitly contain the features 'prediction' or 'future time reference', but is much more general, and allows a future interpretation if it is supported by the context. For instance, among the perfective set, the Tahitian future anterior reading is only possible with a temporal clause that sets up the future time (e.g. *'When you return, I will already be dead'*).

Our evidence also provides support for the hypothesis that the original meaning of a gram determines the particular range of use it will have later in its history (Bybee and Pagliuca, 1987). The particular 'flavor' of future signalled by the imperfective grams in our corpus is different from that signalled by the perfectives. Except for the future anterior of Tahitian, all the perfectives (those of Abkhaz, Baining and Kanuri) signal immediate future only, suggesting that the non-past interpretation of a perfective, which cannot be present, is as close to the present as possible. On the other hand, none of the imperfectives has immediate future as a use. Further, all the 'expected futures' in the database (with the exception of one in Inuit) occur on this imperfective list. Expected or scheduled futures are the type we have in English sentences such as 'I go to Chicago tomorrow.'

The aspectual futures as a class differ from the other futures in their mode of expression: Three of them have zero expression, while none of the other 102 futures have zero as their only means of expression. Moreover, only 15% of the aspectual futures are polysyllabic, which, as we will argue later, suggests a generally high degree of grammaticization for grams in this class; in contrast, 40% of the non-aspectual futures are polysyllabic.[6] We predict a high degree of formal grammaticization for general perfectives and imperfectives because these meanings themselves evolve from more specific meanings through a long process (see Bybee and Dahl, 1989). We return in section 7 to a test of the hypothesis that aspectual futures have the expression properties of grams that have undergone a long evolution. For the moment, we simply display the distribution of expression types for aspectual vs. non-aspectual futures in Table 1, where the numbers indicate the percentage of

Table 1 *Modes of expression of aspectual and non-aspectual futures*

		Future type: aspectual	non-aspectual
Mode of expression:	zero	12	00
	monosyllabic	73	60
	polysyllabic	15	40

futures of each type that are expressed by zero, monosyllabic, and polysyllabic forms.

Table 1 shows the heavy concentration in monosyllabic expression for aspectual futures and their low incidence, relative to non-aspectual futures, of polysyllabic expression. Note also that the following implication holds between mode of expression and future type: If zero expression, then aspectual.

The aspectual path is different from the other paths, which we will discuss next, in that in the other cases, a morpheme or construction originally rich with lexical meaning is eroded or generalized into the grammatical meaning of future, whereas with the aspectual forms, the future meaning is just one interpretation of a much broader semantics. For those grams with zero expression, it is just one interpretation of a default meaning, which might result, for example, following the development of a past or perfective gram, the syntagmatic absence of which would then be interpretable as signalling present or imperfective.

4. MODALITY

It is often remarked that future grams tend to have modal uses (Chung and Timberlake, 1985). Our diachronic theory makes certain predictions about the nature of such modal uses, and how and when they appear upon the future paths. Before discussing the tests of these predictions on the cross-linguistic database, we first propose a division of the modality domain into three linguistically significant types.

If one works from a cross-linguistic database, one finds modalities of various sorts expressed grammatically in languages, but the classification of modalities usually put forth by philosophers (cf. Lyons, 1977), largely on the basis of English, is not adequate to describe their meanings, to classify them

into types or to understand how they change. Consequently, in Bybee and Pagliuca (1985) and Bybee (1985), we have proposed that the following three types of modality be distinguished:

i. *agent-oriented* (desire, obligation, ability, root possibility, permission)
ii. *epistemic* (possibility, probability)
iii. *speaker-oriented* (imperative, hortative, optative)

The usual distinction among modalities is between deontic and epistemic modalities, and it is often said that diachronically, deontic modalities develop into epistemic. We do not include deontic in our typology because it cuts across the modality domain in a way that is not cross-linguistically valid.

Deontic modality has to do with obligation and permission, and some authors associate deontic modality exclusively with directives, that is, utterances by which the speaker is doing something — issuing a command or granting permission — rather than with utterances in which the speaker is reporting on a situation (Palmer, 1979). While this supplies us with a classification of utterances, it does not supply us with a classification of the meanings of individual morphemes, since many of them, e.g. *must*, *have to*, *can*, *may*, are used both in directives and in reports with the same meaning. For instance, one can use *have to* to report a fact, as in (4), or to issue a command, as in (5).

(4) Todd has to go home at 5 on Thursdays.

(5) Todd, you have to go home now.

Such elements should be contrasted with grammatical elements which are used *only* in directives, such as imperatives, optatives and permissives. We might even want to say that the latter type, which we call *speaker-oriented* (because the *speaker* is doing something with the utterance) have little inherent semantic content, and that their principal function is to signal the type of speech act in which they occur.

Agent-oriented modality, in contrast, includes those elements that may be used to report on situations in which a modality is in effect. They predicate certain conditions on the agent with respect to the completion of the action or event described in the rest of the predicate, as in (4) and (6). Secondarily, as indirect speech acts, agent-oriented modalities may be used as directives for imposing obligations or granting permission, as in (5) and (7).

(6) I *can't* check books out. Only registered students *can* take out books

(7) You *can* come in now

Speaker-oriented modalities have the whole proposition in their scope and, rather than relating an agent to a predicate, relate the utterance to the speech situation by signalling what the speaker wants to accomplish with the utterance.

The linguistic reason for making this distinction is that agent-oriented modalities typically have a different expression type than speaker-oriented modalities. Specifically, as reported in Bybee (1985), agent-oriented modalities are not usually found as inflections on verbs, but speaker-oriented modalities commonly are. This is an indication that the agent-oriented modalities are less grammaticized than the speaker-oriented modalities.

A second problem with the term 'deontic' is that it includes only obligation and permission, that is, only the social modalities, those describing the social conditions on the agent. Many languages, however, also mark in similar ways other types of conditions on the agent, in particular the internal conditions of ability and desire. In addition, ability and desire undergo semantic change that is parallel to that which obligation undergoes.[7] In order to class all of these conditions together, we have proposed the broader term of agent-oriented modality.

We adopt the usual definition of epistemic modality as indicating 'the degree of commitment by the speaker to what he says' (Palmer, 1986:51).[8] Some examples from English include the use of *may* to indicate epistemic possibility in (8), and of *must* to indicate inferred probability in (9).

(8) I *may* have that book at home, but I'm not sure

(9) Jake *must* have missed the bus and that's why he's not here

Note that epistemic modality has in common with the speaker-oriented modalities propositional scope and the function of indicating what the speaker is doing with the utterance. Bybee (1985) finds epistemic modalities expressed inflectionally, often in a similar fashion to the speaker-oriented modalities, and proposes that these two should be called moods in contra-distinction to the agent-oriented modalities, which are not usually inflectional.

The prediction sense of future grams may also be considered an epistemic modality because it also has propositional scope and indicates 'the degree of commitment by the speaker to what he says'. That is, a prediction is an

assertion by the speaker that the state of affairs in the proposition will be true at some future time, or will be known to be true at some future time, as shown in (10) and (11).

(10) We shall no doubt live to see stranger things

(11) There's someone coming up the steps now with a blue coat — Oh, that'll be Martin

The third reason for choosing this classification is that from a diachronic standpoint it is more accurate to say that agent-oriented modalities develop into the other two than to say that deontic becomes epistemic.

There are a number of reasons for hypothesizing that agent-oriented modalities occur in the early stages of grammaticization: First, there are many documented cases where it is clear that agent-oriented modalities precede the other types. For instance, the earliest documented senses of the English modal auxiliaries are all agent-oriented: *Shall* meant 'to be obliged or destined'; *will* meant 'to want or desire'; *may* meant 'to be physically able'; and *can* meant 'to know (how to)'. Second, the semantics of agent-oriented modality is closer to lexical semantics, and, like lexical meaning, agent-oriented modality contributes to the propositional content of the utterance rather than operating on the whole proposition. Finally, agent-oriented modalities tend to have more co-occurrence restrictions than the other two types, since they require an animate agent and usually also an active verb.

Our data link three agent-oriented modalities to futures: Desire and obligation, which are richly attested, and ability, which is much less common. We are claiming that these modalities co-occur with future not because they develop later out of the future sense, but because they are present in the lexical antecedent of the future gram. Since we hypothesize that the only way for a future gram to have agent-oriented meaning is to have retained it from its original source, each future gram may have only *one* agent-oriented meaning. This prediction is borne out in our data — the future grams with agent-oriented meaning clearly fall into separate classes, that is, no one future gram has more than one agent-oriented sense.

On the basis of the classification of modality presented in this section and the documented diachronic development of future grams in languages such as English, we can postulate four stages in the semantic development of futures from modalities, which we will refer to as *future ages* (or FUTAGEs) 1, 2, 3 and 4. The semantic characteristics of each of these stages will be presented in the next section.

5. SEMANTIC AGE FOR FUTURES BASED ON MODALITY USES

In this section we will present and justify the four semantic ages for futures based on modality uses. All of the grams that are being classified for FUTAGE have 'future' as one their uses. That is, they are all used for prediction, or future time reference. Their classification into the four semantic ages is based on the uses they have in addition to their future use. In the following, we justify the criteria for the FUTAGEs using English examples; later on, we will test the hypothesized FUTAGEs on our cross-linguistic sample.

We have already mentioned that agent-oriented modalities are common sources for future grams. In Bybee and Pagliuca (1987) we show that developing grams often retain traces of their earlier lexical meaning, losing it only very gradually. The implication of this fact for futures from agent-oriented modalities is that if a future gram has a use expressing an agent-oriented modality, we may postulate that it is a young gram. Thus at FUTAGE 1 we place all the future grams that have agent-oriented uses. In this group are four grams expressing desire (found in Danish, Nimboran, Bongu and Tok Pisin), six grams expressing obligation (in Basque, Danish, Slave, Bari, Haka and Buriat), and two grams expressing ability (both in Cantonese). We consider these forms to be the youngest type of future. That is, we regard future grams with agent-oriented uses as being at the beginning of their development as futures.

The next stage also involves agent-oriented modality, but the particular modalities — intention, willingness and root possibility — have generalized from the more specific source meanings. In tracing the development of *shall* (from a verb meaning 'to owe' — an obligation source) and *will* (from a verb meaning 'to want' — a desire source) in English, Bybee and Pagliuca (1987) find that intermediate between the full agent-oriented senses of these auxiliary verbs and the prediction or future sense is a stage at which they are commonly used to express intention, especially of a first person agent. Such uses can still be found in the modern British English examples (12) thru (15) (from Coates, 1983).

(12) I'*ll* get you a map

(13) I'*ll* keep an eye open for it

(14) ...and I *shall* get to London as soon as I can

(15) She finally said 'don't argue with me, if you argue with me I *shall* put this phone down'

The intention sense may derive from either type of agent-oriented modality. This use develops because when a speaker states in first person what s/he *wants to* do or *has to* do, it can usually be inferred that the speaker *intends to* do just this. In fact, the use of modalities to state intentions is an important link in the chain of developments leading to future, because once the intention use becomes well-established, a further inference can be made: One can predict that the subject *will* do what s/he intends to do. Since the intention use appears after agent-oriented uses and before the prediction use, we classify grams with intention as a use as occurring at FUTAGE 2.

Root possibility refers to a type of agent-oriented possibility that commonly develops out of the ability sense, and is more general than simple ability. While ability describes enabling conditions internal to the agent, root possibility predicates more general enabling conditions which include those internal to the agent, but also conditions external to the agent, which may be physical or social. Modern English *can* is commonly used to express root possibility, as in (16) and (17).

(16) I know a place where I *can* get tickets, if you are willing to pay a little more

(17) *Can* you pick me up at three-thirty?

Since root possibility also encompasses social enabling conditions, it usually expresses permission also (see examples (6) and (7) above). In English, the permission and root possibility senses of *can* (and also *may*) developed out of the earlier ability sense. Thus we catergorize future grams with a root possibility sense as belonging to FUTAGE 2. Our database contains only two future grams with root possibility uses, in Tanga and Trukese, and in the latter case, the root possibility sense occurs only in the negative.

FUTAGE 3, the next stage in the semantic development, contains all those grams that have future as their only reported use. This stage represents the period during which the agent-oriented uses have been eroded, and no further uses have yet developed. Of course, at this point we can no longer tell if a particular gram has developed from a modality or from some other source, but at this stage, we claim, it does not matter. If the gram has evolved so as to signal prediction or future and nothing else, then semantically it belongs at this point.[9]

At the final stage postulated here, FUTAGE 4, are those grams with future as a use which also have epistemic uses (other than prediction) or

speaker-oriented modalities as uses. We have already mentioned that epistemic modalities develop out of the agent-oriented ones (Shepherd, 1982; Bybee and Pagliuca, 1985). Thus *may* used for epistemic possibility, as in example (8), develops out of earlier uses of *may* which were more like the uses of *can* in (16) and (17) (Shepherd, 1982; Bybee, 1988). Similarly, the use of *must* for epistemic probability is much later than its use for obligation.

We also propose that the use of future for imperative develops out of the prediction sense, and thus is a late use. One piece of evidence for this sequence of events is the fact that futures with imperatives as other uses in our database come from all the possible sources for futures — from desire, obligation, copula and movement verb futures. If imperative were an early use and developed only from a particular source (such as obligation), then we would expect it to be restricted to futures with that source. We propose that the imperative use of futures develops out of the reinterpretation of a secondary speech act: In a situation in which the speaker has authority over the addressee, a second person prediction, such as (18), is interpreted as a command:

(18) You will go to bed!

Thus futures with imperative uses are classified as belonging in FUTAGE 4.

Also in FUTAGE 4 are futures that are used in certain types of subordinate clauses, i.e. complement clauses and the protases of conditional sentences. Since the development of future sense from modalities involves the main clause functions of stating intentions and making predictions, future uses develop first in main clauses; only after a long period of development do they come to be used obligatorily in subordinate clauses. In fact, modalities that develop into futures often retain their older meanings in subordinate clauses, as, for example, *will* in Modern English protases, which conveys the sense of willingness rather than future:

(19) If you *will* help me, we can finish faster

Thus, in order for a future gram to come to be used obligatorily in such clauses without conveying any agent-oriented meaning, it must have already lost that meaning in other contexts, and must be quite old semantically.

To summarize, then, we postulate the paths of development shown in Figure 1, and classify the parts of the continuum according to four semantic ages.

Note that all of the grams we are classifying have future as a use, so that the ones with other uses as well actually occupy a stretch of the continuum rather than a single numbered position.

FUTAGE 1		*FUTAGE 2*		*FUTAGE 3*		*FUTAGE 4*
obligation desire	>	intention	>	future	>	probability imperative
ability	>	root possibility	>	future	>	possibility

Figure 1 *Stages of development and FUTAGE for modality senses*

6. FUTURE GRAMS FROM MOVEMENT VERBS

In this section we assign FUTAGE values to the future grams in our database that derive from movement verbs. We begin with some comments about the lexical semantics of the source constructions.

In the database are seventeen future grams whose lexical antecedents are identifiably verbs of movement; in eight cases, this verb is glossed as 'go', in nine as 'come'. The 'go'-derived futures occur in:

003 Margi	(Afroasiatic, Chadic, East Chadic)
011 Cocama	(Andean-Equatorial, Equatorial, Tupi)
018 Maung	(Australian, Iwaidjan)
027 Atchin	(Austronesian, Oceanic, N-Western New Hebrides)
043 Abipon	(Ge-Pano-Carib, Macro-Panoan)
075 Mano	(Niger-Kordofanian, Niger-Congo, Mande)
081 Zuni	(Penutian, Penutian isolate)
088 Nung	(Sino-Tibetan, Tibeto-Burman, Bodo-Naga-Kachin)

The 'come'-derived futures occur in:

003 Margi	(Afroasiatic, Chadic, East Chadic)
014 Tucano	(Andean-Equatorial, Tucanoan)
051 Danish	(Indo-European, Germanic)
072 Mwera	(Niger-Kordofanian, Niger-Congo, Benue-Congo)
073 Tem	(Niger-Kordofanian, Niger-Congo, Gur)
074 Engenni	(Niger-Kordofanian, Niger-Congo, Kwa)
075 Mano	(Niger-Kordofanian, Niger-Congo, Mande)
080 Tojolabal	(Penutian, Mayan)
089 Cantonese	(Sino-Tibetan)

Taken together, then, in this sample, futures with sources in verbs of movement outnumber futures from desire, obligation or temporal adverb sources. Given the number of citations in the literature for *go-* and *come-*futures (Ultan, 1978; Marchese, 1986; Fleischman, 1982b; Heine and Reh, 1984; Bybee and Pagliuca, 1987), this is probably not remarkable. That a genetically balanced world-wide sample has resulted in an almost identical number of futures from each source, however, suggests that the sometimes alleged source asymmetry — that 'go' is more frequently a source of future than is 'come' (cf. Givón, 1973; Traugott, 1978; Fleischman, 1982b) — may be illusory. However, it should be observed that four of the languages with a 'come' source are Niger-Congo languages.

Since we find movement verbs as sources for future, pasts and progressives (see Bybee, Pagliuca and Perkins (In preparation)), we hypothesize that the semantics of movement is not sufficient in itself to give rise to the future sense. Rather, movement constructions that are sources for future grams actually signal that the subject is in the process of moving towards a goal. That is, along with movement as a component of meaning, the source of such futures includes an imperfective (or progressive) component and an allative component. Neither allativity nor imperfectivity need be overtly marked (as they are, for instance, in the English *be going to* construction), but we would not expect a motion verb with an ablative component or perfective marking to develop a future reading.[10]

Evaluating the hypothesis that the semantic components of allativity and imperfectivity are present in *come-* and *go*-derived futures is made difficult by the fact that relatively few of the lexical sources in our database are glossed with an explicit allative element such as 'towards', though *come*-derived futures seem more obliging than *go*-futures in this regard. Whether we can infer from this that the 'default' directional component for a generalized movement verb is allative, with ablative typically requiring an explicit marker, is not yet clear; we would prefer to evaluate such a possibility only after we know more about directional marking in general in the sample languages.[11] However, relatively clear cases include the Margi *come*-future auxiliary from a verb glossed as 'to approach/come near to', and Danish, with 'come' plus an allative preposition as parts of the *come*-future construction. Somewhat less clear is the case of the so-called Future Progressive in Tojolabal, the form of which may descend from a verb glossed as 'to enter'. The *come*-derived future in Mano, *nu pèá*, which is used as a future progressive and a venitive, consists of 'come' followed by the progressive marker *pèá*, which is 'perhaps identical' with the allative postposition *piè*, *pià* (Becker-Donner, 1965:37);

Mano's *go*-derived future, on the other hand, is simply identified with the verb 'go'. As a case in which evidence for an explicit allative component is lacking for a *come*-derived future but the expected aspectual restriction seems to be in force, we may cite Tem, whose future is built on *kɔni* 'to come', inflected for agreement and for non-past (Present or Remote Future) tense, which is expressed by *n*-, the full expression being:

[person/number + n + kɔni] [person/number + main verb].

Movement-derived futures do not provide the richness of semantic nuance that modality-derived futures do, and thus do not allow us the opportunity of making fine distinctions in semantic age. However, we can map movement-derived futures onto the FUTAGEs established in the last section by examining their other uses.[12]

It will be recalled that grams which express intention are classified as being at FUTAGE 2, and that intention uses can come from any of the agent-oriented modalities. Intention can arise in movement-derived futures as well, as can be seen in (20) and (21), examples of the *be going to* future from Coates (1983):

(20) We're not going to let you walk home on your own

(21) I'm going to give you something

Note that this use comes fairly directly from the literal meaning of 'the subject is on a path moving towards a goal'. The only necessary change is for the movement and path to be taken figuratively for the intention sense to arise.

Thus all futures used for intention are classified as FUTAGE 2, no matter what their lexical source.

Six out of the nine *come*-futures in our database have 'immediate future' rather than simple future as their use. We regard this as a more specific, semantically richer type of future than the simple future sense, and therefore classify it as being younger than the simple future. So all instances of immediate future (which are much more common for *come*-futures than for futures from any other source) are classified as FUTAGE 2, just one step younger than simple futures.

Other uses found with movement-derived futures, such as speaker-oriented, epistemic and subordinate uses, classify a gram as FUTAGE 4, just as they do for modality-derived futures. As we mentioned earlier, grams whose sources are unknown are classified according to this same scheme, given the meanings expressed in their other uses. Also classified in this way are the

Table 2 *Meaning components for each FUTAGE*

FUTAGE 1	*FUTAGE 2*	*FUTAGE 3*	*FUTAGE 4*
obligation	intention	future	probability
desire	root possibility		possibility
ability	immediate future		imperative
			use in complements
			use in protases

grams from the less common sources — the temporal adverbs, and a few grams derived from the meaning 'enter a state'. Table 2 displays the meaning components associated with each of the FUTAGEs.

To summarize, all of the non-aspectual future grams in our database can be classified as belonging in one of these four groups on the basis of the other meanings that they express, no matter whether we have information about their lexical sources or not. The aspectual futures are not classified in these four categories, because, on the basis of the other meanings they express, we hypothesize that the aspectual futures develop in an entirely different way: As we argued in Section 3, aspectual futures develop as imperfectives or perfectives, which happen to accomodate future readings, but they do not develop explicit future semantics. The grams that we have included in FUTAGEs 1 through 4 are grams that are undergoing similar developments, and include grams that are on the same paths. The aspectual futures are on an entirely different path. However, the range of meanings expressed by aspectual futures suggests to us that these grams are very advanced in their semantic development. We therefore predict that they are more grammaticized than explicit future grams, and thus assign them a semantic age as well; we will here call it FUTAGE A.

Having established a means of classifying individual grams according to how advanced their semantic development is, we turn now to the classification of the formal development of these grams in order to test our hypotheses concerning the universality of the paths of development for future and the parallel reduction of meaning and form.

7. FORM/MEANING COVARIATION

In Bybee (1985) it is argued that certain formal properties of affixes are related to the interaction of their semantic properties with those of the stem.

In particular, the degree of semantic relevance of the meaning of the affix to that of the stem correlates with the position of the affix and its degree of fusion with the stem. While in Bybee (1985) this principle is formulated in terms of the supercategories of aspect, tense, mood, person, number and so on, Bybee, 1986 shows that within a language affixes are not positioned according to membership in these supercategories, but rather correlations of meaning with form apply to individual affixes. That is, it is not true that languages tend to have all of their tense affixes in a single position, all of their aspect affixes in a single position, and so on. On the contrary, it is just as common to find, for example, a future affix in a different position from the past affix in the same language. The reason for this is that in most cases such affixes have developed independently of one another, and their positioning is governed by the source constructions from which the affixes arose diachronically, which in turn is governed by the general typological features of the language at that time.

Taking this historical perspective on the development of affixes (or grammatical morphemes in general) means that in addition to the formal properties predicted by semantic relevance, there may be formal properties predictable by the stage of development a gram has reached in any particular synchronic slice. Thus, intersecting the form/meaning correlation whose basis is the relevance principle will be a form/meaning correlation due to the degree of grammaticization of the form. We would then expect grams that are older — i.e. that have undergone more development — to be closer to the stem, more fused and shorter or more reduced in segmental material than younger grams of equal relevance.

Some findings of this nature are presented in Bybee (1985) and in Bybee and Pagliuca (1985), where it is reported that the agent-oriented modalities of obligation, permission, desire and ability, which are relatively close to lexical meanings, tend to occur in periphrastic expressions, while the more grammaticized epistemic and speaker-oriented moods can be either bound or free. Similarly, Dahl (1985) reports that in the tense/aspect domain the following correlations are found in his sample of 64 languages:

periphrastic	*bound*
progressive	imperfective
perfect	perfective
	past

Since there is evidence that progressives develop into imperfectives and perfects develop into perfectives or pasts, Bybee and Dahl (1989) argue that the process of grammaticization accounts for these correlations as well.[13]

In Dahl's (1985) data, the future grams are split between periphrastic and bound — 27 are periphrastic, 23 are bound. If form and meaning change in parallel in grammaticization, we would predict that the periphrastic ones are 'younger' and the bound ones 'older' semantically. In order to test this hypothesis, in the next section we establish three measures of 'degree of grammaticization of form' which we attempt to correlate with the semantic ages of futures established above.

8. DEGREE OF FORMAL GRAMMATICIZATION

The GRAMCATS database contains a number of variables that are designed to indicate varying degrees of grammaticization of form. For present purposes, we have chosen to divide these variables into three groups: Those that are indicative of the degree of fusion with the verb which the gram has undergone (FUSION), those that are indicative of the extent to which the gram has lost its autonomy and become dependent (DEPENDENCE), and those that indicate the extent to which the gram's phonological bulk has been reduced (SHORTNESS). In the next three subsections we explain how each of these measures is calculated.

8.1. Fusion with the verb

In Dahl's study, and to a large extent in Bybee (1985), whether a gram is considered an affix or not depends upon whether the orthographic conventions employed by the questionnaire respondent (in Dahl's study) or the author of the reference grammar (in Bybee's study) treat the gram as an affix (written together with the verb stem) or a free morpheme (written separately from the verb). Unfortunately, orthographic conventions only partially reflect the phenomenon of affixation, which is diachronically gradual and complex, involving several different factors. We have therefore attempted to construct a much finer measure of the degree of fusion of the gram to the stem by taking into account in our coding procedure for each form the following factors:

1. *Written bound*: Here we note the orthographic conventions of the author, by indicating Y for yes, N for no, H for hyphenated and S for sometimes written bound.

2. *Open class intervening*: Here we note whether (Y) or not (N) open class material can come between the gram and the main verb.

3. *Phonological process conditioned by stem*: Y is entered if the stem causes allomorphy of a phonological nature in the gram; N is entered if it does not.

4. *Lexical conditioning*: Y is entered if the lexical class of the verb (the conjugation class) conditions allomorphy in the gram; N is entered if it does not.

5. *Conditions stem change*: If the gram conditions a change in the verb stem, the following are entered: T for a change in tone, S for a stress change, V for a vowel change and C for a consonant change.

These five factors together allow us to calculate for each gram a figure that we call 'fusion with verb' by assigning numerical values to the variables entered for each field, as follows:

FUSION:

1.	Written bound:	N=0; H or S=1; Y=2
2.	Open class intervening:	N=1; Y=0
3.	Phon. process cond'd by stem:	N=0; Y=1
4.	Lexical conditioning:	N=0; Y=2
5.	Conditions stem change:	N=0; T or S=1; V or C=2

The numerical values are chosen so that a higher degree of grammaticization will be reflected in a higher score. Thus, for each variable, the higher value indicates greater grammaticization, with the following justification.

Assuming that grams develop from free lexical items, those that can be and are written bound are more grammaticized than those that are not, with hyphenated and variable representation assigned an intermediate value. 'Open class intervening' refers to the possibility of non-grammatical elements occurring between the gram and the verb stem. Such possibilities usually decrease as a gram fuses with the verb, although languages with incorporation are exceptions to this tendency. Phonological processes in the gram conditioned by the stem will occur only as the gram starts to become more fused with the stem. This would be an early indication of fusion. 'Lexical conditioning' refers to allomorphy conditioned by the lexical classification of the verb stem (i.e. conjugation class). Such allomorphy usually arises only after the affix has been attached to the stem long enough for phonological processes to lose their productivity and for reanalysis to occur. Under 'conditions stem change', tone or stress changes count less than consonant or vowel changes, since the

former may be predictable on a phonological basis, while we exclude that possibility for the latter. This scale ranges from 0 to 9, and, as we said above, a higher number indicates a greater degree of grammaticization.

To illustrate how this scale works, we apply it to two well-known English grams, which display typical differences in expression properties, the English Past Tense (*-ed*) and the English Future, *will*.

English Past Tense:

1. The Past Tense gram is written bound (score 2);
2. It allows no open class items to come between it and the verb stem (score 1);
3. It is affected by a phonological process conditioned by the stem, i.e. the devoicing found after voiceless consonants (*helped* [hɛlpt]) (score 1);
4. It has lexically conditioned allomorphy, since we would consider the stem changes of irregular verbs to be lexically conditioned (score 2); and
5. It conditions stem changes in some verbs, such as *kept* (score 2).

Total FUSION score: 8

Will:

1. *Will* is not written bound to the verb (score 0);
2. Open class items do intervene between *will* and the verb, in particular, manner adverbs (*I'll gladly help.*) (score 0);
3. There are no phonological processes conditioned by the verb stem (score 0);
4. There is no lexical conditioning — *will* does not vary according to which main verb it is used with (score 0);
5. *Will* does not condition any changes in the verb stem (score 0).

Total FUSION score: 0

The procedure produces the desired results, since *-ed* emerges with a much higher score than *will*, reflecting the fact that the former is a suffix and the latter is an auxiliary.

8.2. Dependence

The DEPENDENCE factors are kept separate from FUSION because there are cases in which a gram undergoes changes in form during grammatici-

zation which reduce its autonomy without resulting in fusion with the verb. We believe the lack of fusion is usually due to the position of the gram. This scale is called 'dependence' since it measures the extent to which the gram is losing its autonomy independently of whether or not it is fusing with the verb. For this measure, the following criteria are considered.

1. *The number of allomorphs which are not purely phonetically conditioned*: Entered as 0, 1, 2, 3 or >3.
2. *Phonological conditioning for allomorphs*: Possible values are Y and N.
3. *Morphological conditioning for allomorphs*: This refers to allomorphy conditioned by other grammatical morphemes, and the responses are Y and N.
4. *Suprasegmental reduction*: Here we note whether (Y) or not (N) the gram has stress or non-neutral tone; if stress alternates, A is recorded.

The numerical values are as follows:

DEPENDENCE:

1. Other allomorphs: 0=0; 1 or 2=2; 3=3; >3=4
2. Phonological conditioning: N=0; Y=1
3. Morphological conditioning: N=0; Y=2
4. Stress: N=2; A=1; Y=0

This scale ranges from 0 to 9, and again the higher number indicates a greater degree of grammaticization. The values assigned are justified as follows:

As a gram becomes more dependent upon surrounding elements, it develops allomorphs, especially due to phonological processes, some of which will be conditioned by other grammatical morphemes. Thus the number of allomorphs and the presence of phonological and morphological conditioning are indicative of greater dependence. Note that phonological conditioning could be controlled by the verb stem, indicating fusion with the verb, but since such conditioning is not necessarily verb-induced, we have included phonological conditioning under dependence. Since lexical items are usually stressed and grammatical morphemes usually not, we consider the lack of stress to indicate higher grammaticization. For tone languages we assume the comparable development would be the loss of high or non-neutral tone.

To illustrate this scale, we will again calculate the values for English *-ed* and *will*.

English Past Tense:

1. We count the Past Tense as having more than 3 allomorphs because it has [d] in addition to [id], as well as several different vowel and consonant changes internal to the stem (score 4);
2. The [id]/[d] alternation is phonologically conditioned (score 1);
3. There are no allomorphs conditioned by other grammatical morphemes (score 0);
4. The gram is always unstressed (score 2).

Total DEPENDENCE score: 7

Will:

1. If we count the reduced version of *will* [əl] as an allomorph, and also its form in contraction with *not*, then it has two allomorphs (score 2);
2. They are not phonologically conditioned, as far as we know (score 0)
3. The allomorphy in *won't* is morphologically conditioned (score 2);
4. *Will* is normally not stressed (score 2)

Total DEPENDENCE score: 6

On the DEPENDENCE scale, then, the Past Tense and *will* come out with similar scores, demonstrating that at least in some cases, FUSION and DEPENDENCE are separable measures.

8.3. Shortness

The third scale measures the extent of the segmental reduction of a gram, by starting with a value of 10 and subtracting points for each segment in the gram as follows:

SHORTNESS:
C = −1
V = −2
V: = −3

The resulting scale shows a higher number for a shorter form, and a lower number for a longer form. This non-iconic representation was decided upon so that all of the scales have values running in the same direction — a higher number for a higher degree of grammaticization.

The SHORTNESS scale includes another variable. Our coding procedure

requires us to record the longest allomorph as basic in most cases, but we also provided for the existence of a reduced allomorph, that is, one which appears to be the same as the basic allomorph, but with some segment(s) or feature(s) absent. Thus we add 2 points if a reduced allomorph exists. This scale ranges approximately from 0 to 10, although higher and lower values are possible.

Again, applying this procedure to the English Past Tense and to *will*, we obtain the following results:

English Past Tense:
The longest allomorph is [id], which consists of one vowel (score −2) and one consonant (score −1); 3 from 10 leaves 7. This gram also has a reduced allomorph, [d], so we add two points, yielding a total SHORTNESS score of 9.

Will:
The longest allomorph is [wIl], which consists of a consonant (−1), a vowel (−2) and another consonant (−1); 4 from 10 leaves 6. Since there is also a reduced allomorph, we add 2, yielding a total SHORTNESS score of 8.

9. CORRELATIONS

Given these three measures of degree of grammaticization, we calculated values on each of these scales for each future gram in our database. We then attempted a correlation of these values with the division of future grams into the four semantic ages we presented in sections 4 and 5.

We predict that forms with higher FUTAGEs have significantly higher scores for the FUSION, DEPENDENCE, and SHORTNESS variables. The statistical test of this hypothesis may be seen in Tables 3 and 4. Table 3 shows the mean values for DEPENDENCE, FUSION and SHORTNESS for grams at each of the four FUTAGEs. Here we observe that the numerical measures of formal grammaticization rise as the FUTAGE rises. This is true for each FUTAGE in DEPENDENCE and FUSION, but for SHORTNESS, FUTAGE 2 shows a slightly lower score than FUTAGE 1, and FUTAGE 4 lower than FUTAGE 3. However, if FUTAGEs 1 and 2 are considered together and FUTAGEs 3 and 4 together, it is clear that 3 and 4 are more grammaticized than 1 and 2.

In Table 4 we show two ways to assess the significance of the correlations shown in Table 3, Kendall's Tau B and Gamma. The Tau-B statistic was used

Table 3 *Correlations of DEPENDENCE, FUSION and SHORTNESS by FUTAGE*

Variable	Value Label	Mean	Std Dev	Cases
Summaries of DEPENDENCE by levels of FUTAGE				
For Entire Population		3.1569	2.398	102
FUTAGE	1.00	2.0000	2.4863	12
FUTAGE	2.00	2.3750	1.9134	32
FUTAGE	3.00	3.7143	2.6521	35
FUTAGE	4.00	4.0000	2.7634	23
Summaries of FUSION by levels of FUTAGE				
For Entire Population		2.5098	1.7952	102
FUTAGE	1.00	1.5000	1.7838	12
FUTAGE	2.00	2.3750	1.5811	32
FUTAGE	3.00	2.5143	1.9000	35
FUTAGE	4.00	3.2174	1.7309	23
Summaries of SHORTNESS by levels of FUTAGE				
For Entire Population		6.3529	2.2719	102
FUTAGE	1.00	5.7500	1.8153	12
FUTAGE	2.00	5.5000	2.7001	32
FUTAGE	3.00	7.1429	2.0167	35
FUTAGE	4.00	6.6522	1.7738	23

Table 4 *Significance figures for DEPENDENCE, FUSION and SHORTNESS by FUTAGE*

Statistic	Value	Significance[14]
Statistics for DEPENDENCE by FUTAGE		
Kendall's Tau B	.25144	.0004
Gamma	.31420	
Statistics for FUSION by FUTAGE		
Kendall's Tau B	.22163	.0018
Gamma	.29530	
Statistics for SHORTNESS by FUTAGE		
Kendall's Tau B	.17975	.0062
Gamma	.22705	

because it has a statistical significance measure corresponding to it. The significance figure tells us what the probability is that the observed correlation occurred by chance. In the social sciences, any figure less than .05 (which means that the probability is 5 in one hundred) is considered statistically significant. By this measure the correlations of FUTAGE with each of the three types of formal grammaticization is highly significant.

Gamma was used because it has a proportional reduction in error interpretation; that is, the Gamma figure tells us how well we can predict the pairing of meaning and form if we know the FUTAGE and DEPENDENCE, FUSION and SHORTNESS scores for each gram. Whereas random pairings would produce approximately 50% error, knowing, for example, the FUTAGE and DEPENDENCE scores for a gram reduces that error by .31420, i.e. 31%.

Tables 3 and 4 show, then, that there is, as predicted, a highly significant correlation between the semantic classification we have imposed on the future grams and their formal properties. Moreover, these correlations are in the predicted direction — the semantic properties hypothesized to belong to older grams correlate with the formal properties that are accrued in the process of grammaticization. We conclude, then, that there is good reason to believe that future grams develop in very similar ways across unrelated languages, and that semantic and formal changes move in parallel in grammaticization.

10. THE ROLE OF LANGUAGE TYPOLOGY

It has long been observed that languages differ in their morphological types along just the formal parameters we have been discussing; in some languages grams are typically more fused, more dependent and shorter than in other languages. It is possible, then, that the form/meaning correlations we have presented in the preceding section may be affected by language type. There are two ways in which language type may interact with these correlations:

(1) Typology may interfere with the correlations. Thus, for example, if a language tends toward high fusion, then a semantically young gram may show higher fusion in that language than a semantically older gram in a less fusional language. In such a case, if we control for language type, then our form/meaning correlations will be stronger.

(2) Typology may contribute to the correlations. Thus, for example, a particular language may not only have more fused grams, but may also tend to have older grams. In this case, if we control for language type, then the correlations between form and meaning will not be as strong.

In order to control for language type, we standardize the FUSION, DEPENDENCE and SHORTNESS scores for each future gram. That is, we recalculate the scores so that e.g. a long gram in a language with mostly long grams will not be treated as long as a long gram in a language with mostly short grams. In order to calculate this standardized score, we derive the means for FUSION, DEPENDENCE and SHORTNESS for each language, giving us a quantification of the language type along our three parameters. Then a standardized score for each parameter for each future gram is derived by subtracting from the gram's raw score the mean score for the language and dividing by the standard deviation for the language. We can then apply the same methods used above for the raw scores to the standardized scores, as shown in Table 5.

The standardized scores look different from the raw scores because a gram that is less grammaticized than the mean for the language will have a negative value, and a gram that is more grammaticized than the mean a positive value. These values are often less than 1.0 because they are divided by the standard deviation. Thus they are actually measures of deviation from the mean for the language.

Note in Table 5 that, in general, the degree of formal grammaticization increases with increasing FUTAGE. There is a deviation from this trend only in FUTAGE 4 in DEPENDENCE and SHORTNESS.

Table 6 shows the tests of significance for these standardized scores. Note that the form/meaning correlation is still highly significant statistically, but that in comparison with Table 4, the degree of significance has gone down, indicating that language typology is contributing to the correlation. One way of putting it would be to say that some languages tend to have older grams — both semantically and formally — than other languages.

If it is the case that some languages tend to have grams that are more highly grammaticized than other languages, and if grammaticization involves both form and meaning, then it will also be the case that the traditional division of languages into isolating, agglutinative and fusional involves meaning as well as form. Of course, Sapir proposed that meaning was involved in morphological typology in 1921, but it has been impossible until now to test

Table 5 *Standardized DEP, FUSE and SHORT by FUTAGE*

Variable	Value Label	Mean	Std Dev	Cases
Summaries of DEPSTD by levels of FUTAGE				
For Entire Population		.0284	.9941	102
FUTAGE	1.00	−.2900	.9001	12
FUTAGE	2.00	−.2693	.7510	32
FUTAGE	3.00	.2720	1.0648	35
FUTAGE	4.00	.2382	1.1227	23
Summaries of FUSESTD by levels of FUTAGE				
For Entire Population		−.1104	.9390	102
FUTAGE	1.00	−.4488	.6517	12
FUTAGE	2.00	−.2171	.9318	32
FUTAGE	3.00	−.1573	.9328	35
FUTAGE	4.00	.2861	1.0104	23
Summaries of SHORTSTD by levels of FUTAGE				
For Entire Population		.0133	.8472	102
FUTAGE	1.00	−.2819	.6668	12
FUTAGE	2.00	−.2418	.8718	32
FUTAGE	3.00	.3363	.8824	35
FUTAGE	4.00	.0306	.7002	23

Table 6 *Significance figures for standardized scores*

Statistic	Value	Significance
Statistics for DEPSTD by FUTAGE		
Kendall's Tau B	.15384	.0104
Gamma	.18079	
Statistics for FUSESTD by FUTAGE		
Kendall's Tau B	.13985	.0161
Gamma	.16465	
Statistics for SHORTSTD by FUTAGE		
Kendall's Tau B	.15341	.0105
Gamma	.18026	

this hypothesis. However, since we have found a way of quantifying both meaning and form, we have been able to find support for this hypothesis. Another way of testing the hypothesis is to attempt to derive a typology of languages by correlating the formal parameters of DEPENDENCE, FUSION and SHORTNESS with the semantic age of the future grams of each language. Thus Table 7 shows the correlations between the semantic age of a gram and the means for each formal parameter for the languages in which the grams occur. Table 8 shows that we find statistically significant correlations such that semantically older future grams tend to occur in languages with more dependent, fused and short grams overall.

If what we find for futures is true of other gram-types as well, it would mean that morphological typology is not just a matter of form, but also of semantic content. That is, we might conclude that grams in isolating languages are *less* grammaticized — in both meaning and form — than grams in

Table 7 *Correlations of the means for DEPENDENCE, FUSION and SHORTNESS for each language with FUTAGE*

Variable	Value Lable	Mean	Std Dev	Cases
Summaries of DEPMEAN by FUTAGE				
For Entire Population		3.1435	1.2273	102
FUTAGE	1.00	2.7055	1.3699	12
FUTAGE	2.00	3.0128	1.1419	32
FUTAGE	3.00	3.1692	1.2065	35
FUTAGE	4.00	3.5146	1.2710	23
Summaries of FUSEMEAN by FUTAGE				
For Entire Population		2.6710	1.0673	102
FUTAGE	1.00	2.0475	1.2568	12
FUTAGE	2.00	2.5991	.8463	32
FUTAGE	3.00	2.7282	1.1422	35
FUTAGE	4.00	3.0093	1.0344	23
Summaries of SHRTMEAN by FUTAGE				
For Entire Population		6.3983	.9989	102
FUTAGE	1.00	6.3059	.8453	12
FUTAGE	2.00	6.1986	1.0267	32
FUTAGE	3.00	6.5143	1.0118	35
FUTAGE	4.00	6.5477	1.0243	23

Table 8 *Significance figures for the language means by FUTAGE*

Statistic	Value	Significance
Statistics for DEPMEAN by FUTAGE		
Kendall's Tau B	.12095	.0277
Gamma	.14270	
Statistics for FUSEMEAN by FUTAGE		
Kendall's Tau B	.14647	.0134
Gamma	.17290	
Statistics for SHRTMEAN by FUTAGE		
Kendall's Tau B	.11833	.0298
Gamma	.13983	

agglutinative languages, which are in turn less grammaticized than grams in fusional or inflectional languages.

The significant correlations between the means for DEPENDENCE, FUSION and SHORTNESS for each language and the semantic age of the future gram or grams in the language points to a dimension in grammaticization that needs much further investigation. However, it is important to remember that even when we control for language type the correlations between meaning and form predicted by our grammaticization hypotheses are still highly significant.

11. ASPECTUAL FUTURES

In section 3 we proposed that aspectual grams that have a future use are semantically very old, having gone through a long development before reaching the stage of signalling imperfective or present (Bybee and Dahl, 1989; Bybee, Pagliuca and Perkins, In preparation). We can test this hypothesis by comparing the measures of formal grammaticization for aspectual futures with those of the other futures. This comparison is in Table 9, where the aspectual futures are labelled as FUTAGE A. The significance figures are given in Table 10.

From Tables 9 and 10, it can be seen that aspectual futures do indeed display the properties of older grams, being more dependent, more fused and shorter than all the other FUTAGEs.

Table 9 *DEPENDENCE, FUSION and SHORTNESS by 5 FUTAGES*

Variable	Value Label	Mean	Std Dev	Cases
Summaries of DEPSTD by FUTAGE				
FUTAGE	1.00	−.2900	.9001	12
FUTAGE	2.00	−.2693	.7510	32
FUTAGE	3.00	.2720	1.0648	35
FUTAGE	4.00	.2382	1.1227	23
FUTAGE	A	.3379	1.0567	27
Summaries of FUSESTD by FUTAGE				
FUTAGE	1.00	−.4488	.6517	12
FUTAGE	2.00	−.2171	.9318	32
FUTAGE	3.00	−.1573	.9328	35
FUTAGE	4.00	.2861	1.0104	23
FUTAGE	A	.5590	1.0483	27
Summaries of SHORTSTD by FUTAGE				
FUTAGE	1.00	−.2819	.6668	12
FUTAGE	2.00	−.2418	.8718	32
FUTAGE	3.00	.3363	.8824	35
FUTAGE	4.00	.0306	.7002	23
FUTAGE	A	.8252	.9475	27

Table 10 *Significance figures for 5 FUTAGEs*

Statistic	Value	Significance
Statistics for DEPSTD by 5 FUTAGEs		
Kendall's Tau C	.19695	.0015
Gamma	.20241	
Statistics for FUSESTD by 5 FUTAGEs		
Kendall's Tau C	.23631	.0003
Gamma	.24384	
Statistics for SHORTSTD by 5 FUTAGEs		
Kendall's Tau C	.28829	.0000
Gamma	.29619	

12. CONCLUSION

We hope to have demonstrated how formal and semantic information on grammatical material in the languages of a large, carefully constructed sample of the world's languages can be profitably used in evaluating and testing hypotheses about grammaticization. In particular, analysis of the information in our database has allowed us to infer that the hypothesis of unidirectionality with respect to the evolution of future grams is, in large measure, confirmed. It has also allowed us to identify language typology as a significant factor in the grammaticization process. While the role of typology needs more investigation, even when the influence of typology is controlled for, the results provide strong corroboration of the hypothesis that there are certain universal pathways for the development of futures from different lexical sources. Some senses or uses, in particular, agent-oriented modality uses, occur early on in the evolution along a particular path, and are unique to that path. These uses, which are comparatively close to the lexical meaning of the source material, eventually erode and are gradually lost, giving way to uses which are semantically more general, and are the points of convergence for futures from distinct sources. Moreover, futures with only or predominantly 'late' uses are morphologically reduced relative to 'young' futures, suggesting a dynamic correlation of generalization of meaning with concomitant reduction in the formal expression of future grams. This correlation is shown to have strong statistical support, which suggests in turn that a rigorous evaluation of a general hypothesis, by which formal and semantic reduction proceed in parallel, is not only feasible but may in fact turn out positive.

NOTES

* We are very grateful to Daniel Devitt, Martin Haspelmath, Bernd Heine, Soteria Svorou, Sandra Thompson, Elizabeth Traugott, and Nancy Woodworth for detailed comments on this paper, and to the participants at the Eugene conference for helpful advice and information. The research on which this report is based was supported in part by NSF grant BNS 83–18262 and by a grant from the John Simon Guggenheim Foundation to the first-named author.

1. For a detailed description of the sampling algorithm, see Perkins (1984).

2. Two qualifications are necessary here. First, in languages in which much of the verbal morphology is fused into an auxiliary that is fixed in position with respect to the clause, i.e. in second position (e.g. Papago), we relaxed the rule that the grams had to be fixed in position with respect to the verb. Second, since there is a continuum between closed

and open class, we arbitrarily decided not to code any grams belonging to a class with more than twelve members.

3. A language entered our database with no future gram if it happened that the author of the grammar did not specifically mention 'future' in connection with any of the grams that met our coding criteria.

4. We follow the convention of capitalizing the first letter of the names of language-specific grams (which are proper names), while writing in lower case the 'meaning components' which are defined independently of any particular language. For definitions of meaning components referred to in this paper, see Appendix 1.

5. In our data, only one language has a future gram that is reported to occur in protases. However, since this is not something that would necessarily be mentioned in a reference grammar, it is good to have confirmation of this tendency from Dahl's questionnaire survey, where fewer than 20% of the futures in his 64 languages could occur in *when*, *if*, *even if* or *whatever* clauses.

6. The figure of 15% for polysyllabic aspectual futures includes the two aspectual futures of Alyawarra, a central Australian language, one in Baluchi, and an affirmative form in Udmurt (the Udmurt negative form, because it is built on the affirmative, was not counted as a separate form for Table 1). All four are polysyllabic, but the Alyawarra forms — the disyllabic suffixes *-ima* and *-iyla* — are perhaps only marginally so, since in this language, all words end in *a*, and this vowel is elided in connected speech (Yallop 1977:30).

7. Permission seems to be a special case, since in many known cases, it develops as just one of several readings of root possibility, usually by the generalization of a gram signalling ability (see Bybee and Pagliuca, 1985; and Bybee, 1988).

8. We are not, however, including here the closely-related domain of evidentials, which identify the information source the speaker relies on for his/her assertion. For a treatment of the grammaticization of evidentials cross-linguistically, see Willett (1988).

9. Positing a FUTAGE with 'future' as the only use does not imply that future grams are ever completely devoid of modal nuances. We would, however, suggest that the modal nuances are less salient at this stage. Furthermore, for some of our languages, a gram might be classified as having a future use but no other because the reference material did not attend to semantic detail. Such cases detract from the accuracy of our classification, and consequently from the correlations we seek to establish in the preceding sections. However, the fact that these correlations are still highly significant suggests that if the quality of data were higher, our results would be even better.

10. The case of Catalan is sometimes cited as a problem for a theory that makes universal predictions about the relation of grammatical to source meaning, since in Catalan a perfective past has developed from the verb 'to go' plus the infinitive. This periphrastic construction differs from the *go*-futures of Spanish and French, however, in lacking the allative preposition *a*. Furthermore, it does not use the regular present forms of the verb *anar*, but special forms, in all but 1st and 3rd Singular. The long versions of these special forms quite transparently consist of a stem *va* (the 3SG form) plus the inflectional perfective suffixes (Badía Margarit, 1951: 326ff.).

11. For information on directional marking in a 26-language subset of the GRAMCATS sample, see Svorou (1988).

12. We have one case each of a future gram also used for andative and venitive. An andative

conveys the sense that the subject is going towards some location away from the speaker in order to perform some activity, that is, it conveys the literal meaning of 'be going to'. A venitive conveys the meaning that a subject is coming to a location near the speaker in order to perform some activity, 'x is coming to' do something. However, most of the andatives and venitives in our database do not have future as a use. Bernd Heine (p.c.) reports that andative and venitive are usually derivational in African languages and occur in different types of constructions than movement-futures (see also Heine and Reh, 1984:142). Our data on andatives and venitives are thus in accord with this report. The situation parallels that of 'desire' components, which may become either futures or derivational desiderative markers.

13. For a finer-grained analysis of the correlation of meaning and form in the aspectual domain, see Bybee, Pagliuca and Perkins (In preparation).

14. Since our hypothesis predicts the direction of the association, the applicable statistical significance is a one-tailed significance figure and is so reported here and in the balance of the paper.

15. Three of the grams listed under FUTAGE 3 below are problematic. At the Eugene conference, Phil Young informed us that Alphonse (1956) is unreliable as a source for Guaymi, and graciously supplied us with more accurate source material. Although we recoded this language, the statistical correlations were, unfortunately, run on a version of the database that did not include the recoded future. Similarly, language 047 in the original sample and coding was Umbrian, which we later, for reasons of data quality, abandoned in favor of Latin, but, again, the statistics were run on the version of the database which contained the older forms. Finally, we have been unable to verify that the Kadugli form is in fact Kadugli rather than that of a closely-related dialect, and therefore cannot regard the form as representing an accurate coding. To our knowledge, there are no other errors of this type in the data presented here.

REFERENCES

Allen, Jerry and Janice Allen. 1965. *Halia Language Course*. Port Moresby, Papua New Guinea: Department of Information and Extension Services.

Alphonse, Ephraim S. 1956. *Guaymí Grammar and Dictionary with some Ethnological Notes*. Washington, D.C.: U.S. Government Printing Office [*Bureau of American Ethnology Bulletin* 162].

Anceaux, J.C. 1965. *The Nimboran Language: Phonology and Morphology*. 'S-Gravenhage: Martinus Nijhoff [*Verhandelingen van het Koninklijk Instituut voor Tall-, Land- en Volkenkunde* 44].

Badía Margarit, Antonio. 1951. *Gramática Histórica Catalana*. Barcelona: Noguer.

Barker, Muhammad Abd-al-Rahman and Aqil Khan Mengal. 1969. *A Course in Baluchi*, Vols. 1 and 2. Montreal: Institute of Islamic Studies, McGill University.

Barnard, J.T.O. 1934. *A Handbook of the Rawang Dialect of the Nung Language*. Rangoon: Government Printing.

Becker-Donner, Etta. 1965. *Die Sprache der Mano*. Wien: Hermann Böhlaus [*Österreichische Akademie der Wissenschaften, Philosophisch-Historisch Klasse*, Sitzungsberichte 245, Band 5].

Bell, F.L.S. 1977. *Tanga-English, English-Tanga Dictionary.* Sydney: University of Sydney [*Oceania Linguistic Monographs* No. 21].

Benton, Richard A. 1971. *Pangasinan Reference Grammar.* Honolulu: University of Hawaii Press.

Bright, William. 1957. *The Karok Language.* Berkeley: University of California Press [*UCPL* Vol. 13].

Bybee, Joan L. 1985. *Morphology: A Study of the Relation between Meaning and Form.* Amsterdam: John Benjamins.

Bybee, Joan L. 1986. "On the nature of grammatical categories." *ESCOL '85*:17–34.

Bybee, Joan L. 1988. "Semantic substance vs. contrast in the development of grammatical meaning." *Berkeley Linguistics Society* 14:247–64.

Bybee, Joan L. and Östen Dahl. 1989. "The creation of tense and aspect systems." *Studies in Language* 13:51–103.

Bybee, Joan L. and William Pagliuca. 1985. "Cross-linguistic comparison and the development of grammatical meaning." In *Historical Semantics and Historical Word Formation*, J. Fisiak (ed.), 59–83. The Hague: Mouton.

Bybee, Joan L. and William Pagliuca. 1987. "The evolution of future meaning." *Papers from the 7th International Conference on Historical Linguistics*, A.G. Ramat, O. Carruba, and G. Bernini (eds), 109–122. Amsterdam: John Benjamins.

Bybee, Joan L., William Pagliuca and Revere D. Perkins. In preparation. *The Origin and Development of Verbal Categories in the Languages of the World.*

Capell, A., and H.E. Hinch. 1970. *Maung Grammar: Texts and Vocabulary.* The Hague: Mouton.

Capell, A. and J. Layard. 1980. *Materials in Atchin, Malekula: Grammar, Vocabulary and Texts.* Canberra: The Australian National University [*Pacific Linguistics Series* D No. 20].

Caughley, Ross Charles. 1982. *The Syntax and Morphology of the Verb in Chepang.* Canberra: Australian National University [*Pacific Linguistics Series* B No. 84].

Chung, Sandra and Alan Timberlake. 1985. "Tense, aspect and mood." In *Language Typology and Syntactic Description*, Vol. III, T. Shopen (ed.), 202–58. Cambridge: University Press.

Coates, Jennifer. 1983. *The Semantics of the Modal Auxiliaries.* London: Croom Helm.

Cortade, Jean-Marie. 1969. *Essai de grammaire Touareg (dialecte de l'Ahaggar).* Alger: Institut de Recherches Sahariennes, Université d'Alger.

Dahl, Östen. 1985. *Tense and Aspect Systems.* Oxford: Basil Blackwell.

Der-Houssikian, Haig. 1980. *Tem Grammar Handbook.* Brattleboro, VT: For ACTION/Peace Corps.

Dyen, Isidore. 1965. *A Sketch of Trukese Grammar.* New Haven: American Oriental Society [*American Oriental Essay* No. 4].

Faust, Norma. 1971. "Cocama clause types." In *Tupi Studies* I, D. Bendor-Samuel (ed.), 73–105. Norman, OK: Summer Institute of Linguistics [SIL Publications in Linguistics and Related Fields No. 29].

Fleischman, Suzanne. 1982a. *The Future in Thought and Language.* Cambridge: University Press.

Fleischman, Suzanne. 1982b. "The Past and the Future: Are they *coming* or *going*?" *Berkeley Linguistics Society* 8:322–34.

Foreman, Velma M. 1974. *Grammar of Yessan-Mayo.* Santa Ana: Summer Institute of Linguistics [*Language Data Asian-Pacific Series* No. 4].

Fortescue, Michael. 1984. *West Greenlandic.* London: Croom Helm.

Furbee-Losee, Louanna. 1976. *The Correct Language: Tojolabal: A Grammar with Ethnographic Notes*. New York: Garland.
Givón, Talmy. 1973. "The time-axis phenomenon." *Language* 49:890–925.
Greenberg, Joseph H., Charles A. Ferguson and Edith Movavcsik (eds). 1978. *Universals of Human Language, Volume 3: Word Structure*. Stanford, CA: University Press.
Hall, Robert A. Jr. 1943. *Melanesian Pidgin English: Grammar, Texts, Vocabulary*. Baltimore: The Linguistic Society of America.
Hanke, A. 1909. *Grammatik und Vokabularium der Bongu-Sprache (Astrolabei, Kaiser-Wilhelmsland)*. Berlin: Georg Reimer [*Archiv für das Studium deutscher Kolonialsprachen*, Band VIII].
Harries, Lyndon. 1950. *A Grammar of Mwera*. Johannesburg: Witwatersrand University.
Heine, Bernd and Mechthild Reh. 1984. *Grammaticalization and Reanalysis in African Languages*. Hamburg: Helmut Buske.
Hewitt, B.G. 1979. *Abkhaz*. Amsterdam: North-Holland [*Lingua Descriptive Studies* Vol. 2].
Hoffmann, Carl. 1963. *A Grammar of the Margi Language*. London: Oxford University Press.
Householder, Fred W., Kostas Kazasis and Andreas Koutsoudas. 1964. *Reference Grammar of Literary Dhimotiki*. Bloomington: Indiana University [*IJAL* 30.2 Part II].
Iorss, Martial T. 1963. *Grammaire Tahitienne*. Papeete, Tahiti: Imprimerie Officielle.
Koefoed, H.A. 1958. *Danish*. London: The English Universities Press [*Teach Yourself Books*].
Kwok, Helen. 1971. *A Linguistic Study of the Cantonese Verb*. Hong Kong: Center of Asian Studies, University of Hong Kong [*Center of Asian Studies Occasional Papers and Monographs*, No. 3].
Lane, George M. 1898. *A Latin Grammar for Schools and Colleges*. Reprinted, New York: Greenwood Press.
Leslau, Wolf. 1945. *Short grammar of Tigre*. New Haven: American Oriental Society [*Publication of the American Oriental Society, Offprint Series* No. 18].
Leman, Wayne. 1980. *A Reference Grammar of the Cheyenne Language*. Greeley: University of Northern Colorado Museum of Anthropology [*Occasional Publications in Anthropology, Linguistic Series* No. 5].
Li, Paul Jen-Kuei. 1973. *Rukai Structure*. Ann Arbor: University Microfilms.
Lister-Turner, Rev. R. and Rev. J.B. Clark. n.d. *A Grammar of the Motu Language of Papua*. 2nd edn, Percy Chatterton (ed.). Sydney: Government Printer.
Love, J.R.B. 1932. "An introduction to the Worrora language, Part II." *Journal of the Royal Society of Western Australia* 18:13–22.
Lukas, Johannes. 1937. *A Study of the Kanuri Language: Grammar and Vocabulary*. London: Oxford University. (Reprinted 1967 by Dawsons of Pall Mall, London, for the International African Institute.)
Lyons, John. 1977. *Semantics*. 2 vols. Cambridge: University Press.
Marchese, Lynell. 1986. *Tense/aspect and the Development of Auxiliaries in Kru Dialects*. Dallas: Summer Institute of Linguistics.
Milne, Mrs. Leslie. 1921. *An Elementary Palaung Grammar*. Oxford: Clarendon Press.
Nadzhip, E.N. 1971. *Modern Uigur*. Translated by D.M. Segal. Moscow: Nauka.
Najlis, Elena Lidia. 1966. *Lengua Abipona*, Tomo I. Buenos Aires: Centro de Estudios

Lingüísticos, Universidad de Buenos Aires [*Archivo de Lenguas Precolombinas* Vol. 1].

N'Diaye, Genevieve. 1970. *Structure du Dialecte Basque de Maya*. The Hague: Mouton.

Newland, A.G.E. 1897. *A Practical Hand-book of the Language of the Lais as Spoken by the Hakas and other Allied Tribes of the Chin Hills (commonly the Baungshè dialect)*. Rangoon: Government Printing.

Newman, Stanley. 1965. *Zuni Grammar*. Albuquerque: University of New Mexico Press [*University of New Mexico Publication in Anthropology* No. 14].

Palmer, F.R. 1979. *Modality and the English Modals*. London: Longman.

Palmer, F.R. 1986. *Mood and Modality*. Cambridge: University Press.

Perevoshchikov, P.N., et al. 1962. *Grammatika sovremennogo udmurtskogo jazyka: fonetika i morfologija*. Izhevsk: Udmurtskoe Knizhnoe Izdatelstvo.

Perkins, Revere D. 1984. "GRAMCATS sampling procedure." ms, SUNY at Buffalo.

Platt, J.T. 1972. *An Outline Grammar of the Gugada Dialect, South Australia*. Canberra: Australian Institute of Aboriginal Studies.

Poppe, Nicholas N. 1960. *Buriat Grammar*. Bloomington and The Hague: Indiana University and Mouton [*Indiana University Publications Uralic and Altaic Series* Vol. 2].

Prost, Gibert R. 1962. "Signalling of transitive and intransitive in Chacobo (Pano)." *IJAL* 28:108–18.

Rascher, Matthäus. 1904. "Grundregeln der Bainingsprache." *Mitteilungen des Seminars für Orientalische Sprachen* 7:31–85.

Renck, G.L. 1975. *A Grammar of Yagaria*. Canberra: Department of Linguistics, Australian National University [*Pacific Linguistics Series B* No. 40].

Rice, Keren. 1989. *A Grammar of Slave*. The Hague: Mouton.

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

Sharpe, Margaret C. 1972. *Alawa phonology and grammar*. Canberra: Australian Institute of Aboriginal Studies [*Australian Aboriginal Studies* No. 37].

Shepherd, Susan C. 1982. "From deontic to epistemic: an analysis of modals in the history of English, creoles, and language acquisition." In *Papers from the 5th International Conference on Historical Linguistics*, Anders Ahlqvist (ed.), 316–23. Amsterdam: John Benjamins.

Shipley, William F. 1964. *Maidu Grammar*. Berkeley: University of California Press [*UCPL* Vol. 41].

Sorensen, Arthur P. Jr. 1969. The Morphology of Tucano. Columbia University dissertation.

Spagnolo, L.M. 1933. *Bari Grammar*. Verona: Missioni Africane.

Stevenson, Roland C. 1956–7. "A survey of the phonetics and grammatical structure of the Nuba mountain languages, with particular reference to Otoro, Katcha and Nyimang." *Afrika und Übersee* 40:73–84, 93–115; 41:27–65, 117–52, 171–96.

Subrahmanyam, P.S. 1971. *Dravidian Verb Morphology (a Comparative Study)*. Annamalainagar, Tamilnadu: Annamalai University [Department of Linguistics Publication No. 24].

Svorou, Soteria. 1988. The Experiential Basis of the Grammar of Space: Evidence from the Languages of the World. SUNY at Buffalo dissertation.

Taylor, Douglas. 1956. "Island Carib II: Word-classes, affixes, nouns and verbs." *International Journal of American Linguistics* 22:1–44.

Thomas, Elaine. 1978. *A Grammatical Description of the Engenni Language.* Dallas: Summer Institute of Linguistics [*SIL Publication in Linguistics* No. 60].
Traugott, Elizabeth C. 1978. "On the expression of spatio-temporal relations in language." Greenberg et al. (eds), 1978:369–400.
Turner, Glen D. 1958. Jivaro: Phonology and Morphology. Indiana University dissertation.
Ultan, Russell. 1978. "The nature of future tenses." Greenberg et al. (eds), 1978:83–123.
Vandame, R.P. Charles. 1963. *Le Ngambay-Moundou: phonologie, grammaire et textes.* Dakar: Ifan [*Mémoires de l'Institut Français d'Afrique Noire* No. 69].
Wacke, K. 1930. "Formenlehre der Ono-sprache (Neuguinea)." *Zeitschrift für Eingeborenen-Sprachen* 21:161–208.
West, Birdie. 1980. *Gramática Popular del Tucano.* Bogotá: Ministerio de Gobierno, Instituto Lingüístico de Verano.
Willett, Thomas. 1988. "A cross-linguistic survey of the grammaticization of evidentiality." *Studies in Language* 12:51–97.
Wilson, W.A.A. 1961. *An outline of the Temne Language.* London: School of Oriental and African Studies, University of London.
Winfield, W.W. 1928. *A Grammar of the Kui Language.* Calcutta: The Asiatic Society of Bengal [*Biblioteca Indica* Work No. 245].
Yallop, Colin. 1977. *Alyawarra: An Aboriginal Language of Central Australia.* Canberra: Australian Institute of Aboriginal Studies.
Yates, Warren G., and Souksomboun Sayasithsena. 1970. *Lao Basic Course* I. Washington, D.C.: U.S. Foreign Service Institute.

APPENDIX 1: GRAMCATS Meaning Components

ability:	The agent of the verb has the mental or physical ability to complete the action of the main verb.
admonitive:	The command constitutes a warning; "you had better not..."
almost:	Nearly; only a little less than.
andative:	The agent moves away from the deictic center in order to do something, literally, "to be going to do something".
anterior:	The situation occurs prior to reference time, and is relevant to the situation at reference time.
certainty:	The speaker is emphasizing that the proposition is true.
continuous:	A single situation is viewed as in progress, as maintained over a period of time.
counterfactual:	The proposition describes an unreal or imagined situation that could have been true but was not.
desire:	The agent of the verb desires or wants to complete the action of the main verb.

expected: The situation is to occur in the near future; what is scheduled to happen; qualifies *future*.

future: The situation takes place after the moment of speech; the speaker predicts that the situation in the proposition will hold.

gnomic: The situation described in the proposition is generic; the predicate has held, holds and will hold for the class of entities named by the subject.

habitual: The sitution is customary or usual, repeated on different occasions over a period of time.

hypothetical: The situation is unreal or imagined, but one that could be true.

immediate: A meaning component that may be combined with other tense components to indicate a situation not simultaneous with the moment of speech, but very close to it:
immediate future: What is about to occur.
immediate past: Occurring immediately before the moment of speech.

imperative: The utterance is a direct command.

inceptive: The action or event begins.

indefinite: The situation took place or will take place at some non-specific time or place.

inferred certainty: The speaker infers from evidence that the proposition is true.

intention: The agent intends to carry out the action of the main verb.

iterative: The action is repeated on one occasion.

narrative time: The use of a form with no other *past* uses for reporting a past narrative.

past: The situation occurred before the moment of speech.

perfective: The situation is viewed as bounded temporally. It cannot be simultaneous with the moment of speech; in the non-past it is sometimes interpreted as future.

permission: The agent is allowed to complete the action of the main verb.

possibility: The speaker is indicating that the situation described in the proposition is possibly true.

present: The situation occurs simultaneously with the moment of speech.

probability: The speaker is indicating that the situation described in the proposition is probably true.

progressive: The action takes place simultaneouly with the moment of reference; 'to be in the process of...'.

purpose: The clause states the purpose for which the main clause action is taking place.

root possibility: It is possible for the agent to carry out the action of the main verb; i.e. s/he is able *and* external conditions allow it.

sequence-first: A marker occurring on the first verb of a two-verb sequence indicating that the situation described by the first verb is prior to that described by the second.

venitive: The agent moves towards the deictic center in order to do something, literally, "be coming to do something".

APPENDIX 2: Future grams in the GRAMCATS database grouped by assigned semantic age (FUTAGE).[15]

FUTAGE 1

lng	lng – name	form	shape	source
002	Basque	19	-tu bear	N'Diaye 1970:211f.
051	Danish	4	skulle	Koefoed 1958:192
051	Danish	5	ville	Koefoed 1958:192
058	Nimboran	6	d	Anceaux 1965:58f.,71,79,112
059	Bongu	9	aʁ	Hanke 1909:52f.;66;87
059	Bongu	10	aʁ...m	Hanke 1909:53;87
068	Slave	50	gó?o	Rice 1989:415f
076	Bari	4	kɔ	Spagnolo 1933:105
089	Cantonese	3	hó nàng	Kwok 1971:77f.
089	Cantonese	4	wui	Kwok 1971:77
093	Buriat	21	xa	Poppe 1960:61f.
094	Tok Pisin	7	lajk	Hall 1943:32f.

FUTAGE 2

lng	lng – name	form	shape	source
001	Inuit	17	niar	Fortescue 1984:274f.,325
001	Inuit	44	qqajar	Fortescue 1984:285
001	Inuit	65	sussaa	Fortescue 1984:292
001	Inuit	89	llarumaar	Fortescue 1984:296
003	Margi	42	ləgəri	Hoffman 1963:222
008	Tigre	22	STCH	Leslau 1945:7f.
011	Cocama	6	utsu	Faust 1971:55
014	Tucano	28	gi+ti+gi+wé	West 1980:74f.
014	Tucano	32	gi + ti	West 1980:36
014	Tucano	35	gi + s a	Sorensen 1969:270
026	Motu	8	ai	Lister-Turner and Clark n.d.:12,16f.
026	Motu	36	b	Lister-Turner and Clark n.d.:17
028	Halia	17	moto	Allen and Allen 1965:31

FUTAGE 2 (*continued*)

lng	lng–name	form	shape	source
028	Halia	19	ow	Allen and Allen 1965:31
034	Pangasinan	10	na	Benton 1971:134
052	Yagaria	25	o	Renck 1975:97f.
060	Yessan-Mayo	32	nin	Foreman 1974:144f.
068	Slave	11	de+y+0	Rice 1989:588
068	Slave	51	gha	Rice 1989:417
072	Mwera	34	ci	Harries 1950:98
072	Mwera	35	cika	Harries 1950:98
073	Tem	14	ngɔni	Der-Houssikian 1980:64f.
075	Mano	3	nu pèá	Becker-Donner 1965:37f.
076	Bari	5	tu	Spagnolo 1933:105
076	Bari	23	dc'ya	Spagnolo 1933:186
080	Tojolabal	4	ʔoh	Furbee-Losee 1976:129,135
085	Chepang	38	dhaŋ	Caughley 1982:108,101,124
085	Chepang	41	kheʔ	Caughley 1982:199
086	Haka	10	lai	Newland 1897:34f.,40,468
086	Haka	16	de mung	Newland 1897:36,378
089	Cantonese	20	jau	Kwok 1971:114
091	Udmurt	4	u-...-y	Perevoshchikov et al. 1962:200-202

FUTAGE 3

lng	lng–name	form	shape	source
001	Inuit	18	jumaar	Fortescue 1984:275,325
003	Margi	37	ra	Hoffman 1963:212–217
006	Touareg	15	ed	Cortade 1969:38ff.
009	Cheyenne	63	hta	Leman 1980:191
009	Cheyenne	66	stse	Leman 1980:191
011	Cocama	2	á	Faust 1971:42
013	Jivaro	29	tat	Turner 1958:64,71
019	Worora	14	STCH	Love 1932:14
030	Tanga	5	rokok	Bell 1977:xvi–xxvii
030	Tanga	12	ku	Bell 1977:xvi–xxi
033	Trukese	4	saq	Dyen 1965:24
035	Rukai	6	ay	Li 1973:156f.
039	Abkhaz	48	zaa	Hewitt 1979:176
040	Guaymi	8	ai	Alphonse 1956:27–41
043	Abipon	7	am	Najlis 1966:37
046	Karok	15	aviš	Bright 1957:124
047	Latin	9	bi	Lane 1898:149,273ff.
051	Danish	6	få	Koefoed 1958:192
052	Yagaria	23	s	Renck 1975:94

FUTAGE 3 (continued)

lng	lng – name	form	shape	source
052	Yagaria	24	g	Renck 1975:95f.
056	Ono	4	ke	Wacke 1930:165
063	Baining	3	i-	Rascher 1904:62f.
069	Kadugli	19	ara	Stevenson 1956–7:60
072	Mwera	36	jiya	Harries 1950:99
074	Engenni	13	TONE	Thomas 1978:70
074	Engenni	25	si	Thomas 1978:68
075	Mano	5	ọ	Becker-Donner 1965:40f.
076	Bari	3	dé	Spagnolo 1933:105
084	Lao	1	si	Yates and Sayasisthsena 1970:19f.,161
088	Nung	8	di	Barnard 1934:19,21,28f.
088	Nung	12	hka	Barnard 1934:21,29,33
091	Udmurt	3	-o-	Perevoshchikov et al. 1962:200ff.
092	Uigur	27	i	Nadzhip 1971:123
093	Buriat	3	uuža	Poppe 1960:56ff.
094	Tok Pisin	1	bajmbaj	Hall 1943:90,27

FUTAGE 4

lng	lng – name	form	shape	source
001	Inuit	16	ssa	Fortescue 1984:274f.,64–7
002	Basque	5	-ko da	N'Diaye 1970:203
012	Chácobo	23	tiari	Prost 1962:116
014	Tucano	13	some	Sorensen 1969:155,184
014	Tucano	34	gi + sa	West 1980:34
017	Alyawarra	9	iya	Yallop 1977:54f.
018	Maung	3	wana	Capell and Hinch 1970:74;67ff.
018	Maung	7	ji	Capell and Hinch 1970: 67ff.,78ff.
020	Alawa	30	yi	Sharpe 1972:87ff.
026	Motu	7	bain	Lister-Turner and Clark n.d.: 11,18
027	Atchin	3	p-	Capell and Layard 1980:75ff. et passim
027	Atchin	10	ro	Capell and Layard 1980:75ff.
033	Trukese	3	qe	Dyen 1965:24
039	Abkhaz	49	p'	Hewitt 1979:176f.
050	Modern Greek	17	θa	Householder et al. 1964:105f.
051	Danish	18	komme til at	Koefoed 1958:18;183;185
060	Yessan-Mayo	14	iti	Foreman 1974:40f.
063	Baining	4	di-	Rascher 1904:62f.
077	Ngambay	6	a k-	Vandame 1963:97–9
081	Zuni	3	ʔanna	Newman 1965:37

FUTAGE 4 (*continued*)

lng	lng – name	form	shape	source
081	Zuni	5	šuk^{w}a	Newman 1965:38
082	Maidu	42	mak	Shipley 1964:47
085	Chepang	12	ca$^{?}$	Caughley 1982:49,104–111, 92f.

FUTAGE A

lng	lng – name	form	shape	source
003	Margi	32	a	Hoffman 1963:190–6
006	Touareg	2	STCH	Cortade 1969:29f., 189f.,90
008	Tigre	21	STCH	Leslau 1945:6f.
010	Island Carib	16	-ba	Taylor 1956:6,23f.,20, 42f.,xxiv
015	Gugada	3	ninj	Platt 1972:29
017	Alyawarra	4	ima	Yallop 1977:49,52f.
017	Alyawarra	5	iyla	Yallop 1977:51f.
023	Palaung	7	di	Milne 1921:35,47,68
023	Palaung	8	chang	Milne 1921:68f.
025	Tahitian	3	ua	Iorss 1963:62f.
025	Tahitian	4	e	Iorss 1963:63
034	Pangasinan	5	on	Benton 1971:133ff.,196f.
035	Rukai	4	REDUP	Li 1973:156ff.,267f.,177,208
039	Abkhaz	44	ZERO	Hewitt 1979:173
040	Guaymi	11	tau	Alphonse 1956:36f.
042	Kui	2	in	Winfield 1928:60–85; Subrahmanyam 1971:278f.
049	Baluchi	1	ZERO	Barker and Mengal 1969:129f.
049	Baluchi	9	əga bu	Barker and Mengal 1969:233f.
050	Modern Greek	7	ZERO	Householder et al. 1964:115ff.
051	Danish	1	-r	Koefoed 1958:180f.,185
063	Baining	5	sa	Rascher 1904:62ff.
071	Temne	10	tə	Wilson 1961:25
078	Kanuri	1	in	Lukas 1967:35f.,48,50f.
078	Kanuri	7	i	Lukas 1967:45ff.
091	Udmurt	1	-içk-o-	Perevoshchikov et al. 1962: 197–200
091	Udmurt	2	u-...-içk-y	Perevoshchikov et al. 1962: 197–200
091	Udmurt	5	-i-	Perevoshchikov et al. 1962:202ff.