

The History of the Concept 'Phrase' in linguistics Part 2: The modern concept of the phrase and its development

1 Introduction

Last week, we took a historical view of the notion of the phrase (or the absence of it) by starting with the earliest grammatical investigations, around 600 BCE, and working our way forwards towards the present. This week, we will do the opposite. We start with the present views on the phrase and work our way backwards in time to discover the origins of the notion.

Space and time limitations mean that we will not be able to go into very much detail about each of the stages of development. But I will endeavour to provide enough information to allow a basic understanding of the issue under discussion.

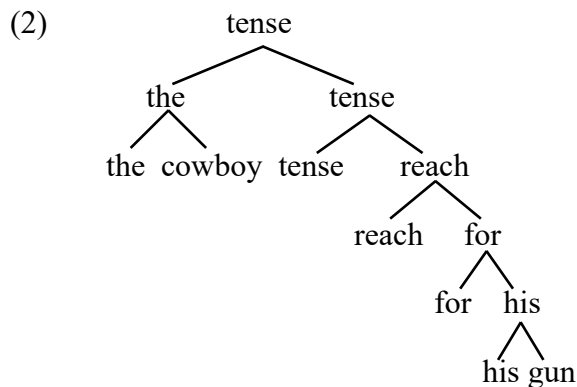
2 1990s to the present: the decades of minimalism

The Minimalist Programme, introduced in the last decade of the 20th century, aimed to reduce syntax to its bare minimum. The purpose of this was that, if successful, the explanatory nature of the theory achieved would be maximised. Only restrictive theories offer explanations for observable phenomena, as can be seen if we consider what a completely unrestricted theory provides. Without restriction, the theory predicts that anything can happen. Thus, no matter what happens, the theory will never be disconfirmed. But at the same time, we gain no understanding of why things happen as they do. By adding restrictions, we start to gain an understanding of what is happening. If the restricted theory proves wrong, we have learned something and if it is confirmed, then we can say that the proposed restrictions are part of the nature of what it is that we are observing and thus we have gained an understanding of it. The greater the restrictions we impose on the theory, the greater the possibility that it will be wrong, but also the greater our understanding is if it is confirmed.

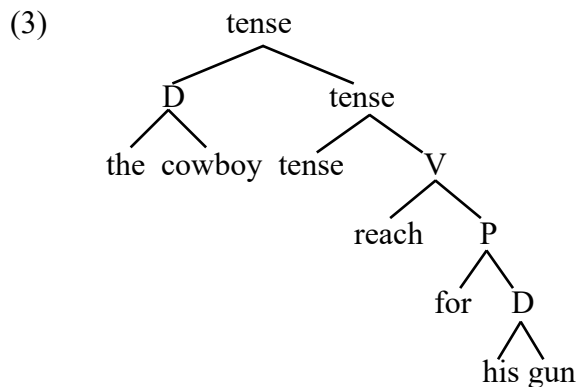
In 1995, Noam Chomsky, a linguist who will figure prominently in this lecture, published his minimalist approach to phrase structure entitled 'Bare Phrase Structure'. This assumed the following:

- (1) a Structures are built out of lexical elements and add nothing more than these.
- b The structure building process is as follows:
 - i) take two elements (lexical items or structures already built)
 - ii) put them together to form a new element
 - iii) choose one to be the label

Here I will give a brief demonstration of this using the sentence ‘the cowboy reached for his gun’, which we have already seen represented as a phrase structure in last week’s lecture. In Bare Phrase Structure, this would be rendered as follows¹:



Admittedly, at first sight, this looks very strange. But with a few instructions concerning how to read this structure, things turn out to be quite straightforward. First, while it looks like ‘words’ are repeated in this structure, the only elements which have any phonetic and semantic reality are those elements at the far ends of the branches. The other elements are simply to be taken to be the information from lexical items which is relevant for the structure. The main piece of relevant information in this context is the category of the element. Thus, it would be possible to represent the structure as follows:



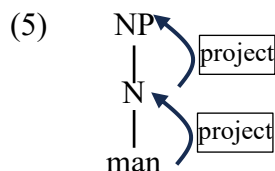
The reason why Chomsky represented structures as in (2) as opposed to (3) was to highlight the fact that in building the structure, nothing else than the lexical items themselves has been added. In a traditional phrase structure representation, one might see labels such as ‘NP’ and ‘VP’, etc. to be extra elements of the structure which go beyond the lexical items, which would go against the minimalist approach.

Still, one might feel that there is something lacking in a representation such as (3), as there appear to be no phrases indicated. Again, this would be the wrong way to look at the structure. Chomsky argued that notions such as words and phrases (and things in between these, which we will look at later) are perfectly well represented under the following assumptions:

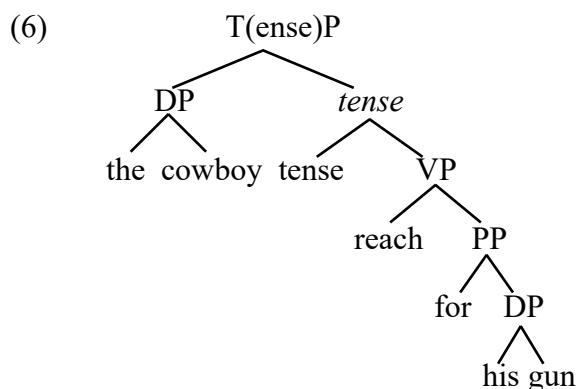
¹ I am ignoring important aspects of the analysis here in order to focus purely on issues to do with the phrase. There are many other aspects involved in structure building and analysis which would take us too far afield to go into here.

- (4) a a word is something which is not projected.
 b a phrase is something which does not project.

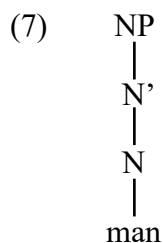
The notion of projection used here, is one which dates back to the 1970s in which structural information, such as category, is seen to originate in lexical items and *project* upwards to the syntactic nodes on a tree. Thus, a word such as *man* inserted into a tree would project its noun status to position above it and this would then be further projected to the phrase node above that:



Applying the definitions in (4) to the structure in (2), we can see that the elements at the bottom of the structure are words because they are not projected from anything underneath them. On the other hand, the labels above most of the words qualify as phrases as they do not project their properties further. So, we could represent the structure as follows, though Chomsky would argue that to do so would be a needless complication:



Note that there is one element in (6) which fails to be defined as either a word or a phrase. The label in italics *tense* is not a word as it is projected from the *tense* element below it. On the other hand, it is not a phrase because it projects to the TP above it. This is in line with previous theories starting in the 1970s, which we will outline a little later, which admits elements which have statuses between the word and the phrase. These are known as 'X-bar' elements and are seen to be bigger than words, but smaller than phrases. They are projected from words and they also project to the phrase level. (7) provides a standard analysis of what we had in (5):



To some degree, it was representations such as (7) that the Bare Phrase Structure account was reacting to as it clearly includes elements which are added other than the lexical item itself.

The BPS approach claimed that this was totally unnecessary and that more minimalist assumptions are therefore superior.

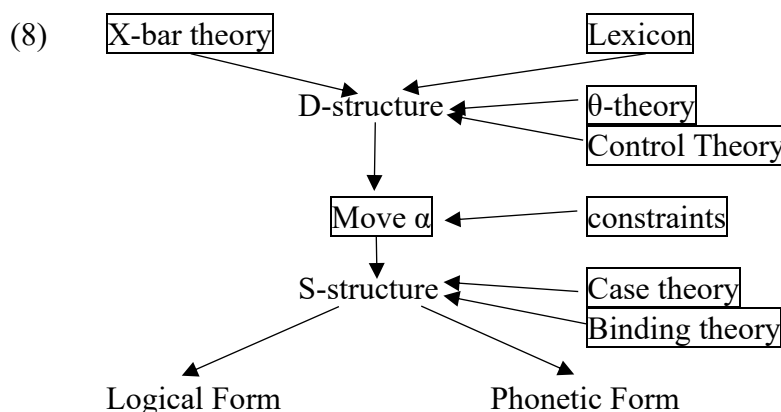
There are other observations of interest we can make about the structure in (2) which relate to what we were discussing last week about phrases being bigger than words. Note that in (2) the words *cowboy* and *gun* are not associated with any projection but just sit inside the phrase associated with the accompanying determiner (*the* and *his*). However, by the definitions given in (4), these elements are both words (as they are not projected) and phrases (as they do not project) simultaneously. In this way, they represent exactly the same information as is represented in something like (7) in a much simpler way. While (7) overtly represents that the element *man* is a word by being associated with the word level label 'N', and that it is at the same time a phrase, being the sole element associated with the label 'NP', in (2) exactly the same information is represented by including the word *cowboy* and *gun* in a particular structural context.

This is as much as we need to go into about the current views on phrase structure. We now take one step backwards in time to the ideas from which the BPS directly originated.

3 1980s: the decade of generalisation

The 1980s was the decade of generalisation. Not the radical minimalism of the following decade, but rather generalising the system of rules which had been proposed over the previous 20 years or so. It was seen at the time as a period which greatly furthered our understanding of linguistic matters, particularly in syntax.

One thing which helped the generalisation process was a novel view of the grammatical system which saw it as being made up of phenomena specific 'modules'. Such modules contained a small number of very general and simple rules. Complexity in the system was the result of the interaction between these simple modules. Here, I give a picture of this modular grammar, but will not go into detail about most of the modules named.

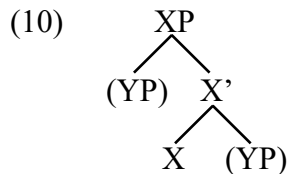


The things in boxes are the modules of the theory, and they are applied to numerous levels of structural representation (D-structure, S-structure, etc.). The module we are going to be concerned with here is the X-bar module, which is the part of the grammatical system which dealt with phrase structure.

The X-bar module consisted of two basic rules:

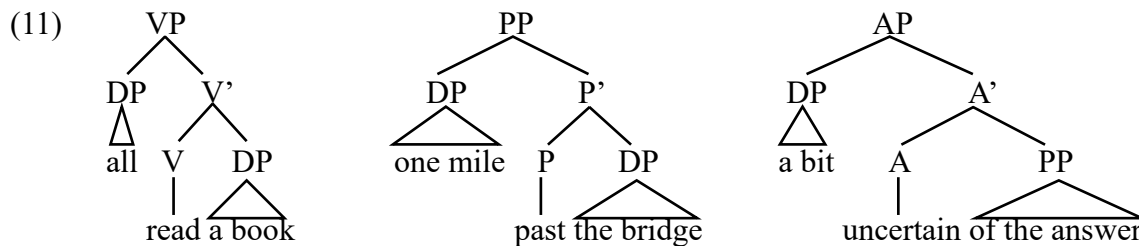
- (9) $XP \rightarrow (YP) X'$
 $X' \rightarrow X (YP)$

As can be seen, these are phrase structure rules of the sort we briefly mentioned last week. They involve an element to the left of the arrow, which is the thing to be expanded by the rule, and some other elements to the right which are what are expanded into. So, the two rules together form a structure which looks like this:



To understand what is being claimed here, one first needs to understand the significance of the X and Y symbols. These represent categories. But unlike category symbols such as V or N, etc., which represent specific categories, the symbols of X-bar theory are variables, which means that they stand for any category. This is why only two rules are needed; these are rules about the structure of phrases in general, not specific phrases. Second, the X stands for a word and XP stands for a phrase. The X' stands for a structural unit bigger than a word, but smaller than a phrase. This, then, there are three levels of structure admitted here. Note that there is only one word position in the phrase. This is known as the 'head'. The use of the same category variable on both sides of the arrows in the rules in (9) means that the category of the head will be the same as that of the phrase (and the X' level). What this means is that phrases will always contain a head: a word of the same category as the phrase. While this may seem intuitive for some phrases – VPs should contain a verb, for example – it has not always been assumed that all phrases are headed, as we will see later on. This, then, is a result of the generalisation of these rules. A third point also concerns the observation that there is only one word position in a phrase, all other positions are phrasal. The consequence of this is that all words must be the head of a phrase. The brackets around the two phrases which may accompany the head indicate that these are not obligatory. The phrase which accompanies the head inside the X' is called the complement and the phrase preceding the X' is the specifier. Besides determining the category of the phrase, the other role of the head is to restrict the nature of the complement: different heads allow for different complements and some heads do not allow a complement at all. Thus, the complements presence and its category will depend on the lexical properties of the word which sits in the head position. This can be most easily seen with verbal heads, as it is traditionally recognised that different types (subcategories) of verbs are accompanied by different phrases. Intransitive verbs have no complement, transitive verbs have a nominal complement and prepositional verbs have a PP complement, etc.

A number of things follow from this theory. The most obvious thing is that with only two rules which are applicable to all phrases, all phrases will have the same structure (i.e. that of (10)). This might be seen as one of the major findings of the previous decade in phrase structure investigation. When one looks at phrases of different sorts, one sees the same pattern over and over:

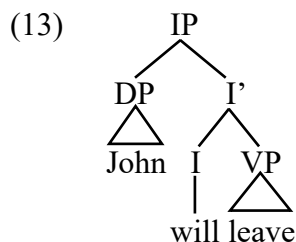


Obviously, there are differences to be seen between these phrases, but the structural similarity is particularly striking.

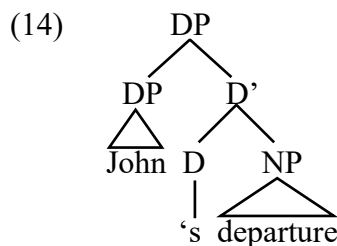
Another consequence of the theory is that sentences must be considered as phrases and moreover they must have a head. The element which is best suited for this role is what has come to be called 'inflection' since the 1970s. This category consists of modal auxiliary verbs, the infinitival morpheme 'to' and the tense morphemes. The complementary distribution between these elements is what leads us to consider them as a single category:

- (12)
- a * John can left.
 - b * I want John to can leave.
 - c * I want John to can left.

For those cases where the inflectional element is obvious, its position is between the subject and the VP, meaning that the supposed sentence structure fits the X-bar framework perfectly:



You may have noticed that in (11) and (13) the term 'DP' is used, where you might have expected NP. This is a further consequence of the generalised nature of the phrase structure theory. Recall that under these assumptions all words must be heads. This means that determiners must also head their own phrase, as DP. In 1986, Abney argued convincingly that the determiner was in fact the head of the whole nominal phrase and that the noun heads a phrase within it. The analysis provides us with an innovative treatment of the possessor within the nominal phrase as the specifier of the DP, making the long observed connection between subjects and possessors a consequence of the structure:



Again, there is no time or space to go into the justifications and consequences of these analyses. All I will say is that at the time they were proposed the insights that they offered to

a lot of well-known but until that point mysterious observations was a compelling argument for them and they were readily adopted by most linguists. Their

assumptions still hold today, even if the general rules that gave rise to them have been replaced by the principles of BPS.

4 1970s: the decade of constraints

Just as generalised X-bar theory was proposed in reaction to what was seen as the over specific treatment of phrases of the previous approaches, the constrained nature of the proposals of the 1970s were in reaction to the freedom of analyses of the 1960s.

We have already discussed how placing limits on a theory improves its explanatory nature, but it was also discovered that by placing constraints on grammatical processes we can simplify those processes too and thus achieve more elegant analyses. This stems from work by Ross in 1967. I will exemplify this with a brief discussion of one of Ross's observations. Grammatical rules often relate different positions in a sentence. For example, passives relate subjects to objects, in that the subject of the passive is interpreted as the object of the active:

- (15) a John was seen
b I saw John

Another grammatical process, termed 'raising' relates the subject of a higher clause to the subject of an immediately embedded non-finite clause:

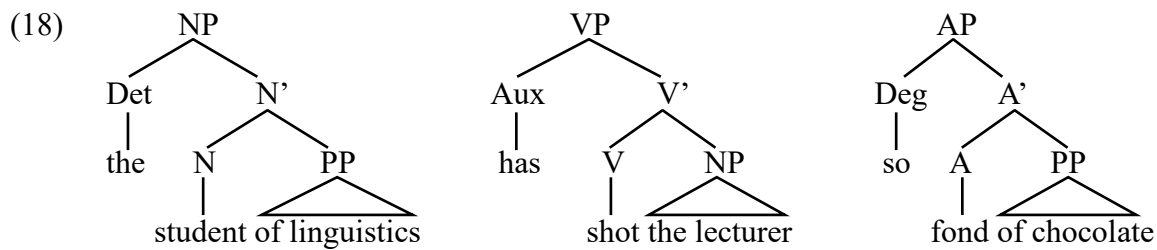
- (16) a John seems [to be smart]
b It seems [John is smart]

Ross's observation was that we never get such associations between positions separated by a finite clause. Thus, the object of one clause can never be associated with the subject of a higher clause and the subject of a raising construction can never be associated with the subject of an embedded finite clause:

- (17) a * John was believed [Mary liked] (c.f. It was believed [Mary liked John])
b * John seems [is smart]

If the processes which link object to subjects in passive constructions and those which link subjects to subjects in raising constructions are specific to those constructions, then we will have to build into each of these processes (and any other process to which the restriction applies) virtually the same set of restrictions. This adds both complexity and redundancy into the grammar. Ross's solution was to extract the limitations from each process and state it as a simple restriction (what Ross called a 'constraint') relevant to all the individual processes. It was later discovered that if we extract all the complex restrictions from all processes, the processes themselves become simpler and, surprisingly, more similar to each other. In the end, many of these processes collapse into one, which is then seen to have a very general application. Ultimately, it was this line of research that led to the modular theory which emerged in the 1980s. The modules of simple rules were either what was left after complex restrictions had been factored out of them, or they were the constraints themselves, which being stated independently of the complexities of specific structures could be stated in far simpler terms.

X-bar theory was first proposed by Chomsky in 1970 as a method of extracting out some common features shared by several phrases, (VP, NP and AP at first, though PP was added a few years later).



However, not everything fell within the scope of the theory. As is obvious from (18), words such as determiners, auxiliary verbs and degree adverbs were just words with no associated phrase. Thus, these were words which were not heads and, as such, stood outside of the theory. The other thing that did not conform to X-bar theory was sentences. At this point, sentences were assumed to be one of those structures which did not have a head. As X-bar theory is essentially the theory of heads, this meant that the rules governing sentences stood outside the X-bar framework. For this reason, there were separate rules for phrases and sentences, a complication that the generalised version of the theory sought to overcome:

- (19) a $S \rightarrow NP\ INFL\ VP$
 b $XP \rightarrow Spec\ X'$
 $X' \rightarrow X\ Comp$

At the start of the 70s, some took the X-bar rules to be constraints on a set of more structure specific phrase structure rules. Thus, differences between phrases were to be accounted for by these rules and similarities between them were due to the restrictions that the X-bar framework imposed on them. However, throughout the decade, as other constraints were extracted from the complex rules, it came to be realised that differences between phrases were due to considerations not directly due to the phrase structure rules themselves. This enables grammarians to replace the specific phrase structure rules with the X-bar rules and this was the situation established at the start of the 1980s, allowing X-bar theory to be identified as one of the grammatical modules.

5 1950s and 60s syntax: the origins of phrase structure grammar

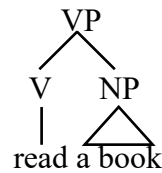
In 1957 Chomsky wrote a book which changed linguistics and put in on the course it is currently on. This book was a very short one which compared three different types of grammatical rule, as a demonstration of how to proceed in developing a grammatical theory. The general idea was that we should proceed by comparing different grammars, considering the kinds of syntactic phenomena they were capable of accounting for.

One type of rule, which we do not need to go into here and which Chomsky quickly showed to be particularly inadequate, was one that assumed a linear organisation of words in a sentence. In other words, a grammar which dispensed with the notion of a phrase.

It is the second type which we are interested in here. This was a phrase structure grammar which was made up of the kinds of rules that we have been considering above. That is, the

kind of rule which defines a small grammatical structure involving a phrase and its immediate content:

(20) $VP \rightarrow V NP$



Interestingly, Chomsky proposed this type of grammatical rule in an attempt to formalise the ideas that preceded Chomsky's work in order to show that it was inadequate. We will get on to what those ideas were in due time, but the important point to note is that they assumed the notion of a phrase, but did not attempt to define these in a rule based system as they had no theory of phrases at all. It is very difficult to attack a position which has no theoretical basis on empirical grounds as such a position makes no predictions which can be demonstrated to be false. You can attack the position for its lack of theoretical basis, which makes it untestable, but Chomsky wanted to show that its underlying ideas themselves were not right for basing any theory of syntax on. Therefore, he set about providing a theory based on these underlying ideas and the result was a phrase structure grammar.

Of course, one might claim that Chomsky was putting words into the mouths of his predecessors and building a straw man in order to knock it down, and to some extent this would be true. But Chomsky's demonstration of the inadequacy of phrase structure grammar was profound enough to cast doubt on the previous generation of linguists' ideas and ultimately the school that they belonged to, the school of American Structuralism, collapsed shortly after the publication of Chomsky's book.

The phrase structure grammar that Chomsky came up with consisted of a set of specific phrase structure rules for various phrases. This was to be considered a 'fragment' of the grammar, and many more rules would be needed to capture all the possible structures in any given language. The rules were both language and structure specific, meaning that different languages would require different rules to capture the nature of their phrases and different phrases in each language would need to have different rules. We can see how X-bar theory was a reaction to this approach to syntactic structure.

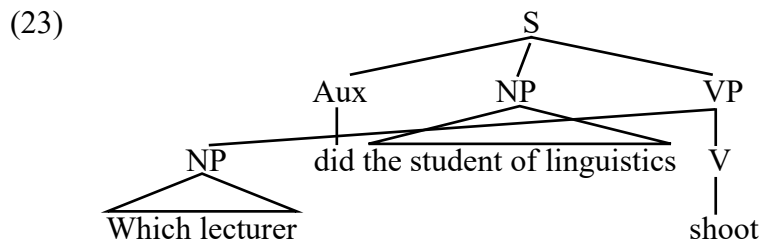
An example fragment of the rules for English might look like this:

(21) $S \rightarrow NP VP$
 $VP \rightarrow V$
 $VP \rightarrow V NP$
 $VP \rightarrow Aux V NP$
...
 $NP \rightarrow N$
 $NP \rightarrow Det N$
 $NP \rightarrow Det A N$
...

Let us turn to what Chomsky thought to be an inadequacy of this sort of grammar. This concerns discontinuous constituencies. These are constituents in a sentence which appear to be split in two by some other constituent which is independent from them. For example, consider the following:

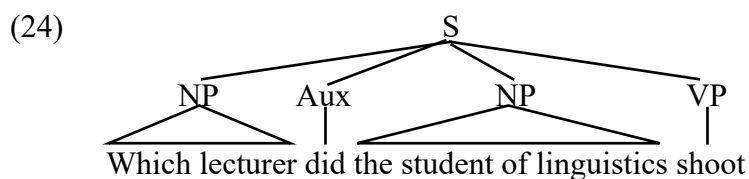
(22) Which lecturer did the student of linguistics shoot?

In this sentence, the object of the verb, *which lecturer*, appears at the front instead of the normal position for objects after the verb and inside the VP. But this object cannot be considered to be part of the VP, as the VP phrase structure rule ($VP \rightarrow V NP$) would have it, as a very strange structure would result:



Not only does this structure look odd, but Chomsky showed that it is impossible to form using phrase structure rules. The reason for this is that such rules concern only structural nodes and their immediate constituents – the node to the left of the arrow and the nodes to the right. They state the order of the elements to the right of the arrow. But to achieve a structure such as (23), we would need a rule which imposed an order on a constituent of the VP (the object) with respect to a constituent of the S (the auxiliary and subject). But no phrase structure rule can do this.

The other possibility would be to consider this to be a special type of sentence, in which the object is seen as a part of the sentence node and not part of the VP:



But again, there is a problem. This structure would necessitate the addition of another rule for the sentence:

(25) $S \rightarrow NP Aux NP VP$

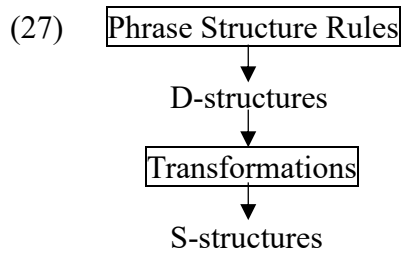
However, there is a condition on the use of this rule. It can only be used if the rule for the VP which does not have an object is selected ($VP \rightarrow V$). If this were not so, we would be able to generate sentences such as:

(26) * Which lecturer did the student of linguistics shoot the professor?

But there is no mechanism in phrase structure rules to ensure that the application of one rule for one part of a structure entails that another rule must be applied for another part of the structure. Thus, a phrase structure grammar is simply not able to cope with fairly normal sentence such as (22).

Despite this demonstration that phrase structure rules, by themselves, were unable to cope with natural language phenomena, Chomsky did not argue that they should be abandoned completely. Instead, he proposed a grammar that included phrase structure rules as one of its components and another kind of rule was to be added to supplement this component and

enable discontinuous constituents to be accounted for. These new rules, transformations, would operate on the structures produced by the phrase structure rules, called ‘deep structures’ (later D-structures) and provide other structures, called ‘surface structures’ (S-structures):



We see here the origins of the intricate theory which emerged in the 1980s.

I will not go into the nature of the transformational rules which Chomsky proposed here, just as I have not gone into their development during the other periods of the theory discussed above. Though they have certainly changed over this time, just as the phrase structure rules have, they are still very much a part of current theory and so phrase structure rules and transformations can be seen as ideas that have endured since their introduction in the late 1950s.

There is not much more to be said about the phrase structure rules themselves and we have seen how they have developed over the last 50 or so years. What remains is to talk about the ideas that Chomsky invented phrase structure rules in order to criticise. This will be the topic of the next section.

6 1940s and 50s: Immediate Constituent Analysis

The school of American Structuralism, founded in the early 1900s by Leonard Bloomfield, who we will say more about in the next section, had become the virtually the only school of linguistic investigation in America since the beginning of the first world war. After the second world war, a generation of Bloomfield’s students worked to further the philosophy that Bloomfield had introduced, though as mentioned in the previous section, no actual theory of syntax was established. However, unlike what had been happening in Europe at the end of the 19th and beginning of the 20th centuries, the notion of a phrase was used in syntactic analysis.

We will mention two of these linguists, Zellig Harris, who is important here mainly because he was Chomsky’s teacher, and Charles Hockett, who worked particularly on the analysis of sentences into their ‘immediate constituents’, i.e. phrases.

Harris was a formalist and more than any other linguist of his generation he was concerned with stating grammatical rules in highly formalised mathematical ways. This was obviously a great influence on Chomsky, who continued to adopt formal ways of representing grammatical rules, and his phrase structure grammars were an obvious example of this. Harris, like other structuralists, never developed his rules into any theory and the rules were just seen as a linguist’s device to describe linguistic phenomena rather than a way to explain it. Hence, there were no restrictions on the formal rules that Harris proposed; as long as they described the phenomena, that was enough. Interestingly, the X-bar formalism can be attributed to Harris, though he did not propose them as a way of describing phrase structure. Harris used the ‘bar’ notation to describe a morphological observation. There are some

morphemes which attach to a certain category and do not change that category. However, not all of these morphemes are equivalent and some have to be applied before others. For example, the morpheme *-ship* can be added to a noun to form another noun:

(28) friend – friendship; fellow – fellowship; relation – relationship

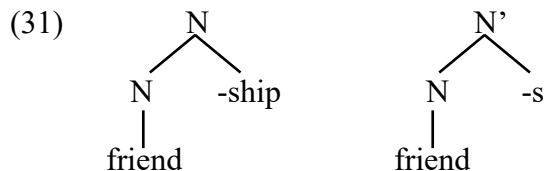
Another morpheme which does the same thing is the plural:

(29) friend – friends; fellow – fellows; relation – relations

However, while a plural morpheme can be added to a noun formed with the *-ship* morpheme, the *-ship* morpheme cannot be added to a noun formed with the plural:

(30) friendships *friendship

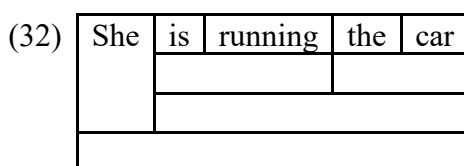
To account for this, Harris suggested that the plural morpheme when added to a noun produces an N', not an N. *-ship* on the other hand can only attach to N and when it does it produces another N:



I am using tree diagrams here for ease of exposition rather than for historical accuracy. It was Chomsky who invented the tree diagram as a way to represent the structures that his phrase structure rules produced. Harris never represented this idea in this way. Thus, while Harris may have introduced the notation, Chomsky used the notation to propose a completely different theory of phrase structure. So Harris's contribution is rather minimal.

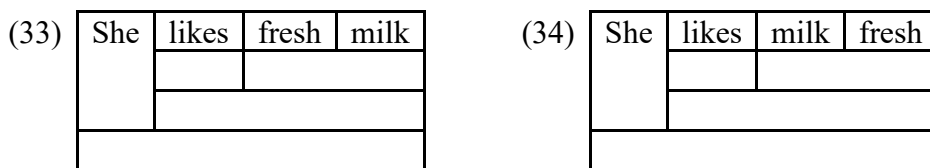
Turning to Hockett, it was his Immediate Constituency Analysis that Chomsky created phrase structure rules to demonstrate their inadequacy. The ideas behind ICA were straightforward. We take a sentence, break it into its immediate constituents and then break them into their constituents and so on, until we get down to words. It is fairly easy to see how the tree diagrams which Chomsky's rules produce relate to this idea. The big difference between Hockett's notion of hierarchical structure and Chomsky's was that Hockett did not propose any rules which governed how hierarchical structures are defined to be well formed. This follows from the perspective of the American Structuralists that they were not trying to explain why a given language ordered things the way it did. This was just taken to be a basic fact. It was the linguist's job to describe this basic fact and no more.

Hockett did, however, devise a way to represent hierarchical structures, which is often referred to as the 'Chinese box' representation, a Chinese box being a set of boxes of successively diminishing sizes so that each one can fit inside the next biggest. For example, here is Hockett's representation of the structure of the sentence 'She is running the car.':



This is taken from his 1958 book *A Course In Modern Linguistics*, which I will be referencing throughout this section. It is fairly intuitive how this analysis works. We start with each word in its own box and underneath are other boxes which cover more than one of those above. This represents that the boxes covered by the ones underneath form a constituent together. These boxes are then covered by bigger ones until there is just one box covering the whole sentence. Clearly, then, the bigger boxes represent phrases in a similar way to how the nodes on a tree diagram represent phrases.

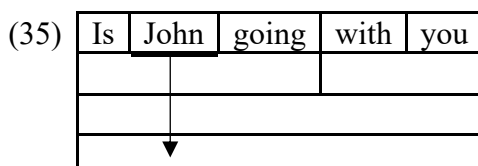
One large difference between tree diagrams and Chinese box representations is that the latter do not label the phrases. Apparently, Hockett saw no need for terms such as ‘noun phrase’ or ‘verb phrase’. This is problematic, as can be seen in the following two analyses he provides:



As we can see both of these sentences receive the same structural analysis. But while it might be that the words group together in the same way, the sentences are far from being structurally identical and the adjective plays a different role in both. In (33), the adjective is attributive, a role it gets from being a modifier inside the nominal phrase. But in (34), the adjective is predicative as it is a predicate head of an adjectival phrase which has a subject. Thus, the difference between these sentences is not in the grouping of its constituent words, but in the structures they form, differentiated by their categories. This is therefore not a distinction that Hockett could easily make.

Hockett did however make a distinction between ‘endocentric’ and ‘exocentric’ phrases. These notions he took directly from Bloomfield, who we will discuss in the next section and they have to do with Bloomfield’s notion of the head of a phrase. Essentially an endocentric phrase is one with a head and an exocentric one lacks a head. Unlike the Chomskyan notion given in X-bar terms, Bloomfield’s head was not defined as the word which provided the phrase with its category, but the word which could replace the phrase without losing grammaticality. Essentially, this is a distributional definition as the head of the phrase is a word which has the same distribution as the phrase itself. Therefore, Hockett was able to identify endocentric phrases, essentially those types of phrases which have heads, without necessarily connecting the category of the phrase to that of the head.

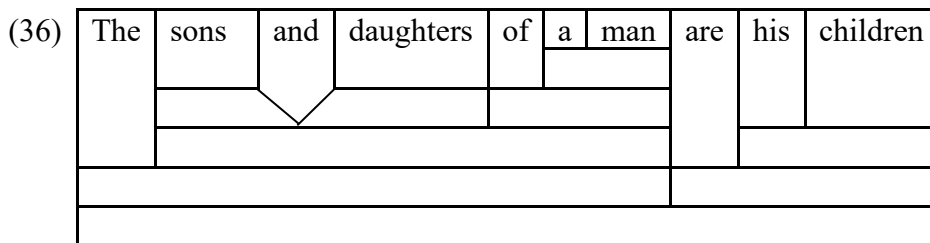
It is interesting to note that Hockett had a way of representing discontinuous constituents using Chinese boxes. Here is the analysis of ‘Is John going with you?’, where the auxiliary verb forms a discontinuous constituent with the verb, separated from it by the subject:



The idea seems to be that while the subject is placed linearly between the auxiliary and the verb, it is not incorporated into the structure until the last box is formed, allowing the auxiliary and verb to be boxed together without the subject. While this may well be a representation of a discontinuous constituent, it highlights exactly the problem that Chomsky

raised with the American Structuralist approach. The reason why it is possible to represent discontinuous constituencies using Chinese boxes is because there are no restrictions placed on them. One can represent virtually anything using such an analysis. It is the lack of a theoretic underpinning of the representation which means that it is unconstrained and therefore limitless. It also makes it incapable of explaining anything. Note, this is very different from Chomsky's phrase structure rules. The rules that Chomsky proposed are not unlimited. They are restricted to referring to a single node (the one to the left of the arrow) and the nodes that are immediately contained in it (the ones to the right of the arrow). This is exactly why a phrase structure grammar cannot cope with discontinuous constituents.

Another interesting point arises from the fact that Hockett offers no support for his analyses. One might wonder, therefore, on what basis did he decide to analyse hierarchical structures in the way he did. Looking at his analyses, it looks as though they are based on a number of different things. The fact that the main division in the sentence is between the subject and the predicate suggests that Hockett follows the traditional way of analysing sentences that has its roots in Greek linguistics. Of course, while the traditional notion is associated with the noun, Hockett's subjects are phrasal, as in his example:



As can be seen here, the first division of the sentence in the second to last line splits 'the sons and daughters of a man' from 'are his children', indicating that the initial *phrase* is associated with the subject function.

The Chinese box analysis in (36) also indicates an odd aspect of Hockett's claims about structure. He considered the coordinating particle, *and*, not to be part of the constituent that it participates in. For Hockett, *sons and daughters* is made up of the two noun constituents and the conjunction merely marks a relationship between them; specifically, they form a constituent together. Why such an indication of constituency is necessary in this case, where simply putting two elements together into the same box does the same job elsewhere, is not a question Hockett goes into. It appears that he considered such markers to be elements which make no contribution to the semantic interpretation of the sentence:

Some morphemes ... serve not directly as carriers of meaning, but only as markers of the structural relationships between other forms. (p. 153)

If the conjunction does not carry meaning, however, it is hard to account for why the meaning of the sentence would change if a different conjunction, such as *or*, had been used.

Another observation from the analyses he exemplifies is that Hockett groups verbs and auxiliary verbs together as a constituent, before the object is included in the VP, as can be seen in (32). Based on syntactic evidence, these days, it would be claimed that the object is merged with the verb before auxiliary verbs are. From this point of view, Hockett seems to be reliant on the semantic relatedness of elements to guide his syntactic analysis. Hockett is not at all unusual in this and the habit of analysing verbs with their auxiliaries into a constituent

normally labelled the 'verb group' was still common practise in the 1970s with descriptive linguists. Chomsky, however, had from the beginning argued for an independent syntactic system, related to the semantic component but still separate from it. Therefore, notions of 'semantic relatedness' are not enough to justify a structural analysis. We need syntactic evidence for that.

It is not clear why Hockett does not propose at least some syntactic evidence for his analyses, after all he is clearly aware of notions such as distribution. He talks about similarities of phrases which might have slight deviations from each other in terms of a shared configuration of boxes in his Chinese box representation. He also talks about 'substitution', which is his term for replacing a phrase with a pronoun or having a pronoun take its reference from a phrase. This is a standard syntactic test for constituent structure which is also based on the notion of distribution. It is possible that he just did not see the need for precision, given the relatively low goal of description, rather than explanation, that the linguists of the time strived for. In the introduction to the section on syntax, he writes:

Some linguists like to believe that grammatical analysis has become a completely objective operation, but this is not true. ... grammatical analysis is still, to a surprising extent, an art. ...

Consequently, the reader will find in these sections many an example which the writer has handled in one way, but which might also be handled in some other way. (p 147)

It is clearly this attitude that Chomsky was reacting to. If no attempt is made to provide a sound foundation for syntactic analysis, it is not very likely that one will ever develop. In this respect, Chomsky has been proven correct as we currently have a rather large battery of empirical ways of justifying and testing syntactic analyses.

7 Bloomfield: the father of American Structuralism

In all the above sections we have been reviewing ideas about phrases. Clearly at each time discussed the notion of the phrase was present and considered a central part of syntactic analysis. But we also know that just 50 years before the work of Hockett and Harris, there was no awareness of this structural notion. We now come to the period in which we can definitely say was where the phrase itself was discovered.

In fact, we can put that time somewhere in the 1920s. This is because, the linguist who first came up with the notion wrote two very influential textbooks on linguistics: one in 1914 and one in 1933. In the first, *Introduction to the Study of Language*, he mentions the word 'phrase' three times, all three referring to what we would call an idiom these days rather than a structural unit. In an article also published in 1914 called *Sentence and Word*, he does not mention phrases at all, as the title of the paper seems to promise. It is clear that at this time, Bloomfield, like all other linguists that preceded him, had no notion of the concept. However, in the second textbook, *Language*, he discusses the analysis of sentences into immediate constituents (i.e. phrases). Thus his, by now famous, analysis of the sentence *Poor John ran away*, identifies two main parts: *Poor John* and *ran away*. Obviously, by this time the notion of phrase has entered linguistics and is sufficiently established to be included in a textbook on the subject.

It is quite difficult to determine exactly when the notion was discovered. In a paper published in 1916, *Subject and Predicate*, Bloomfield apparently used the term to refer to what would these days be considered to be a VP:

If, in the course of a philosophical discussion, there occurs the statement, *homo mortalis est*, ... To the logical subject, that talked about and underlying the predication, corresponds the word *homo*, and to the logical predicate, that said about the subject, corresponds the **phrase** [my emphasis – MN] *mortalis est*. (p. 15)

However, it is odd that this only occurs once in the paper and moreover in the same sentence Bloomfield adopts the traditional notion that the subject is a word. Such observations cast doubt on a conclusion that the term phrase has entered syntactic analysis, in the way we understand it today, at this point.

By contrast, in a paper published in 1926, *A Set of Postulates for the Science of Language*, Bloomfield has a section titled ‘Morpheme, Word and Phrase’ in which he offers the following definition of a phrase:

12. Def. A non-minimum free form is a *phrase*.
E.g., *the book*, or *The man beat the dog*; but not, e.g. *book on* (as in *Lay the book on the table*), for this is meaningless, hence not a form; and not *blackbird*, which is a minimum free form.

There is no need to explain this definition here. For our purpose it is enough to note that the notion is being defined and therefore obviously part of linguistic investigation.

Thus, we can say that the phrase was identified in linguistic analysis probably some time in the 1920s. There is also reason to believe that it was Bloomfield who came up with the notion, given that he was the leading figure in American linguistics at the time, and from the fact that he did not attribute the notion to anyone else in his writing.

Yet, Bloomfield did not write extensively on the subject. Even in his textbook, meant to be introducing all aspects of linguistics to students of the subject, the chapter on Syntax amounted to 22 pages of a 566 page book.

One striking thing absent from Bloomfield’s writing on phrases is terms such as ‘verb phrase’ or ‘noun phrase’. The reason for this appears to be Bloomfield’s dislike for the notion of ‘word-classes’ (i.e. categories) because he felt they were ‘impossible to set up’ as a ‘coherent scheme’ as they ‘overlap and cross each other’ (p. 196). He therefore used terms such as *substantive expression* for phrases headed by nouns, though he also uses *nominative expression* for those nominal phrases in the subject position of a finite sentence. This appears to confuse various aspects of syntactic analysis, but the apparent lack of precision is not something he discusses. Elsewhere, he refers to (*finite*) *verbal expressions* which are those phrases headed by a verb.

However, there are problems to be seen with this treatment, which boil down to Bloomfield’s notion of a head. As mentioned above, Bloomfield considered there to be two types of phrases: those with heads (endocentric) and those without (exocentric). Both what we refer to these days as NPs (or DPs) and VPs have instances of both endocentric and exocentric phrases, according to Bloomfield’s definition. While *poor John* and *ran away* are both endocentric, being able to be replaced by *John* and *ran* respectively, phrases such as *the man*

and *hit John* are exocentric as they could not be replaced by any of their contained words. Yet both would be eligible for being called ‘nominative expressions’ and ‘finite verbal expressions’ on the basis of their functions in a sentence such as *the man hit John*. Thus, it appears that we have two ways of categorising phrases, one which is based on properties of heads and one which is based on the function that the phrase has in a bigger phrase. Yet these things are separate notions and should not be confused as being instances of the same sort of thing. For example, would a preposition phrase be eligible for the category ‘nominative expression’ if it appears as the subject of a finite clause, as in *out here is nice*, and would an adjectival phrase be considered as a ‘verbal expression’ despite not containing a verb but because of its predicative usage in *she likes milk very fresh*.

Bloomfield’s notion of a head is also problematic in and of itself. He acknowledges that there is some sort of connection between a head of a phrase, if it has one, and the nature of the phrase itself. But a head is defined only in distributional terms and not in terms of its projective function in the phrase, as it is with X-bar theory. As there are exocentric phrases which have the same distribution as endocentric phrases, these too should have the same category. But if category is determined by heads, an exocentric phrase should have no category at all, let alone the same one as an endocentric phrase that it distributed like.

There is not much said about such issues in Bloomfield’s chapter on syntax, and indeed there is not much said about phrases in general. He did not develop a representation of hierarchical structure, as Hockett did, and in fact he gives relatively few examples, sticking to very simple cases which are unlikely to challenge his ideas. He gives no discussion about how we can know which elements go together to form a phrase, though, like Hockett, he is obviously aware of distributional phenomena and how this links elements in one sentence to elements in others. So, while we can say that Bloomfield came up with the notion of a phrase, it cannot be said that he developed it to any great extent.

The burgeoning questions at this point, though, are where did Bloomfield get the notion of a phrase from and why was he able to come up with it while thousands of years of previous scholarly work had failed to do so? To understand this, we need to go into Bloomfield’s background.

Bloomfield became interested in linguistics as an MA student at Wisconsin university, specifically in German philology and before the first world war, Bloomfield was in Germany studying with the comparative linguists at Leipzig. In America, at that time, the leading influential figure in linguistics was Franz Boas and in Germany Bloomfield attended lectures by the psychologist Wilhelm Wundt. Both of these were instrumental in forming Bloomfield as a linguist.

Boas was originally a geologist, but after going to America to study geology there, he became fascinated by the culture and languages of native Americans and so switched to anthropology. The situation facing native Americans at that time was dire: centuries of oppression and displacement had had its toll and the culture and language of these people was rapidly disappearing. Boas set himself the huge job of recording these for posterity. But the question was how to write the grammars of languages that he didn’t speak? Learning each language was not feasible. On top of this, his anthropological side had turned him against the current extremely racist view that cultures were ranked from ‘primitive’ to ‘advanced’, advanced of course being deemed to be western European. This was the view that had influenced Adolph Hitler and led to the situation that arose in Germany in the 1930s. Against this, Boas

developed the idea of cultural relativism, which meant that cultures could only be understood in their own terms rather than in comparison to other cultures.

As we have seen, linguistics in Europe had, since Roman times, been based on providing analyses of one language based on the grammars developed for others. The linguistic concepts that the Greeks had proposed were recycled first for the study of Latin and then other European languages. Of course, comparative linguistics took this to the highest level, finding connections between languages which had previously been thought to be unrelated. To Boas, however, native American languages were so different from European ones, that he doubted that linguistic notions relevant to Indo-European languages would be relevant at all and so he extended his notion of cultural relativism to his linguistic work. From this perspective, one could not study native American languages from the point of view of Indo-European, and only concepts which were based on these languages could be used in their analysis. Thus, the job of recording their grammars is even more difficult as one has to start from scratch to discover the relevant basic concepts and principles.

In order to achieve this aim, Boas came up with a set of ‘discovery procedures’ which would enable a researcher to apply simple tests to data provided by a native speaker and these would yield the basic units of the language, such as phonemes, morphemes, word categories. Notice that the phrase was not one of these, so as far as that notion was concerned, Boas was like all his other contemporaries. Most of the discovery procedures were based on the notion of distribution and are basically the same as we still use today. For example, Boas was the one who proposed the use of minimal pairs to discover phonological distinctions.

It was this work that Bloomfield would have been brought up on as a student of linguistics in the early part of the 20th century and he has credited Boas as one of his main influences in his work.

Wilhelm Wundt was a psychologist. He is known as the father of experimental psychology and his main aim was to make psychology an accepted member of the scientific studies. The way he saw of doing this was to make the field based on experimentation. Although Wundt was interested in the psychological aspects of language and had a number of things to say about it, this was not so much an influence on Bloomfield, especially as most of Wundt’s experiments were to do with introspection, which tends not to yield very much about a person’s linguistic system as most people are generally unaware of technical aspects of their language. Instead, what Bloomfield took from Wundt was the general desire to make linguistics a scientific study. This, incorporated with Boas’s discovery procedures led Bloomfield to conceive of linguistics in a structured way and when he returned to America, just before the outbreak of the 1914 war, he produced his classic textbook which established him as one of the main linguistic in the US. The school that he founded, American Structuralism, soon became the dominant one on that side of the Atlantic and Bloomfield was very much at its head.

However, as we have seen, at the start Bloomfield had still not come up with the notion of the phrase, and that wasn’t to happen until the 1920s. It hardly seems coincidental that at about this time Bloomfield abandoned his affiliation with Wundtian psychology and became enamoured with the newly emerging school of psychology: Behaviourism. This based its affiliation to scientific study by adopting an extreme form of empiricism called operationalism. This is the view that a scientific theory should not entertain concepts for which it does not have direct evidence for. This was a radical perspective to take in psychology, and in retrospect a rather odd one, as it led to the rejection of the concept of

mind. Instead, the only things we have direct evidence of is actual human behaviour and the environmental conditions which 'trigger' them. Thus, they viewed psychology as the study of Stimuli-Response patterns in human behaviour, thus bypassing the concept of the mind altogether.

While Bloomfield did attempt to incorporate Behaviourist ideas into linguistics, and the first chapter in his 1933 text book talks about language from the perspective of stimuli and response, this has mostly to do with some rather ill informed ideas concerning language acquisition (basically children learn to speak by initially producing random sounds, some of which are rewarded by adult responses and so these turn into words which are associated with aspects of the environment in which the adults reacted positively to the child's vocalisations). Incidentally, Chomsky also brought down the school of Behaviourism at the same time as he was dispensing with American Structuralism via a paper he wrote in 1959 criticising the work of the then leading behaviourist working on language B. F. Skinner.

I think we are now in a position to understand how these influences combined in Bloomfield to give rise to the idea of the phrase. The operationalist stance taken by Behaviourists posed a bit of a problem for Bloomfield as a linguist as much of the linguistic system is not directly observable. We can observe speech directly, but more abstract notions such as phonemes, morpheme, words, etc. cannot be directly observed. This is where Boas's ideas come into play. The discovery procedures, Bloomfield realised, offer a way to base the less observable aspects of language on that which is directly observable. Thus, by the use of discovery procedures based on observable phones, we can observe the phonemes in a language. Also, by observing phonemes and applying distributional tests, we can come to observe morphemes. Thus, Bloomfield envisaged a way of bootstrapping linguistic investigation into the more abstract levels of language based on phonetic observations and the application of discovery procedures. This gives rise to a common view of linguistic levels which was very influential until generative linguistics overwhelmed the area:

(37) phonetics → phonology → morphology → syntax

So, how did this give rise to the discovery of the phrase? The basic step in linguistic investigation according to the picture given in (37) is to discover the units of each level on the basis of the behaviour of the units of the previous level. Thus, we discover the phonemes of a language through the study of that language's phones and how they distribute. We discover the morphemes of a language through the study of its phonemes and how they distribute. We discover the morphemes by studying the distribution of the phonemes and we discover the word from the distribution of morphemes.

Where do we go to from here? The natural step is to see what we get if we study the distributional patterns of words, which of course does not lead us directly to the sentence, but to sub-units of the sentence, i.e. phrases. And there we have it. It was the particular methodology of looking at language as a successive set of levels all built on each other which leads to the notion of a phrase. Clearly, as this only came into being by Bloomfield taking on the ideas of Boas, Wundt and Behaviourism, this was such a specific set of ideas that it is not surprising that it had not occurred previously. But once the notion of the phrase had been proposed, it became so obvious that it didn't require Bloomfield's particular stance on language structure to be maintained. Thus, while modern linguistics has long dropped the structural view of language and the empiricist philosophy that spawned it, the phrase lives on.

8 Conclusion

The notion of the phrase has been about for about 100 years. If we were to fit the roughly 2600 that the serious study of language has been being undertaken into the space of a day, this means that the phrase would have been discovered at about 5 minutes to midnight. Yet today the phrase is taken to be one of the most basic units of linguistic analysis that most people think of it as having been around since time immemorial. However, it took the combination of a specific set of ideas to come together in one place for the notion to become visible and history is such that this only happened relatively late in linguistic investigations. Up to that time, no one was really looking for units that went beyond the word and the units that had been discovered up to that point were so obvious that no one suspected that there would be something else beyond them.

We have seen that since its discovery, the phrase has undergone several bouts of development and re-conceiving. The earliest notions were simply groupings of words, which might have had properties, but they were not prominent enough to categorise. Phrase structure grammars gave phrases properties in the same way that words had been seen to have properties and although there must have been some idea that word and phrase properties were similar, otherwise terms such as 'verb' phrase and 'noun' phrase would not have been used, it wasn't until the proposal of X-bar theory, 50 years after the first proposal of the phrase, that the relationship between word categories and phrase categories began to be understood. In the 50 years that have followed this, more development had taken place, mainly generalising the notion so that by now one cannot move without bumping into a phrase: they are everywhere. The current view has not so much changed the view of the phrase so much as changed the view of how they come into being. Instead of being specific rules which create them for the purposes of putting words together, it is the process of putting words together which give rise to them. This view has been with us for just over 20 years and not so much development has changed it since. Perhaps then, we have reached our final understanding of what a phrase is. But in science, it is hard to say where the end to understanding lies.