PREDICATE AS A UNIVERSAL SYNTACTIC CATEGORY

Eloise Jelinek University of Arizona

0. INTRODUCTION

Relating categories across languages is the crucial question in the study of language universals.¹ It will be argued here that the syntactic categories (primary sentential constituents) of a language are not projections of lexical categories, and that identifying categories across languages as equivalent, as Steele (1981) has for instantiations of the category AUX, does not rest upon a language internal correspondence between these syntactic categories and particular lexical categories. A set of language independent definitions of the syntactic categories SUBJECT, AUX, PREDICATE and ADVERBIAL in terms of the functional properties (role in function/argument structure) of sentential constituents is proposed, and the instantiation of these categories in the unrelated languages Egyptian Arabic and English is shown. This set of category definitions suffices for an economical account of sentence structure in these configurational languages, and the definitions are shown to be useful in cross-language comparisons. The claim is made here that PRED-ICATE is a universal syntactic category: that is, all (complete) sentences of all languages necessarily have some constituent that we may label PREDICATE. This is not true of the other syntactic categories to be identified here, nor is it true of any lexical category, including verb.

1. ON DEFINING CATEGORIES

In generative grammar, syntactic categories have traditionally been assumed to be projections of lexical categories. Recently, \overline{X} theory has made explicit the claim that each syntactic category has some unique morphological class (i.e. noun, verb) as its 'head' constituent.

Bresnan (1980, p. 12) gives the following formalization of correspondences of this kind (Bresnan's numbering):

(6) Partial Syntactic Encoding (English)

(a)	S> NP	VP	VP>	V (NP)	(PP)
(b)	(↑Subj)=↓			(^OBJ)=↓	(†P Case)= BY
					(↑By Obj)=↓

The rules in (6a) give the C [constituent]-structure of the simple English sentence. The sets of equations in (6b) are assigned to C-structure positions as indicated by the vertical lines. Thus, they specify that the subject of a simple sentence is the NP dominated by S and preceding the VP; the

optional post-verbal NP immediately dominated by VP is the object; and that a PP whose preposition is by may be the oblique BY OBJ.

Some problems connected with this assumed correspondence between lexical categories and sentential constituents are as follows:

- 1) NP, as in Bresnan's (6), subsumes:
 - a) Nouns (a morphologically defined class) and noun phrases (endocentric constructions with a head noun)
 - b) Pronouns: person, demonstratives, Q-words, locative elements such as <u>here</u>, <u>there</u>; some quantifiers; pleonastic "it".
 - c) Proper names; cited material.²
 - d) Finite sentences; infinitives, gerunds.
 - e) Topicalized material, including prepositional phrases, etc.

If we label whatever may function as the subject of a sentence \underline{NP} , then this designation reflects the functional role of these items rather than their lexical category affiliation.

- 2) Some languages, including English, have auxiliary verbs. Steele et al (1981) argue that AUX is a syntactic category not dominated by VP. Thus, in some languages verbs occur in more than one sentential constituent. AUX verbs have a syntactic function that is distinct from that of non-AUX verbs.
- 3) NPs also have more than one syntactic function: subject, object, predicate noun, etc. In some languages, objects are not dominated by VP.
- 4) There are complete sentences in some languages without nouns, or verbs, or both. These sentences are nonetheless used to perform acts of reference and predication. (Examples will be considered below.)
- 5) Some languages may lack a noun/verb distinction. This has been repeatedly argued for certain languages of the Northwest Coast area. (See Kuipers, 1968; Kinkade, 1978; Thompson and Thompson, 1980; Demers and Jelinek, 1982)

If we are to generalize over these cases, it would appear to be useful to have some language independent definitions of syntactic categories that can be employed in making comparisons across languages. Sentential constituents are defined by syntactic rules (anaphoric relations, movement rules, etc.) and phonological rules (word boundaries, clisis, etc.) that are language particular. These language particular constituents need not correspond uniquely to lexical categories, which are also language particular, defined by the morphological component of a grammar. But sentential constituents may be related to certain roles in the function/argument structure of sentences, which provide for the use of sentences in acts of reference and predication; these are language universals. Thus the properties that sentential constituents share across languages follow from their functional properties, rather than from the properties of lexical categories.

Akmajian, Steele and Wasow (1979) observe that in order to relate categories across languages, reference must be made to semantic features associated with these language internal categories. Since by definition the categories defined by different grammars are not identical, it is some semantic correspondence between categories across languages which prompts us to relate them. However, categories are not related on the basis of lexical meaning, but on the basis of grammatical "meaning" or function. In postulating language internal categories, we would not expect to group items meaning "ear" and "listen" in one category as against "oven" and "bake" in another, for example. It is for this reason that "semantic" definitions of categories are unsatisfactory. Neither lexical meaning nor lexical inventory determines the syntactic categories of a language.

Steele in Steele <u>et al</u> (1981) proposes that categories across languages may be termed <u>equivalent</u> when they may be recognized as instantiations of a category defined in language independent terms. Steele defines AUX as follows:

> Given a set of language internal analyses, in terms of constituents, those constituents which may contain only a specified (i.e. fixed or small) set of elements, crucially containing elements marking tense and/or modality will be identified as non-distinct. (p. 21)

Note that this language independent definition of AUX does not make reference to lexical categories. Across languages, AUX has a small, often mixed constituency; these constituents are sometimes 'auxiliary' verbs, but may be particles or clitics, as Akmajian, et al (1979) show for Luiseño. Steele emphasizes that this definition of AUX does not depend upon the semantic criterion alone -- the syntactic criterion that the constituency of this category be fixed and small is of equal importance. But in order to show that the inventory of a category is fixed and small, it is necessary to list that inventory. Listing the members of a set is the least efficient means of defining that set; some more economical means of characterizing the set is to be preferred. The importance of the work of Steele and Akmajian, et al has been to show that a syntactic category cannot always be defined as corresponding to some unique lexical category, and that more than one lexical category may be represented in the inventory of a syntactic category. My proposal here is that there is a basis for language independent definitions of AUX and other syntactic categories that is more economical than, and therefore preferable to, listing the inventories of the categories. I propose an alternative language independent definition of AUX that is a refinement of the definition proposed by Steele, and show that this definition selects the same class of elements across languages as does the definition proposed by Steele. Language independent definitions of the syntactic categories SUBJECT, PREDICATE, and ADVERBIAL are also given. These definitions and the definition of AUX form a coherent set of language independent category definitions that are all defined with reference to certain functional properties of sentential constituents. Before proposing these definitions, some comments on Kaplan and Bresnan's (1980) "functional structure" will be helpful.

2. FUNCTIONAL ATTRIBUTES AND SYNTACTIC STRUCTURE

Bresnan (1980) and Kaplan and Bresnan (1980) distinguish two levels of syntactic description that may be assigned to sentences. The first of these is constituent structure (c-structure) which is represented in a conventional phrase-structure tree, and provides the input into phonological structure. The second level of description is functional structure (f-structure) which "provides a precise characterization of such traditional syntactic notions as subject, "understood" subject, object, complement, and adjunct. The f-structure is the sole input to the semantic component, which either translates the f-structure into the appropriate formulas in some logical language or provides an immediate model-theoretic interpretation for it." (Kaplan and Bresnan, p. 3). A sample f-structure is as follows (p. 4; Kaplan and Bresnan's numbering):

(5)	SUBJ	SPEC A NUM SG PRED 'GIRL'
	TENSE	PAST
	PRED	'HAND ((† SUBJ) († OBJ2) († OBJ)) '
	OBJ	SPEC THE NUM SG PRED 'BABY'
	OBJ2	SPEC A NUM SG PRED 'TOY'

A girl handed the baby a toy.

The functional structure of a sentence encodes the meaningful grammatical relations of a sentence in "a set of ordered pairs, each of which consists of an attribute [which is] the name of a grammatical function or feature (SUBJ, PRED, OBJ, NUM, CASE, etc.) and a specification of that attribute's value for this sentence." (p. 4).

C-structure and f-structure correspond as follows (p. 13; Kaplan and Bresnan's numbering):

(23)



Some comments on this correspondence are as follows: 1) Recall that Bresnan's (6b) above specifies that the Subject of a simple sentence is "the NP dominated by S and preceding the VP", and that the equations given in (6b) provide no designation for VP corresponding to the designation "Subject" for NP of S. Similarly, in (23) VP is the only syntactic constituent that is not assigned an f-attribute, that is, assigned some meaningful functional role. The f-structure of a sentence encodes its meaningful grammatical relations; if NP of S bears the SUBJECT relation to S, then we need a corresponding designation for the meaningful grammatical relation that VP bears to S. PREDICATE is a traditional term for this relation. In (23), all nouns and the verb are given the f-attribute PRED, since they are true of individuals. Similarly, the attribute specified by VP is true of the individual specified in SUBJECT. 2) In f-structure, as shown in Kaplan and Bresnan's (5), the verb hand is labeled PRED, and identified as the unique item that lexically determines the argument structure of the sentence. TENSE is also an fattribute assigned to hand in (23), although it is a major entry in fstructure. 3) Tree nodes may carry more than one f-attribute label; that is, some f-attribute distinctions are marked in the morphology (by lexical decomposition) rather than by decomposition into syntactic constituents. Also, not all the meaningful grammatical relations are signaled by any overt lexical material; SG is marked for the word girl in (23) by the absence of -s. Therefore, in the case of a one-word sentence, all the f-attributes are given values in that word (by lexical decomposition) or by some zero constituent. An example of a one-word Equptian Arabic sentence is as follows:



(6)

c-structure

- CITR T	[NUM SG]
5010	PERS 3
TENSE	PRESENT
ASPECT	IMPERFECTIVE
PRED	'WRITE ((^SUBJ)(^OBJ))'
OBT	NUM SG
-	PERS 3

f-structure

(Compare <u>kaan biyiktibu</u> 'He was writing it.') This sentence is composed of a verb form and zero tense marking in AUX. In Lummi, a Coast Salish language, there are one-word sentences composed of a nominal:



(Compare <u>sway'qa</u>'-sx^W "You're a man"; see Demers and Jelinek, 1982.)

In Egyptian Arabic, some full sentences have no verbs, although they have SUBJECT, AUX, and PREDICATE constituents:

(8) Samiir huwwa il-mudiir NP PRO NP SUBJ AUX PRED "Samiir is the director."

Others have a single constituent that is neither NP nor VP:

(9) ma9andiis NEG:PREP P:NEG PRED "I haven't (any)." (Neg:with:me)

The analysis of these sentence types will be discussed in Section 5.

A sentence may have many items that carry the f-attribute PRED, as shown in (23); but it has a unique item with the f-attribute PRED that lexically determines the function/argument structure of the sentence, as identified in Kaplan and Bresnan's f-structure. Across languages, this item is not necessarily a verb. This item appears in some sentential constituent where other f-attributes may also be given values. A suitable term for this minimal syntactic constituent is PREDICATE -- a language universal.

3. PREDICATIONAL VS. SENTENTIAL FUNCTIONS

In the preceding section, I defended the view that syntactic categories are to be defined with reference to their functional properties, rather than with reference to lexical categories, if these definitions are to have cross-language utility. In this section, I will contrast predicational and sentential functions. The language independent definitions of syntactic categories to be presented in the next section will exploit this contrast.

3.1 Predicational Functions

We have seen that every sentence has some unique element with the f-attribute PRED that lexically determines the function/argument structure of that sentence. This class of elements has a second unique trait: at least some of them are intransitive, that is, mark polyadic functions. This class of elements may be said to mark predicational functions. In a simple transitive sentence such as

(10) John gave Mary a book. F (a,b,c)

Gave marks a predicational function with three arguments. Consider the following set of sentences:

- (11) a. It was John who gave Mary a book.
 - b. A book was given to Mary by John.
 - c. Mary was given a book by John.
 - d. Who John gave the book to was Mary.

- e. What John gave Mary was a book.
- f. A book was what John gave Mary.
- g. Mary was who John gave a book to.
- h. What John did was (to) give Mary a book.
- i. John's gift to Mary was a book.

These sentences do not mean quite the same, and their differences in meaning are related to their differences in syntactic structure. All the sentences describe the same happening, and for certain purposes may be considered paraphrases of one another; they have the same truth conditions. Their differences in syntactic structure serve to bring one or another aspect of the event described into focus. In (10) and (11a), the agent is in focus; in (11b-g) either the patient or the recipient is in focus; and in (11h) and (11i) the act of giving is given prominence. Note that the sentences in (11) are intransitive, while (10) is transitive. It is (rough) paraphrase relations of the sort seen among these sentences that underlies the claim that sentences of the kind seen in (11) are derived syntactic structures that share an underlying structure with (10). In the transitive sentence (10), the agent argument of the verb is marked in the subject of the sentence; when any other element is made subject of the sentence, the resultant derived sentence is intransitive, as in (11). Thus the fact that some marked or derived intransitive syntactic structures have the same truth conditions as certain simple or unmarked transitive sentences follows from the fact that in these marked syntactic structure, the "underlying" or "logical" subject -- the agent argument of the corresponding transitive sentence -- is not the subject of the derived intransitive sentence.³

Similarly, an intransitive sentence that is unmarked in syntactic structure has the single non-oblique argument of the predicational function as subject.

- (12) John swims. F (a)
- (13) John swims before breakfast. $(F_2 (F_1))$ (a)
- (14) Before breakfast is when John swims.

In the formula given for (13), \underline{F}_2 refers to the adverbial prepositional phrase before breakfast, which takes the predicational function swim (\underline{F}_1) of the sentence under its scope. Sentences (13) and (14) have the same truth conditions, and (14) may be said to be a marked or derived syntactic structure as compared to (13), where the single argument of the verb is marked in the subject of the sentence. In sum, in non-ergative languages, the agent argument of a transitive predicate or the single argument of an intransitive predicate (the "logical subject") is marked in the subject of the sentences with simple or unmarked syntactic structure.

3.2 Sentential Functions

Some elements of a sentence mark functions that take the predicational function and its argument(s) under their scope. TENSE is such a function, sentence MOOD is another. Kahn (1973, p. 187n) describes the role of the indicative or declarative mood as follows: Compare a sentence with a map or drawing. The map or drawing may be said to have truth conditions; it shows how things stand if it is a faithful representation. But it does not <u>claim</u> to be faithful; there is nothing in the picture that corresponds to the indicative mood. A declarative sentence, on the other hand, not only describes a possible state of affairs but says that it is realized.

Kahn also cites Wittgenstein's more cryptic formulation of the same point:

A proposition shows how things stand \underline{if} it is true. And it says that they do so stand. (Tractatus 4.022)

It is of course the speaker who makes a claim; by employing a declarative sentence, he executes a particular speech act, that of asserting or claiming. But there are certain dependencies between the mood of a sentence and the kind of speech act which may be executed by uttering that sentence: declarative sentences are associated with the execution of assertions; interrogative sentences with the execution of questions; and imperative sentences with the execution of commands and requests. By using a declarative sentence, a speaker both states something and claims that what he states is true. This self-referring feature of declarative sentences lies at the root of paradoxes such as

(15) This sentence is false.

and other versions of the Cretan paradox. We may record sentence mood in a notation of the function/argument structure of sentences as follows:

- (16) He cited Wittgenstein. \vdash (T (F (a,b)))
- (17) Did he cite Wittgenstein? ? (T (F (a,b)))
- (18) [you] cite Wittgenstein! ! (F (a,b))

Here <u>F (a,b)</u> shows the function/argument structure that these sentences have in common, and tense is recorded in (16) and (17) by (T). Sentence mood is recorded by the sentential function markers (\vdash), (?) and (!).

Modality, aspect, and negation are other sentential functions that may be marked or given values in an AUX constituent across languages. The occurrence of these functions in English and Egyptian Arabic will be described in Section 5.⁴

We turn now to the language independent definitions of syntactic categories, which will exploit the distinction between predicational functions, some of which are transitive and have more than one argument, and non-predicational functions that are necessarily intransitive. Some of the latter are sentential functions; other non-predicational functions are adverbial, and differ in scope, but are also necessarily intransitive.

4. LANGUAGE INDEPENDENT DEFINITIONS OF THE SYNTACTIC CATEGORIES

The definitions in terms of functional properties are as follows:

(19) PRED: A sentential constituent where at least some transitive functions are marked.

- (20) AUX: A sentential constituent where intransitive only functions are marked, that embed a function marked in PRED and its arguments.⁵ That is, the functions marked in AUX are all sentential in scope.
- (21) SUBJ: A sentential constituent where the single argument of an intransitive function marked in PRED is marked, and where the agent argument of a transitive function marked in PRED is marked.
- (22) ADV: A sentential constituent where only intransitive functions are marked, that embed a function marked in PRED, or embed such a function and its arguments. That is, some of the functions marked in ADV are sentential in scope, and some are not.

These definitions, like the definition of AUX proposed by Steele, depend upon both syntactic and semantic criteria. The syntactic notion of sentential constituent is presupposed here, as in Steele's definition; and the crucial semantic notion is that of <u>transitivity</u>. Whether a sentence is transitive or not depends upon the <u>lexically</u> determined function/argument structure of that sentence. Once PREDICATE is defined as a sentential constituent where at least some transitive functions are marked, the other categories can be defined off the notion of transitivity.

Transitivity is a central notion in grammatical analysis. We have seen how certain kinds of derived intransitive sentences, including passive and cleft sentences, are related to simple transitive sentences. Transitivity thus participates in the definition of derived structure. Similarly, the clause may be defined with reference to transitivity, as follows: a <u>clause</u> is a domain where a single transitive function may occur. Not all clauses have a transitive function, but a clause may have no more than one. In a complex sentence, there may be several transitive functions, corresponding to the number of clauses:

- (23) I don't want John to play the piano.
- (24) John likes to play the piano and sing "Old Man River".

Intransitive functions may be marked by adjectives, prepositions, AUX, adverbs, etc.; there may be many intransitive functions marked in a clause. In main clauses, the agent argument of a transitive predicational function and the single argument of an intransitive predicational function is marked in the category SUBJECT.

The syntactic category ADVERBIAL is always optional:

(25) Possibly, John gave Mary a hard time.

Some functions that may be marked in ADV are modal, and are sentence operators, like the functions marked in AUX. Some functions that may be marked in ADV are not sentence operators, and embed only the predicational function:

(26) Quickly, John gave Mary a kiss.

These non-modal functions, and some modal functions, may also be marked in the PREDICATE constituent. Thus, while all the functions marked in ADVERBIAL are intransitive, they differ in scope; some are sentential and some are not -- and ADV are invariably <u>optional</u> sentence constituents.

In some languages, those with AUX verbs, the analytic problem is that of differentiating AUX from PREDICATE. In other languages, particularly 'isolating' languages with little morphological structure, the problem may be one of distinguishing AUX from ADV particles. But if a language has a sentential constituent that marks intransitive functions only, and these functions are all sentential in scope, then we may label this constituent AUX.⁶

The language independent definition of the category AUX proposed here is a refinement of the definition proposed by Steele on the following grounds: 1) it does not require that the inventory of the category be listed; 2) it is one of an integrated set of language independent category definitions that are all defined in the same terms; 3) it is more economical in that it presupposes only the notions of sentential constituent and function/argument structure.

5. INSTANTIATIONS OF THE CATEGORIES

Arguments were presented in Akmajian et al (1979) for the category AUX in English, and in Jelinek (1981) and Steele, et al (1981) for the category AUX in Egyptian Arabic. I will not repeat those arguments here, but will briefly summarize the claims as follows: AUX in English is a sentential constituent where tense and modality are marked, that has the following small, fixed inventory: the auxiliary verbs be, have, and do, and the AUX modals. I will add to this inventory, optionally, the particles marking sentential negation: not, n't. AUX in Egyptian Arabic is a sentential constituent where tense is marked, that has the following small, fixed inventory: the copular verb kaan, 'be', certain pronominal elements marking person subject, and, optionally, the particles marking sentential negation (mis, ma...s). (There are no AUX modals in Egyptian Arabic.) Elements marking person subject are a recurring non-definitional feature of AUX across languages, as Steele has shown.⁷ These sentential constituents in English and Egyptian Arabic thus meet Steele's definition of AUX. They also meet the definition proposed here, since all the functions marked in these constituents are sentential functions.

The instantiations in Egyptian Arabic of the remaining category definitions was demonstrated in Jelinek (1981; in press), where it was shown that these categories suffice for an account of the primary constituents of all sentence types in the language, and that a salient feature of the categories is their co-variance across sentence mood.

The syntax of Egyptian Arabic may be briefly summarized as follows: <u>tense</u> in finite sentences (declarative and interrogative, which together constitute the indicative mood) is always and only marked in the AUX constituent. Past and future tense are marked by some finite inflection of the copula, <u>kaan</u>, to which the negative particles may attach: <u>makan</u>s.⁸ There is no present tense copula, as is not uncommon across languages. But present tense indicative sentences have a constituent where the particle marking sentential negation (mis) may occur alone.



The examples in (27) and (28) show the AUX constituent, marking tense, preceding the inflected verb <u>biyiktib</u>, which marks aspect. There are three such verbal paradigms in the language, all of which are preceded by AUX marking tense contrasts: the <u>bi</u>-Imperfect, which marks imperfective aspect; the <u>Ha</u>-Imperfect, which marks prospective aspect; and the Perfect, which marks perfective aspect.¹⁰ There are also the Active and Passive Participles, which do not mark person subject, as verbs do:

(29)	huwwa	kaan Haykuun	naayim Act Part
	he	is was will-be	sleeping ms
(30)	id-dar:	s { kaan Haykuur) maktuub Pas Part
	the- lessor	n (is was will-be	written ms

These examples show that the Active and Passive Participles are also preceded by AUX.¹¹ There are in addition a number of non-verbal predicate classes: nouns, adjectives, prepositional phrases, etc., which do not mark person subject or aspect, and also occur with tense contrasts in AUX:

(31)	huwwa	kaan	tabiib N
	he	(Haykuun)	doctor
		was will-be	ms

34

$$(32) \quad \text{huwwa} \left\{ \begin{array}{c} \\ \hline \text{kaan} \\ \text{Haykuun} \end{array} \right\} \begin{array}{c} \text{fil-beet} \\ \text{PP} \\ \text{Haykuun} \end{array}$$

$$\text{he} \quad \left\{ \begin{array}{c} \text{is} \\ \text{was} \\ \text{will-be} \end{array} \right\} \begin{array}{c} \text{at-home} \\ \end{array}$$

Let us call the class of predicates that do not mark person subject $PRED_1$, and the class of predicates that mark person subject $PRED_2$:

- (33) PRED1: participles nouns adjectives prepositional phrases
- (34) PRED₂: verbs nouns prepositional phrases

Non-verbal PRED₂ include certain nouns and prepositions that mark person subject by means of pronominal suffixes that mark possession elsewhere:

 $(35) \quad \left\{ \begin{array}{c} \hline \\ kunt \\ Hakuun \end{array} \right\} \begin{array}{c} 9andi kitaab \\ PP \\ N \\ \\ Hakuun \end{array} \\ \left\{ \begin{array}{c} is \\ was \\ will-be \end{array} \right\} with-me book \\ \hline \\ with-me book \\ with-me book \\ \hline \\ with-me book \\ with-me book \\ \hline \\ with-me book \\ with-me$

I have (had, will have) a book.

(36) (_____) biddi kitaab (kunt) N-PRO N (is) wish-my book (was) I want (wanted) a book.

Note that in these examples, and in (37) and (38) below, AUX agrees in person, gender, and number with the preposition or noun predicate.

(38) (<u>kaanit</u>) biddaha kitaab

(is) wish-her book (was)

She wants (wanted) a book.

(Sentences with the future tense and a noun of volition are considered semantically anomalous.) These non-verbal $PRED_2$ are transitive; some verbal $PRED_2$ are transitive; and some $PRED_1$ (participles) are transitive.¹²

We may also recognize two varieties of AUX elements, those that do not mark person subject (AUX_1) and those that do (AUX_2) :

(39) a. AUX₁: ____ (NEG) b. AUX₂: KAAN (NEG) PRO (NEG)

Some present tense sentences have AUX_1 , where person subject is not marked in AUX. All past and future tense sentences have some inflection of the AUX verb <u>kaan</u>; therefore, they have AUX_2 , where person subject is marked. Some present tense sentences also have AUX_2 , as follows: a pronoun marking person subject appears in the AUX constituent of the sentence when the predicate of the sentence is a definite noun. This pronoun, traditionally called the "pronoun of separation", appears between the subject and predicate:

- (40) axuuya tabiib my-brother doctor My brother is a doctor.
- (41) axuuya huwwa it-tabiib my-brother he the-doctor My brother is the doctor.
- (42) axuuya it-tabiib my-brother the-doctor my brother the doctor

The examples in (40) and (41) are sentences, while (42) is an NP. In (40), the underlined space represents the sentential locus where past and future tense marking and the particles marking sentential negation may occur; in (41), the "pronoun of separation" occupies that same locus. These pronouns may appear with the negative particles attached:

- (43) axuuya mahuwwaas it-tabiib my-brother NEG-he the-doctor My brother isn't the doctor.
- (44) huwwa mahuwwaas it-tabiib he NEG-he the-doctor He isn't the doctor.

These constructions are known as "negative pronouns". Their distribution is severely limited, as follows: they occur only in present tense sentences where person subject is not marked elsewhere. Identifying the "negative pronouns" as AUX constituents enables us to account for their extremely limited distribution. The following paradigms result:

(45) axuuya (huwwa kaan Haykuun) it-tabiib

My brother is (was, will be) the doctor.

(46) axuuya (mahuwwaas it-tabiib makans maHaykuns (or mis Haykuun)

My brother isn't (wasn't, won't be) the doctor.

In sum, person subject is marked in either AUX or PRED or both in some sentences; in other sentences, person subject is not marked in either AUX or PRED. In the former, an independent subject is optional; in the latter, it is required. Where the independent subject is optional it may precede or follow AUX:

- (47) axuuya kaan naayim my-brother was sleeping My brother was asleep.
- (48) kaan axuuya naayim was my-brother sleeping My brother was asleep.

Where the subject is required it is sentence initial. This sentential constituent that in some sentences is sentence initial and in other sentences follows AUX is an instantiation of the category SUBJECT, since the agent argument of transitive functions marked in PRED is marked there. The sentential constituent that follows AUX and SUBJECT, where some transitive functions are marked, is an instantiation of the category PREDICATE. The fourth syntactic category in Egyptian Arabic differs from SUBJECT, AUX and PREDICATE in that it is always optional to the sentence. This is the category ADVERBIAL. Most adverbs in Egyptian Arabic are dominated by the PREDICATE node; some are not. Some are modal, some are not:

- (49) ta9ala bisur9a! IMP Adverb Come quickly!
- (50) ta9ala, min fadlak! IMP ADV (PP) Come, please!

In (49) an <u>adverb</u> is dominated by PREDICATE; in (50), the prepositional phrase min fadlak 'please' appears in the sentential constituent ADVERBIAL.

This sentential constituent is an instantiation of the category ADV, since the intransitive functions marked there are not all sentence operators:

(51) lissa _____ xarag just gone-out ADV AUX PRED (Perf 3ms) He has just gone out.

The optional category ADVERBIAL may precede or follow any other category; however, certain adverbial elements are limited to certain positions in the sentence.

There are non-finite Imperative and Subjunctive sentences in Egyptian Arabic that have no AUX constituent. These sentences differ from certain present tense indicative sentences that have no inflection of the AUX verb <u>kaan</u> in that they also have a non-indicative Imperative or Subjunctive verbal inflection in the PREDICATE:

- (52) iktib ig-gawaab! IMP 2ms Write the letter!
- (53) yiktib ig-gawaab! SBJT 3ms Let him (may he) write the letter!

Sentence types in Egyptian Arabic are displayed in the following sentence schemata:

(54)	S _{Indicative} >	SUBJ	AUX1	PRED	(ADV)
		(SUBJ)	AUX1	PRED ₂	l
			AUX ₂	(PRED1)	
			l	PRED ₂	ļ
		Optiona AUX.	al SUBJI	ECT may pre	cede or follow
(55)	S _{Imperative} >	(SUBJ)	PREDI	np (ADV)	

(56) SSubjunctive ----> (SUBJ) PRED_{Sbjt} (ADV)

The optional category ADVERBIAL occurs in various sentential loci. These schemata account for all sentence types in Egyptian Arabic except for existential sentences, which are <u>sui generis</u> in syntactic structure, as is not uncommon across languages. Conditional sentences, which are highly variable in structure, are also not included here. Interrogative sentences differ from declarative sentences in Egyptian Arabic in certain optional features and in intonation. See Jelinek (1981) for an account of existential, conditional, and interrogative sentences in the language.¹³

There are finite and non-finite embedded clauses in Egyptian Arabic; the former have AUX and the latter do not. Non-finite embedded clauses also have a non-indicative verb in the category PRED, as do Subjunctive and Imperative sentences.

The schemata shown in (54) through (56) show that the language independent category definitions proposed here are instantiated in Egyptian Arabic, that they suffice for an account of the primary sentential constituents in the language, and that they provide for an economical account of the structure of sentences in the language.

Instantiations of the category definitions in English are as follows: Arguments are given in Akmajian, Steele, and Wasow (op cit) and Steele, et al (op cit) for the category AUX in English. This syntactic category is required in some finite sentences and excluded in non-finite sentences. AUX is sentence initial in yes/no interrogative sentences and in second position in declarative sentences. The sentence initial constituent in declarative sentences with unmarked syntactic structure has an inventory of nouns, pronouns, names, sentences etc.; the agent argument of transitive functions is marked in this sentential constituent, and therefore it qualifies as an instantiation of the category SUBJECT. Transitive functions are marked by verbs and participles, and intransitive functions are marked by nouns, prepositional phrases, adjectives, etc., in the sentential constituent that follows SUBJECT and AUX; this constituent therefore counts as an instantiation of the category PREDICATE. There is an optional category in some sentences where some modal and non-modal functions are marked; this constituent we may label ADVERBIAL.

Sentence types in English are displayed in the following sentence schemata:

(57)	SDeclarative>	SUBJ	AUX	PRED ₂	(ADV)
· ·				pred ₁	
(58)	SInterrogative>	AUX1	SUBJ	PRED2	(ADV)
(59)	S _{Imperative} >	AUX2	(SUBJ)	PRED ₂	(ADV)

The varieties of AUX and PREDICATE in English differ with respect to tense marking, as follows:

- (60) AUX₁: finite AUX verbs modals optional NEG
- (61) AUX₂: non-finite <u>do</u> optional NEG
- (62) PRED₁: finite verbal predicators
- (63) PRED₂: non-finite verbal predicators non-verbal predicators optional NEG

On this non-transformational account, some declarative sentences in English have an AUX constituent and some do not; the latter have PRED1. Tense

may be marked either in AUX or in PRED, but AUX and PRED cannot be collapsed, since SUBJECT may intervene. Wh- questions are not included in the schemata given above; I assume that they may be derived from yes/ no questions. The NEG particle may attach to or follow some AUX element. Imperative sentences are said to have AUX₂ on the following grounds: 1) NEG attaches only in AUX; 2) do marks emphasis only in AUX. I assume that do! is non-finite, on the analogy of be!. SUBJECT is optional with don't! but is excluded with do! in modern English; (59) requires a more detailed statement. English also has both finite and non-finite embedded clauses. The non-finite embedded clauses have non-finite predicates in PRED, and no AUX constituent.

This outline of the syntactic structure of English sentences across sentence mood is intended merely to show the instantiation of the language independent definitions of the categories. The questions of whether there are (a) declarative sentences in English without an AUX constituent, as shown in (57), or (b) an AUX constituent in imperative sentences, as shown in (59), would require separate argumentation.¹⁴

6. CONCLUSIONS

The following table exhibits the properties of the categories in English and Egyptian Arabic surveyed above:

		SUBJECT	AUX	PRED	ADV
1.	Defining functional properties	agent or single argument of func- tion marked in PRED	intransitive functions embedding PRED func- tion and its arguments	at least some transitive functions	intransitive functions; some embed PRED function and some embed PRED function and its arguments
2.	<u>Non-defini-</u> <u>tional</u> <u>functional</u> <u>properties</u>	subject person number gender modality negative	mood tense modality aspect (E) voice (E) subject person number gender (A) negative	mood tense (E) modality aspect voice subject object person number gender (A) negative	modality non-modal secondary attributes negative
3.	<u>Lexical</u> inventory	nouns, names, pronouns, sentences, etc.	small, closed: particles, verbs pronouns (A)	large, open: VP, NP, ADJ, PART, PREP no class excluded	adverbs, particles, PREP P
4.	<u>Distribu-</u> tion	optional in some sentence types; required in others	required in some sen- tence types; excluded in others	required	optional

<u>Table 1</u> Syntax and Semantics of the Categories

The traits listed in Table 1 are shared by both languages, except as follows: traits marked <u>A</u> are peculiar to Egyptian Arabic, and those marked <u>E</u> are peculiar to English. The table shows that certain functional properties serve to define the categories, whereas others are non-definitional. The defining properties are most directly related to the role of the predicate in determining the function/argument structure of the clause. AUX and PREDICATE in these languages share many of the non-definitional functional properties shown in (2). In some languages (Walbiri for example) object as well as subject is marked in AUX.¹⁵ English marks both tense and modality in both AUX and PREDICATE.

All the categories differ in <u>distribution</u> -- in required/optional status. But it is possible that SUBJECT, for example, may be a required constituent in all complete sentences in certain "isolating" languages where person subject is not marked in verb morphology or AUX clitics. And it is possible that AUX may be a required constituent of all independent sentence types in a language such as Walbiri, where imperatives have an AUX. More information across languages is required, but it seems likely that distribution of the categories across sentence type would fail as a single defining feature.

All the categories differ in <u>inventory</u>, and differences across the categories in inventory are the basis of the traditional assumption that syntactic categories are projections of lexical categories. Since lexical categories are not unique to particular syntactic categories, distinctions among the categories in terms of inventory must be made by means of cumbersome lists; and inventory is not an economical means of defining any set.¹⁶

The definitional functional properties shown in (1) of the table suffice to distinguish among the categories, and seem to provide the most economical basis for the category definitions. Given these differences in functional properties as the defining feature, the other differences among the categories shown in Table 1 remain as linguistically interesting non-definitional properties.¹⁷ For example, AUX is defined as marking sentential functions. Tense is a modality; all sentential functions are modal. Therefore, modality is necessarily marked in AUX, if a language has an AUX.

Steele (1981) lists the features marked in AUX across languages. I propose that there is a semantic unity among these features: they are all either modal (sentential functions) or deictic (person markers) or both (tense).¹⁸ By determining the mood and modality of a sentence, AUX gives information as to how a proposition relates to the world; the deictic features marked in AUX point to things in the world that the sentence is used to refer to. In this sense, AUX may be said to specify the relationship between some proposition and reality.

What I would like to emphasize here is the fact that there is no simple correspondence between lexical and syntactic categories. Across languages, AUX has a small mixed inventory with respect to morphological class membership. In most languages, there is a clear association between the morphological class <u>noun</u> and the syntactic category SUBJECT, and between the morphological class <u>verb</u> and the syntactic category PREDICATE; but this association is not a unique correspondence. Lexical categories are constituents of syntactic categories, not of sentences. We could term the syntactic categories identified here <u>functional</u> categories, as long as we do not lose sight of the fact that they are defined by the syntax of the language, in addition to playing specified roles in the functional structure of sentences. Syntactic categories have two important properties; they correspond to parts of sentences to which the rules of the syntax refer, and they correspond to certain elements (functions or arguments) in the logical form assigned to that sentence. These two properties of syntactic categories are independent of one another, and mark an interface between syntax and semantics. In defining syntactic categories, and in matching syntactic categories across languages, we make reference to both of these properties. The syntax tells us what are the sentential constituents; given these constituents, we match syntactic categories across languages on the basis of correspondences in functional properties.

I have assumed here, but not defined, the morphological classes and lexical categories of Egyptian Arabic and English. Cross language comparisons of morphological classes and lexical categories must be based on a procedure similar to that employed here in matching syntactic categories across languages. The derivational/inflectional component of a grammar selects the morphological classes of a particular language. Some language independent definition of a verb -- perhaps as a morphological class where transitive functions are marked -- would be required to compare verbs across languages. In both the languages under consideration here, there is more than one morphological class where transitive functions are marked. (English: verbs and participles; Egyptian Arabic: verbs, participles, nouns of volition, prepositions in possessive sentences.) In such cases, we must choose where to apply the label verb. In English, we apply the label verb to that morphological class where transitive functions are marked and where there is agreement with the subject. In Equptian Arabic, we apply the label verb to those paradigms where transitive functions are marked, person subject is marked, and where suffixes marking person object may attach. (Non-verbal transitive predicates do not occur with these affixes.) Such a procedure enables us to decide which morphological classes we will consider equivalent across languages.

There has been considerable controversy over the status of AUX as an available syntactic category across languages; see Steele, <u>et al</u> (1981). Some arguments against recognizing AUX as a category have run as follows: the elements that have been assigned to AUX in some languages are in reality either verbs or clitics, etc.; therefore, they cannot be AUX constituents. Such arguments are specious because syntactic categories are not projections of lexical categories, but are sentential constituents whose inventories cut across or include various lexical categories.

Sentence mood is the critical factor in the use of sentences to perform particular speech acts -- claiming, asking, ordering, requesting. SUBJECT is linked to referring, PREDICATE to predication; AUX most directly to the marking of sentence mood. These uses of language are universal, as is the decomposition of sentences into constituents where functional properties are given values. Lexical categories are language particular; and although, for a particular language, there are associations between lexical categories and sentential constituents, this correspondence is not exact; and most importantly, this association varies across languages. But syntactic categories may be related across languages in terms of the universal pragmatic properties of sentences. I have shown that the set of syntactic categories proposed here are adequate for the analysis of sentence types in English and Egyptian Arabic, and that they are useful in comparing these languages, which are both configurational in type (Hale (1981, 1982)). The set of categories given here would obviously need modification for some languages, for example ergative languages, and languages with an OBJECT category not dominated by the PREDICATE. These categories would not apply to a non-configurational language such as Walbiri. The distinguishing feature of non-configurational languages may be that they lack a SUBJECT syntactic category. For an analysis of a non-configurational language without a SUBJECT category, see Demers and Jelinek (1982).

SUBJECT is not a required constituent in some or all sentence types in some languages, because the subject may be marked in the PREDICATE or in AUX. AUX may not be a required constituent in some or all sentence types in certain languages. Languages may lack a SUBJECT or AUX category; these are typological dimensions along which languages may vary. This is not the case with PREDICATE. All sentences necessarily have some overt constituent where the predicational function that lexically determines the function/argument structure of the sentence is marked, and no other functional attribute is necessarily marked in some overt constituent that is independent of the predicate. In this sense, PREDI-CATE is a universal syntactic category, and the only one that is universal.

FOOTNOTES

¹I thank Adrian Akmajian, Susan Steele and Dick Oehrle for very helpful criticisms of earlier versions of this manuscript. The remaining errors are mine. The framework employed here for comparing syntactic categories across languages was developed by Steele in Steele, <u>et al</u> (1981). I am grateful to the Wenner-Gren Foundation and National Science Foundation for support during my research on Egyptian Arabic. I benefited from consultations with Prof. Adel S. Gamal on Egyptian Arabic, and from comments on the Arabic analysis by Trandil El Rakhawy and Nagwa Younes.

²Proper names do not belong to the morphological class <u>noun</u> in English and Egyptian Arabic. Plurals are different:

i. The Childs (Julia and Paul) have invited us to dinner.

- ii. *The Children (Julia and Paul) have invited us to dinner.
- iii. He plays with the Toronto Maple Leafs.
- iv. *He plays with the Toronto Maple Leaves.

In Egyptian Arabic, a family name such as <u>nabiil</u> 'Noble' cannot be used in the plural, <u>nubilaa'</u>, when referring to more than one person of that name. An expression such as <u>9eelit</u> <u>nabiil</u> 'the family <u>nabiil'</u> is used instead.

By cited material, I mean material in quotes:

v. "Abracadabra" is nonsense.

³Using Church's lamba calculus, whereby the subject/predicate relation is more directly exhibited in a logical representation (see for example Sag, 1977), the contrast between a simple transitive sentence and related derived intransitive sentences would be equally clear, as the following will roughly show:

i. John gave a book to Mary.

John, $\lambda \times$ (x gave a book to Mary)

ii. Mary was given a book by John.

Mary, λx (John gave a book to x)

The landa calculus succeeds in bringing the representation of logical form closer to that of syntactic structure, and therefore demonstrates more perspicuously the relationship between them.

⁴Woisetschlaeger (1976) presents an interesting treatment of aspect as a modality. See Jelinek (1981) for an analysis of modal constructions in English and Egyptian Arabic.

⁵The linguistic term <u>embed</u> is used here since it is more convenient than speaking of functions taking certain material under their scope.

⁶The definitions of syntactic categories proposed here apply to the constituents of main clauses, and thus exclude complementizers, which are constituents of embedded clauses. Complementizers also may represent a sentential constituent where sentential functions are marked. Some embedded clauses have both AUX and COMP.

Those dialects of English which allow <u>Have you any wool?</u> would require an analysis that differed from the one given here, which assumes the American usage <u>Do you have any wool?</u> In an earlier stage in the development of English, all finite verbs could appear in sentence initial position in <u>yes/no</u> questions. In modern English, only AUX verbs appear in this position, and possessive <u>have</u> has been preserved in this position in certain dialects.

⁷Elements that mark person-subject do not mark sentential functions, but rather the arguments of predicational functions, and therefore are non-definitional features of AUX across languages. Person-subject is not marked in AUX in some languages, for example, Hopi (see Steele, <u>et al</u>, 1981) or Dyirbal (see Jelinek, 1981).

⁸The transcription of Egyptian Arabic employed here is standard. \underline{t} and \underline{d} represent pharyngealized stops; <u>H</u> is the voiceless pharyngeal fricative, and 9 is the voiced pharyngeal fricative.

⁹Some Egyptians say <u>maHaykuns</u>; most educated Cairenes prefer <u>mis</u> <u>Haykuun</u>, where unattached NEG precedes the inflection of <u>kaan</u>. The point here is that the NEG particle attaches to inflections of the copula. ¹⁰Comrie (1976) provides a valuable analysis of these varieties of aspect across languages.

¹¹Some Active Participles mark imperfective aspect, and others mark perfective aspect. All Passive Participles mark perfective aspect. See Jelinek, 1981; in press.

¹²<u>biddaha</u> in (37) is a transitive predicator, not the subject of the sentence. Compare:

- i. kaanit biddaha tifaaHa
 was wish-her apple
 (3 fs)
 She wanted an apple.
- ii. giddaha kaan Hakiim grand- was doctor father-her (3 ms) Her grandfather was a doctor.

NEG-attachment for PREP varies with their function:

- iii. 9andina ik-kitaab with-us the-book We have the book.
 - iv. ma9andinaas ik-kitaab NEG-with-us the-book We don't have the book.
 - v. il-walad mis 9andina the-boy NEG with-us The boy isn't with us (at our house).

NEG does not attach in (v) since the PREP P predicate does not mark person subject. The PREP 9 and in (iii, iv) functions as a PRED₂: in (v) as a PRED₁.

¹³Conditional sentences in both English and Egyptian Arabic employ past tense as irrealis. Compare:

i. If somebody gave me a million dollars....

ii. suuf iza kaan fil-beet see if was at-home See if he's at home.

See Steele (1975) for a discussion of past tense and <u>irrealis</u> across languages.

¹⁴Emphasis and negation in the AUX node in imperative sentences in English are both sentential functions, and would pose no problem for the definition of AUX proposed here.

¹⁵See Hale (1973) for an analysis of AUX in Walbiri.

¹⁶Steele's language independent definition of AUX requires that the inventory of AUX be listed, but this inventory is then <u>characterized</u> as fixed or small.

¹⁷I have not included here certain non-definitional properties of AUX identified by Steele, such as her important findings on the restrictions on the position of AUX in sentences across languages.

¹⁸Partee (1973) points out some interesting structural analogies between tenses and pronouns in English.

REFERENCES

- Akmajian, Adrian, Susan Steele and Thomas Wasow (1979). "The Category AUX in Universal Grammar", Linguistic Inquiry, 10:1-64.
- Bresnan, Joan (1980). "The Passive in Lexical Theory", Occasional Paper #7, Center for Cognitive Science, MIT.
- Comrie, Bernard (1976). Aspect. Cambridge University Press.
- Demers, Richard A. and Eloise Jelinek (1982). "The Adjoined Clause in Lummi" (ms.)
- Hale, Ken (1973). "Subject Marking in Walbiri" in S.R. Anderson and P. Kiparsky, eds., <u>A Festschrift for Morris Halle</u>, Holt, Rinehart and Winston, New York.

(1981). "On the Position of Walbiri in a Typology of the Base." Indiana University Linguistic Club, Bloomington, Indiana.

(1982). "Preliminary Remarks on Configurationality". MIT ms.

- Jackendoff, R. (1977). X Syntax. Linguistic Inquiry Monograph 2. MIT Press, Cambridge, Massachusetts.
- Jelinek, Eloise (1981). On Defining Categories: AUX and PREDICATE in Colloquial Egyptian Arabic. University of Arizona dissertation.

(in press). "Person-Subject Marking in AUX in Egyptian Arabic." to appear in Linguistic Categories: Auxiliaries and Related Puzzles, Frank Heny and Barry Richards, eds., Reidel.

- Kahn, Charles H. (1973). <u>The Verb 'be' in Ancient Greek</u>. (Part 6 of J.W.M. Verhaar, ed., <u>The Verb 'be' and its Synonyms</u> = Foundations of Language Supplementary Series, 16. Dordrecht: Reidel.
- Kaplan, Ronald M., and Joan W. Bresnan (1980). "Lexical-Functional Grammar: A Formal System for Grammatical Representation." Occasional Paper #13, Center for Cognitive Science, MIT.

- Kinkade, M. Dale (1978). "Salish Evidence Against the Universality of 'Noun' and 'Verb'." Ms.
- Partee, Barbara (1973). "Some Structural Analogies Between Tenses and Pronouns in English", <u>The Journal of Philosophy</u>, Vol. LXX, No. 18, 601-609.
- Sag, Ivan (1977). <u>Deletion and Logical Form</u>. Indiana University Linguistics Club, Bloomington.
- Steele, Susan (1975). "Past and Irrealis: Just What Does It All Mean?", IJAL 41:200-217.
- Steele, Susan, with Adrian Akmajian, Richard Demers, Eloise Jelinek, Chisato Kitagawa, Richard Oehrle, and Thomas Wasow (1981). <u>An En-</u> cyclopedia of AUX: A Study in Cross-Linguistic Equivalence. Linguistic Inquiry Monograph 5. MIT Press, Cambridge, Massachusetts.
- Thompson, Laurence C. and M. Terry Thompson (1980). "Thompson Salish //-xi//", IJAL 46:27-32.

(Ms.) The Thompson Language. (To appear)

- Wittgenstein, L. (1953). <u>Philosophical Investigations</u>. Blackwell, Oxford.
- Woisetschlaeger, Erich F. (1976). <u>A Semantic Theory of the English</u> <u>Auxiliary System</u>. Indiana University Linguistics Club, Bloomington.