

В А П Р О С Ы Я З Ы К О З Н А Н И Я

МЕГАСБОРНИК НАНОСТАТЕЙ

13 СЕНТЯБРЯ

2020

УДК 81-13
ББК 81
В12

В12 ВАПросы языкознания: Мегасборник наностатей. Сб. ст. к юбилею В. А. Плуногяна / Ред. А. А. Кибрик, Кс. П. Семёнова, Д. В. Сичинава, С. Г. Татевосов, А. Ю. Урманчиева. — М.: «Буки Веди», 2020. — 684 с. — ISBN 978-5-4465-2882-0.

В сборник, приуроченный к юбилею Владимира Александровича Плуногяна, вошли краткие этюды, посвящённые разным областям лингвистики — среди них морфология, стиховедение, семантика и прагматика, корпусная лексикография, ареальная лингвистика, типология, разнообразные глагольные категории, синтаксис и дискурс. Все эти сюжеты так или иначе связаны с многосторонней научной деятельностью юбиляра.

ББК 81

В оформлении переплётa использована репродукция картины Клода Моне «Антиб, старая крепость» (1888)

ISBN 978-5-4465-2882-0



9 785446 528820 >

© Коллектив авторов, 2020

On the difference between a clause and a sentence

Martin Haspelmath

Max Planck Institute for Evolutionary Anthropology / Leipzig University;
haspelmath@gmail.com

1. A definition of ‘sentence’

Clause and *sentence* are two terms that linguists use all the time, but what exactly do they mean by these terms? An informal discussion on Facebook a few months ago confirmed to me that many linguists are unsure how these terms should be understood and defined. These terms and the possible differences are almost never discussed, but there is widespread agreement that they mean different things. But what exactly?

There may not be any need to agree on precise meanings of highly general words like “language” or “question”, but terms like “clause” or “affix” clearly belong to the technical part of our vocabulary. And technical terminology should be uniform across a discipline, so here I present some thoughts about how to define these terms. I will also comment on the term “main clause” (below in §5), because this term also causes confusion.

So first of all, here is my proposal for the definition of the term *sentence*:

- (1) A sentence is a maximal clause, i.e. a clause that is not part of another clause.

It seems to me that this definition will not cause controversies, but it crucially relies on the term *clause*, which we need to understand in order to understand what a sentence is. It is less straightforward to give a definition of *clause* that satisfies most linguists, i.e. that conforms to our intuitions. Here are two possibilities:

- (2) a. A clause is a combination of a predicate (full verb or nonverbal predicate) and its arguments, plus modifiers.
- b. A clause is a syntactic unit that can be independently negated (cf. Haspelmath 2010, Haspelmath 2016)

If we are not happy with either of these definitions, we can take “clause” as a primitive notion. As Anna Wierzbicka has reminded us (in many works,

e.g. in Wierzbicka 1996), any attempt at consistent definition must ultimately work with undefinable (= primitive) words. So it may be that “clause” will have to be one of these undefinable concepts that we simply need to take for granted in linguistics. I will not propose a decision here, and limit myself to noting that the definition in (1) provides a clear answer to the question about the difference between a sentence and a clause.¹

However, in actual practice, linguists almost never define a sentence as in (1) (I do not remember having seen a definition along these lines). Let us consider four possible alternatives.

2. Four possible alternatives

2.1. Is a sentence a possibly complex clause?

One intuition that one often finds is that a clause is “simple”, while a sentence may be “complex”, i.e. consist of multiple clauses. This is reflected by the definition in SIL’s *Glossary of Linguistic Terms*:²

(3) “A sentence is a grammatical unit that is composed of one or more clauses.”

However, this will not work as a definition, because a clause may itself be “complex”, as everyone agrees. For example, a relative clause may contain an adverbial clause, as in (4a), and an adverbial clause may contain a relative clause, as in (4b).

(4) a. *the pen [that I bought [after my pencil broke]]*

b. *[when I bought the pen [that I am now using]]*

Someone might suggest that the situation is different when two clauses are conjoined, because in that case, one clause does not contain the other. However, conjoining is usually thought to create constituents of the same type — two conjoined nominals are still a nominal, two conjoined adpositional phrases are still an adpositional phrase, and so on. Thus, two conjoined clauses should still be a clause. So the definition in (3) does not conform to the way we actually use the term “sentence”, while the definition in (1) does.

¹ This also means that I am ignoring here the issue of whether discourse can be divided into clauses in the same way in all languages (cf. Mithun’s (2005) discussion, prompted by a study of Hualapai discourse).

² <https://glossary.sil.org/term/sentence> (2020 May 26)

2.2. Is a clause a CP?

Many linguists think that the technical vocabulary that is taught in many syntax courses provides a solution — after all, why have this technical vocabulary if it does not add precision? Unfortunately, however, in the case at hand it certainly does not help to say, for example, that a clause is a “CP”, while a sentence is an “S” (as is done in Larson 2010: 293). In that textbook, Larson uses the term “sentence” in the first half of the book, but switches to “CP” in the second half. At some point, he replaces “S” by “TP”, but then later he keeps using “sentence” for what in his analysis is really a CP.

Larson’s book would have been less confusing if he had used only “clause” or only “sentence”, but this would not have conformed to general terminological usage. All linguists sometimes use the term “clause” (a combination such as “subordinate sentence” sounds quite strange), and most also use the term “sentence” in some situations. But in generative syntax, neither “CP” nor “TP” correspond closely to “clause” or “sentence”. By contrast, the definition in (1) above corresponds quite well to standard usage.

2.3. Is a sentence a discourse unit?

In the Facebook discussion, Alexandre François made the proposal that while “clause” is a syntactic concept, “a sentence is a unit of discourse: It is a self-contained unit of speech endowed with one speech act / illocutionary force (declarative, interrogative, imperative...)” This definition is similar to the one given by Dictionary.com.³

- (5) **sentence**: a grammatical unit of one or more words that expresses an independent statement, question, request, command, exclamation, etc., and that typically has a subject as well as a predicate, as in *John is here.* or *Is John here?* In print or writing, a sentence typically begins with a capital letter and ends with appropriate punctuation; in speech it displays recognizable, communicative intonation patterns and is often marked by preceding and following pauses.

This sense of “sentence” is actually quite similar to the definition that I proposed in (1) because illocutionary force (i.e. the difference between the speech act types: statements, questions, and directives) is not usually taken as a property of a unit that can be part of a sentence. A unit that expresses a directive cannot be a part of another sentence: In the example *She told me to go*

³ <https://www.dictionary.com/browse/sentence> (2020 May 26)

home, the expression *to go home* is part of another sentence, but it is not thought to be a speech act by itself. And embedded questions (as in *He asked me where I lived*) may be said to be questions, but not in the sense of speech acts. Similarly, it is not usually possible to conjoin units with different illocutionary forces, e.g. questions and statements (*?*It is raining and are we going out?*).

The main reason why I think that the definition in (1) is better than the one in (5) is that it specifies that a sentence is a kind of clause. By contrast, a definition in terms of a discourse unit and illocutionary force would include expressions such as *hello*, or *ouch* when used as complete utterances, and these are not normally treated as clauses. Likewise, they are not normally treated as sentences, so I feel that a definition of a sentence as a maximal clause corresponds best to ordinary usage.

2.4. Is a sentence a clause plus dislocated elements?

Van Valin and LaPolla (1997: 38) say that a sentence consists of a clause plus a pre-clausal position for left-dislocated elements, and a post-clausal position for right-dislocated elements, as in (6).

(6) *I have not seen them in two weeks, the Smiths.*

Again, this definition is quite similar to my proposed definition, because left dislocation and right dislocation are not normally thought to be possible in embedded clauses. But as a universally applicable definition, it faces the problem that the notions of “left dislocation” and “right dislocation” are not easily applicable to different languages. Van Valin and LaPolla say that English allows a left-dislocated element before the position of the question word, as in (7).

(7) *Yesterday, what did Robin show to Pat in the library?*

But in German, the counterpart of this would be completely impossible. So does German have “left dislocation” in the same sense as English? This is unclear, but a definition of “sentence” must be valid for all languages and cannot make reference to concepts that are specific to particular languages.

3. “Complex sentences” are really “complex clauses”

The idea of a sentence as a possibly complex clause is very widespread, and the term “complex sentence” has been used very widely, both in comparative studies of the world’s languages (e.g. Austin (ed.) 1988, Bybee, Noonan (eds.) 2002) and in psycholinguistic studies of English (e.g. Diessel 2004). It seems

that in all these cases, a “complex” clause/sentence really means “clause that includes at least one other clause”.

Some authors (e.g. Culicover 2009: Chapter 7) use the more accurate term *complex clause*. And this is probably what the authors mentioned in the previous paragraph actually mean. Clauses such as those in (4a–b) above are presumably not excluded just because they are not maximal clauses. Thus, the term “complex sentence” should be avoided unless it actually refers to a maximal clause which contains a clause.

4. “Sentence” has no counterpart in nominal syntax

It has often been suggested that the internal structure of clauses and the internal structure of nominal phrases show similarities, so one may ask whether there is a counterpart to the term “sentence” in nominal syntax. The answer is no. We have no special term for a maximal nominal, i.e. a nominal that is not part of another nominal. Just like clauses (which may contain other clauses), nominals may be recursive and contain other nominals. Everyone knows this, but nobody seems to miss a term for a maximal nominal.

Thus, I suspect that we wouldn’t miss the term *sentence* if it did not exist, because every sentence is a clause, and we rarely need to talk about maximal clauses in contrast to non-maximal clauses.

In fact, German does not make the distinction between “clause” and “sentence” and uses *Satz* for both notions. When I speak about grammar in German, I do not miss the distinction. In Russian, too, the term *кляуза* was added to the grammatical terminology only fairly recently, and it is still quite normal to use *предложение* also for subordinate clauses (i.e. non-maximal clauses which are never described as sentences in English).

5. What is a “main clause” and a “matrix clause”?

5.1. Main clause

Many authors distinguish between subordinate (or embedded) clauses and main clauses, but what exactly is a main clause? Of course, we all know stereotypical examples such as (8).

(8) *She e-mailed me [because she didn’t have my phone number].*

The sentence in (8) contains a causal subordinate clause, and if a student in an examination said that the part “*She e-mailed me*” is the main clause, we would not say that this is wrong. But in fact, we also all agree that a subordi-

nate/embedded clause is part of another clause, so there are really only two clauses here: the subordinate clause, and the full clause (which contains the subordinate clause).

In the Facebook discussion, Rob Truswell noted that linguists often talk in a sloppy way, as if in such complex sentences there were two non-overlapping clauses. For example, in a discussion of example (9a), linguists would typically say that the interrogative pronoun *what* has been fronted from the subordinate clause into the main clause.

(9) a. *What do you think [that Mary ate]?*

b. *I think [that Mary ate pasta].*

But of course, in (9b), the nominal *pasta* is not only part of the subordinate clause, but also of the full clause. It is therefore not accurate to say that *what* “has been fronted into the main clause”. So what do we call the part of the clause that the question word *what* has been moved into in this example?

Before getting to an answer to this question, we need to consider yet another term: the *matrix clause*.

5.2. Matrix clause and matrix-clause fragment

What is the difference between a “main clause” and a “matrix clause”? In Trasks’s *Dictionary of English grammar* (2000), we read:

(10) “A matrix clause is often a main clause ..., but it need not be: it can itself be a subordinate clause. In the sentence *The victim told the police [that the man [who attacked her] had had a beard]*, the subordinate clause *who attacked her* is contained within the subordinate clause *that the man ... had had a beard.*”

I am not sure how frequently the distinction between *main clause* and *matrix clause* is made, but it could be made, and in fact it should be made unless one wants to allow for the possibility of a main clause that is itself a subordinate clause — and this does sound contradictory.

Thus, in (9a) above (*What do you think [that Mary ate]?*), the question word *what* has been fronted out of the subordinate clause and is now included in the part of the matrix clause that is not the subordinate clause. For this part, it is useful to resort to the special term *matrix-clause fragment* (= the part of the matrix clause that is not the subordinate clause).

This may sound like a strange term, but if we want terminological consistency (as we should if we aspire to rigorous science), then I see no way around it.

Note that I did not invent the term: The grammatical glossary of the Institut für Deutsche Sprache in Mannheim⁴ contains the term *Hauptsatzfragment* (*Haupt-satz-fragment* [main-clause-fragment]), which is defined in roughly this sense.⁵

5.3. Main clause again

So we now have a definition of *matrix clause* and *matrix-clause fragment*, and we can describe the fronting in (9a) more accurately: The interrogative pronoun *what* was fronted into the matrix-clause fragment. But we still have no definition of *main clause*. Is a *main clause* the same thing as “a *matrix clause* that is not part of another clause”? Is it an “independent matrix clause”, i.e. a sentence that contains another clause? I do not think so.

The Wikipedia entry *independent clause*⁶ equates *main clause* with *independent clause*, and it seems to claim that an independent clause is not part of another clause (“can stand by itself”). The SIL Glossary’s definition of *main clause* is similar (“may stand alone as a complete sentence”).⁷

I think that this is correct, and I would simply define a *main clause* as “a clause that is not a subordinate clause”. (This ignores the distinction between subordinate and dependent clauses that I made in Haspelmath 1995, but this would take us too far afield.) This does not accord perfectly well with current usage (where a main clause is usually contrasted to a subordinate clause that is contained in it), but it is the only coherent sense that I was able to give to this term.

6. Summary

The definitions proposed here are summarized in the following:

- **clause**: A clause is a combination of a predicate (full verb or nonverbal predicate) and its arguments, plus modifiers.
- **sentence**: A sentence is a maximal clause, i.e. a clause that is not part of another clause.
- **complex clause** (= “complex sentence”): A complex clause is a clause that contains at least one other clause.

⁴ https://en.wikipedia.org/wiki/IDS_%E2%80%93_Institute_of_German_Language (2020 May 26)

⁵ <https://grammis.ids-mannheim.de/terminologie/97> (2020 May 26)

⁶ https://en.wikipedia.org/wiki/Independent_clause (2020 May 26)

⁷ <https://glossary.sil.org/term/main-clause>

- **subordinate (= embedded) clause:** A subordinate/embedded clause is a clause that is an argument (“complement clause”), and adnominal modifier (“relative clause”), or an adverbial modifier (“adverbial clause”) (but see Haspelmath 1995).
- **matrix clause** (of subordinate clause S): The matrix clause of subordinate clause S is the minimal clause that contains it.
- **matrix-clause fragment:** A matrix-clause fragment is the part of a matrix clause that is not the subordinate clause.
- **main clause:** A main clause is a clause that is not a subordinate clause.

It required some effort to arrive at these definitions, but I think that this exercise has shown that it is possible to give reasonably simple and clear definitions of grammatical terms. Part of my inspiration for this sort of work came from Mel’čuk (1982), who was an important member of the Moscow school of lexical semantics.

Acknowledgements

I am grateful to several colleagues who commented on my Facebook query (<https://www.facebook.com/martin.haspelmath/posts/10217026915437463>), as well as on a subsequent blogpost (<https://dlc.hypotheses.org/1725>), on which the present paper is based.

References

- Austin P. (ed.) 1988. *Complex sentence constructions in Australian languages*. Amsterdam: Benjamins.
- Bybee J. L., Noonan M. (eds.) 2002. *Complex sentences in grammar and discourse: Essays in honor of Sandra A. Thompson*. Amsterdam: Benjamins.
- Culicover P. W. 2009. *Natural language syntax*. Oxford: Oxford University Press.
- Diessel H. 2004. *The acquisition of complex sentences*. Cambridge: Cambridge University Press.
- Haspelmath M. 1995. The converb as a cross-linguistically valid category. In: M. Haspelmath, E. König (eds.). *Converbs in cross-linguistic perspective*. Berlin: Mouton de Gruyter, 1–56. <https://zenodo.org/record/227108>.
- Haspelmath M. 2010. The interplay between comparative concepts and descriptive categories (Reply to Newmeyer). *Language*, v. 86, n. 3, 696–699.
- Haspelmath M. 2016. The serial verb construction: Comparative concept and cross-linguistic generalizations. *Language and Linguistics*, v. 17, n. 3, 291–319. <https://doi.org/10.1177/2397002215626895>.
- Larson R. K. 2010. *Grammar as science*. Cambridge, MA: MIT Press.
- Mel’čuk I. A. 1982. *Towards a language of linguistics: A system of formal notions for theoretical morphology*. Munich: Fink.

- Mithun M. 2005. On the assumption of the sentence as the basic unit of syntactic structure. In: Z. Frajzyngier, A. Hodges, D. S. Rood (eds.). *Linguistic diversity and language theories*. Amsterdam: Benjamins, 169–183.
- Trask R. L. 2000. *The Penguin dictionary of English grammar*. Harmondsworth, Middlesex: Penguin Books.
- Van Valin R. D. Jr., LaPolla R. J. 1997. *Syntax: Structure, meaning, and function*. Cambridge: Cambridge University Press.
- Wierzbicka A. 1996. *Semantics: Primes and universals*. Oxford: Oxford University Press.