

## A STUDY OF SYNTACTIC BOOKKEEPING DURING SENTENCE PRODUCTION

GERARD KEMPEN

Experimental Psychology Unit  
University of Nijmegen

**SUMMARY.** It is an important feature of the human sentence production system that semantic and syntactic processes may overlap in time and do not proceed strictly serially. That is, the process of building the syntactic form of an utterance does not always wait until the complete semantic content for that utterance has been decided upon. On the contrary, speakers will often start pronouncing the first words of a sentence while still working on further details of its semantic content. An important advantage is memory economy. Semantic and syntactic fragments do not have to occupy working memory until complete semantic and syntactic structures for an utterance have been computed. Instead, each semantic and syntactic fragment is processed as soon as possible and is kept in working memory for a minimum period of time. This raises the question of how the sentence production system can maintain syntactic coherence across syntactic fragments. Presumably there are processes of "syntactic bookkeeping" which (1) store in working memory those syntactic properties of a fragmentary sentence which are needed to eliminate ungrammatical continuations, and (2) check whether a prospective continuation is indeed compatible with the sentence constructed so far. In reaction time experiments where subjects described, under time pressure, simple static pictures of an action performed by an actor, the second aspect of syntactic bookkeeping could be demonstrated. This evidence is used for modelling bookkeeping processes as part of a computational sentence generator which aims at simulating the syntactic operations people carry out during spontaneous speech.

The cognitive processes underlying sentence production may roughly be categorized under the headings of content, form and sound. One group of activities determines the conceptual (semantic) content for language utterances. They select to-be-verbalized conceptual structures in such a way that they are 'digestible' for the listener, that is, comprehensible, interesting, not too redundant, etc. The total conceptual content is split up into a sequence of messages which are each expressible in a complete or partial sentence. These and related activities may be termed conceptualizing. A second group of processes takes care of translating conceptual messages into syntactic form. This I call formulating. Finally, syntactic structures built by the formulating system are handed over to the mechanisms of speech for overt articulation.

### CONCEPTUALIZING, FORMULATING AND ARTICULATING: PARALLEL PROCESSES

What do we know about the temporal relationships between the three sub-processes of speaking? The common view, implicitly held by many students of sentence production, is that they are ordered strictly serially in time. First the conceptual content is fully specified by the conceptualization process. Next, the syntactic structure is built for the whole utterance. Finally, this structure is realized phonetically (cf. Figure 1a). This view is too narrow, though. Most

speakers will often experience situations where they initiate overt speech production after having worked out only a fragment of the conceptual content of the resulting utterance. Such cases force us to give up the strictly serial view in favor of the position that the three subprocesses run in parallel (Figure 1b). As soon as a fragment of conceptual content has been computed it is passed over to the formulator which attempts to translate it into a sentence fragment that is then articulated. In the meantime work on further conceptual and syntactic fragments continues. Figure 1b also shows

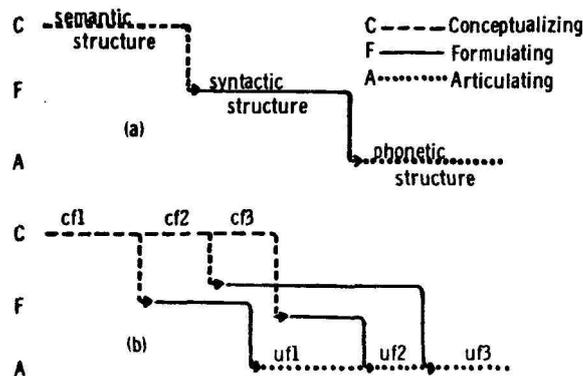


Figure 1. Two theoretically possible alignments of conceptualizing, formulating and articulating processes ( cf=conceptual fragment, uf=utterance fragment).

that the order of conceptual fragments does not always correspond to order of utterance fragments. This is caused by rates of syntax which may call for a reversal. Below (section 3) we'll see an example.

An important advantage of the parallel setup of the sentence production process is memory economy. Conceptual and syntactic fragments don't have to occupy working memory until complete conceptual and syntactic structures for an utterance have been computed. Instead, each conceptual and syntactic fragment is processed (formulated and articulated, respectively) as soon as possible and is kept in working memory for a minimum period of time. However, it also has a drawback since special processing is needed for maintaining syntactic coherence across successive sentence fragments. I call these processes syntactic bookkeeping.

#### SOME THEORY ON SYNTACTIC BOOKKEEPING

The computational model of sentence production that is being developed at the University of Nijmegen generates sentence (1) by constructing a tree

(1) Tonnie baked two cakes which looks roughly like Figure 2. Starting out from a conceptual representation of the meaning underlying (1), it looks up in the mental lexicon a number of words which each express part of the meaning. This process is called lexicalization. Lexical entries possess an internal structure akin to case or functional grammar (cf. DIK, 1978). That is, they contain a head and zero or more dependent cases. E.g. a verb entry (V) displays at least a subject (Subj) and.

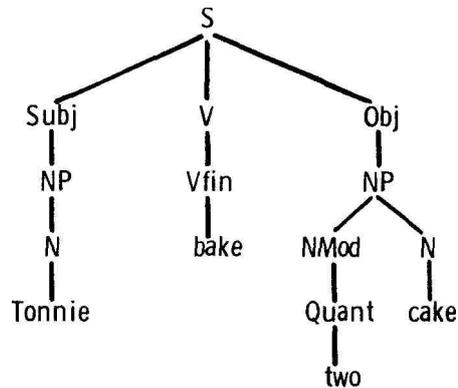


Figure 2. Tree corresponding to sentence (1), with syntactic morphology left out.  
NMod=noun modifier, Vfin=finite verb.

if transitive, also a direct and possibly an indirect object (Obj, IObj). Lexical entries come in various flavors depending on the part of speech the head word belongs to (Verb, Noun, Adjective, Preposition, Conjunction; V, N, Adj, Prep, Conj, etc.).

The case structure in a lexical entry guides the tree construction process to build all constituents which are needed in the syntactic environment of the head word. E.g. if the lexicalization process under an S selects an entry of type V, it is the Subj and Obj listings in this entry which cause the tree construction process to provide S with proper subject and object branches.

Two properties of the tree construction process are critical in the context of this paper. First, the various branches of a tree (e.g. Subj, V, and Obj in Figure 2) are constructed basically in parallel and do not wait for each other to finish. This also applies when the meaning becomes available in bits and pieces. Suppose the meaning underlying sentence (1) is conceptualized by the speaker in two conceptual fragments: first the actor, next the action he performs. Then the formulator would first work on the Subj branch in Figure 2 and hand it over to articulatory mechanisms for overt pronunciation. Work on the V and Obj branches begins as soon as the action fragment is ready, irrespective of whether or not the Subj branch got finished in the meantime. I also make the assumption that overt pronunciation of words belonging to a syntactic fragment is initiated only after the underlying conceptual fragment has been fully formulated. In terms of

the example, pronunciation of baked will have to wait till both the V and Obj branches are complete. This assumption also applies to conceptualized structures which translate into whole sentences. (The rule is almost certainly too strict, though, witness occasional word finding difficulties most speakers have. But this is irrelevant here.)

Second, when suspending branches from nodes, the tree construction process is subject to certain syntactic constraints. Hanging a V-entry directly from an S-node is permitted but not a N- or an Adj-entry, for example. In order to determine whether a looked-up lexical entry is a legal dependent ("daughter") of a given node, a checklist such as in Figure 3 is consulted. To the right of each colon a list of legal daughter nodes is specified for the node at the left-hand side. Notice that the lists are not exhaustive. For instance, the first row does not mention Obj as a legal daughter of S. The explanation is simple: if a candidate lexical entry is of a legal type (e.g. V under S), then not only the head word is suspended but also the cases which accompany the head (e.g. Subj and Obj if V is transitive). The central feature of the checklist, however, is the following.

S:	<u>V</u> , SMod*, Subj
SMod:	<u>Adv</u> , <u>Prep</u> , <u>Conj</u>
Subj:	NP
NP:	<u>N</u> , NMod*
NMod:	AdjP, <u>Quant</u> , <u>V</u>
AdjP:	<u>Adj</u> , SMod*
Obj:	NP
Iobj:	NP
etc.	

Figure 3. Rules for checking syntactic appropriateness of candidate lexical entries. Asterisks mark branches which may occur more than once. 'Terminal' categories (not allowing a deeper search) are underlined.

Consider what happens when the formulator embarks on an S but looks up a lexical entry of type N - perhaps because only a person was being conceptualized, not yet the action performed by that person. Then the tree construction process will try to find a path which begins at S and ends in N, traversing the checklist in a depth-first, left-to-right manner. In terms of the example, since V is inapplicable SMod is tried. SMod offers several possibilities (second row of Figure 3) but N isn't one of them. Only via Subj (first row) and NP (third row) is it possible to bridge the gap between S and N. At that point the tree construction process has discovered a role for N to play in the sentence, and the whole chain (Subj-NP-N) is hung off S. Thus N has been "booked" as Subj and later stages of tree construction will have to reckon with the fact that a

certain piece of conceptual structure has already been formulated as a subject.

In passing we have already touched upon two processes of syntactic bookkeeping. First, we have seen how certain syntactic features of a sentence fragment are kept in working memory as nodes and branches of a tree. (Although I cannot work out this point here, I'd like to remark that such trees do not represent the generation history of a sentence but only the syntactic information necessary to eliminate ungrammatical sentence continuations.) Second, determining whether lexical material retrieved from the lexicon offers opportunities for arriving at a grammatical continuation is done by manipulating the checklist in the way indicated. I should add here that the checklist is used in conjunction with the (partial) syntax tree kept in working memory. (At the beginning of planning an utterance the tree is empty or only consists of the node S.) E.g. an S-rooted tree with a Subj branch cannot receive a second subject branch because the first row of Figure 3 says that only one Subj is allowed.

Section 3 deals with the second kind of bookkeeping operations. It presents some data which seem to support the foregoing analysis of how the checklist is used for assigning syntactic roles (functions) to lexical entries. The experiments, which were conducted jointly with Ben Maassen and Pieter Huijbers, are only briefly reviewed (detailed report in preparation). The reader who wishes to learn more about the sentence production system we are working on may consult KEMPEN (1978) for a description of an early version (dating back to 1976), and KEMPEN & HOENKAMP (in preparation) for recent developments.

#### **DESCRIBING STATIC PICTURES UNDER TIME PRESSURE: EVIDENCE FOR SYNTACTIC BOOKKEEPING**

In order to link up the theory of Section 2 with reaction time (RT) data I make the assumption that traversing the checklist is a time consuming process. It then follows that bridging the gap between S and N takes more time than between S and V.

Native speakers of Dutch were instructed to describe (as quickly as possible) simple static pictures showing a person (man, woman, boy or girl) engaged in some action (kicking, slapping, greeting or teasing). Natural descriptions contained a subject noun and an intransitive verb in present tense, e.g. *het meisje slaat, de man slaat, het meisje plaagt* (the girl slaps, the man slaps, the girl teases), etc. The pictures were presented on a TV screen. Reaction times, measured through a voice key, were defined as the interval between the moment a picture appeared on the screen and the onset of the speaker's descriptive utterance. No articles were allowed in the subject noun phrases.

In four experimental conditions speakers produced four types of descriptions:

- Subj: subject noun (actor) alone
- V: main verb (action) alone

SubjV: subject followed by main verb

VSubj: main verb followed by subject

Descriptions of a certain type were elicited by a one- or two-word "frame", which preceded the pictures by one second. The frame consisting of the words "OMDAT HIER ..." ("because here ...", "weil hier ...") would elicit SubjV word order. The frame "WANT HIER ..." ("for here ...", "denn hier ...") conditioned VSubj order; "ZELFST.NAAMW." ("noun") implied condition Subj and "WERKWOORD" ("verb") implied V. Here we have exploited a word order rule of Dutch and, for that matter, German. Dutch and German word order in main clauses differs from that in subordinate clauses. After a subordinating conjunction followed by an adverbial phrase Subj always precedes V. But VSubj is the only appropriate order after a coordinating conjunction plus adverbial phrase. For instance, the subordinate clause because here the boy is greeting translates into (Du.) omdat hier de jongen groet and (Ger.) weil hier der Junge gruesst, with SubjV order. On the other hand, when because is replaced by the synonymous coordinating conjunction for, Dutch and German word order changes to VSubj: (Du.) want hier groet de jongen, (Ger.) denn hier gruesst der Junge.

The interpretation of the results (which I admit is in part post hoc) hinges upon the assumption that actor and action are two conceptual fragments which are fed into the formulating process one after the other: actor first, then the action. This is not too surprising because in general it is much harder to draw a clearly recognizable static picture of an action than of a person or animal. Indeed, the average RT in condition Subj was more than 100 ms shorter than in condition V (843 vs. 950 ms), confirming that the actors could be recognized ("conceptualized") earlier than the actions.

What does the theory of Section 2 have to say about SubjV and VSubj trials? Presumably the speakers attempt to construct an S but are confronted with a lexical entry of type N (the result of lexicalizing the actor, the picture fragment recognized first). This state triggers a time consuming search through the checklist. This effectively delays the onset of overt response production because Subj opens the sentence. As soon as the action part of the picture is recognized it is lexicalized and connected to the tree. Traversing the checklist is not necessary now because V is immediately identified as a permissible daughter node of S (first row of Figure 3).

All this holds true for the VSubj condition, too, except for length of the response delay. Since V is to be uttered first and checking V is easier than checking N, we may expect VSubj responses to suffer less from bookkeeping delays than SubjV responses. (The N-check for Subj takes place in the VSubj condition as well but cannot affect RT, Subj being in second position.) This is indeed what seems to happen: the difference between VSubj and SubjV is only 23 ms, much smaller than the 107 ms difference between Subj and V; a statistically very significant result (mean latencies were 974 (SubjV) and 997 (VSubj) ms; cf. Figure 4).

But couldn't the interaction of Figure 4 be explained in simpler terms? As part of a long series of picture description experiments, Lindsley (1975,

1976) checked the alternative Interpretation that in the SubjV condition speakers are simply waiting for the action to be recognized. This would indeed explain why the three conditions containing V (SubjV, VSubj and V) yield similar latencies. In order to see whether this could be true he varied the number of different actions which could occur in an experimental session. The V latencies were indeed sensitive to this variable: they increased with increasing number of different actions. However, SubjV latencies failed to covary with the number of actions, and remained the same irrespective of how hard it was to recognize the action on a picture. This effectively rules out an interpretation in terms of "waiting for the action".

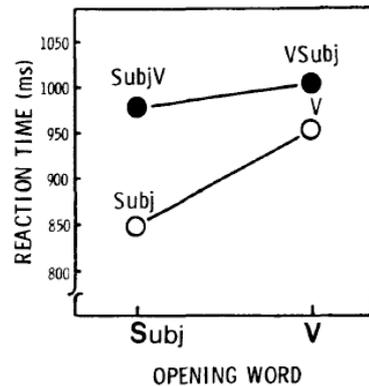


Figure 4. Interaction between length of utterance and syntactic category of the word opening the utterance.

Our interpretation in terms of syntactic bookkeeping is corroborated by the results of related studies which used the same experimental paradigm but a different set of pictures. We prepared 20 "stereotype" pictures depicting typical actions of well-known persons, animals or objects, e.g. birds singing, girls biking, planes flying. There were 20 different actors and actions, and each actor always did the same action. The experimental subjects had to inspect the set before RT measurement began, and during the session they saw each picture at least four times. The point I want to make is that in these studies the actor-action combinations formed well-integrated structures of conceptual information and I venture the hypothesis that they were conceptualized and formulated as a whole.

It follows that the tree construction process can always start with a V lexical entry and that SubjV responses will no longer incur a delay as a result of N-checking. Instead of the VSubj-SubjV < V-Subj effect shown in Figure 4, we now expect the two differences to be equal. This turned out to be the case in two experiments with stereotype pictures.

In conclusion, I feel that syntactic bookkeeping is a demonstrable and measurable part of sentence production. I hope improved versions of the picture description paradigm will reveal some of its inner structure. In particular we would like to know whether the specific details of the checklist rules and the depth-first, left-to-right search we have assumed are realistic proposals. This question we are currently working on.

Part of the research described in this paper was supported by grants from ZWO, the Dutch National Foundation for the Advancement of Pure Research.

#### REFERENCES

- Dik, S.C. (1978). *Functional grammar*. Amsterdam: North-Holland.
- Kempen, G. (1978). Sentence construction by a psychologically plausible formulator. In Campbell, R. & Smith, P. (Eds.), *Recent advances in the psychology of language: Formal and experimental approaches*. New York: Plenum Press.
- Kempen, G. & Hoenkamp, E. (in preparation). An incremental procedural grammar for sentence formulation.
- Lindsay, J. (1975). Producing simple utterances: how far ahead do we plan? *Cognitive Psychology*, 7, 1-19.
- Lindsay, J. (1976). Producing simple utterances: details of the planning process. *Journal of Psycholinguistic Research*, 5, 331 - 353.