A Time-Based Approach to Agrammatic Production

HERMAN KOLK

Nijmegen Institute for Cognition and Information, Nijmegen, The Netherlands

A time-based approach to agrammatic speech is presented. The paper consists of three parts. In the first part, the literature which deals with agrammatic comprehension as a problem of disrupted timing, that is, as a slow-down of syntactic computation and/or a rapid decay of the results of syntactic processing, is reviewed. In a second part, the hypothesis that similar timing problems cause difficulties in production as well is discussed. Two possible ways in which this can happen are described. First, slow down or rapid decay can lead to desynchronization within the process of syntactic tree formation. Second, a slow down of syntactic processing can cause asynchrony between the production of a syntactic slot and the retrieval of the proper grammatical morpheme from the mental lexicon. This hypothesis predicts that morphemes which are dependent on a relatively complex part of the syntactic tree will elicit relatively many errors. Results from the literature which seem to confirm this prediction are discussed. In the third part of the paper, the possible ways in which a patient can adapt to the reduced temporal window that would result from a timing deficit are discussed. Message simplification will reduce the size of the required temporal window and will therefore have a beneficial effect on the error rate. Restart of the computational process will profit from previously reached activation levels so that synchrony is easier to reach and error rate is reduced. Empirical work which appears to support these hypotheses is reviewed. © 1995 Academic Press, Inc.

DESYNCHRONIZATION AS A POSSIBLE SOURCE OF AGRAMMATIC COMPREHENSION

Language comprehension and production are exceedingly complex tasks in which numerous pieces of information have to be juggled within fractions of seconds. All parts have to fall in the right place at the right time. The disruption of this temporal fine-tuning could well be the deficit responsible for the difficulties aphasics have in expressive and receptive language tasks. This has been the claim of a number of authors in the so-called chronogenetic tradition in aphasiology (see Kolk & van Grunsven, 1985, for references). As early as 1885, Grashey proposed that lack
of synchrony between the representations of word meaning and word form can lead to word-finding difficulties. He reported on a patient with severe naming difficulties. This patient was very impaired in naming objects presented to him by means of pictures. But he had another problem as well. If the patient was given a picture and the picture was taken away and immediately replaced by the correct picture and one or more other pictures, he was unable to identify the correct one. Grashey suggested that the pictorial representation was subject to very rapid decay and that, as a consequence, activation of the word form had too little time to develop, hence the word-finding difficulties (Grashey, 1885). The great Carl Wernicke commented very favorably upon this hypothesis stating that “...I would not hesitate to suggest that it may be considered the most significant contribution in the development of the doctrine of aphasia during the past decade. In his work Grashey pursues an entirely new and very fruitful notion by postulating the temporal factor as an important consideration in the formation of the spoken word, which can be likewise applied to reading and writing” (Wernicke, 1885).

Not only semantic information but also word-form (phonological) information itself could be subject to pathological decay. Just recently, Martin, Dell, Saffran, and Schwartz (1994) simulated the speech errors of a jargon aphasic by means of Dell’s (1986) interactive spreading activation model of language production. By assuming a high decay rate in the semantic-lexical-phonological network, they were able to simulate various qualitative error patterns in the speech of this patient, not only as they varied over tasks (naming versus repetition), but also as they varied over time (spontaneous recovery). In a follow-up study, Schwartz, Dell, Martin, and Saffran (1994) investigated the naming errors of 17 fluent aphasics. They found that patients fell into three groups. For one group, the error pattern could be simulated by means of a fast decay, just as with the jargon aphasic described above. For another group, however, a different change in the temporal parameters was necessary. The behavior of this subgroup could only be simulated by assuming a slow activation of word-form information within the network. Still another group seemed to have both kinds of temporal disruption. So the behavior of these 17 patients could be accounted for by either a fast decay or a slow activation, or a combination of both.

In 1985, Kolk and van Grunsven (1985; Kolk, van Grunsven, & Keyser, 1985) proposed a general framework for the explanation of agrammatic sentence processing as a timing disorder. In this framework, it is assumed that every element needed to build a sentence representation has some activation that determines the availability of that element. It takes some time for elements to reach a critical level of activation and after a peak level, the activation is subject to decay. A third assumption is that activations of elements are typically interdependent. This means
that the activity of one element is required for the activation of another element. For instance, information about the subject of the sentence must be active in order for the right form of the verb to become activated. Between these two types of information, there must therefore be computational simultaneity or synchrony. When the parameters, determining the various activation levels, are changed, this delicate simultaneity can be jeopardized, leading to a premature disintegration of sentence structures. Two changes to the normal situation were suggested. In the "slow activation" case, it takes longer for an element to reach its critical level of activation and occasionally, this level is reached too late. "Fast decay" makes elements unavailable when they fall below their critical level too soon to be combined with other elements in the sentence representation.

Haarmann and Kolk (1991a) have constructed a computational model to investigate these possibilities. This model, SYNCHRON, simulates the temporal course of the building up of a hierarchical phrase-structure representation. The nodes, needed to construct a syntactic tree, take some time to reach their "memory time phase," that is, to become available to interact with other nodes. Furthermore, this memory time is limited; if it is exceeded, elements disappear from memory. A particular syntactic category, say a VP, can be retrieved only if all immediate daughter categories (e.g., V, NP, PP) are available. The model successfully reproduces the agrammatic performance profile over reversible active, locative, and passive sentences at two different levels of severity (reported by Kolk & van Grunsven, 1985). It does so both by slowing down computation and by limiting memory time; the fit with the empirical data is equally good in the two cases. The type of elements affected by the temporal deficit do make a difference, however. When function word nodes are affected, the required pattern does not emerge. It appears only when phrasal category nodes are impaired.

The SYNCHRON model has a number of clear advantages over most existing hypotheses on agrammatic comprehension (see Kolk & Weijts, in press, for a discussion of these other hypotheses). First of all, it accounts for the finding that sentences with a more complex constituent structure pose more problems than sentences with a simpler structure, not only in comprehension (cf. Caplan & Hildebrandt, 1988; Kolk & Weijts, in press) but also in the detection of ungrammaticality (Haarmann & Kolk, 1994). Second, it can explain that agrammatic comprehension is a phenomenon that varies in degree. That is, even though for instance passives are generally harder to understand than actives, the absolute level of performance can vary. Thus, one finds patients who perform at chance on both actives and passives and other patients who are above-chance on both actives and passives (Kolk & van Grunsven, 1985; Mitchum, Haendiges, & Berndt, 1993). Third, the model produces a complexity x severity interaction, a trend of which was present in the Kolk
and van Grunsven data and, more convincingly, in Caplan and Hildebrandt’s (1988) findings. Finally, SYNCHRON can, at least in principle, give an account for the well-known dissociation between agrammatic comprehension and grammaticality judgment (Linebarger, Schwartz, & Saffran, 1983). To do so, it must be assumed that comprehension, entailing both role assignment and the selection of the correct picture or the acting-out of the proper event, requires a longer availability of the syntactic representation than grammaticality judgment and is therefore more easily disrupted by rapid decay or slow activation of syntactic information.

In its present form, SYNCHRON also has a number of clear limitations. First of all, there is no stage of role assignment. Adding this process to the model would make it possible to see whether in fact desynchronization within the parser would leave grammaticality judgments relatively unimpaired. Second, there is no chain formation. Chain formation would put additional constraints on synchronization, because the trace and the moved argument, including their coindexations, must be simultaneously available for correct role assignment. If SYNCHRON would be provided with this facility, it would be possible to investigate how much the movement factor contributes to the performance deterioration, independent from constituent structure complexity. Finally, SYNCHRON’s precise assumptions regarding the synchrony requirement may need revision. Haarmann and Kolk (1991a) demonstrate that the predictions SYNCHRON makes with respect to sentence types other than actives, locatives, and passives are far from optimal. These predictions could be improved, the authors argue, by requiring all daughter nodes to be available for a particular mother category to be retrieved, rather than only the immediate daughter nodes (see Cornell, 1995, for a computational model of agrammatic comprehension, which can be regarded as an extension of SYNCHRON).

There are various kinds of empirical evidence to support the claim that desynchronization of language processing contributes to agrammatic sentence comprehension. Swinney and Zurif (1995) discuss data to show that at least some agrammatics suffer from a slowing down of lexical activation and, as a consequence, are not able to reactivate moved arguments at the point in time that such reactivation is needed: at trace positions. Friederici (1995) reviews a number of studies she conducted which led her to conclude that agrammatics are slowed down in the first-pass assignment of syntactic structure to the incoming sentence. Furthermore, she presents ERP data which suggest a very early activity (around 300 ms) over the frontal areas of the left hemisphere, an activity she thinks reflects this early assignment process. Hagoort (1990) reports two studies which are suggestive of a processing slow down. In both experiments, ambiguous words were presented in a sentence context. Research with normals has shown that both meanings of an ambiguous word are active
for a short period of time and then the contextually inappropriate meaning is quickly suppressed (cf. Onifer & Swinney, 1991). Hagoort found evidence that Broca's are slowed down in this respect: they do suppress the inappropriate meaning but only at a relatively late point in time (750 ms). This was true not only when the appropriate meaning of the ambiguous word was biased by the meaning of the sentence but also when it was biased by the syntactic context. The first effect—disambiguation slowed down by the sentence meaning—has very recently been replicated in an ERP study. At an interstimulus interval of 100 ms, Broca's aphasics showed an N400 effect of semantic priming in response to both meanings of the ambiguous noun. Only at 1250 ms was the N400 effect selective for the appropriate meaning (Hagoort, 1994). Haarmann and Kolk (1991b) tested 12 agrammatics in a syntactic priming paradigm in which they varied grammaticality and SOA between context sentence and target word (SOAs of 300, 800, and 1200 ms were employed). Only at 1200 ms was a significant priming effect obtained. In a subsequent word monitoring study, again grammaticality was varied as well as interstimulus interval between the point at which the sentence became ungrammatical and the target word (Haarmann & Kolk, 1994). This time, there was no evidence for slow activation, even though to a large extent the same group of patients was used. However, there was evidence for rapid decay, as the effect of ungrammaticality had disappeared after an ISI of 750 ms, although the elderly controls still showed a significant effect. Why there is this difference between the syntactic priming task and the word monitoring task is unclear. Taken together, the evidence for a slow activation or fast decay looks promising although the issue is far from settled.¹

A TIME-BASED APPROACH TO (PAR)AGRAMMATIC PRODUCTION

In view of the progress that has been made with the timing approach to agrammatic comprehension, it seems reasonable to explore the possibilities of this approach to agrammatic sentence production. A few years ago, one of my doctoral students, Mart van de Kerkhof, and I designed a first-time-based model (Kolk, 1987) of agrammatic sentence production. In this model, we employed the Kempen and Hoenkamp (1987) Incremental Procedural Grammar, an algorithm which generates grammatical sentences in Dutch. It does so in an "incremental" fashion; that is, as soon as a particular part of the syntactic tree is ready, the corresponding part of the sentence is produced (see also Levelt, 1989, for an extensive de-

¹ Tyler, Ostrin, Cooke, and Moss (1995) have recently demonstrated by means of a word-monitoring paradigm, that at least some agrammatic patients can retrieve some types of grammatical information quite rapidly. Tyler et al., however, did not manipulate SOA or ISI. Therefore, there could still be slow activation in some patients and/or for some types of information. There also could be fast decay.
scription and discussion of the IPG model). Similarly to what was later
done by Haarmann and Kolk (1991a), we assumed that for correct pro-
duction of the sentence, synchrony between the various parts of the syn-
tactic tree was necessary. We were able to show that with both rapid
decay and slow activation, the required synchrony often failed to materi-
alize, so that a grammatical sentence could not be produced. The num-
ber of failures was dependent upon (a) the complexity of the sentence and
(b) the activation and decay rates. No attempt was made to actually
generate sentences on the basis of disintegrated syntactic representa-
tions.

In addition to problems with complex sentences, agrammatics have
difficulty producing the correct grammatical morphology. In the well-
known classification of Goodglass and Kaplan (1972), these two symp-
toms are represented by means of different rating scales. It has been
proposed that we are dealing here with two separate and dissociable
deficits, a syntactic and a morphological one (cf. Berndt, 1987; Saffran,
Berndt, & Schwartz, 1989). Although the syntactic and the morphological
component could indeed be independently damaged in some individual
cases, this does not exclude the possibility that a purely syntactic deficit
has consequences for the production of grammatical morphology. In par-
ticular, if one thinks of syntactic and lexical processes to be partially
autonomous production routines, morphological errors could arise as a
result of an asynchrony between these two lines of processing. This is a
possibility I will elaborate below.

What precisely do I mean by "morphological" errors? In our previous
work we have demonstrated that there are, at least in Dutch and German,
large differences in type of output between "spontaneous speech" (that
is, free conversation) and elicited speech (various kinds of picture de-
scription and cloze tasks). Our research has demonstrated that in free
conversations, patients typically produce agrammatic speech: output
that—according to Kleist's original definition of agrammatism (1916)—
lacks much of the required grammatical morphology but contains rela-
tively few erroneously produced morphemes. Elicited speech is typically
more par agrammatic: it contains a high number of wrongly selected mor-
phemes and relatively few omissions (Heeschen, 1985; Kolk et al., 1985;
Haarmann & Kolk, 1992; Kolk & Heeschen, 1992; Hofstede & Kolk,
1994). Our theory of this task effect (see the above references) holds
that elicited speech largely reflects the underlying deficit, but that the
aggrammatic character of spontaneous speech is primarily due to the way
in which the patients adapt themselves to this underlying deficit. I will
first discuss our views on the deficit and explain how a timing disorder
can lead to paragrammatic speech. I will then describe how patients can
adapt to such a deficit in a way that leads to agrammatic output.

There are relatively few models of paragrammatic speech. One impor-
tant processing account of paragrammatism is given by Butterworth and Howard (1987). The paragrammatic output they tried to explain was produced by a group of jargon aphasics. Butterworth and Howard assume that, in this group at least, the sentence production system itself is intact. There is no selective loss of a particular component but rather an impairment of the processes controlling the operation of the various components. These control operations involve the transfer of information between system components, the initiation and termination of component processes, and the checking of the output of components. An omission of a word, for instance, could be due to a failure in initiation of lexical selection or a loss in the transfer from lexical selection to phonology assembly. As to why these control processes are impaired, Butterworth and Howard refer to Freud (1890), who thought there was a transient attentional disturbance leading to a failure of inhibition. Given the assumption of such a global impairment, affecting all components of the sentence production system to some extent, one would expect close similarity between normal speech errors and aphasic paragrammatisms and this is indeed what their data suggest.

My theory is similar to that proposed by Butterworth and Howard, in that I assume no change in the overall structure of the sentence production system but only in the mode of processing. Instead of a transient attentional deficit that can affect any set of components, I propose a slowing down of the syntactic component that affects not only the computation of structure but also the selection of the proper grammatical morphology. How could slowing down have such an effect? In most psycholinguistic theories (cf. Levelt, 1989), sentence production is seen to involve the generation of syntactic frames, with categorized slots (e.g., Det, Adj, N). The slots are filled with morphemes, retrieved from the mental lexicon. Garrett (1975, 1980) has argued, on the basis of speech error data, that there is a fundamental distinction between closed-class elements (function words and inflections) on the one hand and open-class elements (content words) on the other. Whereas open-class words are generated by lexical processes and have to be inserted into the syntactic frame, closed-class items would be generated by the syntactic process itself: they are part of the frame and need no independent retrieval. Stemberger (1984, 1985), however, has shown that one can interpret the critical speech error data in a different way, without making the assumption that closed-class elements and syntactic frame are accessed within a single operation. Furthermore, if the closed-class vocabulary would be

2 Stemberger relates the different behaviors of open- and closed-class elements primarily to frequency differences between the two sets of elements. Evidence for the important role of frequency has been obtained by Stemberger and MacWhinney (1986) and by Dell (1990). Another argument against Garrett's hypothesis is Bock's (1989) finding that "syntactic
FIG. 1. Synchronization of the delivery of a syntactic slot and the activation of a word form. (Top) The growth and decay of activation of the word form is depicted. The dark gray areas indicate that at the beginning and at the end of the activation period, when the amount of activation is relatively small, competition with other lexical candidates is relatively high, and the chance to select the wrong word form is equally high. In the light gray area, competition is low and there is a low chance to make errors. (Bottom) Time at which the syntactic slot is delivered is indicated with simple and complex sentences and with normal and delayed rate of syntactic processing. Optimal syntacto-lexical integration is obtained when the slot is delivered in synchrony with the “safe” period of lexical activation.

generated by the syntax, there is no possibility to get an asynchrony between syntax and lexicon with respect to these items and we have no way to account for paragrammatic errors within our general framework. In the theory I am going to propose, therefore, not only content but also function morphemes are represented in the syntactic representation by means of a categorized slot and both types of morphology have to be integrated with their syntactic slots. This integration requires synchrony and a processing delay can disrupt this synchrony.

In Fig. 1, I have given a pictorial sketch of the theory. The activation curve at the top represents the lexical activation process of one particular item, as it develops over time. With current models of spreading activa-
tion (e.g., Dell, 1986), I assume that both at the beginning and at the end of the activation period, when the amount of activation is relatively low, competition from alternative lexical candidates is relatively high. At these two periods, the selection process is error prone. The middle period, on the other hand, is relatively "safe." In Fig. 1, the safe and the unsafe regions are indicated by different shades of gray.

The lower part of Fig. 1 represents the syntactic process, in particular the point in time at which this process delivers a syntactic slot for integration with a lexical filler. Two assumptions are reflected. First of all, in agrammatism the process of slot delivery is slower than in normals. Second, in both normals and aphasics, a slot that is part of a complex syntactic structure is produced at a later point in time than a slot in a simpler structure. The latter assumption follows from our general line of reasoning, also evident in the SYNCHRON model, that a more complex structure requires more computational steps and therefore more time to be computed.

As can be derived from Fig. 1, the theory predicts that complex structures will elicit more morphological errors than simple structures. I will follow the Kempen and Hoenkamp (1987) model discussed above in assuming that for a particular morpheme to be produced, it is not necessary that a syntactic tree for a full sentence be available. Production of morphology is possible, as soon as a minimally required syntactic representation has been made available. For the production of plural inflections of nouns, for instance, the minimally required structure is the NP. For lexical prepositions to be produced, a PP would have to be available. Of course, what will count as a minimally required structure will change from language to language.

In three studies evidence in support of this prediction has been obtained. The first study was conducted by Goodglass and Berko (1960). These authors presented 21 unclassified aphasics with a sentence completion task. In this task, they elicited 10 different grammatical morphemes. Of particular interest is the contrast they observed between the plural, verb, and possessive inflections, because the same phonological forms are involved in each case (−s, −z, or −z). First of all, they found the plural inflection to elicit significantly fewer errors than the verb inflection. This is what one would expect, assuming that for the production of a plural inflection, only an NP has to be available, whereas an S-node is required for a verb inflection to be produced. Goodglass and Berko also found the possessive inflection (e.g., "the horse's blanket") to be significantly harder than both the plural and the verb inflections. That possessive inflections are harder than plurals is to be understood, given the fact that possessives require the availability of a compound NP. Why the possessives are harder than the verb inflections is less obvious, since the two critical structures are of comparable complexity (a compound
NP, consisting of two NPs on the one hand, and an S, consisting of an NP and a VP). Perhaps possessives have a lower frequency. If this is the case, it could be argued that the structure takes more time to develop and as a consequence leads more readily to errors in lexical selection.

A second relevant finding was reported by Friederici (1982). She elicited prepositions in a cloze task. Prepositions that could occur in different functional roles were employed. Thus, the German preposition *auf* can occur not only as a lexical preposition (*Peter steht auf dem Stuhl*—*Peter stands on the chair*) but also as a subcategorized preposition (*Peter hofft auf dem Sommer*—*Peter hopes for the summer*). With her group of Broca's aphasics, Friederici found that subcategorized prepositions were much harder to produce than the lexical ones. With respect to our hypothesis, this result makes sense, since for the production of the subcategorized preposition, the integrity of a complex VP seems to be required (a VP with an embedded PP), whereas the lexical preposition can be produced solely on the basis of a PP being available.

In a recent study, Haarmann and Kolk (1992) compared various types of grammatical morphology in a cloze task. It is important to realize that, unlike in the previous two studies, the phonological form of the grammatical morpheme was not kept constant, so variation in the properties of these morphemes, such as frequency and the number of morphological options in a given category, could have played a role (Menn & Obler, 1991). Three types of inflectional morphology were studied. (a) The plural-noun inflection; here the minimally required structure is the NP. (b) The inflection of adjectives in attributive position; here the minimally required structure is a complex NP, in which the head noun is modified by an adjective. (c) The inflection of verbs in present or past tense; here the minimally required structure is the S. There is a clear order of structural complexity of these three structures: NP < NPadj < S. On this basis, we predict a corresponding order of difficulty for the three types of inflection: Plural inflection < adjectival inflection < verb inflection. This was indeed the result obtained by Haarmann and Kolk.

Besides inflectional morphology, three types of function words were tested. (a) Determiners (articles and possessive pronouns); these elements require the availability of an NP. (b) Lexical prepositions (locative, directional, or instrumental); for these elements, minimally a PP must be

---

3 The Haarmann and Kolk study was set up primarily to investigate the difference between Broca's and Wernicke's. For this reason, the various sets of morphemes tested were not as homogeneous as one would want. This was particularly true for the category "pronouns," which we will therefore not consider. Furthermore, the category "adjective inflections" contained a few items in which the comparative and superlative inflections were elicited, and in the category "prepositions," 4/12 items were not lexical prepositions. Leaving these elements out, however, did not change the overall order of difficulty reported by Haarmann and Kolk.
available. (c) Auxiliaries; for these elements, the required structure is the S (the task asked for the production of an inflected auxiliary). Again, the structural complexity is clearly ordered: NP < PP < S. So the predicted order of difficulty is determiners < prepositions < auxiliaries. Haarmann and Kolk observed this order of difficulty in their task.

Taken together, these data seem to indicate that structural complexity at least partially determines the number of errors made in the production of morphology. Of course, I have only loosely defined what "minimally required" complexity is; whether the complexity order that I have presented is the right one remains to be seen. Furthermore, the role of other factors such as frequency and number of morphological options still has to be worked out.

So far, I have assumed that paragrammatisms were brought about by syntactic delay. However, the general scheme, depicted in Fig. 1, allows for another possibility. Selection of the wrong lexical alternative would also take place if the syntactic slot would be delivered in synchrony with the early phase of lexical selection (the dark area in the left part of the activation curve). If lexical activation would be delayed, integration with the syntactic slot would be attempted too early and errors of selection could arise. This is the kind of paragrammatism one could expect from a patient with a lexical deficit. Obvious candidates, of course, would be Wernicke's aphasics, who show paragrammatism in their spontaneous speech and are also sometimes claimed to have a lexical impairment (cf. Berndt & Caramazza, 1980). This hypothesis would predict that for Wernicke's aphasics, there should be a reversed complexity effect: since complex structures are produced more slowly, their slots would become available later, at a point in time where there is relatively little response competition and integration with the correct morpheme will succeed more often (see Fig. 1). Haarmann and Kolk (1992), however, have given their cloze task also to a group of Wernicke aphasics. Contrary to the prediction of the lexical hypothesis, the Wernicke group shows the same order of difficulty as the Broca group. Given these data then, the most parsimonious hypothesis is that the Wernicke group also suffers from a slowdown in the syntactic part of their production system.

ADAPTING TO A REDUCED TEMPORAL WINDOW IN PRODUCTION

According to the argument given above, a syntactic slow down will lead to desynchronization, both within the process of syntactic encoding and in integrating syntactic slots with lexical fillers. Such asynchronies will also occur in normals, but only under more extreme circumstances, that is, with very complex syntactic structures. The difference between aphasics and normals could be described as one in size of the temporal
window available for syntactic computation and syntacto-lexical integration. This makes it understandable why aphasics make more grammatical errors than normals, but not qualitatively different ones, as was shown by Butterworth and Howard (1987). What it does not explain, however, is why the normals make as few errors as they do. Since the syntactic system, given its recursive nature, is in itself capable of producing a structure of any given complexity, why is there not a continuous overload, leading to a constant flow of speech errors? This argument, of course, is equally valid if one believes that the capacity limitation common to aphasics and normal speakers is not of a temporal nature but rather one characterized in terms of parsing-work space, as has been assumed by Caplan and Hildebrandt (1988).

The answer to this question must be that somehow the sentence production system as a whole is adapted to its limited capacity. In terms of current models (e.g., Garrett, 1975; Levelt, 1989), such adaptation could occur at the message level, the level at which the “what-is-to-be-said” is represented in conceptual terms. Complexity at this level leads to structurally complex sentence forms. If these message level representations are fine-tuned to the capacity of the syntactic system, a situation of continuous overload is prevented. All this fine-tuning, of course, would be the effect of a long learning process and would therefore, in the adult speakers at least, be highly automatized.

Now back to the aphasic speakers. If they suffer from a reduction in the size of their temporal window, the fine-tuning between message level and syntactic process is disrupted and the patients are confronted with continuous overload. There is little reason to suppose that the pressure towards fine-tuning that was assumed in the normal speakers would no longer be present in the aphasic. This means that there will be a tendency for simpler message level structures to be selected up to a point at which the system is no longer overloaded and relatively error-free output is possible. However, such selection can no longer occur automatically, as it does in the normal speaker. Readaptation will start as a controlled process that can gradually become automatic again. If it succeeds, it will have a double effect. Not only will it eliminate errors due to overload, it will also create simpler sentence forms. This is basically the hypothesis that my colleagues and I have been defending over the last 10 years. Agrammatic spontaneous speech results from message simplification and message simplification is the adaptive reaction from the aphasic speaker to capacity overload.

The hypothesis that agrammatic speech results from message simplification was first proposed by Isserlin (1922). "It is not damage to sentence finding so much as damage to sentence creation that lies at the root of the grammatical disturbance in aphasia. There is not a derailment of ordinary
speech: the telegram style is laid down, a priori; and we must assume
that there is a schema of telegram thought that anticipates the creation
of telegram speech” [English translation, p. 336].

Isserlin obviously limited his message simplification hypothesis to tele-
graphic speech: the output pattern characterized by a tendency to omit
grammatical morphology. I will have more to say about telegraphic
speech below. First, I will discuss another property of speech produced
by Broca's aphasics, to which the hypothesis can be extended. This is
the symptom of simple but complete sentence form, known as "reduced
variety of grammatical form” (Goodglass & Kaplan, 1972) or "structural
simplicity" (Berndt, 1987; Saffran et al., 1989). Its most important fea-
tures are paucity of phrasal elaboration and relative absence of sentence
embedding. Research by Saffran et al. (1989) has shown that both features
are reliably present in nonfluent speakers, classified as Broca's aphasics.
Hofstede (1992) looked at sentence embedding in a group of 16 Dutch-
speaking Broca's aphasics and found that the aphasics produced signifi-
cantly fewer embedded sentences than the controls (6% vs. 22% for the
controls). However, there is considerable variation between patients: one
patient scored just below the control level (21%) whereas 5 patients pro-
duced 0% embedded sentences. In the theory we have proposed (cf. Kolk
et al., 1985; Kolk & van Grunsven, 1985; Kolk, 1987; Kolk & Heeschen,
1990), structural simplicity results from message simplification on the part
of the aphasic speaker in an attempt to prevent computational overload.
In Kolk et al. (1985), a model is sketched which demonstrates how simpli-
fied messages lead to structurally simple sentences.

Besides structural simplicity, we have telegraphic speech: the so-often
described pattern of output that, at least for languages like English, Ger-
man, and Dutch, is characterized by a tendency to omit grammatical
morphology. Hofstede and Kolk (1994) have made an attempt to quantify
this phenomenon in their recent study with a group of 16 Broca's and
three German speaking aphasics. Both for prepositions and determiners
an average omission rate of 49% for the Broca's was observed, with a
range from 5% to 97%. The normal controls omitted about 4% of these
morphemes, significantly less than the Broca's. It is important to observe
the gradual variation between patients in the amount of telegraphic
speech: from 4% up to 97%, almost every decile contains one or more
patients. Contrary to what is implicitly assumed in the bulk of the agram-
matism literature, agrammatic speech is not an all-or-none phenomenon
but varies in degree. Following Isserlin, we have proposed that tele-
graphic speech results from message simplification, though a more aus-
tere form of it than underlies structurally simple speech. In the message-
simplification model referred to above, we have indicated how this could
occur.

Most of our research has been devoted to the question of whether
telegphic speech in Broca's aphasics is indeed due to adaptation. We have provided two types of evidence to suggest an affirmative answer. The first type of evidence is that the constructions used in telegraphic speech are part of the normal repertoire. The second kind of evidence is that it seems possible for at least a number of telegraphic speakers to shift to a different, more paragrammatic type of output.

Let us first discuss the argument that telegraphic speech is normal. Isserlin's hypothesis is that by means of message simplification, aphasics select sentence forms that their impaired processing system is still able to handle. This implies that these sentence forms are a subset of the normal repertoire. My colleague Claus Heeschen and I have proposed that this subset is the set of elliptical constructions normal speakers use under particular circumstances (1992). We have made a list of eight properties of normal ellipsis in German and Dutch and have shown that these properties are significantly more characteristic of the spontaneous speech of German speaking Broca's aphasics than of Wernicke's aphasics. The properties are the following.

(A) Normal ellipsis allows for the omission of function words. Accordingly, Broca patients showed an omission rate of 56%, and Wernicke patients omitted only 13% of their function words.

(B) Normal ellipsis does not allow for the omission of bound morphology. Therefore, if a normal speaker wants to be very short, he can say "two tables here" instead of "I want you to put two tables here." But he is not allowed to say "*two table_ here." Here the theory makes a nontrivial prediction. Counter to the standard view that agrammatic speech consists of "omission of free and bound grammatical morphology," it is predicted that agrammatics do not omit this bound morphology. The prediction was confirmed, not only for the German-speaking but also for a group of Dutch speaking Broca patients studied by Kolk and Heeschen. It was again confirmed by Hofstede and Kolk (1994) for a different group of Dutch agrammatics. The group of Wernicke patients also showed a near absence of inflection omission. The impression that Broca's omit inflections probably derives from the fact that in English, the infinitive of the verb has the same form as the verb stem and as we will show, Broca patients do overuse infinitives, also in German and Dutch.

(C) Normal ellipsis does not allow for morphological substitutions. Although elliptical speech is deplete of "errors" of omission, substitution errors would make the utterance ungrammatical, just as is the case for full sentences. Taking free and bound morphemes together, Kolk and Heeschen found a 3% substitution rate for the German speaking Broca's and one of 13% for the Wernicke's. The Dutch agrammatics made 1% substitutions. Hofstede and Kolk (1994) observed a similarly low substitution rate in their Dutch agrammatics.
(D) Normal ellipsis allows for the use of nonfinite verb forms (e.g., "Me leaving showbusiness? Never!"). If agrammatics overuse ellipsis, we predict a higher incidence of infinitive use—without accompanying auxiliary—for Broca's than for Wernicke's. Kolk and Heeschen observed percentages of 53 and 9, respectively.

(E) Normal ellipsis allows for main verb omission. So besides "buying flowers for Sally," we have "flowers for Sally" as a regular form of ellipsis. Overuse of ellipsis would therefore lead to a high omission rate for main verbs. The percentages found by Kolk and Heeschen were 41 and 9, for Broca and Wernicke patients, respectively.

(F) Normal ellipsis allows for omission of the grammatical subject (see the examples in the previous paragraph). The percentages for the two groups of German speaking patients were 50 and 18, respectively.

(G) Normal ellipsis prescribes the word order of the subordinate (S-O-V) rather than that of the main clause (S-V-O). S-O-V word order was observed in 41% with Broca's and in 10% with Wernicke's.

(H) Normal ellipsis prescribes, at least in German, a strong adjective inflection. So a normal speaker can say "Der blaue Engel" (the blue angel). He is allowed to omit the article, but then he has to inflect the adjective differently (e.g., "Blauer Engel"). The German-speaking agrammatics had this strong inflection with 74% of their adjectives, and Wernicke's had it in 34%.

Another strategy to demonstrate the normality of agrammatic speech has been followed by Hofstede (1992). Instead of looking at the relative frequency with which elliptical features are present in agrammatic speech, he selected telegraphic utterances from both agrammatic and normal speech and compared the two sets of constructions. A "telegraphic utterance" was defined as any utterance in which finiteness was omitted. Finiteness was considered to be omitted whenever (a) the main verb was omitted or (b) a nonfinite form of the main verb was used, without there being an additional auxiliary. This criterion was chosen because finiteness omission appears to represent the clearest break with the full sentence. It turned out to be a good choice: most of the omissions Hofstede observed in the speech samples he studied were covered by this criterion (92% of the agrammatic omissions and 83% of the normal omissions). Three samples of speech were obtained: (a) Speech obtained in conversations with normal adults, matched in age and education with the aphasics. During these conversations, an attempt was made to create an informal atmosphere, which we know stimulates the production of elliptical speech. (b) Speech produced by municipal officials talking to foreigners, the so-called "foreigner talk." (c) Speech obtained from Broca's aphasics during free conversations.

The normal telegraphic utterances were categorized first. The most frequent category was labeled "isolated predicates" (91%). The following subtypes were observed.
(1) Nonfinite verbs, e.g., "in China geboren" (born in China).
(2) Noun phrases, e.g., "mooi huis" (beautiful house).
(3) Prepositional phrases, e.g., "naar een feestje" (to a party).
(4) Adjectives, e.g., "heel normaal" (very normal).
(5) Adverbials, e.g., "nog een keer" (one more time).

The other main category was called "subject-predicate constructions" (9%). The same five subtypes occur here as in the category "isolated predicates," this time preceded by an NP.

(6) Noun phrases followed by a nonfinite verb, e.g., "ik nooit gezien" (I never seen).
(7) Noun phrases followed by another noun phrase, e.g., "rode wijn tien gulden" (red wine ten guilders).
(8) Noun phrases followed by a prepositional phrase, e.g., "zoon twee dagen in bed" (son in bed for two days).
(9) Noun phrases followed by an adjective, e.g., "mensen in Spanje vriendelijk" (people in Spain friendly).
(10) Noun phrases followed by an adverb, e.g., "mijn vrouw terug" (my wife back).

The samples of foreigner talk showed the same 10 categories of telegraphic expressions. Furthermore, more or less the same relative frequencies were observed in the two samples: isolated predicates by far outnumbered the subject-predicate connections and within the first category, isolated NPs were the most frequent, followed by the isolated nonfinite verbs; the other categories had about the same (low) frequency.

The interesting question, of course, was what do the aphasic utterances look like? As it turned out, all categories of expressions present in the two normal samples were also present in the aphasic samples and vice versa. Furthermore, as was the case with foreigner talk, the relative frequencies of the two main categories and of the two most frequent subcategories (1 and 2 in the above list) were the same. The main difference between the aphasic and the normal samples concerned the complexity of the telegraphic constructions. The aphasic showed less phrasal elaboration (e.g., adding a determiner or an adjective to a noun) than was present in the normal conversations, although the amount of elaboration of foreigner talk was less than in the aphasic speech. Furthermore, the aphasics produced fewer grammatical objects (direct, indirect, or prepositional object) with the nonfinite verbs (e.g., "op Jan wachten"—waiting for John).

It is interesting to note that Saffran et al. (1989) describes a subcategory of agrammatic utterances they call "topic-comment structures." The five examples they present (Saffran et al., 1989, p. 471) correspond exactly with our subcategories 6 to 10: "dancing Cinderella and prince" (6); "Cinderella washerwoman" (7); "Cinderella in the house" (8); "Cinderella very pretty" (9); "party over" (10). Isolated predicates were observed as well; they were put into the class of "Other utterance catego-
ries, NP, VP, etc." Unfortunately, no examples or frequencies of occurrence were provided.

I think it is no exaggeration to say that, as far as structural similarity is concerned, the evidence for the ellipsis hypothesis is pretty strong. Let us now turn to the other type of evidence, that at least some agrammatics can shift from agrammatic to more paragrammatic speech. We have assumed that agrammatic speech results from message simplification and that this simplification is under control of the subject. This implies that a patient should be able to at least partially refrain from message simplification. This will result in more overload for the grammatical system, synchrony will be harder to achieve, and more errors of substitution will occur. On the other hand, the omission rate for function words will go down (omissions of inflections are true errors and will therefore show a rise).

Heeschen (1985) asked a group of German-speaking Broca and Wernicke patients to describe a picture to an experimenter who could not see the picture. He observed that, in comparison to spontaneous speech, Broca's produced fewer errors of omission and more of substitution. The Wernicke patients showed no such effect. In terms of omissions and substitutions of grammatical morphology, the output patterns of the two groups were indistinguishable. Kolk and Heeschen (1992) analyzed a larger data set from Heeschen's original study and came to the same conclusion. Hofstede and Kolk (1994) demonstrated with a large group of Dutch-speaking agrammatics that, at least in a number of patients, one can obtain a shift toward paragrammatism simply by giving the patients pictures to describe, without hiding the pictures from the experimenter. Hiding the pictures gave an additional effect for a number of patients. Bastiaanse (1994) described a case of a Dutch-speaking agrammatic woman, who—in the middle of an interview—demonstrated a complete change in type of output. Whereas in the beginning her sentences were complete but with many morphological as well as constructional errors, later on she shifted to a classical form of telegraphic speech. A final demonstration comes from Kolk and Hofstede (1994). In a case study of an agrammatic patient, they compared spontaneous speech to a condition in which the patient simply was requested to "talk in complete sentences." Each time the patient produced a telegraphic utterance, he was reminded of the instruction. A dramatic decrease in omission rate resulted from this manipulation, accompanied by a significant rise of the percent of substitutions.

Two recent studies (Goodglass, Christiansen & Gallagher, 1993; Hesketh & Bishop, in press) appear to indicate that there is only a limited shift toward paragrammatism in English-speaking agrammatics. On the one hand, it could be that a language-specific factor is responsible for this finding, having to do with the existence and form of an elliptical
register in English. On the other hand, there are a number of important procedural differences between the two studies and ours, having to do with baseline, elicitation procedure, response parameters, and the like (see Kolk & Heeschen, in press, where the Hesketh and Bishop study is discussed in detail). The question whether substantial shifts toward paragrammatism can also be obtained with English-speaking agrammatics is therefore still unanswered.

It does appear, therefore, that at least some patients can make a shift from agrammatic to a more paragrammatic type of output, as predicted by our theory.

Let us now return for a moment to the Kolk and Hofstede (1994) study, where it was shown that a shift from agrammatism to paragrammatism could be produced "upon request." In this study, we also looked at speech rate and number of word repetitions. We observed that the shift to more complete utterance forms was accompanied by (a) a sharp reduction in speech rate and (b) a large increase in number of word repetitions. How do we account for these changes?

In 1985, Marianne van Grunsven and I proposed that besides simplification of sentence form, there is another way a patient could react to the newly acquired capacity limitation (see also Kolk & Heeschen, 1990, where this point is elaborated in more detail). Instead of trying to prevent premature disintegration of sentence representation by means of message simplification, a patient could try to repair the disintegrated representation. We called these two reactions preventive adaptation and corrective adaptation, respectively. Whereas preventive adaptation leads to structural simplicity and telegraphic speech, corrective adaptation leads to the symptom of nonfluency, one of the hallmarks of the syndrome of Broca's aphasia. Increased nonfluency is what we observe in Kolk and Hofstede's (1994) patient, when he gives up telegraphic style. Because the patient tries to produce more complex utterance forms, he finds himself more often in a situation of capacity overload. As a consequence, more errors occur at what Levelt (1989) calls the level of "inner speech" (also referred to as the "phonetic plan"). The patient can allow these errors to come out and produce paragrammatic errors. He can also, by means of "covert repairing" (Levelt, 1989) edit these errors out and produce relatively error-free speech. This will take time, hence the nonfluencies.

How does this covert repairing take place? In the normal speaker, one could conceive of this repair process as involving a sequence of steps: (a) error identification; (b) computation of the correct element; (c) replacement of the erroneous for the correct element. If this were to be the case, it is hard to see how an aphasic patient with a severely limited capacity would be able to do any repairing at all. I do not think, however, that all this processing is necessary. After step (a), the normal speaker can simply restart the production of the sentence, or part of the
sentence. Since the likelihood of making a speech error is very low, it is highly unlikely that he will make two errors in a row. Restarting therefore automatically leads to repairing. In the aphasic speaker, however, the situation is different. Here the likelihood of making errors is high and, other things being equal, it will remain high, even if a second attempt is made. How then can restarting lead to improvement?

For an answer to this question, it is necessary to consider again the nature of the capacity limitation we have proposed. Suppose the critical limitation would consist in a reduction of the size of a syntactic buffer, as has been suggested by Caplan and Hildebrandt (1988). If a particular construction would be too complex for this small buffer, restarting would have no benefit whatsoever: a structure that does not fit into the buffer will never fit. When the limitation is a temporal one, however, the situation is different. The basic difficulty that results from a timing deficit, as I have defined it, is that a particular representational element decays before other elements, with which it has to be in synchrony, are activated. Now restarting does offer an advantage. After decay, the element may still have a relatively high level of activation, because of the fact that it has just been activated. Reactivation can occur from this level, rather than from the rest level, and the critical element will reach the threshold sooner. In this way, restarting leads to faster processing and to improved performance.

We can conclude therefore that there is evidence for both preventive and corrective adaptation. Is it possible for a patient not to adapt, by neither simplifying nor restarting the sentence? In such a case, we would expect fast and complex speech, full of errors. The speech of Wernicke patients is like this and consequently, my colleague Claus Heeschen and I have proposed that the difference between Broca's and Wernicke's, as far as sentence production is concerned, is a difference in adaptation, not in type of grammatical impairment. Evidence that both groups of patients have the same production deficit comes from two studies, already described above. Kolk and Heeschen (1992) have shown that in the constrained picture description task, overall omission and substitution rates are the same for the two groups. Haarmann and Kolk (1992) gave a cloze task to the two groups. They compared the error profiles over seven types of grammatical morphemes and found no significant difference. The same was true for the profile of response rates, as well as for the average rates overall. So the hypothesis that there is a common production deficit in the two groups receives some support.

Suppose for the moment that the two groups do have the same impairment; then another question arises: why do the Wernicke's aphasics not adapt? This is a tough question that we have not yet been able to answer satisfactorily. We have considered the possibility that Wernicke's aphasics are less aware of or less concerned about their errors and therefore
do not bother to avoid them (Heeschen, 1985). This led us to expect that
the behavior of the two groups in the cloze task would be very different.
The Broca's would be expected to be slow and accurate, the Wernicke's
to be fast and sloppy in their performance. What we found, however,
was that both groups were equally slow and equally accurate (Haarmann
& Kolk, 1992). The fact that the two groups were equally slow was re-
markable, because in their free conversations, the Wernicke patients
spoke three times faster than the Broca patients. An alternative explana-
tion that we suggested in Kolk and Heeschen (1992) is that, as a side
effect of their brain damage, Wernicke's suffer from a press of speech
and have therefore little time to notice their errors. But this leaves an-
other question unanswered: why is no press of speech observable in the
cloze task?

CONCLUSION

It will be clear from the above that the timing approach to agramma-
tism, both in production and comprehension, has its problems and open
ends. However, I believe it also has a number of virtues. (a) It gives
an explicit description of the capacity limitation underlying agrammatic
behavior. (b) It suggests a model for the genesis of paragrammatic errors.
(c) It also suggests a model for the genesis of agrammatic errors. The
strategies that are assumed to underlie these errors make sense within
the time-based approach as a whole. They are not postulated in an ad
hoc fashion. (d) Finally, an integrated account is given of an aspect of
agrammatic speech that is often considered to be less interesting: nonflu-
ency.

REFERENCES

Bastiaanse, R. 1994. Broca's aphasia: A syntactic and/or a morphological disorder? A case
Berndt, R. S. 1987. Symptom co-occurrence and dissociation in the interpretation of agram-
matism. In M. Coltheart, G. Sartori, & R. Job (Eds.), The cognitive neuropsychology
Applied Psycholinguistics, 1, 245-299.
MA: MIT Press.
Cornell, T. L. 1995. On the relation between representational and processing models of
asymptatic comprehension. Brain and Language, 50, 304-324.
Dell, G. S. 1986. A spreading-activation theory of retrieval in sentence production. Psycholog-
ical Review, 93, 283-321.
Dell, G. S. 1990. Effects of frequency and vocabulary type on phonological speech errors.
Language and Cognitive Processes, 5, 313-349.


Natural language generation; new results in artificial intelligence, psychology and linguistics. Dordrecht: Martinus Nijhoff.


[Abstract]