

On Compounds

1 Compounds, packaging two or more words in one, are a problem of modern linguistics.

2.1 However, in some ways the problem of compounds also lies at the roots of modern linguistics. Usually, the birth of modern linguistics is dated around 1800, with the rise of comparative linguistics and the decline of the 'cartesian' linguistics, i.e., the linguistics of Port-Royal where language was embedded in logical structure. Under the influence of Locke's philosophy, some philosophers of the late 18th century tried to remold the achievements of Port-Royal linguistics. Instead of logical relations, a combinatory structure of 'primitives' that were available to the senses were sought for.¹ A combination of 'primitives', of course, is nothing less - grammatically spoken - than a compound of simpler words. It was J. Horne Tooke, who in 1786 first thought of applying this principle to all non-primitive words in language.² Therefore, he tried - sometimes in rather imaginative though phantasmagorical ways - to reconstruct older versions of non-decomposable words, especially of function words³, that he could decompose into 'primitives'.

2.2 This idea was taken up by German philosophers as Fichte⁴, who imagined a scheme of language development in which original aoristic compounds of 'doing' (verbs) and 'things' (nouns) gradually decomposed into more analytical and abstract words such as function words. A far more detailed and powerful theory along these lines was written down by A.F. Bernhardi.⁵ A next generation, better schooled in the classical languages due to F.A. Wolf, more familiar with historical strands of language that got more or less critically edited and also with notions of Sanskrit as the 19th century started⁶, took up these philosophical musings and provided the first conjectures and derivations based on factual evidence.

¹See, e.g., J. H. Lambert (1764): *Neues Organon oder Gedanken über die Erforschung und Bezeichnung des Wahren und dessen Unterscheidung vom Irrtum und Schein*. Johann Wendler, Leipzig.

²J. H. Tooke (1786): *ETIENNA PUTEOLANA or, the Diversion of Purley. Part I*. Printed for the Author, London.

³Function words, e.g., *the, he, can, do, before, etc.*, can be characterised by two correlative properties: 1) They occur much more frequently at large than other words in a coherent piece of language; 2) they belong to a so called closed class of words, i.e., one can expand other word classes by derivation, compounding etc., but function words all belong to strictly limited classes.

⁴J. G. Fichte (1795): Von der Sprachfihigkeit und dem Ursprunge der Sprache. *Philosophisches Journal* (1), 255-273 & 287-326.

⁵A. F. Bernhardi (1805): *Anfangsgründe der Sprachwissenschaft*, Frlich, Berlin.

⁶Knowledge of Sanskrit was in 1800 still limited to England and Paris (Jesuits), only in the beginnings of the 19th century did German scholars either go to Paris (Bopp), or could they consult a grammar (e.g. the one by Colebrooke).

2.3 A simple though important example of this figure of thought is the following, consider the Latin conjunction of *videre* (to see):

vide-o; vide-s; vide-t; vide-mus; vide-tis; vide-unt

and the corresponding forms of *esse* (to be):

sum; es; est; sumus; estis; sunt

It was conjectured that the endings of conjunction historically evolved of compounding a root with forms of *esse*⁷:

vide-sum; vide-es; vide-est; vide-sumus; vide-estis; vide-sunt

Taking into account some old forms and hoping the holes in the reasoning would eventually be closed by new evidence, some of the basic ideas of comparative linguistics stem from these roots, such as, sound change, derivation of roots etc.

3.1 In the other half of the 20th century, compounds return, as a tricky problem for structural linguistics. The problem is already diagnosed in a footnote in the brilliant, though often forgotten (in the shadow of the fast coming Chomskyan linguistics) study on the structure of English by C.C. Fries. One of the few studies, as Harvey Sacks remarked, that seriously enter into conversation as the fundamental environment of language, *the Structure of English*⁸ brings clearcut description of the analytical - if not pidgin-like - structure of English, bringing to the foreground the combinatorics applied to the atomic words together with their restrictions. Ordering word classes by their combinatory properties, Fries, however, stumbles over the Class 1 Words (nouns) with Class 1 modifiers, i.e., compounds, whose grammatical bond he semantically describes as "identification". In the note though:

The label "identification" is not a very satisfactory term for the somewhat diverse meanings generally signalled by this structure. I have not been able to find formal features that differentiate the more precise meanings included here (p. 224)

The matching of the structural patterns with the semantic patterns - in Fries's terms - therefore remains particularly vulnerable.

3.2 More prominent, and - if I may stress - near symptomatic does this problem become in N. Chomsky's *Generative Grammar*. Conceived as a scheme to generate linguistic structures - and it must be remarked: structures, not meaning - Chomsky developed in successive steps a grammatical calculus that can account for the complex combination and positioning of elements (words or clauses).⁹ The fundamental idea is a finite set of words and a finite set of basic grammatical structures that can be transformed according to algebraic rules and thus form complex patterns. A simple example is the following:

⁷Traditionally this idea is credited to Bopp with his knowledge of Sankrit roots, but is mentioned years before Bopp by Boeckh and Grimm. Compare A. Boeckh (1808): *Von dem Uebergange der Buchstaben in einander. Ein Beitrag zur Philosophie der Sprache, Studien* (4), 365-396. J. Grimm (1813): *Grammatische Ansichten, Altdeutsche Wälder* (1), 179-187. Bopp (1816): *Über das Conjugationssystem*. Berlin.

⁸C. C. Fries (1952): *The structure of English*. Longman, London.

⁹N. Chomsky (1956): *Three Models for the Description of Language. IRE Transactions on Information Theory* Vol. 2 Nr. 3, S. 113-124; (1957) *Syntactic Structures*. Mouton, The Hague, Paris.

a man eats a banana → a banana is eaten by a man

An example that handles subclauses:

It is late. It is dark → It is late and dark.

He saw her. She was beautiful. → He saw her, who was beautiful OR He saw the beautiful girl OR He saw her, because she was beautiful etc.

It may be seen from this example that for every set of elements, many transformations can be made. In the latter two alternatives, I must admit to cheating a little, since "He saw the beautiful girl" introduces the element 'girl' corresponding to 'she' which is a matter of lexicon; in the same vein, I added a causal relationship in the last transformation by altering the neutral 'who' relative clause into a causal clause.

3.3 Compounds can of course be handled by this grammatical calculus:

a dancing girl → a girl, who dances

outcome → what comes out

ice box → a box of ice

However, the transformation of a compound into the right clause is by no means trivial. Whereas the above examples became many and the transformation ambiguous because I implicitly introduced new elements, a compound always carries an implicit element in it. Fries summarised it as a relationship of 'identification', but indeed, many more semantic relationships are possible.

ice box → a box of ice

ice age → age of ice

ice pick → a pick to grind ice

ice top → a top of an ice mountain etc.

This shows indeed clearly the intrusion of semantics in the structural field of formal patterns. Of course, one may want to solve the problems by constructing lexical classes, that, when combined in a compound, generate certain clauses, but this task has shown to be near endless. Excepting some large and easy to resolve groups, the classes amounted to being but mere lists without giving much insight or suggesting structure. In the end, the classifying of compounds into classes led to so many classes that it may be said that near every compound is a class in itself, the classes mere individuals.

4.1 Jumping backwards and forwards in history, we arrive in 1897. In this year the linguist Jacobi published a controversial booklet: *Compositum und Nebensatz*.¹⁰ In the booklet he tackled the hitherto neglected problem of Indo-European syntax, i.e., the construction of sentences and clauses through grammatical signals, among others, word order. He remarked that the order of words in a compound corresponds to the word order in a subclause. Thus, 'tire-bouchon' corresponds to 'qu'il tire le bouchon' and 'Korkenzieher' to 'dass er

¹⁰H. Jacobi (1897): *Compositum und Nebensatz. Studien ber die Indogermanische Sprachentwicklung*. Friedrich Cohen, Bonn.

den Kork zieht'.¹¹ On the basis of these observations, Jacobi conjectured that compounds are the historically older grammatical device, and that subclauses gradually were derived from complex compounds.

4.2 Jacobi's theory is now generally considered false, but he probably came to it through a historical phenomenon in the Sanskrit language. Namely, the grammatical device of compounding came to a particular popularity in classical Sanskrit texts, combining upto 5 or even more words in one compound and thus avoiding subclauses and conjugated verbs.¹² This compounding helped to phrase the content of the Vedic sutras in the most succinct manner possible. In the oral transmission of the sutras, this entailed two advantages: First it optimised memory requirements, second, there was also an intricate system of noun substitution, several nouns denoting aspects of the same to be transmitted semantic entity¹³, so that the fine tuning of the metrum (also a memory aiding device) got easier. This aspect of compounds, as a part of message encoding for long time transmission, might be one of the possible roads linguistics still has to explore.

4.3 However, Jacobi's observations do not lack any truth. Near a hundred years later, his theories were re-discovered by W.P. Lehmann as a study into the universals of syntax.¹⁴ In the light of Chomsky's theory, the study of syntax (instead of phonetics and lexicon) became a new focus for linguists, and after Ginsberg's modification, became a quest for universals of language on the level of syntax. Therefore, Jacobi's observations on word order in compounds and subclauses got re-integrated into linguistics.

5.1 What should be the conclusion of these ramblings into linguistics, its history and problems? As Roman Jakobson repeatedly stated: Meaning is and remains the central problem of linguistics. The 19th century focussed on lexicon, later on sounds and phonemes, the late 20th century on syntax. Semantics remained at the borders, and it seems that upto the present day, we have no viable theory, nay, not even a clear statement of what should be such a theory, nor a clear formulation of the conditions under which meaning functions.

¹¹Remark: the word order in a subclause need not necessarily correspond to the word order in a main clause.

¹²Cfr. W. D. Whitney (1889): *Sanskrit Grammar including both the Classical Language, and the older Dialects, of Veda and Brahmana*. Reprint 2004, Munshiram Manoharlal, Delhi.

¹³In this framework, it is also noteworthy, that in Indian mathematics this memory aiding device was used for numbers, who could be expressed through several nouns, thus admitting more variation in the phrasing, according to the metrum.

¹⁴W. P. Lehmann (1974): *Proto-Indo-European Syntax*. University of Texas Press, Austin and London.

[There are at present] no techniques for the structuring of meaning that would allow it to be related usefully to the formal structure of language.¹⁵

Statement, which, unfortunately, still is valid today.

5.2 Returning to the compounds, what makes them appear irritatingly at crucial moments in the history of linguistics? It may be because they question our fundamental concept of a word. Be it syntax, a combinatorics of words, be it sound changes, in words, be it lexicon, a list of words, thus far, all of linguistics could not do without the word. Fries, and many before him, starts his *Structure of English* with a radical questioning of the concept word, but in the end cannot do without. His definition amounts to a 'block' consisting of sounds that functions independently within speech, i.e., speech signals that have a grammatical function affect these 'blocks' in a singular fashion and differently than other 'blocks'. E.g. they have a stress pattern differentiating it from other 'blocks', they can be moved around in a sentence without taking with them other 'blocks', they change differentially due to grammatical devices such as case, conjunction. Though neither of these characteristics alone suffice to have a word, the combination of all of them points to a word.

5.3 Exactly these properties serve also to discern compounds from words next to each other. Thus this question becomes fundamental to a classification of combinatory rules and to discerning elements – two formalised features of language that are intrinsically and perhaps unavoidably linked with each other. One such a contested essay was Biggs's analysis of Moari, discarded today, but still interesting for the research program proposed:

To isolate all significant units that enter in to the composition of the code we call Maaori, and to state all possible sequences of these units.

Meaningful units ... are classified according exhaustively according to the positions they may occupy with reference to other units of the same level. This classification plus the statement of the combinatorial possibilities, insofar as it is successful, allows for the generation for the near infinite generation of a near infinite number of utterance partials.¹⁶

This program is in so far near to Chomsky's issues, that it attempts to construct a finite basis from which to generate utterances – which is the difference with the linguists of the 19th and early 20th century, who focussed on description, not generation. It differs, however, from the generative grammars, because it tries to attack not only grammatical or sequential combinatorial rules *an und für sich* (and use the lexicon as a static database), but attempts at merging combinatorial rules with the build-up of the lexicon. This being no trivial task, it does not surprise that linguists now consider Biggs's effort to be too subjective, too arbitrary.

¹⁵B. Biggs (1961): The Structure of New Zealand Maaori. *Anthropological Linguistics* 3 (3), 1-54, p. 5.

¹⁶Biggs, Structure, 5.

6 Conclusion? Rub out the word – and we still await fresh insights in language. Compounds remain a problem of modern linguistics.

17.09.2006

Maarten Bullynck

www.kuttaka.org