The time course of conceptualizing and formulating processes during the production of simple sentences

Gerard Kempen
and
Ben Maassen

Experimental Psychology Unit
üniversity of Nijmegen, The Netherlands

June 1977

Paper presented at the Third Prague Conference on The Psychology of Human Learning and Development, July 1977

Abstract The psychological process of producing sentences includes conceptualization (selecting to-be-expressed conceptual content) and formulation (translating conceptual content into syntactic structures of a language). There is ample evidence, both intuitive and experimental, that the conceptualizing and formulating processes often proceed concurrently, not strictly serially. James Lindsley (Cognitive Psych.,1975, *7,* 1-19; J.Psycholinguistic Res., 1976, 5, 331-354) has developed a concurrent model which proved succesful in an experimental situation where simple English Subject-Verb (SV) sentences such as "The boy is greeting","The girl is kicking" were produced as descriptions of pictures which showed actor and action. The measurements were reaction times defined as the interval between the moment a picture appeared on a screen and the onset of the vocal utterance by the speaker. Lindsley could show, among other things, that the formulation process for an SV sentence doesn't start immediately after the actor of a picture (that is, the conceptual content underlying the surface Subject phrase) has been identified, but is somewhat delayed. The delay was needed, according to Lindsley, in order to prevent dysfluencies (hesitations) between surface Subject and verb. We replicated Lindsley's data for Dutch. However, his model proved inadequate when we added Dutch Verb-Subject (VS) constructions which are obligatory in certain syntactic contexts but synonymous with their SV counterparts. A sentence production theory which is being developed by the first author is able to provide an accurate account of the data. The abovementioned delay is attributed to certain precautions the sentence generator has to take in case of SV but not of VS sentences. These precautions are related to the goal of attaining syntactic coherence of the utterance as a whole, not to the prevention of dysfluencies.

The psychological processes which underly sentence production may be grouped under three beadings: (a) conceptualization (selecting to-be-expressed conceptual content), (b) formulation (translating conceptual structures into syntactic structures), and (c) articulation (pronouncing the words of a syntactic structure). In the present paper we are concerned with some temporal relationships between these three aspects of speaking, in particular with the order in which they are initiated and to what extent they can overlap in time (i.e. proceed in parallel).

We start out from the assumption that conceptualization makes use of the limited processing capacity of the central channel of attention. In previous research it has been established for two subprocesses of speaking that they do not occupy the central channel of attention and thus can overlap with conceptualization activity. Lindsley (1976) showed that lexicalization, that is, retrieving from the mental lexicon a name for a concept, does not interfere with central (conceptual) activity . Van Galen and Ten Hoopen (1976) demonstrated this for articulation, i.e. sending motor commands off to the articulatory organs and executing the commands. These results imply that articulation and at least one aspect of formulating (lexicalization) are peripheral processes which may proceed in parallel with the central process of conceptualizing. Parenthetically, these conclusions are valid for fluent speaking only. The study of hesitations during spontaneous speech has made it clear that word finding problems can give rise to considerable delays (so-called lexical pauses) and use up central processing capacity (cf. Butterworth, 1976). Similar things may undoubtedly happen to the articulation process.

Limiting ourselves to the case of fluent speaking, we do not know of any good evidence with regard to central or peripheral status of other aspects of formulating. For instance, it is unknown whether the processes which regulate order of lexical items and the insertion of syntactically required morphemes consume central capacity or are carried out peripherally. However, chances are that large portions of them are indeed peripheral, seeing the conclusion Goldman-Eisler reached in her studies of hesitational pausing: "Syntactical operations ... seem to be organized at the level of skills" (1968, p.76). Much detailed experimentation is needed to settle this question.

The specific issue we address in this paper is the initiation point of formulating activity. At what stage of the conceptualization process are the formulation processes initiated? Lindsley (1975, 1976) has developed a set of ideas on this topic in the context of reaction time experiments.

Subjects had to describe simple static pictures of a person who is performing some action. The person is either a man, a woman, a boy or a girl; the action is kicking, greeting, touching or simply standing beside another person. The sentences produced were of the type The subject is verbing, e.g. The boy is greeting, The woman is kicking, etc. Reaction times were defined as the interval between the onset of picture presentation and the onset of the vocal response by the subject. In addition to the Subject-Verb (SV) condition there were trial blocks in which only the actors or only the actions were named: S and V conditions respectively.

The V reaction times were on average some 100 ms longer than the S latencies. This difference has no theoretical implications since it is at least partially due to the way of drawing actors and actions. Lindsley's first important finding was that the SV latencies exceeded the S latencies by about 100 ms. This delay remained constant over variations in recognizability of the action. That is, if the number of different actions which could occur in a series of trials varied from 2 to4, then the V latencies increased but not the SV latencies. These data (together with some other results) imply that formul-ating does not wait until both actor and action have been identified, i.e. until the complete conceptualization process is over. Nor does the formulating process for SV sentences start immediately after identification of the actor.

Lindsley's solution consists of three elements. First, he makes the assumption that the experimental subjects first work on recognizing the actor of a picture and only then start processing the actions. Second, the recognition process is divided into an attention and an identification stage. The attention stage serves to extract perceptual features from the picture. During the identification stage the identity of actor or action is derived from those features. The essential assumption Lindsley adds is that attention time is constant or, more precisely, independent of number of possible actors/actions. On the other hand, identification time does increase with number of alternatives. Third, the constant delay which SV sentences show in comparison with S utterances is attributed to "attending to the action". The formulating process (lexicalizing the actor) is supposed to wait until after the first stage of action recognition. What is the cause of this delay? Lindsley argues that it prevents dysfluencies (hesitations) in the middle of SV sentences: the delay brings the moment of S articulation closer to the point where V is ready for being pronounced.

The sentence production model which is being developed by Kempen (1977) also predicts Lindsley's constant delay, but for totally different reasons. The model attaches much importance to the problem of establishing syntactic coherence between sentence fragments which are constructed one after another, as formulations for successive

fragments of conceptual content. In Lindsley's experiments, actor and action may be considered as two conceptual fragments which are fed into the formulating process one after the other. Kempen assumes that each conceptual fragment is handled immediately and that the resulting sentence fragment will be passed down to the articulatory mechanisms as soon as is allowed by rules of syntax. Thus the period of time during which conceptual and sentential fragments occupy central processing mechanisms is minimized.

However, the formulating process will have to keep some information about the sentence fragment it is working on, since the shape of further sentence fragments may depend on it. For instance, if the first fragment of a sentence is the syntactic subject then at least its number and person must be remembered so as to enable the ensueing verb to agree with it in number and person. Hence, before uttering a sentence fragment, the formulating process will make up a "syntactic summary" which contains all its syntactically relevant features and which is then stored in some register. Not earlier than after these precautions have been taken can articulation of the sentence fragment begin. Notice that the precaution of storing a syntactic summary enables efficiënt management of working memory: without the summary the whole sentence fragment would have to be kept in store.

If we make the further assumption that putting together the syntactic summary for a sentence fragment is a time consuming process, then we have an explanation for the constant delay Lindsley observed for SV as compared to S utterances. It is caused by the process of making a summary for the subject of SV sentences. The delay is constant (i.e. independent of ease of identifying the action) simply because it has nothing to do with the action in the picture. As a consequence we no longer need the distinction between attention and identification stages of actor and action recognition.

Devising an experimental confrontation between the two theories, we 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, the Subject always precedes Verb. But Verb-Subject is the only appropriate order after a main-clause 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 grüsst, with SV order. On the other hand, when because is replaced by the synonymous coordinating conjunction want (Eng. for,) Dutch and German word order changes to VS: (Du.) want hier groet de jongen, (Ger.) denn hier grüsst der Junge.

Lindsley's theory predicts a difference between VS and SV which is equal to the difference between V and S. We assume that in the VS condition the recognition processes will take place in the order action-actor, and that the attending stages for actors and actions are equally long. These seem the most reasonable extensions of Lindsley's model.

(Remember that the difference between V and S is caused by the greater difficulty of identifying V; the very same factor will increase VS latencies in comparison with SV. Parenthetically, we have considered another extension of Lindsley's model which adheres to the original actor-action sequence of recognition processes. But its predictions would be totally out of line.)

Kempen's model predicts a smaller difference between VS and SV than between V and S. The time needed for picture recognition is assumed to be equally divided between actor and action (no priority for either of them). The actors, which are easier to recognize than the actions, will usually be the initial input to the formulating process, both in the SV and VS conditions. But since in condition VS the Subject cannot be pronounced immediately, it is stored in some register and waits there until after V has been articulated. The important point is that no syntactic summary needs to be prepared for the Subject. Nor for the Verb, since it is formulated immediately after the conceptualization processes have done their job. We finally predict faster responses in conditons V and S than in conditions VS and SV, because in the single-word conditions the recognition efforts can be concentrated on one aspect of the picture (actor or action).

Procedure

A set of 48 pictures was prepared, each containg an actor (man, woman, boy, girl), an action (kicking, slapping, greeting, teasing) and a patiënt (man, woman, boy, girl; never identical to the actor). The actor was always on the lefthand side, the patiënt on the right. The pictures were recorded four times, in random order, on 192 tracks of an Ampex video disk. Each track containing a picture was preceded by a track containing a "frame" which served to identify the experimental condition. The frame consisting of the words 'OMDAT HIER ...'' ("because here ...") would elicit SV word order. The frame "WANT HIER ..." ("for here ...") conditioned VS order; "ZELFST. NAAMW." ("noun") implied condition S; and "WERKWOORD" ("verb") implied V. Each of the four frames occurred 48 times in the total series of 192 stimuli. The stimuli were presented to the subjects (20 students of the University of Nijmegen participating in individual sessions) on a TV monitor. Each trial consisted of 3 seconds of gray foliowed by a frame (1 second) foliowed by a picture (4 seconds). As soon as a picture became visible, the subject produced a verbal description in the syntactic form suggested by the accompanying frame. No articles were allowed in the descriptive utterances. Reaction times (in milliseconds) were measured via a voice key (as in Kempen, 1976) and noted down by the experimenter. The percentage of erroneous responses was about 10, equally distributed over conditions. Each subject received a few training trials in order to get familiar with pictures and apparatus.

Results

We carried out an analysis of variance on the untransformed data. The factor Frames proved highly significant ($F_{3,57}$= 23.02, p<.001). The mean latencies were 843 (S), 950 (V), 974 (SV) and 997 (VS) ms. The t-test for a priori comparisons gave the following results. The comparison VS-SV<V-S was significant (t=3.14; p<.005). VS was not significantly different from SV (t=1.22; p>.20). Finally the difference between V and S was highly significant (t= 5.56; p<.0005).

Discussion

The average latencies we obtained in the S, V and SV conditions display a pattern very similar to Lindsley's (1975) Experiment I (597, 710 and 715 ms respectively), although they are considerably higher. The latter is probably a combined effect of having four possible actors and actions (as against three in Lindsley's study), of randomizing experimental conditions (Lindsley used blocks of trials in one condition), and of picture quality. So we have been able to replicate Lindsley's basic finding.

However, the fact that we obtained a VS - SV difference of slightly more than 20 ms, whereas V took more than 100 ms longer than S, rejects Lindsley's theory in favor of Kempen's sentence production model. The implication is that speakers have been shown to spend a measurable amount of processing time, not on trying to prevent hesitations between successive fragments of a sentence, but on precautions which can establish syntactic coherence between sentence fragments. In the absence of such precautions these fragments would have been formulated totally independent of each other.

We take this evidence as supporting our attempts at developing a sentence production model which, in contrast to the usual practice of transformational and computational linguistics, views conceptualizing and formulating as concurrent, not as serially ordered processes.

References

Butterworth, B. Semantic planning, lexical choice and syntactic organisation in spontaneous speech. Manuscript, U. of Cambridge, 1976.

Goldman-Eisler, F. Psycholinguistics. London: Academie Press, 1968.

Kempen, G. Syntactic constructions as retrieval plans. British Journal of Psychology, 1976, 67, 149 - 160.

Kempen, G. Sentence construction by a psychologically plausible formulator. In Campbell, R. & Smith, P. (eds.) Proceedings of the Psychology of Language Conference, Stirling, June, 1976. NY: Plenum Press, 1977.

Lindsley, J. Producing simple utterances: how far ahead do we plan? Cognitive Psychology. 1975, 7, 1-19.

Lindsley, J. Producing simple utterances: details of the planning process. J. Psycholinguistic Research, 1976, 5, 331 - 353.

Van Galen, G and ten Hoopen, G. Speech control and single channelness. Acta Psychologica, 1976, 40, 245 - 255.