

6. Functional categories: an evolutionary perspective

Pieter Muysken

Radboud University, Nijmegen

Abstract

In this chapter I will focus on the status and origin of functional categories in the languages of the world, and explore various ideas, with reference to proposals by researchers such as Bickerton and Jackendoff, concerning three possibilities: (a) that in the course of language evolution, language developed from a functional category-free to a functional category-rich state; (b) that language started out as a system consisting mostly of syntactic patterns and functional categories (albeit of a different nature), and only then acquired a content lexicon. Ultimately, I tentatively propose that neither scenario is the most likely one, but that we need to consider the possibility of co-evolution of the syntactic and the lexical subsystems or modules in language, with functional categories at the interface between the two.

6.1 Functional categories

Functional categories can take various shapes. Consider for instance, the Andean language Cuzco Quechua. In this Quechua variety, the original language of the Inka Empire, sentences generally have the form given in (1):

- (1) *Tusu-yu-spa* *taki-ya-spa-n* *asa-ru-n-ku* *kargo-yoq-kuna-qa.*
dance-AUG-SS sing-AUG-SS-AF pass-EXH-3-PL charge-POS-PL-TO
'The charge holders [patrons of the feast] pass dancing and singing.'
(Cusihuamán, 1976: 223)

In this example we find two kinds of elements: lexical roots (two of which, *kargo* and *pasa-*, are loans from Spanish) and affixes (in bold in (1)).

These two classes of elements differ on a number of dimensions:

Phonology and morphology

- (a) Roots often consist of two syllables, while affixes are generally monosyllabic;
- (b) In roots we sometimes find aspirated and glottalized consonants; these are absent in affixes.

Semantics and pragmatics

- (c) The meaning of the affixes is very often much more abstract than that of the roots;
- (d) Some of the affixes play a role in organizing the information structure of the clause, such as affirmative marker *-n* and the topic *-qa*.

Syntax

- (e) Many of the affixes play a role in the syntactic organization of the clause; thus *-spa* marks temporal subordination (here simultaneity) with identical subjects; *-n-ku* marks the subject-verb agreement relation.

Then there are a few affixes in (1), often occurring closer to the root, which have properties intermediate between those of affixes and roots. The elements *-yu-* (full form *-yku-*) and *-ru-* (full form *-rqu-*), glossed here as 'augmentative' and 'exhortative', but with much more complex meanings, as well as the possession marker *-yoq*, have somewhat more concrete meanings, are often lexicalized, and play a less central role in the organization of the clause. Their shape is also slightly more complex than that of other suffixes.

Thus we find a rather sharp division in Quechua between roots and affixes, with some of the affixes having root-like characteristics. These types of distinctions do not exist only in languages with a complex morphology, such as Cuzco Quechua, but also in languages with more simple morphologies, often termed "isolating".

Consider now an example from the Surinam creole Saramaccan, the language of one of the Maroon societies in the tropical inland forest:

- (2) *a bi tús wan sitónu gó a mi báka*
 3s PST throw one stone go LOC 1s back
 'He threw a stone at me (from behind)' (De Groot, 1981: 118).

Here all elements are separate words (although there is some morphology in Saramaccan as well). Again, considering the different dimensions, the elements in Saramaccan can be split up into more lexical and more functional ones.

Phonology and morphology

- (a) Functional elements in Saramaccan generally consist of only one syllable, while lexical elements often are longer;
- (b) The functional elements in this Saramaccan example do not carry specified high tone (marked with a $\acute{\text{}}$), but there are some elements that do (Norval Smith, p.c.).

Semantics and pragmatics

- (c) Again, we find more concrete meanings such as *sitónu* 'stone' and *túe* 'throw (away)', and more abstract meanings, such as *a* 'third singular pronoun', *bi* 'past tense', and *a* 'locative'.

Syntax

- (d) The functional elements, such as the person clitic *a* and the tense marker *bi*, play a role in organizing the clause syntactically.
- (e) In contrast with Cuzco Quechua, there is no overt marking of agreement.

We also find a few elements with intermediate status, such as the serial directional verb *gó* and location specifier *báka*. These may carry a high tone, have a somewhat concrete meaning, and in phrase structure have a status similar, although not identical, to that of lexical heads.

Not surprisingly, other languages, such as Dutch, may have a set of distinctions similar to the ones made in Cuzco Quechua and Saramaccan. Consider an example such as (3):

- (3) *Vergeef-0* ***me*** ***toch*** ***mijn*** *ont-nuchter-end-e*
 forgive-IM 1s.AC though 1s.GE dis-illusion-ing-AG
en *scept-isch-e* *toon.*
 and scept-ical-AG tone.
 'Please forgive me my disillusioning and sceptical tone.'

In this sentence again we find some functional elements (in bold), as well as lexical elements. In addition, there are some elements intermediate between the two, and heavily lexicalized, which in this case are affixal (prefix *ont-* 'dis-' and adjectivizing *-isch* 'ical').

Phonology and morphology

- (a) Many functional elements in Dutch are pronounced with a schwa (even though sometimes written differently, as in the case of *mijn* 'my');
- (b) Some morpho-syntactic categories in Dutch receive a null expression, such as imperative;
- (c) Some morpho-syntactic distinctions are suppletive, e.g. the *me* 'me'/*mijn* 'my' distinction;
- (d) Some functional categories appear to be separate words, and others appear as affixes.

Semantics and pragmatics

- (e) Some affixes have a purely formal function, such as adjective agreement *-e*, without a clear meaning;
- (f) There are elements with a purely discourse organizing functions, such as *toch* 'though'.

Syntax

- (g) Various elements play a role in signalling the grammatical organization of the clause, such as the agreement marker and the case forms of the first person pronouns.

Thus the lexical/functional distinction is valid for various languages, but with different morpho-lexical manifestations.

While Cuzco Quechua, Saramaccan, and Dutch all have reasonably large inventories of functional elements, in some languages there may be many fewer. This has been claimed for Riau Indonesian by David Gil (e.g. 2001). Consider his often cited sentence (4):

- (4) *ayam makan*
 chicken eat

This may mean 'the chicken eats', but also 'the chicken is making somebody eat', or 'somebody is eating where the chicken is', and a host of other things. In Riau Indonesian many aspects of sentence meaning are not expressed through formal markers or rigid structural patterns, but through general interpretive conventions, heavily relying on the context.

Still, Riau Indonesian has functional elements, such as the relator *yang* and tense markers such as past proximate *tadi*, but there are fewer of them than in, say, Cuzco Quechua, and they are relied on less frequently in actual language use. Parkvall (2005), using data from the World Atlas of Linguistics Structures (Haspelmath et al., 2005), has tried to calculate the number of

morpho-syntactic distinctions that different languages make, and concludes that there are large differences in these numbers. There have been claims that there are some languages (e.g. Old Chinese as it appears in poetic texts) without functional categories, but no such language appears in Parkvall's sample.

From this brief survey we can perhaps draw the tentative conclusion in (5):

- (5) a. All languages are characterized by a lexical-functional distinction in their inventory of elements;
- b. In most languages there are some elements intermediate in status on the lexical functional/dimension;
- c. There are various ways that the functional categories of a language may be realized: as affixes, as separate words, as null forms, etc.
- d. Languages may differ in the richness of the inventories of functional elements.

It should be noted that sign languages also have developed functional categories (cf. Muysken, in press), similar to those of spoken languages. This supports the idea that these categories are universal.

The question addressed in this paper is what could be the evolutionary basis for the lexical/functional distinction? Botha (2003: 11) puts the justified demand of 'ontological transparency' on work in this area: if we are discussing the paths of evolution of a certain element or property of language, we should know what that element or property is. This is not simple in the case of functional categories. So how can we define functional categories, given their different manifestations? I propose the following definition (borrowing from Anderson's definition of inflection in his paper *Where's morphology?* (1982)):

- (6) A functional category is an element in the inventory of items in a language that plays a clear role in the grammatical rules of the language.

This definition clearly locates functional categories at the interface of the lexicon and the grammar. What is the reason that human languages have developed this interface category? Before turning to this question, I will turn to further discussion of the question of whether the lexical / functional is sharp or gradual.

6.2 A gradualist perspective: category models and semantic maps

There is fairly widespread recognition, as mentioned in section 1, that among the functional categories not all elements are equal. Some adpositions are more

clearly 'functional' than others (compare French *de* 'of' to *dessus* 'on top', clitic pronouns show special behaviour when compared with strong pronouns (compare *le* '3SG.MS.OB' to *lui* 'him'), copulas (*be*) are more restricted than aspectual auxiliaries (*finish*), etc. Roughly, we get a number of chains converging, from more lexical categories, roughly as in figure 6.1, where four points of reference are taken. Departing from locational and directional adverbs, verbs marking boundaries, discourse markers, and descriptive nouns, several chains can be construed that converge on a class of definiteness and finiteness markers, the most grammaticalized categories perhaps in the nominal and clausal domains:

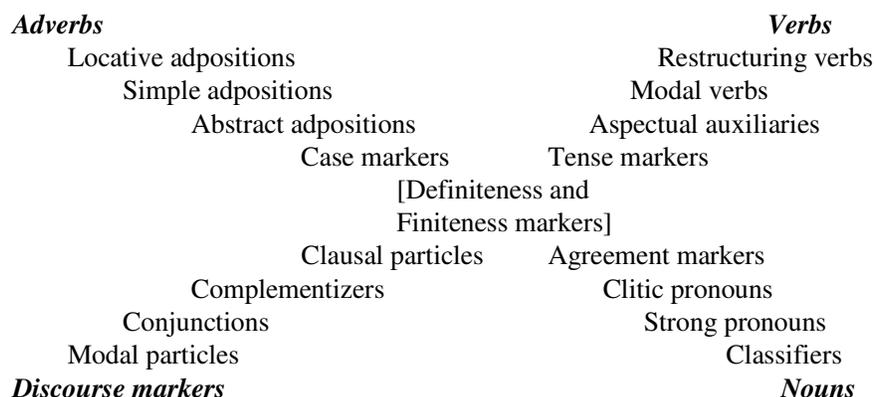


Figure 6.1: Chains of related categories, arranged from lexical to functional from the outside inwards

In order to properly deal with this array of notions expressed in categories, several models can be envisaged, which correspond to different approaches to grammatical categories. Currently there are at least four main models for categorisation (some of which, to be sure, have not been given very precise definitions yet, and may be better labelled 'views' or 'perspectives'):

- Prototype models
- Scale and Hierarchy models
- Mono-dimensional models
- Multi-dimensional models, including Multi-level and Chain models.

The **Prototype** model (e.g. Croft, 1991) assumes that each category has a typical meaning or use (e.g. nouns are typically used to refer), expressed by core members of the category, while other words may belong to a category without expressing this core meaning. A typical noun would be *table*, a less typical one *size*. Thus, one could envisage a proto-typical functional category such as 'the' at the centre of the definition (highly specific morpho-lexical properties, specialised syntax, reduced phonological shape, abstract meaning) and other elements more or less distant from this prototype. The prototype model does not seem to help us much in accounting for lexical/functional differentiation.

The **Scale** model (Ross, 1972; Sasse, 1981) likewise assumes that the boundaries of a category may be fuzzy, but makes the additional assumption that categories can be arranged on a linear scale. There is a large literature on gradience in grammatical categories (cf. the summary in Sasse, 2001), e.g. the adverb ... preposition cline or the noun ... verb cline. The **Hierarchy** model (cf. e.g. Comrie, 2001: 34, who makes this relevant distinction) is a scale model which has a high/low dimension. This asymmetry could be due to historical change, as postulated in grammaticalisation theory, to cognitive development (from simple to complex), to language evolution, etc. Thus adpositions could be on a scale with adverbs on the lexical end and case markers on the functional end. Modals could be on a scale with auxiliaries on the functional end and full verbs on the lexical end, etc. Clearly, figure 6.1 is an instance of this type of scale model.

The **Mono-dimensional** model (e.g. Baker, 2003) assumes that categories are not squishy and that they consist of one-to-one pairings of forms and meanings. Possible disparities between form and meaning are solved through special adjustment rules at either the syntax/phonology or the syntax/semantics interfaces. With respect to the issue at hand this model would assume that there is a true set of functional categories, and a number of other elements which might share features of functional categories but are really lexical in nature. The discussion then would be whether a certain class is 'truly' functional or not. This type of model would require an absolute boundary somewhere in the scales introduced in figure 6.1.

The **Multi-dimensional** model (Plank, 1984; Sadock, 1991; Jackendoff, 2002) assumes that categories lie at the interface of different representations – morpho-lexical, syntactic, phonological, and semantic. The **Multi-level** model as one instance (Cann, 2000: 58) would assume that functional categories can be distinguished, in absolute terms, at one level of analysis, in this case E-language, but not at another level, I-language. The **Chain** model assumes that various categories may be part of a chain of some kind, as in the T-chain

proposal (Guéron and Hoekstra, 1988), where the Verb, Tense, the Inflection, and the Complementizer nodes may be part of a syntactically coherent subsystem. Conceptually, it can be seen as a type of multi-dimensional model, since the chain represents only one dimension.

Independently of where one draws the line exactly in figure 6.1, it appears that there are at least a number of elements that fall under the loose definition of 'functional category', a class of elements which is relatively limited in size and does not readily admit new members. The elements typically have a fairly abstract meaning and serve to link different expressed notions, such as predicates and arguments. In (7) seven sets are distinguished, arranged in terms of their degree of integration into the clause.

(7) *Rough classification of functional categories*

- a. Interjections are not at all integrated into syntactic structure, and often have deviant morpho-phonological shapes;
- b. Discourse markers and adverbs, are only loosely integrated into the clause;
- c. Evidential markers, conjunctions and adpositions;
- d. Classifiers generally fit into more or less closed classes, but these are often not paradigmatically structured;
- e. Pronouns are generally tightly organised in paradigms, but often play a relatively independent role, and often do not have special morphology;
- f. Modals and auxiliaries are often closely linked to tense markers, but vary somewhat in their morpho-phonological features;
- g. Tense and agreement markers, case markers and determiners tend to have reduced phonological shapes and are tightly linked to the syntactic skeleton of the clause.

The classification in (7) could be made more or less fine-grained, following the chains in figure 6.1, to be sure. However, it is clear that we are not dealing with a unified phenomenon here. Rather, it seems that lexical elements are 'recruited', to a greater or lesser extent, into syntactic processing.

This syntactic processing is linked to some kind of semantic map, involving major syntactic categories and the various functional notions linked to these. An approximation of such a map is given in figure 6.2.

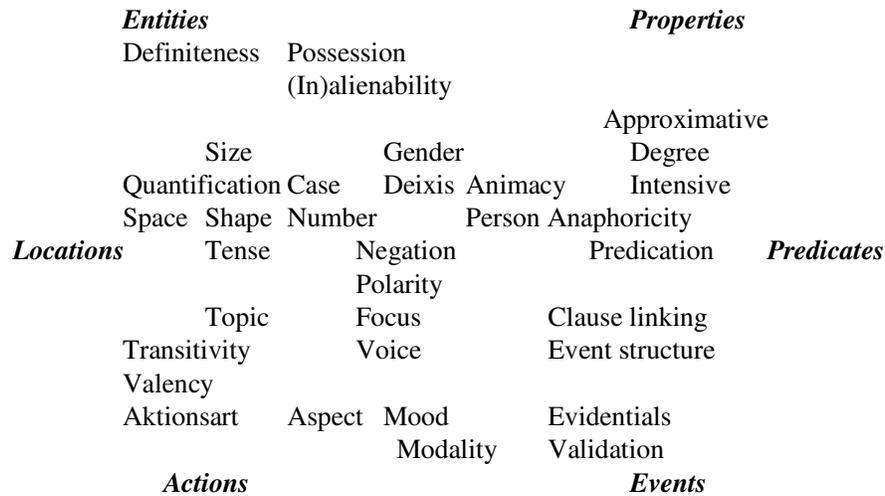


Figure 6.2: Rough semantic map of a number of notions likely to be expressed by functional categories

The map sketched in figure 6.2 involves six 'anchor points': Entities, Properties, Locations, Predicates, Actions, and Events, and is indicative of the richness of the concepts expressed in the functional domain.

The seemingly gradual nature of the classification of functional categories is the result of their being an interface category, and hence we must adopt a modular perspective to deal with this classification, taken up in the next section.

6.3 A modular perspective

Linguists can contribute to the study of language origins by considering the very nature of language itself. What is it about language that might inform us about its possible genesis? The most serious problem with most structural approaches to language, particularly those within the generative tradition (but already immanent in the Sausurean dictum that language is a system *où tout se tient*), is that they have tended to view the language capacity as a single monolithic whole. This monolithic view stands in the way of the gradualist perspective of language evolution that is required if we bring the origin of language in line with the genesis of other human capacities, both cognitive and more generally neurological, adopting a Darwinian perspective.

A more promising approach takes the modular organization of the language capacity as its starting point. In human language, (at least) four essentially different modules intersect: (a) the structure-building and -processing capacity (*syntax*), (b) the sign forming and using capacity (*semiotics*), (c) the capacity to engage in sustained exchange of information (*interaction*), and (d) the capacity to form complex representations of information (*cognition*) in our mind. This claim of four different interacting modules remains empty unless we manage to isolate the formal properties of these modules. I will briefly sketch those of the two of the modules involved that relate most directly the concerns of this paper: syntax and semiotics.

Two crucial features of language are part of the module of *syntax*. These are not found outside of language: *endocentricity* plays a role in sentence grammar (through X-bar theory), in word formation (headedness), and in phonology (e.g. in syllable structure). The property sometimes confusingly labeled displacement by Chomsky (e.g. 2000), I will term '*movement*' here (without any of the derivational claims often associated with this term): the fact that in language elements do not always appear in the place in the sequence where they are interpreted ([where] do you live < you live [what place]).

Berwick (1998) assumes that the following five properties are characteristic of the syntactic capacity: digital infinity (including recursion), displacement (see above), structure dependence, core grammatical relations (subject, object), locality constraints (e.g. subjacency). Ultimately, he assumes the property of Merge (Chomsky, 1995) to account for these properties, interfacing with pronounceability and interpretability requirements. However, it is not clear to me how endocentricity follows directly from Merge.

The module of *semiotics* contributes a number of properties to language; these principles or properties are however also found in non-linguistic semiotic systems (such as traffic signs). The first principle is that of *distinctiveness*: lexical elements must be sufficiently distinctive to contrast with other elements. A second principle is *transparency*: new lexical elements ideally are transparently derived from existing elements. A third principle, *elementarity*, refers to the requirement that a lexical element ideally functions as a coherent whole, as an atom which can be combined with other elements. This principle is often referred to as the lexical integrity principle in the case of language. Fourth, the principle of *analogy*, which causes new forms to be built parallel to already extant forms.

The organization of the lexicon as a semiotic system has several effects. A well-known semiotic effect is that of *blocking* in morphology: the availability of one form (e.g. *thief*), no matter how it arose, will block the acceptance of a possibly more productive other form (e.g. *steal-er*). A second

well-known effect is that of *suppletion*: within a tightly organized paradigm, a non-transparently derived or etymologically unrelated form may still be lexically related (e.g. *bad / worse*). Third, the principle of analogy produces lexical subsystems characterized by paradigmaticity.

Bearing distinctions such as the above in mind, I argue for a multi-dimensional, modular approach to the human language faculty, and subsequently, to grammatical categories, as the examples given in section 6.1 show. This approach implies that syntactic computation, interactive communication, sign building (semiotics), cognition, and perception/production are assumed to cooperate conjointly in what appears to be a single phenomenon: the human language capacity. These capacities include, as stated, syntactic computation, interactive communication, sign building (semiotics), cognition, and perception/production. The modular perspective implies that functional elements can and should be viewed as multi-dimensional. Not only do they have a form and a meaning (the traditional Saussurean notion of sign), but they may or may not play a separate role in syntactic computation (through their feature content); in addition, they may have an interactive function. This multi-dimensional, interface, character is also responsible for the fact that categories are perceived as gradient. Thus, we can perceive the distinction between lexical and functional as gradient due to the interactive character of the different modules. In these modules, the functional categories have different definitions.

1. Syntactically, functional categories are those elements that define the skeleton of the clause and its constituents.
2. Semantically, functional categories are those elements that have an abstract, non-ontologically defined meaning.
3. Phonologically, functional categories are those elements that do not form the nucleus of a phonological word.
4. Morpho-lexically, functional categories are closed class, paradigmatically structured elements.

Different dimensions of functional categories overlap but do not coincide: typically the semantic set of functional categories is a proper superset of the syntactic one, and typically the phonological set is a proper subset of the syntactic one. The morpho-lexical set and the syntactic set overlap only partially; etc. A modular view of functional categories opens the way for a gradualist view on language evolution, I will argue.

6.4 Functional categories from an evolutionary perspective

How can we explain the emergence of functional categories from an evolutionary perspective? We need to ask ourselves what the specific advantage of functional categories would be in human language. I will explore various ideas in this respect, with reference, of course, to proposals by researchers such as Bickerton (1990) and Jackendoff (2002). These proposals depart from the assumption that in the course of language evolution, language went from a functional category-free to a category-rich system, feeding upon the expressive possibilities of a rich lexicon. This I will term the *syntactic enrichment option*. It is also the option explored, albeit with a twist, in work on grammaticalization by Heine and Kuteva (2007).

An alternative would be that language started out as a system consisting mostly of structural patterns and functional categories (albeit of a different nature), and only then acquired a rich content lexicon. It is true that its extremely rich vocabulary is as much a feature setting human language apart from possibly antecedent animal communication systems as its rich syntax and inventory of functional categories. This could be termed the *lexical enrichment option*. However, if we take functional categories to be the original set of elements, it is unlikely that these are the continuation of the limited set of fairly concrete and generic alarm calls etc. of animal communication systems, given the highly abstract and category-bound nature of functional categories.

A third possibility would be to assume that the lexical and the syntactic possibilities of human language co-evolved: a rich lexicon allowed further development of functional categories, and a rich syntactic system allowed further possibilities for lexical enrichment. I will term this the *co-evolution option*.

Since the syntactic enrichment option has dominated the (admittedly syntactic-centric) theoretical literature this will occupy most of the following discussion. I will begin by sketching the preliminary assumptions in this direction of Bickerton (1990), and then those Jackendoff (2002: 236-7), building on Bickerton: in Jackendoff's view, language evolution involved a number of steps of syntactic enrichment towards a more adequate communicative system, and involved the evolution of language-specific cognitive abilities.

6.4.1 *From a functional element-free to a functional element-rich system*

The primary original proponent of the syntactic enrichment or 'functional expansion' scenario is Bickerton (1990), who proposed an early protolanguage, with little syntax, followed by a qualitatively very different stage with Full Language. Bickerton assumes that with *Homo erectus*, about one million years ago, protolanguage emerged, a system with a lexicon, but without syntax. Clausal organization would be in term of pragmatic principles first, then semantic principles, and finally syntactic principles. There would be flat structure rather than layered structure, and no agreement markers in stead of the complex agreement characterising some natural languages. Finally, there would be no recursion.

As to the lexicon, there would be only lexical categories, rather than lexical and functional categories as in full languages. There would not be the complex morphology of full languages. Also, protolanguage would show little diversity in categories, rather than the high diversity in categories of at least some full languages, and vocabulary would be poor.

As far as the pragmatics of information processing is concerned, interpretation would be context dependent rather than context independent, and processing would not yet be fully automatic, as with full languages.

As regards the phonology, protolanguage would have purely syllabic articulation, rather than the articulation based on more complex phoneme combinations. Phoneme inventories would be simple rather than the complex inventories of some contemporary languages, and likewise syllables would be simple. There would be no *lento/allegro* styles, presumably, rather than the complex *lento/allegro* and *sanddhi* rules of contemporary languages.

Bickerton adduces evidence (p. 180) that whatever vocabularies primates manage to acquire in human captivity only involve content words. However, Bickerton (1990: 181-5) assumes that certain functional pre-syntactic categories developed during the protolanguage period, since certain abstract notions are central to any basic communicative system. These would include:

(8) Negation

Wh-words

Possibly pronouns

Modal operators such as *can* and *must*

Aspectual operators to indicate completed or earlier actions, or later actions

Direction and location markers

Quantifiers such as *many* and *few*

For Bickerton, following his work on pidgin and creole genesis in Bickerton (1981), it is crucial to adopt a non-gradualist position: the transition between protolanguage and full language was radical and did not involve any intermediary stages. However, the evidence from pidgin / creole genesis for this position is weak, and in other domains, such as child language and aphasia adduced by Bickerton (1990: 105-129), the evidence for a sharp break is not very strong either, if existent at all.

Jackendoff (2002), building on Bickerton's work, assumes a much more complicated pattern of incremental development, involving a number of logically independent and yet in part logically subsequent separate steps. There is generally no independent evidence for these steps, but certainly they correspond to mostly separate phenomena. I will first present them and then turn to the question how these interact with the possible emergence of functional categories. The steps are (2002: 238-260):

- (9) a. Pre-existing primate conceptual structure
- b. Use of symbols in non-situation-specific fashion
- c. Use of an open, unlimited class of symbols
- d. Development of a phonological combinatorial system
- e. Concatenation of symbols
- f. Use of symbol position to convey basic semantic notions
- g. Hierarchical phrase structure
- h. Symbols that explicitly encode abstract semantic relations
- i. Grammatical categories
- j. System of inflections to convey semantic relations
- k. System of grammatical functions to convey semantic relations

After phase (f) [Use of symbol position to convey basic semantic notions], something like Bickerton's protolanguage may have been arrived at, and after the final phase, (k), we can speak of modern language. The sequence of different developmental steps can be portrayed as in figure 6.3.

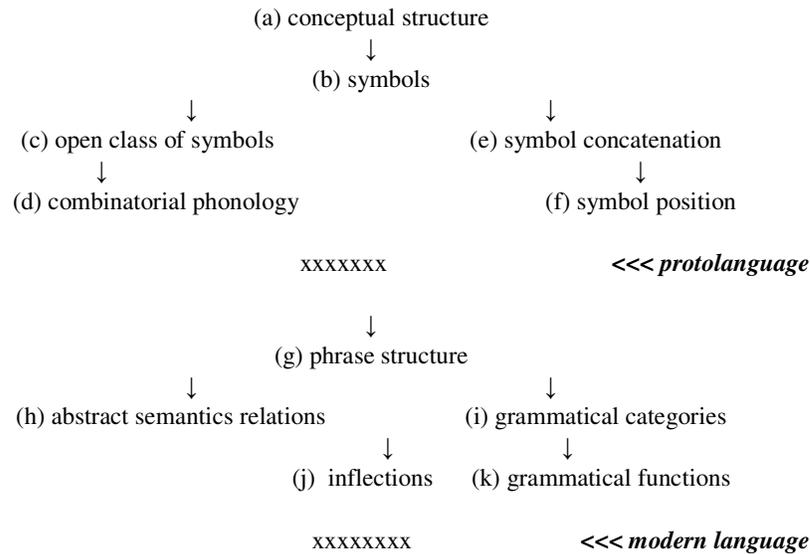


Figure 6.3: Jackendoff's (2002) scenario for the development of grammar

One type of functional categories is introduced in step (h), [Symbols that explicitly encode abstract semantic relations]. Here Jackendoff mentions notions similar to those suggested by Bickerton in (8), enumerating the following types of elements (pp. 253-254):

- (10)
- a. Spatial relation terms: *up, toward, behind*
 - b. Time terms: *now, yesterday, before*
 - c. Markers of illocutionary force or modality: *if, may, can*
 - d. Markers of discourse status: definite and indefinite determiners
 - e. Quantification: *some, all, always*
 - f. Purposes, reasons, intermediate causes: *for ... to, with, because*
 - g. Discourse connectors: *but, however, and so forth*

A second type of functional category is introduced in step (j), [System of inflections to convey semantic relations], referring to case markers, agreement markers, etc., elements whose interpretation is strictly grammatical rather than notional.

Thus Jackendoff splits the functional categories into two sets: elements with more semantic content, often realized as separate words, and more purely formal elements, often realised as inflections. This distinction between two sets

would also correspond to different stages in the process of evolution: content-full elements would evolve before more formal elements.

Pursuing a similar line of argumentation as Jackendoff, Johansson (2006) argues that four different grammatical properties can be distinguished in the syntax of human languages, which to some extent are implicationaly ordered:

- (11) a. Language is **structured**: the ordering and relations of the items in a clause has implications for the system;
- b. Language is **hierarchical**: within the structure there are several layers of ordering;
- c. Language is **recursive**: sometimes categories on one hierarchical layer in the structure re-appear on another one;
- d. Language is **flexible**: sometimes the order and arrangement of elements may alternate in different utterances.

The basic sine qua non property is structure: all other properties imply structure, but not vice-versa. Furthermore, recursivity implies both structure and hierarchy, but not vice-versa. Similarly, flexibility implies structure (and possibly also hierarchy), but not vice-versa. Flexibility and recursivity are logically independent of one another. On the basis of the distinct properties in (11), Johansson arrives at a scenario with the following possible stages in grammatical development:

- (12) one word sentences
- two word sentences (+ 11a)
- sentences with hierarchical structure (+11b)
- sentences with recursion (+11c)
- sentences with full modern syntax (+11d)

6.4.2 *Evolutionary advantages of functional categories*

Assume that there was an earlier stage in the evolution of language, something like Bickerton's or Jackendoff's protolanguage, in which there were no functional categories. Rather, this language only would have the equivalent of content words. What would be the advantage that the emergence of functional categories would confer upon a type of language system, which had them?

Surprisingly perhaps, this question was first posed, as far as I am aware, in an article by Labov written in 1971 and published in 1990: 'The adequacy of natural language I: The development of tense'. Here Labov contrasted the tense marking of the English-lexicon pidgin of Hawaii with the later creole. In the

pidgin temporal reference was indicated with adverbs like 'yesterday' and 'later', in the creole with particles like 'been' and 'go'. Claiming that there is no inherent semantic reason for 'been' rather than 'yesterday', Labov argues that a system with tense particles allows greater stylistic possibilities. Citing the English example of the huge range of variability accompanying the pronunciation of 'I am going to go', which can be reduced to [angnego], but with many intermediary forms. Labov argues that there are at least 27 ways of pronouncing this sequence, and this variability makes a language with a future marker of the 'be going to'-type more attractive than one only with adverbs like 'later'. The problem is of course that theoretically, adverbs might also show the same possibility of being shortened, yielding stylistic options. This is precisely what has happened with the Tok Pisin tense marker 'bye and bye', which is ultimately pronounced like [bə] (Sankoff and Laberge, 1973). However, the reduction of 'bye and bye' went in parallel with its shift to the preverbal position and its incorporation into the emerging Tense/Mood/Aspect particle system.

While Labov's answer focuses on the stylistic dimension, Bickerton (1990: 55) assumes that functional categories provide a cognitive advantage: 'They constitute, as it were, the coordinates of the linguistic map, a kind of topological grid whereby the positions of objects and events can be plotted relative to the observer and to one another.' Bickerton bases himself on work by Leonard Talmy in this respect, who contrasts (2001: 32-33) the open class system, which is engaged in conveying conceptual content, with the closed class system, which conveys conceptual structure. Thus we can think of the emergence of functional categories as the result of the emerging need to provide conceptual structure to messages as they became more complex. Even though the same meaning can be expressed with adverbs as with tense markers, there may be semantic advantages as well to more abstract meaning carrying elements. It may not always be relevant to indicate that something happened 'yesterday' or 'last week', and in such a case a more generic marker like 'before' would be preferable.

Other researchers have sought answers in the domain of formal syntactic patterning. One syntactic possibility, following the work by Hauser, Chomsky and Fitch (2002), was that functional elements emerged because of the development of recursion in the syntax: these elements would help making the links between the constituents visible in structures that were growing to be progressively more complex. The types of elements that could fulfil this task would primarily be complementizers, as in (13a), and adpositions or case markers, as in (13b):

- (13) a. I told you already **that** I had seen **that** Mary had left.
 b. my cousin's neighbour's cat

A second possibility was that functional categories emerged to facilitate efficient marking of head/complement dependencies. As syntactic structures became more complex, there was an increasing need to clearly delineate the relation between the predicate and its arguments, as well as that between a head noun and its dependents. Consider a sentence such as (14) from Cuzco Quechua:

- (14) Mariya-**man** Pedru-**q** wasi-**n-ta** riku-**chi-rqa-ni**
 Mary-DA Peter-GE house-3-AC see-CAU-PST-1
 'I showed Peter's house to Mary.'

In (14) there are agreement markers indicating subject (*-ni* '1st singular') and possessor (*-n* '3rd singular'), and case markers indicating indirect object (*-man* 'dative'), possessor (*-q* 'genitive'), and direct object (*-ta* 'accusative'). These elements generally allow unambiguous identification of the various interactants.

A third possibility is that the functional categories actually define the categorical status of the content words. Consider the words from Quechua in (15a):

- | | | | | | |
|---------|-------|---------|----|------------------|--------------------------------|
| (15) a. | wasi | 'house' | b. | wasi- yki | 'your house' |
| | rumi | 'stone' | | rumi- yki | 'your stone' |
| | riku- | 'see' | | riku- nki | 'you see' |
| | puri- | 'walk' | | puri- nki | 'you walk' |
| | puka | 'red' | | puka- yki | '*you(r) red/your red one' |
| | yana | 'black' | | yana- yki | '*you(r) black/your black one' |

While the lexical forms of Quechua nouns, verbs, and adjectives are non-distinct or very similar, the endings make clear what kind of element is used, as in (15b). Verbs receive forms like *-nki* for the second person, and nouns forms like *-yki*. Adjectives can only be marked with *-yki* when they denote a noun. Thus it is the functional categories that give the lexical ones their clear categorical status.

Thus various syntactic explanations may be given for why functional categories are useful in a slightly more complex grammatical system. However, these various explanations, and other imaginable ones, suffer from two deficiencies, at least.

(A) They each only cover certain sets of cases, not all functional categories. Roughly the following relations hold between these three functions and specific sets of categories:

- (16) a. *recursion marking* complementisers, adpositions, case markers
 b. *head/complement marking* agreement markers, adpositions, case markers
 c. *categorical distinctions* determiners, inflections, case markers

(B) Languages differ considerably in the extent to which these different syntactic functions are realised by functional categories. Everett (2005) claims that the Brazilian language Pirahã has only very limited sentential recursion, if any. Similarly, in some languages head/complement is indicated by strict word order, rather than with agreement and case marking. Finally, some languages show clear lexical distinctions between e.g. nouns and verbs, rather than marking these distinctions with functional categories.

Thus, the syntactic processing advantages of functional categories can not easily be stated in terms of a single aspect of syntax. Rather, it is syntactic processing overall that is facilitated by functional categories. Levelt (1999: 86) writes: 'Syntax develops as "the poor man's semantics" for the child to systematize the expression of semantic roles, just as phonology is "the poor man's phonetics", a lean system for keeping track of the subtle infinitude of articulatory patterns.' In the same vein, we may suggest functional categories are the poor man's lexicon – that part of the lexicon involved in morpho-syntactic patterning. Automaticity of processing is what is involved in the use of highly frequent markers, and automaticity of retrieval of functional elements from the mental lexicon.

6.4.3 *Claims of the linguistic fossil analysis*

Above, particularly in the work of Bickerton (1990), two or more parallel processing systems were postulated, available for language use: a more recently developed syntactic system A, and a fossil paratactic system B (reminiscent or a remnant of the protolanguage). The syntactic processing system A is assumed to use functional categories as part of the functional skeleton, is highly automatised, and has recursion. The primary syntactic structure building takes place through the selection of a specific complement by a specific head, like NP by D(eterminer), or VP by I(nflection).

The paratactic processing system B uses various principles for ordering such as iconicity and information structure, is only partially automatised, and has no recursion. There are no functional categories, and there is no functional

skeleton. The primary syntactic structure building takes place through adjunction.

Evidence for processing system A includes the lexical/functional asymmetries in insertional code-mixing, borrowing, Creole genesis, and mixed languages that I argued for in Muysken (2008). We may also refer to patterns of L1 development versus L2 development, and the results from speech error research.

Possible evidence for system B, cited by Jackendoff and Bickerton, includes the possibility for paratactic speech in Foreigner Talk and other restricted registers, the emergence of early pidgins, the Basic Variety that emerges in early L2 learning (Klein and Perdue, 1997), and the possibility of agrammatic speech by patients with aphasia. In Muysken (2000) I discuss the capacities of bilinguals to create structures in alternational code mixing and the acquisition of complex structures in early L1 through adjunction, as examples of this paratactic processing capacity.

While system B is assumed to be available as a fall-back system in the background, system A is dominant in actual language production for acquired languages, because it is more efficient, rapid, etc. It is tempting to assume that system B is anterior, in evolutionary terms, to system A, making pidgins etc. special windows on biologically earlier stages of human language.

However, the route by which B developed into A is not clear. A first scenario would be that automatised forms would lead to internal restructuring and internally generated structural development, while a second scenario would be that another system (phonology, motor control) was co-opted to provide the functional skeleton.

Also the evidence for a functional-category poor system is mainly indirect and of dubious value. Second language learners may fall back on non-grammatical routines, but where they can transfer functional categories from their L1 they will do so. The data on the second language development of Dutch possessives presented by Van de Craats (2000), concerning both Turkish and Moroccan Arabic learners, are a case in point. To be sure, Klein and Perdue (1997) do not claim that the Basic Variety has no structure or functional categories, but rather that the settings for the values of the functional categories are unmarked. Similarly, pidgins may show many more features of the original native languages of their speakers than is sometimes assumed, and hence contain evidence for the functional categories of these languages. Likewise, the evidence from agrammatic aphasia shows that speakers have to rely on non-syntactic strategies if their syntactic system has been impaired, rather than that such as a system is necessarily anterior to the other system.

Thus the actual evidence for the syntactic enrichment model is not as strong as sometimes made out to be.

6.4.4 *Lexical enrichment models*

The co-optation scenario, by which a non-linguistic system, e.g. the one involved in motor control, is co-opted to serve as the basis of a grammatical subsystem, could also be the basis for a lexical enrichment model. In this model there would be an essential discontinuity between the non-syntactic sign manipulation capacity of primates and human language, assumed to be a control system in its initial stages. However, it would be hard to imagine for such a system to work without a basic lexicon in place. It is also not clear that this lexicon would grow by itself, without the concomitant expansion of the syntactic system, which could become progressively more complex.

6.5 Discussion: towards a co-evolution model

The model proposed by Heine and Kuteva (2007) assumes gradual enfolding of the lexical system, with the emergence of more and more different categories. Thus the syntax and the lexicon would evolve together. A rich lexical system relies on a rich syntactic system as much as the other way around.

Lexical systems, as they become richer, start showing evidence of more and internal structure. This structure is of a different nature, however, from that of syntax. Paradigmatic relationships between items (e.g. analogy and contrastive pairing) start playing an important role, but so do syntagmatic patterns, as words get longer and more complex. Carstairs-McCarthy (2005: 183) reflects upon the evolutionary origin of morphology, following up on his earlier work in Carstairs-McCarthy (2000), and suggests that it involves 'a shift from a domain of grammar in which the syntagmatic dimension is dominant to one in which the paradigmatic dimension has at least equal importance.' In fact, some functional elements (e.g. the series *anybody, anyone, anywhere, anyway*) show complex internal structure, and some combinations of auxiliary elements may also be partly fixed (cf. the combination *ain't* or *didn't (you)*).

Functional categories emerged at the interface of the syntax and the lexicon, as these systems became more complex, and helped structure the interaction between these modules. As lexical forms become part of the syntactic processing system, they tend to become reduced and eventually end up as null, but then new forms come in to take their place. Like pebbles in a mountain stream, the lexical elements get polished, and eventually reduce to sand.

The differences between languages with respect to the extent that the morpho-syntactic categories receive lexical expression reflect their histories and typological make up (see also Kusters, 2003), in that more isolated languages tend to develop more complex morpho-syntactic systems.

Acknowledgements

I am grateful to the editorial team for detailed comments on a previous draft. Some of the material in this chapter is drawn from the first and last chapters of Muysken (2008).

References

- Anderson, S.A., 1982. Where's morphology? *Linguistic Inquiry* 13, 571-612.
- Baker, M.C., 2003. *Lexical categories: Verbs, nouns, and adjectives*. Cambridge University Press, Cambridge.
- Berwick, R.C., 1998. Language evolution and the Minimalist Program: the origins of syntax. In: Hurford et al. (Eds.), pp. 320-340.
- Bickerton, D., 1981. *The roots of language*. Karoma, Ann Arbor, Mich.
- Bickerton, D., 1990. *Language and species*. University of Chicago Press, Chicago.
- Botha, R.P., 2003. *Unraveling the evolution of language*. Elsevier, Amsterdam.
- Cangelosi, A., Smith, A.D.M., Smith, K. (Eds.), 2006. *The evolution of language*. World Scientific, Singapore.
- Cann, R., 2000. Functional versus lexical: A cognitive dichotomy. In: Borsley, R.D. (Ed.), *The nature and function of syntactic categories*. *Syntax and semantics* 31. Academic Press, New York, pp. 37-78.
- Carstairs-McCarthy, A., 2000. *The Origins of Complex Language: An Inquiry into the Evolutionary Beginnings of Sentences, Syllables, and Truth*. Oxford University Press, Oxford.
- Carstairs-McCarthy, A., 2005. The evolutionary origin of morphology. In: Tallerman (Ed.), pp. 166-184.
- Chomsky, N., 1995. *The Minimalist Program*. MIT Press, Cambridge, Mass.
- Chomsky, N., 2000. *New horizons in the study of language and mind*. Cambridge University Press, Cambridge.
- Comrie, B., 2001. Different views of language typology. In: Haspelmath et al. (Eds.), pp. 25-39.
- Craats, I. van de., 2000. *Conservation in the acquisition of possessive constructions. A study of second language acquisition by Turkish and Moroccan learners of Dutch*, Dissertation Tilburg University.

- Croft, W., 1991. *Syntactic categories and grammatical relations*. Chicago University Press, Chicago.
- Cusihuamán, A.G., 1976. *Gramática Quechua Cuzco-Collao*. Instituto de Estudios Peruanos, Lima.
- Everett, D., 2005. Cultural constraints on grammar and cognition in Pirahã: Another look at the design features of human language. *Current Anthropology* 46, 621-646.
- Gil, D., 2001. Creoles, complexity, and Riau Indonesian. *Linguistic Typology* 5, 325-371.
- Groot, A.H.P. de, 1981. *Woordregister Saramakaans-Nederlands*. Bureau Volkslectuur, Paramaribo, Surinam.
- Guéron, J., Hoekstra, T.A., 1988. T-chains and the constituent structure of auxiliaries. In: Cardinaletti, A., Cinque, G., Giusti, G. (Eds.), *Constituent structures*. *Annali di Ca' Foscari, Venice*, pp. 35-99.
- Haspelmath, M., Gil, D., Comrie, B., Bibiko, H.-J., Dryer, M.S., 2005. *The World Atlas of Linguistic Structures*. Oxford University Press, Oxford.
- Hauser, M.D., Chomsky, N., Fitch, W.T., 2002. The faculty of language: What is it, who has it, and how did it evolve? *Science*, 298, 1569-1579.
- Heine, B., Kuteva, T., 2007. *The Genesis of Grammar: A Reconstruction (Studies in the Evolution of Language)*. Oxford University Press, Oxford.
- Hurford, J.R., 2003. The language mosaic and its evolution. In: Christiansen, M.H., Kirby, S. (Eds.) *Language evolution*. Oxford University Press, Oxford.
- Hurford, J.R., Studdert-Kennedy, M., Knight, C. (Eds.), 1998. *Approaches to the evolution of language*. Cambridge University Press, Cambridge.
- Jackendoff, R.S., 2002. *Foundations of language*. Oxford University Press, Oxford.
- Johansson, S., 2005. *Origins of language: Constraints on hypotheses*. Benjamins, Amsterdam.
- Johansson, S., 2006. Working backwards from modern language to protogrammar. In: Cangelosi, A. et al. (Eds.), pp. 160-167.
- Klein, W., Perdue., 1997. The Basic Variety (or: Couldn't natural languages be much simpler?). *Second Language Research* 13, 301-347.
- Kusters, W., 2003. *Linguistic complexity: the influence of social change on verbal inflection*. PhD dissertation, Leiden University.
- Labov, W., 1990. [1971] The adequacy of natural language I: the development of tense. In: Singler, J.V. (Ed.), *Pidgin and Creole Tense-Mood-Aspect Systems*. Benjamins, Amsterdam, pp. 1-58.

- Levelt, W.J.M., 1999. Producing spoken language: a blueprint for the speaker. In: Brown, C.M., Hagoort, P. (Eds.), *The neurocognition of language*. Oxford University Press, Oxford, pp. 83-122.
- Muysken, P., 2000. *Bilingual speech: a typology of code-mixing*. Cambridge University Press, Cambridge.
- Muysken, P., 2008. *Functional categories*. Cambridge University Press, Cambridge.
- Parkvall, M., 2005. The simplicity of creoles in a typological perspective. Presented at the conference on Approaches to complexity in language. Helsinki, August 2005.
- Plank, F., 1984. 24 grundsätzliche Bemerkungen zur Wortarten-Frage. *Leuvense Bijdragen* 73, 489-520.
- Ross, J.R., 1972. Endstation Hauptwort: The category squish. In: Peranteau, P.M., Levi, J.N., Phares, G.C. (Eds.), *Papers of the eighth regional meeting of the Chicago Linguistics Society*, University of Chicago, pp. 316-328.
- Sadock, J.M., 1991. *Autolexical syntax: A theory of parallel grammatical representations*. Chicago University Press, Chicago.
- Sankoff, G., Laberge, S., 1973. On the acquisition of native speakers by a language. *Kivung* 6, 32-47.
- Sasse, H.J., 2001. Scales of nouniness and verbiness. In: Haspelmath et al. (Eds.), I, pp. 495-509.
- Smith, K., 2006. The Protolanguage debate: bridging the gap. In: Cangelosi, A. et al. (Eds.), pp. 315-322.
- Tallerman, M. (Ed.), 2005. *Language origins. Perspectives on evolution*. Oxford University Press, Oxford.
- Talmy, L., 2001. *Toward a cognitive semantics. I. Concept structuring systems*. MIT Press, Cambridge, Mass.