1 Introduction

In early transformational grammar (e.g. Chomsky 1957, 1965) functional categories such as Det(eterminer), Aux(iliary) had a rather peripheral function in the phrase structure system. This minor role was reflected in the terminology where functional categories were labeled “minor lexical categories” (see Jackendoff 1977: 32). In the past sixty years, however, functional categories have come to take a major role. I describe this ascent in three stages – integration, proliferation, and then expansion. First I show how functional categories became structurally equal to the “major” lexical categories such as N(ouns), V(erbs), and A(djectives). At this point, categories such as D and Infl(ection) (an updated version of Aux) become normalized, and new functional categories are added to this group such as Comp(lementizer), Num(ber), K(ase). Soon after this structural shift, the inventory of categories expanded first slowly and then, with the advent of Cinque (1999), explosively. Also, during this period where the structures of functional categories come to resemble those of lexical categories (formerly known as major categories), efforts are made to keep functional categories as a distinct class of category with specific properties. The current state of functional categories can be seen as the extreme end of a pendulum swing. Lexical categories themselves are being put under the microscope and, in some sense, they have become minor or perhaps nonexistent. In this chapter I give a brief description of minor categories, and then track the development of functional categories within X’-theory, their proliferation, and their distinct characteristics. Later I give a glimpse of the far end of the pendulum swing, followed by some concluding remarks.

2 Minor categories

Functional categories for decades had the role of the chorus in syntactic theory. They were important but relegated to the background – considered “minor categories” to be distinguished from “major categories” such as N(oun), V(erb), A(djective). In Aspects of the Theory of Syntax (Chomsky 1965), the structure of the clause places all of the inflectional material that appears between the subject and the verb phrase in one (sometimes branching) node dominated by S, as shown in the tree below.4
The integration, proliferation, and expansion of functional categories

(1) AUX in Chomsky (1965)

```
S
  NP
    N
      sincerity

Aux
  V
    frighten

VP
  Det
    the
  N
    boy
```

The material that appeared in AUX comprises a set of inflectional elements as characterized by the phrase structure rule in (2) below.

(2) Aux → Tense (M) (Aspect) (Chomsky 1965: 107)

In the discussion of the structure in (1), Chomsky divides the elements into category symbols (N, V, etc.) and formatives (the, boy, etc.). The formatives he further subdivides into lexical items (sincerity, boy, etc.) and grammatical items (Perfect, Progressive, the, etc.). Jackendoff, in X'-syntax: a Study of Phrase Structure (Jackendoff 1977), citing Chomsky (1970) as a forerunner, divides categories into two types through a system of features. He proposes that what distinguishes functional (minor) categories from lexical (major) categories is the ability of the latter to take a complement, a property that is represented by the feature [±Comp]. He begins the discussion by distinguishing Adjective and Preposition ([+Comp]) from Adverbial and Particle ([−Comp]) respectively, but continues the discussion to include modals (with respect to verbs), articles and quantifiers (with respect to nouns), and degree words (with respect to adjectives). He thereby creates a system where every lexical [+Comp] category is linked to a set of functional [−Comp] categories as shown in the following table.

(3) Jackendoff’s features (1977: 32)

<table>
<thead>
<tr>
<th></th>
<th>Subj</th>
<th>Obj</th>
<th>Comp</th>
<th>Det</th>
</tr>
</thead>
<tbody>
<tr>
<td>V</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td></td>
</tr>
<tr>
<td>M</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>P</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td></td>
</tr>
<tr>
<td>Prt</td>
<td>-</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>N</td>
<td>+</td>
<td>-</td>
<td>+</td>
<td></td>
</tr>
<tr>
<td>Art</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>+</td>
</tr>
<tr>
<td>Q</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>A</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td></td>
</tr>
<tr>
<td>Deg</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>+</td>
</tr>
<tr>
<td>Adv</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>

At this point, functional categories begin to be part of the larger phrase structural system in that they are part of the featural system that Jackendoff created. They are not, however,
part of the X’-theory of projection. The minor categories remain minor in the creation of trees, mainly appearing in specifier positions either as lexical items (e.g. have, en in (4)) sometimes with category labels (e.g. pres in (4)) or as maximal projections (X’’) with no internal complexity (e.g. Art in (5)).

(4) Jackendoff (1977: 40)

```
(4)  S
    /   
   N'''
     
John

     /     
   have  en
   pres

   V
   prove

   N''
    the theorem
```

(5) Jackendoff (1977: 59)

```
(5)  N'''
     
Art'''
    the

     /     
   N''
     
N'
   king

   P'''
   of England

   P'''
   from England
```

Jackendoff proposes that there are parallel systems within the projections of the lexical categories – N, V, A, and P. The minor categories that appear in the specifier positions of the four lexical categories are seen to share properties across the systems. Not only do the complex specifiers such as N’’ subjects in the verbal system and N’’ possessors in the nominal system share characteristics, but so do specifiers such as Aux, Det and Degree across the three systems of V, N, and A. He acknowledges, however, that specifiers (as opposed to complements) are difficult to study as a class given their relatively small numbers, their idiosyncrasies, and their skewed distribution across the different projections. Many of the themes set up in Jackendoff are taken up in later work on functional categories, as we will see in subsequent sections.

3 Stage I: Equal but distinct

The first important step in the development of a theory of functional categories involves giving them a place in the phrase structure system. It is shown that they behave mechanically
The integration, proliferation, and expansion of functional categories

much the same way as lexical categories such as nouns, verbs, and adjectives – projecting structure, taking complements and specifiers, acting as launching and landing sites of movement. However, they still have a distinct function – not introducing arguments but rather contributing particular feature based semantic content. This section is an overview of the regularization of the structure of functional categories, the establishment of their (distinct) properties, and the additions to their inventory.

3.1 Functional categories and X'-theory

In Stowell’s important PhD thesis, *Origins of Phrase Structure* (Stowell 1981), minor categories join the family of categories structurally. In the earlier versions of X'-theory (Chomsky 1970; Jackendoff 1977), only the major (lexical) categories can be the head of a phrase. The head V projects to S, S projects to S' (Bresnan 1976). Aux is a specifier of VP, Comp is specifier of S', Determiner is a specifier of DP. Stowell (1981: 68), however, proposes that Aux (now labeled Infl) is the head of S, making S now I'' (or IP). He further proposes (Stowell 1981: 388ff.) that Comp is the head of S' (CP). These two proposals change the sentential phrase structure quite dramatically, as can be seen in (6) and (7) below. In (6) the functional categories are truly minor but in (7) the functional categories C and Infl behave like their lexical counterparts, projecting to a phrasal level along the spine of the tree.

(6) Minor categories

```
S'
  |    |
  Comp S
     |    |
    NP Aux VP
         |   |
          VN P
```

(7) Functional categories project

```
CP(S')
  |    |
  Spec C'(S)
     |    |
    (Comp) C IP
         |   |
          NP I'
               |   |
                (Aux) Infl VP
                     |   |
                      V NP
```

It is important to note at this point the type of argumentation that Stowell uses to support his claim that, for example, Comp is a head. He observes that verbs specify what type of
clausal complement they require, (i.e. ± WH). If one assumes that this sort of selection must be local (i.e. a head may only specify the properties of its complement or the head of this complement), both selecting elements and selected elements will be identified as projecting heads. If a verb can specify whether it selects a [–WH] clausal complement (with that or for) or a [+WH] complement (with whether or if), then these lexical items must head the complement of the verb.

(8) a. The children believe that/*whether it will snow.
   b. The children prefer for it to snow.
   c. The children wonder whether/*that it will snow.

The same sort of argument can be used to support Stowell’s proposal for Infl as the head of IP (S). The complementizer that selects for a +finite complement, while the complementizer for selects for a [–finite] complement, as the examples above show. This suggests not only that the selectors that and for are heads but also that the items that are being selected, [+finite] Infl or [–finite] Infl, are also heads.

Stowell’s thesis, by integrating functional categories into the formal phrase structure system, sets the stage for serious research on these categories. This research takes a variety of directions that are discussed below.

3.2 The nominal system

Abney’s thesis, *The English Noun Phrase in its Sentential Aspect* (Abney 1987), represents an important further step in the development of functional categories. Just as Stowell argues that Aux is not the specifier of VP but is the head of its own projection along the clausal spine, Abney, extending proposals of Brame (1981, 1982), argues that Determiner is not the specifier of NP but is a head of its own projection along the nominal spine.

Abney shows that Det is a selector (the way that we saw above that Comp is a selector). Determiners, like verbs, can take a complement obligatorily (*The children wore *(costumes).) or take a complement optionally (*The children sang (a song).). We can see below that the Det the must have a complement while the Det that optionally takes a complement.

(9) a. The *(child) was tired.
   b. That (song) amused the children.

A further test for the head status of the determiner comes from head movement. If it can be assumed that heads may only move into the heads that select them (see the Head Movement Constraint of Travis 1984: 131), then evidence of head movement into a position or from a position can be used to argue that the landing site or the launching site of the movement, respectively, is a head. Abney uses the following paradigm to argue that Det is a head. First we see in (10a) and (10b) that bare adjectives generally cannot follow the N head in English. However, with forms like someone (in (10c)) or everything (in (10d)), these adjectives can appear in the final position of the phrase (data from Abney 1987: 287).

(10) a. a (clever) man (*clever)
   b. a (good) person (*good)
   c. someone clever
   d. everything good
According to Abney, *one* and *thing* are generated in N and undergo head movement from N to D, resulting in this otherwise unexpected word order.

(11) Head movement to D

![Diagram of head movement to D]

Abney’s work is important in the development of functional categories not only because he proposes a nominal structure which contains a functional head Det that parallels the functional head Infl in the verbal domain, but also because he outlines a set of criteria that distinguish functional elements from lexical elements (what he terms “thematic elements”). He argues that functional categories form a natural class with the following characteristics (adapted from Abney 1987: 54–68).

(12) Properties of functional categories

a. Functional elements f(functionally)-select their complement.
b. Functional categories select a unique element.
c. Functional elements are a closed class.
d. Functional elements are morphologically weaker than lexical elements (often dependent, affixes, clitics, and sometimes null).
e. Functional elements are generally not separable from their complement.
f. Functional elements lack “descriptive content”, contributing to the interpretation of their complement often through grammatical or relational features.

At this point in the history of functional categories, they have been incorporated into phrase structure and X’-theory, heading projections along the phrase structure spine in the same way as (major) lexical categories. However, they are still recognized as having distinct characteristics. For Abney, the crucial distinction is that functional categories do not take arguments. In terms of Government Binding Theory (Chomsky 1981), this means that lexical categories can assign theta-roles while functional categories cannot. The functional categories f(functionally)-select their (single) complements rather than setting up a structure that allows theta-assignment to take place.

### 3.3 Parallel structures and extended projections

Abney’s thesis begins to outline the grammatical contribution that functional categories make to syntactic structure. Further work develops and refines this. Ken Hale in class lectures in the 1980s presented a view of phrase structure where Ns and Vs set up parallel systems with parallel domains, in a way reminiscent of Jackendoff’s earlier work discussed in Section 2. Both projections contain a lexical domain (VP and NP) where lexical selection
and theta-assignment occur. Above this domain is an inflectional domain (Inf and Det) which gives reference to the descriptive content of the lexical projection (event or item) and locates it in time and space. The structural layer above this inflectional domain has a purely grammatical function – to provide a formal connection to the rest of the utterance. This layer contains the Comp in the verbal domain and Case (K) in the nominal domain.\(^\text{11}\)

Grimshaw (1991, 2000) also outlines a general view of the phrase structure architecture, introducing the notion of extended projections. As in Hale’s work, she describes the general pattern whereby a lexical head is topped by a complex functional category shell.\(^\text{12}\) As Abney points out, in some sense the functional categories that make up this shell pass on the descriptive content of the most deeply embedded complement. Grimshaw captures this continuity by having one feature shared from the bottom of this projection to the top. In the tree in (13) we see that the V, the I, and the C all share the categorial feature [verbal]. Where they vary is in the level of the F feature. The verbal head that asymmetrically c-commands the other verbal heads in the tree is labeled F2, while the sole verbal head that selects no other verbal head is labeled F0. An extended projection is a projection that shares a categorial feature. Grimshaw labels the more traditional notion of projection a perfect projection. Perfect projections share not only categorial features but also have the same \{F\} value. CP is the perfect projection of C and the extended projection of V.

\[13\] Extended projections

\[
\begin{align*}
\text{CP} & & \{\text{verbal}\}{\{F_2\}} \\
\text{C} & & \{\text{verbal}\}{\{F_2\}} \\
\text{IP} & \{\text{verbal}\}{\{F_1\}} \\
\text{I} & \{\text{verbal}\}{\{F_1\}} \\
\text{VP} & \{\text{verbal}\}{\{F_0\}} \\
\text{V} & \{\text{verbal}\}{\{F_0\}} \\
\text{DP} & \{\text{nominal}\}{\{F_1\}}
\end{align*}
\]

Note that an l-head (lexical head) will never be part of the extended projection of its complement since it will either be selecting a projection with a different categorial feature, or it will select a complement with a higher \{F\} value, as would be the case if V, \{F_0\}, selects CP, \{F_2\}, or both as is the case in (13).

The importance of Grimshaw’s contribution is that she can capture why, for some processes, C acts as the head of CP and sometimes a lower head appears to be visible to processes outside of CP. As an example, we can look at selection. While we have used the locality of selection as an argument for the head status of C in Section 3.1, we can see in (14) below that selection sometimes appears not to be local. Grimshaw calls this semantic selection and posits that the N is being selected.\(^\text{13}\)
If features are allowed to percolate through an extended projection, this apparent counter-example to selection and locality can be accounted for.

3.4 Functional categories as a distinct class

We have seen in the previous section that the functional shells that dominate the projection of the lexical category are different in their content and selectional properties. Now we will look at ways that these differences affect other parts of the grammatical model – in particular, how the distinction between functional categories and lexical categories affects projection and movement, and how this distinction interacts with parameterization.

3.4.1 Projection: Fukui (1986)

Fukui (1986) introduces the term functional category and presents one of the first views of functional categories that treats them as a natural class of category, distinguished from lexical categories in principled ways. According to Fukui, the members of the class of functional categories are similar to lexical categories in that they head a projection that appears along the spine of the tree. He argues, however, that the exact mechanism of this projection differs from that of lexical categories. Among the list of differences articulated in Abney (1987), Fukui proposes that only functional categories can project true specifier positions. He argues that positions within lexical categories are determined by the argument structure of the head (the theta-grid or lexical conceptual structure) and that while this might include external arguments (the VP-internal subject hypothesis of Fukui and Speas 1986, among others), the generation of these positions depends on semantic rather than syntactic considerations. One way of thinking of this in current terms is that the “specifier” position of a lexical head is always created through EXTERNAL MERGE (see Chomsky 2004), that is, through base generation of an element in this position. In contrast, specifiers of functional categories, according to Fukui, were always filled by a moved constituent (INTERNAL MERGE of Chomsky 2004). This distinction foreshadows a distinction made in the Minimalist Program, which will be discussed in Section 5.1, where functional categories consist of formal features, including features that trigger movement.

3.4.2 Movement: Li (1990)

We have seen in Section 3.2 above how head movement has been used to argue that a particular element is a head along the spinal projection of the structure. However, not all head movements are possible. We look at one case of impossible head movement here that highlights a distinction between lexical and functional categories. In Li (1990), a parallel is drawn between head movement and XP movement, and a crucial distinction is made between functional and lexical categories. He investigates the structure of morphologically complex verbal formation using proposals of Baker (1988). In Baker’s analysis of productive verb incorporation structures, the higher verb selects for a sentential complement, CP. For example, a typical causative structure with the relevant movement is shown below. The structure and movement for the data in (15) would be as in (16) (Swahili example taken from Li 1990: 399 and credited to Vitale 1981).
(15) Musa a-li-m-pik-ish-a  mke wake chakula
Musa he-past-her-cook-cause-ind  wife  his  food
“Musa made his wife cook some food.”

(16) Head movement in causatives 1

Li notes, however, that in spite of the movement of the lower V through two functional categories, I and C, material typically found in I and in C is not found within morphologically complex verbal structures. He argues that morphologically complex causatives, in fact, have a simpler structure where the causative V directly selects a VP. The structure for (15), then, should be as shown in (17) rather than as we have seen in (16).

(17) Head movement in causatives 2

Li nevertheless acknowledges that structures of the type in (16) do exist. The example that he gives is presented below where it is clear that the embedded complement selected by the higher verb contains inflectional material and a complementizer (Swahili example taken from Li 1990: 400 and credited to Vitale 1981).

(18) Na-ju-a kama Hamisi a-na-ogop-a giza.
I-know-ind that Hamisi he-pres-fear-ind darkness
“I know that Hamisi is afraid of the dark.”

Li claims, however, that the sort of structure given in (18) would never allow head movement. Basically, movement of a lexical category through a functional category back to a lexical category is ruled out for principled reasons – in particular, this movement would violate Binding Theory. He characterizes functional categories as being the head equivalent of A'-positions and lexical categories being the head equivalent of A-positions. Movement from a lexical category to a functional category to a lexical category would be similar to movement from an A-position to an A'-position back to an A-position, improper movement, constituting a Principle C violation.
The integration, proliferation, and expansion of functional categories

An example of improper movement of an XP is given below.

(19) *John seems that it is [VP t’ [VP considered [ t to be intelligent ]]]

*John has undergone movement from an A-position (the subject position of the most deeply embedded clause) to an A’-position (adjoined to VP) to an A-position (the subject position of the matrix clause). This sort of movement produces an ungrammatical string and one way of accounting for the ungrammaticality is through Binding Theory. An empty category that is A’-bound (e.g. the trace in the lowest subject position) is an R-expression (Chomsky 1981) and R-expressions must be A-free. In this construction, however, this R-expression will be A-bound by *John in the matrix Spec, TP incurring a Principle C violation.

If lexical heads are similar to A-positions and functional heads similar to A’-positions, we can now see why movement of the sort shown in (16) would create the same violation as that in (19). Just as the trace in the embedded subject position in (19) is locally A’-bound, making it a variable, the trace in the embedded V in (16) is locally A’-bound (by the coindexed material in I), making it the head equivalent of a variable. However, this variable will be A-bound by the coindexed material in the matrix V position, in violation of (the head equivalent of) Principle C.

Li’s contribution, and others like it (e.g. Baker and Hale 1990), are important to the development of functional categories because they confirm not only the existence of these categories, but also their distinctiveness, in modules of the grammar other than phrase structure.

3.4.3 Parameters: Borer (1984) and Ouhalla (1991)

Functional categories begin to take a major role in the grammar in Borer’s (1984) work on parameters.17 Before this work, it is not clear how exactly the grammar encodes parameters such as the pro-drop parameter or the choice of bounding nodes. Borer (1984: 29), however, argues that “all interlanguage variation [can be reduced to] properties of the inflectional system”. She claims that grammatical formatives and their idiosyncratic properties are learned the way other vocabulary is learned. Since this learning includes inflectional properties of these formatives, learning these properties is equivalent to acquiring parameters. Borer’s proposal not only changes the view of where parameters are encoded but also gives a central role to functional categories since functional categories are the repository of inflectional information.

Borer concentrates on variation seen in clitic constructions, agreement properties and case assigning properties, but the encoding of parameters in functional categories extends easily to other instances of language variation. While there are many examples of this, I briefly present one of the earliest ones here.18

Ouhalla (1991) extends the range of properties by which functional categories can vary to include the order in which these categories are introduced into the syntactic tree. He looks particularly at variation in word order and ascribes these differences to the selectional properties of functional categories. Since selection determines the order of elements on the phrase structure spine, different selectional properties can vary this order and thereby affect the overall word order of a language. Looking at the differences between SVO and VSO languages, he argues that SVO languages generate Agr above Tense while in VSO Agr is below Tense. The hierarchical order of the functional heads can be seen in the order of the relevant morphemes. In (20) we see an example from Berber, a VSO language where Tense precedes Agr. In (21) we see an example from Chichewa, an SVO language where Agr precedes Tense.
(20) ad-y-segh Moha ijn teddart Tense>Agr: VSO
    fut(TNS)-3ms(AGR)-buy Moha one house
    “Moha will buy a house.”

(21) Mtsuko u-na-gw-a. Agr>Tense: SVO
    waterpot SP(AGR)-past(TNS)-fall-ASP
    “The waterpot fell.”

He examines other languages as well, such as Italian and French (both SVO languages) and Arabic and Chamorro (both VSO languages), to support his claims.

The work of Fukui and Li shows how early in the development of functional categories, they came to be seen as a distinct class. The work of Borer and Ouhalla gave this distinct class a central role in explaining language variation. Next we turn to the explosion of the inventory of functional categories.

4 Stage II: Proliferation

With the regularization of the status of functional categories comes a flurry of interest in both the inflectional (T, D) domain and the grammatical domain (C, K) in both the verbal and nominal extended projections. The sorts of tests that we have seen in previous sections as well as others are employed to uncover additional functional categories. We can highlight five different tests that have been used to test for the presence of a functional head.19

(22) Tests for presence of a (functional) head
    a. The presence of a lexical item
    b. The presence of a morpheme
    c. The landing site of head movement
    d. The presence of a specifier
    e. The presence of semantic material or features

While each of these will be fleshed out with examples below, I note here the important work of Baker (1985, 1988). Baker (1985) points out the tight relationship between morphology and syntax, which can be interpreted as a tight relationship between morphology and the heads along the spine of the phrase structure. Baker (1988) solidifies this relationship with the process of incorporation, or head movement, accounting for why morphology so closely tracks the relative order of syntactic heads. This allows not only the presence of a lexical item (test (22a)) but also the presence of a morpheme (test (22b)) to indicate the presence of a functional head. We will see below how several of these tests may be put to use.20

4.1 Articulation of Infl: Pollock (1989)

Pollock’s work (Pollock 1989) on the articulation of the verbal inflectional domain can be seen as the beginning of a general dissection of these domains that before appeared peripheral to phrase structure – relegated to the morphological component of the grammar. Pollock uses head movement (see test (22c)) of verbs in nonfinite clauses in French to show that the verb can appear in a position that is neither V (the launching site of the head movement) nor Infl (the landing site of head movement of the V in finite clauses). The important data are given below, starting with the relevant template. Pollock shows that the negation marker
The integration, proliferation, and expansion of functional categories

pas and adverbs such as à peine “hardly” or souvent “often” appear on either side of this intermediate landing site and serve as signposts as to whether the verb appears in its merged position (below both: LOW), in the intermediate position (above the adverb but below negation: MID) or in Infl (above both: HIGH).

(23) \[ \text{Infl}^{\text{HIGH}} \] \text{NEG [ MID ] ADVERB [ V LOW ]} \\

As was already well known (Emonds 1978), when the verb is finite in French, it moves to the high position above NEG and the adverb – a position which is assumed to be Infl.21

(24) a. Jean n’aime pas Marie.  
   Jean NEG.like NEG Marie  
   “Jean doesn’t like Marie.”

b. Jean embrasse souvent Marie.  
   Jean kiss often Marie  
   “Jean often kisses Marie.”

Pollock shows, however, that when the verb is nonfinite, it moves above the adverb (shown in (25a)) but must remain below negation (shown in (25b) vs. (25c)).

(25) a. Parler à peine l’italien …  
   to.speak hardly Italian  
   “To hardly speak Italian …”

b. *Ne parler pas l’italien …  
   NEG to.speak NEG Italian  
   “Not to speak Italian …”

c. Ne pas parler l’italien …  
   NEG NEG to.speak Italian  
   “Not to speak Italian …”

In this way, Pollock uses head movement of the verb in French to argue for an additional functional category in the verbal inflectional domain between NEG and V.22 He labels the new category Agr – a labeling that is slightly speculative. Infl at the time was encoding both Tense and Agreement – features of very different types – suggesting that perhaps they should appear in different heads (see test (22e) which specifies that the presence of distinct features can be used to argue for a separate head). Since Pollock connects the difference in the use of this intermediate position to the presence of rich agreement morphology in French and not in English, he labels the intermediate position Agr. While his intention is that this Agr head be used for subject agreement, given that the difference between English and French shows up in subject agreement, the order of morphemes crosslinguistically suggests that subject agreement is outside of tense.23 In a slightly different view of the expansion of Infl, the high position is AgrS – subject agreement – and the intermediate one is T (see Belletti 1990). Another possibility, and the one that becomes fairly standard for a decade, is to have two Agr heads – one above T (AGrS) and one below T (AGrO) (see Chomsky 1991). At this point, inflectional heads begin to proliferate for a variety of reasons. Pollock proposes a new head to account for a landing site of head movement (test (22c)). The presence of both object agreement and subject agreement on verbs suggests two additional functional heads (see test (22b)).24
Articulation of Infl

This articulation of Infl sets the stage for an explosion in the number of functional categories over the next 20 years, two of the best known examples of this being cartography and nano-syntax.

4.2 Cartography: Cinque (1999)

Cartographic syntax is a research program which seeks to map out the details of the functional phrase structure spine using crosslinguistic data (see Belletti 2004; Cinque 2002; Shlonsky 2010; Rizzi 2004). The assumption is that there is a universal template or map. As one of the pioneers of this research program, Cinque (1999), using several of the tests seen in (22), argues for one of the most articulated versions of the verbal spine. What is particularly impressive is the range of evidence brought to bear on his proposals and the number of languages considered. The discussion below only gives a brief overview of this study.

The first type of evidence that Cinque uses, and the one that is best known, is his proposed universal hierarchy of adverbs. He argues that adverbs in many languages appear in specifier positions that are paired with heads that are sometimes unrealized (see test (22d)) and that the arrangement of the specifiers, and therefore the heads, is consistent crosslinguistically. Using adverb ordering from a variety of languages, he fleshes out the details of the hierarchy. Below I have given the relative order for the lower adverbs in Italian and French (other languages discussed by Cinque are English, Norwegian, Bosnian/Serbo-Croatian, Hebrew, Chinese, Albanian, Malagasy).

(27) Relative ordering of “lower” adverbs in Italian and French (Cinque 1999: 11)
   a. solitamente > mica > già > più > sempre > completamente > tutto > bene
   b. généralement > pas > déjà > plus > toujours > complètement > tout > bien

To argue that these adverbs are in specifier positions rather than head positions, he shows that verb movement can place a verb between the adverbs. More specifically, he shows that the past participle may be placed in a variety of positions relative to the lower adverbs in Italian. Given the example below, the past participle rimesso may appear in all of the positions marked by X (Cinque 1999: 45).

(28) Da allora, non hanno X di solito X mica X più X sempre X completamente rimesso tutto bene in ordine.
    “Since then, they haven’t usually not any longer always put everything well in order.”

Assuming that each position represents a possible landing site for head movement, we have arguments for six head positions above the VP (test (22c)).
Using two more traditional tests, the order of lexical items (test (22a)) and the order of morphemes (test (22b)), Cinque continues to both confirm and fine-tune his proposals concerning a highly articulated universal hierarchy of functional categories. Below I give an illustrative example of each. In (29b), we see a case of complex morphology from Korean (Cinque 1999: 53, credited to Sohn 1994: 354) and in (29a) we see a sequence of particles from Guyanese (Cinque 1999: 59, credited to Gibson 1986: 585).

(29) a. ku say-ka cwuk-ess-keyss-kwun-a!
   that bird-NOM die-ANT-EPISTEM-EVALUAT-DECL
   “That bird must have died!”

b. Jaan shuda    bin  kyaan get  fu gu
   J. MOD    epistemic PAST MOD_r MOD_r go
   “J. should not have been allowed to go.”

By lining up relative orders of adverbial elements (arguably appearing in specifier positions), morphemes, and free-standing functional heads, Cinque constructs the very articulated hierarchy given below (Cinque 1999: 106).

(30) \[
\begin{align*}
\text{frankly} & \quad \text{Mood}_{\text{preech}} \\
\text{fortunately} & \quad \text{Mood}_{\text{evaluative}} \\
\text{allegedly} & \quad \text{Mood}_{\text{evidential}} \\
\text{probably} & \quad \text{Mod}_{\text{epistemic}} \\
\text{once} & \quad \text{T(Past)} \\
\text{then} & \quad \text{T(Future)} \\
\text{perhaps} & \quad \text{Mod}_{\text{irrealis}} \\
\text{necessarily} & \quad \text{Mod}_{\text{necessity}} \\
\text{possibly} & \quad \text{Mod}_{\text{possibility}} \\
\text{usually} & \quad \text{Asp}_{\text{habitual}} \\
\text{as} & \quad \text{Asp}_{\text{repetitive(l)}} \\
\text{often} & \quad \text{Asp}_{\text{frequentative(l)}} \\
\text{in} & \quad \text{Asp}_{\text{intentionally}} \\
\text{well} & \quad \text{Voice}_{\text{fast/early}} \\
\text{tutto} & \quad \text{Asp}_{\text{PlCompletive}} \\
\text{characteristically (?)} & \quad \text{Asp}_{\text{generic/progressive}} \\
\text{always} & \quad \text{Asp}_{\text{perfect(l)}} \\
\text{just} & \quad \text{Asp}_{\text{retrospective}} \\
\text{soon} & \quad \text{Asp}_{\text{proximate}} \\
\text{briefly} & \quad \text{Asp}_{\text{durative}} \\
\text{completely} & \quad \text{Asp}_{\text{sgCompletive(l)}} \\
\end{align*}
\]

While this view of the extended projection of the verb may seem extreme, it is, in fact, followed by proposals for still further articulation of the functional spine.


Nano-syntax (see Starke 2009, and references cited therein) is the extreme of proliferation where syntactic heads do not represent lexical items or even morphemes, but rather single features. We have already seen something like this in Pollock’s work. Pollock gives evidence through head movement for an extra head in the verbal inflectional domain and he labels that head Agr. His reason for the label comes from the fact that Inf previously housed two unrelated features, T and Agr. Rather than having two unrelated features in one head, he assigns each feature a separate head. One can further support this move by showing that Tense and Agr are often represented by separate morphemes (-ez in French indicates 2pl Agr while other morphemes – 0, -i, -er – indicate Tense and/or Aspect). Nano-syntax proposes that there is a universal one-feature/one-head mapping. These heads are often sub-morphemic and lexical items often span several heads.

Here I show how nano-syntax leads to a proliferation of heads within the representation of case (for other uses of nano-syntax, see Pantcheva 2011 and Taraldsen 2010). While Travis and Lamontagne (1992) propose that case is a separate head in the syntax (K parallel to C in the verbal projection), Caha (2009) argues that Case has its own feature geometry and that this feature geometry is represented in syntax by separate syntactic heads, expanding K into six distinct heads. One of the generalizations that this hierarchy is created to explain is the pattern of syncretism – only contiguous heads can be realized by the
same forms. Below is a table provided by Caha which shows what sort of syncretism is possible given a case hierarchy of $\text{NOM} > \text{ACC} > \text{GEN} > \text{DAT}$ where shaded cells indicate for which cases the same forms are used.

(31) Table of case contiguity (Caha 2009: 8)

<table>
<thead>
<tr>
<th></th>
<th>NOM</th>
<th>ACC</th>
<th>GEN</th>
<th>DAT</th>
</tr>
</thead>
<tbody>
<tr>
<td>possible</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>possible</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>possible</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>possible</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>possible</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>not possible</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>not possible</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>not possible</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>not possible</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Using data from a variety of languages, he investigates crosslinguistic patterns of syncretism in case forms and creates the following hierarchy, where Nominative is the least complex case and Comitative is the most complex (containing all of the other cases that it dominates).

(32) Caha’s split K

```
Comitative
  F Instrumental
    E Dative
      D Genitive
        C Accusative
          B Nominative
            A DP
```

Morphology realizes subtrees of the spine, so, for example, Accusative case is represented by the constituent that immediately dominates B and Nominative, etc. (see Starke 2009 and Caha 2009 for details).

This is necessarily a very brief overview of nano-syntax, and a full understanding requires much more than this introduction, but what is clear is that this follows from a natural progression of what preceded it. Minor categories such as Determiners became projecting heads in their own right. Since some languages represent these same notions through affixation, it is logical to give affixes the same status, including those that encode Number and Case. And finally, since functional heads are seen as encoding values of
The integration, proliferation, and expansion of functional categories

binary features, it is logical to see functional categories as being features rather than lexical items. Syntax, then, represents an architecture of features rather than an arrangement of lexical items. As a final step in the argumentation, if some notions, such as Case, can be characterized by a system of features with hierarchical dependencies, these features too can be represented in the phrase structure.

4.4 Summary

At this point, while the functional domains along the spine of the tree become progressively complex, one can still maintain the functional/lexical divide. The very complex functional sequences (f_seq) in Starke above the lexical projections still are separate domains. Below we will see, however, that the dividing line between the functional domain and the lexical domain can become easily blurred.

5 Stage III: The expansion

As functional domains of Infl and Comp are articulated, so are the domains of the lexical categories. Larson (1988) argues for VP shells, dividing the verb into different heads in order to create a structure that accounts for the hierarchical relationships of elements internal to the VP (e.g. the asymmetric c-command relationship between that the first object has over the second object of a double object construction) while maintaining a binary branching structure (his Single Complement Hypothesis). As the heads are proliferated within what is known as the lexical domain of the phrase structure (e.g. Grimshaw 1991, 2000), it is no longer clear where the lexical/functional divide occurs. For some, the external argument within the predicate is introduced by a V (e.g. Larson 1988), most likely a lexical category. For some it is introduced by a head that is more arguably functional – for example, Pred (Bowers 1993), Voice (Kratzer 1996), ExtArg (Pylkkänen 2008). The question, then, is where the divide should be between lexical and functional within a projection of a semantic head.

In terms of looking at the phrase structural system, along the lines of Jackendoff (1977) or Grimshaw (1991), one might want to keep just one lexical category at the bottom of every projection. Every subsequent category along the spine of the projection, then, would be functional. However, now there is a mismatch between argument domains and lexical domains. With the advent of VP-internal subjects (e.g. Fukui and Speas 1986; Kitagawa 1986; Koopman and Sportiche 1991), there is a clean divide between thematic domains (VP/NP) and inflectional domains (TP/DP and CP/KP) that coincides with the divide between lexical and functional domains. With the articulation of the VP and the introduction of a new head that introduces only the external argument, one either has to give up the assumption that there is only one lexical head per projection or revise one’s notion of what functional categories are capable of and where the division between the functional domain and the lexical domain can be drawn. This shift in the division between the functional domain and the lexical domain is described in more detail below.

5.1 Chomsky (1995)

Chomsky’s Minimalist Program (Chomsky 1995), while still viewing the distinction between functional (nonsubstantive) and lexical (substantive) categories as being an important one, begins to shift the concept of what counts as a functional category. Parts of the structure that might earlier have been considered lexical categories now are considered to be functional categories according to the criteria that distinguish the two.
In the early Minimalist Program, it is clear that functional categories are central to the theory of movement. Movement is triggered by (strong) features on a head and strong features can only be hosted by nonsubstantive (functional) categories (Chomsky 1995: 232).  

(33) If F is strong, then F is a feature of a nonsubstantive category and F is checked by a categorial feature.

Examples of functional categories with strong features are T in French where a strong V feature forces (overt) V-movement to T, T in English where a strong D feature forces (overt) movement of DP to Spec, TP. C might have a strong feature forcing T to C movement or movement of a wh-phrase to Spec, CP, and D might have a strong feature forcing either head movement of N to D or XP movement of the possessor to Spec, DP. All of these movements involve categories that are uncontroversially functional (T, C, D).

With the introduction of v (little v), the distinction is less clear. This category was assumed to be the highest head within the now articulated predicate phrase (VP). In this position, it was assumed to introduce the external argument into the phrase structure – a.k.a. Pred (Bowers 1993), Voice (Kratzer 1996), ExtArg (Pylkkänen 2008). But the question is whether it is a functional category. The predecessors of v – Larson’s (1988) VP shell and Hale and Keyser’s (1993) “causative” V – might have appeared lexical. Using Abney’s criteria given in (12) to determine whether v is lexical or functional, we get results which are mixed, but tending towards functional. For some, including Chomsky, this head introduces an external argument making it look like a lexical category, since functional categories may only functionally select. For Abney, this test was the most important (see Section 3.2). However, v passes all of the other tests for being a functional category. It may have only one complement. In languages where it has an overt manifestation such as Japanese (see Harley 1995) or Malagasy (see Travis 2000), it is a closed category, it is morphologically weak, and it is generally not separable from its complement. Further, its contributions to the semantic content of the predicate can be seen to be grammatical or relational, suggesting that it is a functional category. For Chomsky (1995) it is crucial that this head be a functional head as it is the possible host of a strong feature which triggers movement of the object to the edge of vP.

If v is now seen as functional, we have shifted where the functional and lexical divide is within the verbal projection. Below we can see that, in a less articulated system (34), the domain divide falls above the external argument, while in an articulated VP structure (35), the divide falls below the external argument.

(34) Thematic vs. inflectional domain = lexical vs. functional domain
The integration, proliferation, and expansion of functional categories

(35) Thematic vs. inflectional domain = lexical vs. functional domain

```
TP
  /\  
 DP  T'
  /\  
 T   vP
   /\  
 DP  v'  
   /\  
 ExtArg   v  
      /\  
    VP    IntArg
```

5.2 Marantz (1997)

The encroachment of functional categories into the thematic domain becomes even more marked in versions of phrase structure where lexical categories themselves are viewed being encoded by functional category, that is, where the verb *destroy* is comprised of a root \(\sqrt{DESTRUCT}\) plus a functional category, \(v\).\(^{30}\)

Marantz (1997), for example, revives and expands on a proposal that lexical items are, in fact, without category. This idea was central in Chomsky (1970) where lexical items like the verb *destroy* and the noun *destruction* are derived from the same categoriless root *DESTRUCT*. It may, then, be the case that the VP is even more articulated and may not contain any Vs at all, only vs, and the functional/lexical divide falls even lower within the predicate.\(^{31}\)

(36) Categoriless roots

```
TP
  /\  
 DP  T'
  /\  
 T   vP
   /\  
 DP  v'  
   /\  
 ExtArg   v  
      /\  
    VP    IntArg
```

In a system such as this, it is no longer clear that there are any truly lexical categories, and if no truly lexical categories remain, it is not clear whether the types of distinctions that are laid out in, for example, Jackendoff (1977), Abney (1987), Borer (1984), Li (1990), and Chomsky (1995) have any validity. In other words, it is not clear there is any place in the current theoretical framework where the distinction between lexical and functional categories is relevant.

6 Conclusion

Since the birth of generative grammar 60 years ago, functional categories have gone from having an important supportive role to being the machinery that drives the grammar. Becoming a full-fledged member of X'-theory is the first step in this progression. At this point, while functional categories remain distinct from lexical categories, most notably having a different sort of relationship to the material in their complement and specifier positions, they nevertheless have the status of projecting heads – heads that can place selectional restrictions on the content of their complements and can act as launching and landing sites for head movement.

As the particular characteristics of functional categories become more understood, additional functional categories are uncovered. If a functional category has a particular meaning, and that meaning is represented in one language by a separate word and in another by an affix, it is a small step to assume that additional affixes indicate additional functional categories. If functional categories can be shown to represent small bits of meaning, perhaps binary features, then additional features can be seen to indicate additional functional categories.

Gradually more and more of phrase structure becomes part of the functional domain of an extended projection. In the end, functional categories have gone from being a niggling detail in syntactic structure to being the defining material of syntactic structure. In an extreme view, all syntactic heads are functional except for the lowest head which is a root.

The central role of functional categories is evident in the syntactically well-formed “Colorless green ideas sleep furiously”. The functional skeleton of the tree, the inflectional heads and the category heads, obey the rules of English syntax. The lexical choice for the adjective (“bright” vs. “colorless”) or for the noun (“ideas” vs “goblins”) etc., while producing semantic mismatches, has no effect on the syntactic machinery. In the end, a view of syntactic structure that is comprised mostly of functional categories is appropriate in that it more accurately represents the formal features that are the tools of syntax and the interest of syntacticians.

Notes

1 This terminology, to my knowledge, first appeared in Fukui (1986). For other works that look at the status of functional categories in more depth, see, for example, van Gelderen (1993) and Muysken (2008). See Hudson (1999) for a critique of the notion that functional words comprise a distinct class. He, however, distinguishes functional word classes from functional position categories. In other work (Hudson 1995) he discusses particular functional position categories.
2 Aux will take on various labels in the course of this paper such as Inf₁, I, and T.
3 P(repositions) sit on the fence between major/minor or lexical/functional categories in many theories. I leave the discussion of P aside for most of this paper.
4 Abbreviations: ACC = accusative; Agr = agreement; ANT = anterior; art = article; DAT = dative; DECL = declarative; EPISTEM = epistemic; EVALUAT = evaluative; ExtArg = external argument;
The integration, proliferation, and expansion of functional categories

... not have a long shelf-life. See Chomsky (1995: 349–355) for conceptual arguments against the existence of Agr heads.

Abney is careful to point out that (12a) is definitional, (12b) follows from (12a), while the others were investigating similar issues at the same time in the same graduate program.

Kornfilt 1984 had already proposed a projecting functional category within the nominal projection in her analysis of nominals in Turkish – an Agr projection that only appears in nominals that contain a possessor.

Care has to be taken with this test, as Abney (1987: 268ff) shows, since there are times when it is a nonlocal head that can be selected. This is discussed again in Section 3.3.

We will see later in Section 4.2 other arguments for proposing new functional structure.

Though see the discussion in the Section 3.4.3 where Ouhalla shows that the order of these two heads might vary.

Agr heads did not have a long shelf-life. See Chomsky (1995: 349–355) for conceptual arguments against the existence of Agr heads.
25 We have seen that Ouhalla argues that order of functional categories might be parameterized; however, it could be that this is restricted to certain functional categories such as Agreement.

26 Rizzi (1997) has done similar work on the articulation of the CP domain.

27 This is reminiscent of Fukui’s observation that the specifiers of functional categories may only be filled by *internal merge*. Chomsky’s view differs, however, since he allows the same head to have a specifier filled by *external merge* and then a second specifier filled by *internal merge* as is the case with v.

28 This may follow from the condition that all structure is binary branching, however (see Kayne 1984).

29 Arguably serial verb constructions have a v that is not weak and that can be separated from the head of its complement. See, for example, Travis (2010).

30 This v is not the same as the v that introduces the external argument. This one represents the categorial signature of the root.

31 Others also have a view of phrase structure comprised mostly of functional categories (Kayne 2011; Borer 2003).

Further reading


References


The integration, proliferation, and expansion of functional categories


