Signed and Spoken Language: Biological Constraints on Linguistic Form, eds. U. Bellugi and M. Studdert-Kennedy, pp. 141-160. Dahlem Konferenzen 1980. Weinheim: Verlag Chemie GmbH.

On-Line Processing Constraints on the Properties of Signed and Spoken Language

W. J. M Levelt
MPI für Psycholinguistik
Nijmegen, Netherlands

Abstract. It is argued that the dominantly successive nature of language is largely mode-independent and holds equally for sign and for spoken language. A preliminary distinction is made between what is simultaneous or successive in the signal, and what is in the process; these need not coincide, and it is the successiveness of the process that is at stake. It is then discussed extensively for the word/sign level, and in a more preliminary fashion for the clause and discourse level that on-line processes are parallel in that they can simultaneously draw on various sources of knowledge (syntactic, semantic, pragmatic), but successive in that they can work at the interpretation of only one unit at a time. This seems to hold for both sign and spoken language. In the final section, conjectures are made about possible evolutionary explanations for these properties of language processing.

SUCCESSIVENESS AND PARALELLNESS IN NATURAL LANGUAGE

A striking design characteristic of spoken language is its serial nature: syllables, words, clauses, and utterances are to a large degree serially ordered. Of course, some parallelism is to be observed as well. There is coarticulation both within and between syllables, and there is prosodic information such as stress and intonation which may spread over words, sentences, and whole paragraphs. And then there are gestures which can be essentially required as parallel information in the deictical use of terms such as "there" and "he," and optionally for various other purposes. Prosody and especially gesture are sometimes called "paralinguistic" variables. This label is often preferred wherever spoken language is not strictly successive, thus making the linguistic structure seem to be even more serial than it is to start with.

Nevertheless, successiveness is doubtless a dominant property of spoken language, and one might ask whether this is a somewhat trivial consequence of the acoustic mode or whether it has its roots in the more central parts of the human apparatus for language processing.

There is much to argue for a mode explanation. Parallelness in spoken language appears quite freely where the articulatory and perceptual equipment allows for it. The musculatures of lungs, larynx, and mouth can be controlled rather independently, which allows for the mentioned forms of prosodic parallelness. The human ear, moreover, shows a remarkable ability to disentangle these parallel sources of information. Though the perceptual mechanisms involved are still not well-understood (see especially Studdert-Kennedy (30) for a recent discussion of some of these issues), they are certainly available.

At the same time, it is clear that parallelness of a more extended sort surpasses the abilities of both the articulatory and auditory apparatus. It is impossible to articulate two words or larger units at the same time, since they involve the same musculature. Moreover, people have great difficulty in identifying two simultaneously sounding words or larger messages.

Still, there is reason to doubt that a peripheral mode explanation of successiveness suffices. People are surprisingly unable to simultaneously speak and write different messages, in spite of the fact that separate musculature is involved. Also, there are strong restrictions on the understanding of two messages which are dichotically presented, i.e., without peripheral auditory interference (see Yates and Thul (38), for a recent summary of this work). These kinds of phenomena are usually explained in terms of central attention limitations which affect both within and between mode processing.

A comparison to sign language adds a new dimension to these considerations, as was already noted by Klima and Bellugi (13). The visual mode allows for larger degrees of simultaneity than the auditory mode. Whereas the ear seems to be primarily designed for the analysis of temporal information, the eye has strong spatial abilities. Parallel computation of simultaneous visual information is normally involved in pattern recognition and spatial orientation. These spatial abilities of the eye could be exploited for parallel processing in sign language. This would, moreover, also be a possibility from the production point of view: the two arms and hands are anatomically equivalent organs of articulation, and although usually one hand is dominant, the two can act simultaneously in rather independent ways, especially if trained to do so. But as Klima and Bellugi (13) observe, much parallel information (e.g., relating to number and aspect) can be packed in a single sign, but two-handed simultaneity of signs is exceptional.[1]

[1] An earlier version of this paper has triggered several observations by myself and others of simultaneous signing. There are, first, the well-known cases of making a lexical sign with one hand and an indexing gesture with the other hand. Second, I observed cases where a lexical sign is made which is then replaced by a classifier in one hand. While this classifier is held, the other hand signs the focal information entertaining some relation to the classifier. In the cases observed, the classifier functioned as a locative in a spatial predication. The use of a classifier in these cases suggests an essentially anaphoric function of the second hand, just as in indexing. I made a third type of observation where this was clearly not the case. This was a relative clause which was made by one hand while the head noun ("the girl") was held by the other. After this latter observation, which was made during the conference, the question arose whether it is essential for one hand to be held during simultaneous signing. This does not seem to be the case; grammatical counter examples could be made in which both hands were moving. So, for instance, one hand could sign a quantified ("all") NP, whereas the other expressed an adjective with a characteristic aspect ("characteristically sick"). Quantification and aspect are both expressed by means of hand movement. It was noticed, however, that these cases were quite difficult to sign. In all observed cases, the two hands participated in a single predication. Simultaneous signing of different propositions seems to be ungrammatical in all cases.

It seems, therefore, that successivity is in the first place a mode-independent design characteristic of human language, and one would expect to find this successivity reflected in the way language is perceived and produced in the one as well as in the other modality. The following pages will deal with the question as to what is successive and what is parallel in the processing of natural language. This will, where possible, be done comparatively for spoken and sign language. The aim is to make a more specific assessment of those features of language processing which underly the serial character of language. Though this is the aim, we will not be able to make more than first steps. The final paragraph lists some of the major psycholinguistic and biological questions which remain unanswered.

THE WORD/SIGN LEVEL

What is the course of recognition of words and signs? Are they analyzed in a sequential component-by-component fashion, or is recognition based on a computation of all features simultaneously? Here, one should distinguish carefully between what is simultaneously there in the sign, and what is simultaneous in the process. Information that is essentially sequential in nature may be assembled in an echoic or iconic buffer and in that way become available for simultaneous analysis. On the other hand, information which is simultaneously given either perceptually or in memory may be attended to in a sequential manner. As long as the perceptual and memory mechanisms are unspecified, all four combinations are possible: sequential information can be analyzed in parallel or sequential fashion, and the same is true for simultaneous information.

It has been claimed that there is more simultaneity of information in the sign than in the word (13,32). The different components of the sign (location, orientation, handshape, and movement) are largely simultaneously visible in adult signing. As compared to signs, phonemic information in words is more successive, especially between syllables. While signs played in reverse are still somewhat recognizable (1), words in reverse are not. And experiments by Tweney, Heiman, and Hoeman (34) give further demonstration of this difference: flickered disruption is far less detrimental to the recognition of signs than an equivalent acoustic disruption is to the recognition of words. This is true for the whole range of interruption frequencies (0.5 to 4.0 interruptions per second) and speech/sign time fractions (25 to 75%). There is more "spreading" of information over the sign than over the word. Bellugi and Fischer (3) have argued that this larger possibility of perceptual co-presence is effectively used in ASL for packaging inflectional information into the sign. Klima and Bellugi (14) show how this facility is used in verb inflection and for some types of adjectival modification.

Still it should be recognized that not all information is simultaneously present in the sign. Grosjean, Teuber, and Lane (9) noticed that the movement information, which is essentially temporal in nature, is only fully given after the other parameters have been established. The delay can be on the order of 200 ms, dependent on the sign. The other components, moreover, are not simultaneous either: orientation and location come first, then comes handshape, after 30-40 ms, and only much later is movement information completed. These findings result from experiments in which native users of ASL had to copy the particular components on the basis of initial stretches of the (gated) sign; these stretches increased in duration on successive trials.

The question then is whether this distribution of information is reflected in the on-line process of sign recognition. As noticed above, this is not necessarily the case, since an iconic buffer may intervene in the recognition of signs. What we would like to claim is that the recognition process for both word and sign is largely sequential in character: different components of information are successively taken into consideration. There is, as yet, only very little support for this claim in the sign language literature. Before turning to that, I should summarize the somewhat more extensive evidence for spoken language which was obtained by Marslen-Wilson and Tyler in our laboratory. They developed what is now called the cohort theory of word recognition. For isolated words the cohort theory says the following:

(a) The acoustic information in the first 100-150 ms of a word is used to activate a so-called "word initial cohort" (WIC). The WIC is the set of items in the listener's lexicon whose initial part is compatible with the acoustic information up to that point.
(b) The members of the cohort monitor the incoming information as it comes in, in bottom-up fashion.
(c) Everytime a mismatch occurs for a member of the cohort, it is excluded. In this way the cohort continually decreases in size.
(d) If only one element is left, a recognition decision is made.

Cohort sizes are not very large: after 200 ms the average cohort for American words has a size of about 29 elements. The single member situation will be reached earlier or later depending on the character of the alternatives in the cohort, not on the number of alternatives. For each word there is a critical "decision point" where all alternatives mismatch the information up to that point.

In one experiment, Marslen-Wilson (21) had subjects perform a word/non-word lexical decision task. The spoken non-words formed the experimental material. They were constructed in such a way that the critical point where the items could be distinguished from all word alternatives varied from the first consonant cluster to the first vowel to the second consonant cluster. The experimental non-words and filler words were presented acoustically to the subjects and word/non-word decision latencies were measured. The main outcome of the experiment was that decision latency was an almost constant 450 ms from the critical point. This was independent of (non-word) length, position of the critical point in the non-word, and number of alternatives in the cohorts.

In another experiment, subjects performed a phoneme monitoring task on words where, again, the decision point was varied. The test phoneme (always a /t/) could occur at various places before or after the decision point. It turned out that the best (and linear) predictor for the monitoring latencies was the temporal separation between test-phoneme and decision point: the later the test-phoneme relative to the decision point, the shorter the latency. These and other results testify to the psychological reality of "decisions points" which are reached by an on-line sequential reduction of the set of possible alternatives in word recognition.

From word monitoring experiments (22), it appears that if subjects monitor a scrambled word list (random text) for the occurrence of a particular target word, identification times (excluding response execution) amount to an average 300 ms for words with an average length of 370 ms. For a normal text this reduces to about 200 ms, i.e., identification can occur long before the word ends. The sequential character of the cohort theory may thus lead to an explanation of the still unsolved problem of word boundary recognition. In fact, the cohort theory seems to mesh beautifully with the Nakatani and Dukes' (27) findings on word juncture.

Let us now return to sign language. Given limited evidence, the cohort theory may turn out to be applicable to the recognition of signs as well. Grosjean, Teuber, and Lane (9) write: "as the presentation time of the gated sign increases in duration, and information concerning location and orientation is obtained, observers start examining the reduced set of lexical alternatives for appropriate signs. Their guesses share the location and orientation parameter with the target sign but differ on two parameters: handshape and movement. As the handshape parameter is identified, the pool of possible signs is reduced even more," etc. This conclusion is based on sign recognition data for gated signs. The copying data mentioned above told us about the temporal order in which component information becomes available perceptually, the recognition data give insight into the order in which parameters are recognized as features of a sign. The data show that in order to use the orientation and location parameters in guessing a sign, some 60 ms more is required than for using them in a copying task. In terms of the cohort theory, this time may have been used to activate the word-initial cohort, or better the sign-initial cohort (SIC): no guess can be made without an activated SIC. The average total time to reach this situation from the onset of the sign is about 400 ms in the Grosjean et al. study, which is roughly the same as the corresponding value for speech. Moreover, the authors observe that the movement parameter is recognized as a feature in almost the same time as is required to recognize the sign as a whole. No more than 51% of the total sign duration is, on the average, required for the correct recognition of the sign, and thus of the movement feature. For spoken words in isolation, Grosjean (8) reports a fraction of 83% needed for correct identification. This value, obtained from a gating study, compares well to the 81% identification time which Marslen-Wilson and Tyler (22) obtained for words in a scrambled list. Though these percentages are clearly higher than the 51% reported for isolated signs, they correspond, as noticed, to an equivalent absolute duration. For both, signs and words frequency effects can be observed, whereas context can effectively reduce both word (see below) and sign identification latencies. One modality difference, brought to my attention by Grosjean, is the existence of a word length effect for spoken words, but the absence of such an effect for signs. In order to theoretically interpret this discrepancy, the cohort theory might be useful: the decision point for words will doubtless vary with word length, but this may be far less so for signs, due to the more simultaneous presence of the different information parameters.

If the cohort theory applies to signs as well as to words, important further questions can be asked with respect to sign recognition in running text. More specifically, one can study how semantic and syntactic knowledge is used in the recognition of a sign. Here it should be remembered that the "classical" picture of the comprehension process is of a rather strict sequential sort: higher-level syntactic and semantic information (which we will, for short, call "context information") will not affect lower level decisions. Lower level analysis strictly precedes higher level analysis: word recognition precedes syntactic parsing (4), syntactic parsing precedes semantic interpretation (4,5), and there are no feedback loops. The claim that context can disambiguate words only after they have been recognized was strongly made by Foss and Jenkins (6), and most recently by Tanenhaus, Leiman, and Seidenberg (31). But in a review of the research on lexical ambiguity, Levelt (15) showed that this model is unnecessarily complicated, and that all known experimental results can be explained by assuming that context works at, not after, the level of word recognition. This is in full agreement with the extensive evidence accumulated over the past years by Marslen-Wilson, Tyler, and co-workers (see especially (20, 22,24)), leading to the conclusion that "the entire range of processing activities are seen as taking place on-line as the utterance is heard" (22). So, for instance, Marslen-Wilson and Tyler (22) found in word monitoring experiments that if the context in which the word appears is not a random word list, but a syntactically structured (though still semantically anomalous) text, word monitoring reaction time decreases by an average of 30 ms. If the text is furthermore made meaningful (i.e., "normal prose"), a further 60 ms reduction of monitoring latency is observed. The syntactic and semantic information, which must have been derived from the previous part of the text, is used immediately where applicable, viz, <u>during</u> the processing of newly incoming words. This often makes it possible to recognize the word before bottom-up acoustic information could have reduced the cohort to a single member. The cohort theory for isolated words, as stated above, therefore has to be extended for words in context in such a way that the cohort members monitor not only the acoustic information for mismatches, but also the syntactic and semantic information. Tyler (36) draws a comparison to the heterarchical models developed in Artificial Intelligence (e.g., Winograd (37)), in which different knowledge sources simultaneously cooperate to produce a single higher-level representation. It should be noted that this is still quite programmatic: a theory of on-line word recognition in context requires, among other things, further specification of the amount of categorizing taking place before (or after) context information is applied. There are alternative theories of word recognition which deal with such issues. As Levelt (15) points out, Morton's logogen theory can easily handle the "simultaneous" effect of context in the recognition of ambiguous words. More detailed specifications of the cohort theory are necessary to find out whether and where it differs from alternative accounts.

The evidence that the cohort theory applies to signs in context is still very limited. An experiment by McIntire and Yamada (25) showed that "close" ASL-shadowers (but not so close as in the Marslen-Wilson (19) study) make on-line semantic substitutions (like OK → FINE), which shows the immediate availability of semantic information. But it is, first, necessary to check whether SIC and WIC are similar in that the onsets of the two signs involved are identical (as it has to be for words which are semantically substituted by close shadowers). And second, the occurrence of a very occasional non-semantic formationally similar substitution (e.g., WAIT → NOW) shows that it is still too early to say that in all cases the available semantic information is in fact used in on-line sign recognition. The obvious approach to study this further is to apply the word, category, and rhyme monitoring paradigm developed by Marslen-Wilson and Tyler to normal, syntactic, and random sign text. Rhyme monitoring is especially interesting since one can define different formational similarities between signs for the subject to monitor. They may be differentially affected by previous context because they perform different functions. For example, while handshape and location and underlying movement stem carry most of the lexical burden, the rich inflectional information is almost exclusively carried by overlaid movement contours.

THE CLAUSE LEVEL

The issue discussed for the word/sign level repeats itself for the clause and sentence level. The "classical" theories (see (4,5)) claimed strong seriality in the processing of clauses. Forster (5) writes: "semantic processing is delayed until intact deep structure units have been isolated, and in general, considerations of meaning are irrelevant to syntactic decisions." The "staggered" model we have seen for the word level reappears for the clause level. The clause is first analyzed syntactically, then semantically. It is not the case that both types of information are used simultaneously in on-line fashion in order to derive a single representation for the clause or sentence.

Tyler (36) reviews various versions of this theory and shows how unlikely they are. Again, her and Marslen-Wilson's experiments show that semantic information is used during clause understanding, before the syntactic structure of the clause can be known. The heterarchical parallel-interactive model seems to apply at the sentence level just as well as at the word level. The syntactic clause in itself has no special status in processing. If the clause boundary is a genuine informational boundary, it will appear to have "psychological reality" in on-line measurements, but if it is informationally incomplete it loses its "integrity" (23).

It should be noted that the literature which seemed to support the clause-by-clause staggered processing theory was mostly based on the use of post hoc measurement. Levelt (15) in a review of the sentence processing literature showed that the click paradigm, as well as immediate recall and other techniques, did not measure what happened during the clause or at the clause boundary. This almost always allowed for alternative explanations since the subject's response could have been (co-)determined by what happened later in the sentence, or even after the sentence. Post hoc measurement paradigms may tell us something about the organization of linguistic information in memory, but they can hardly ever assign the subject's responses to the immediate perceptual processes. Only on-line (or "simultaneous" in the terminology of (15)) measurement can help us sort out the critical issues involved.

It would, similarly, be a waste of time for sign language researchers to study the processing of utterances in context by means of post hoc measurement techniques. Examples are probed and free recall studies (2,11,33,35). These studies may be interesting from the "psychological reality" point of view, and even tell us something about sign memory, but they do not allow for any conclusion with respect to the on-line computation of signed text.

The almost complete absence of empirical "on-line" evidence to support either a successive ("staggered") or a parallel (interactive) theory of clause and sentence processing in sign language reduces our present efforts to the making of hypotheses. The conjecture we would like to make is that observers, just as listeners, will use whatever information is available to interpret the constituents as they come in. They will not postpone semantic interpretation until a full syntactic clause has been assembled. For signers, as well, the clause will have no special status as a syntactic unit in the processing (perception and production) of text. One reason why this should be so for sign language just as for spoken language is that staggered processing requires intensive usage of memory buffers: the next level of processing can only apply to the full output of the earlier level. Therefore, that output has to be assembled and stored to be available in its entirety before the next higher level procedure can be run. However, the usage of memory buffers requires attentional effort, and one would expect human language processing to minimize such effort. This can be done by dismissing information as rapidly as possible. Reef, Lane, and Battison (28) found that gestures which are understood as signs in ASL are more quickly dismissed from iconic storage than gestures that are not taken to be signs. Rapid on-line linguistic interpretation reduces the storage requirements and thus attentional effort in both spoken and sign language.

THE LEVEL OF DISCOURSE

As far as the production and perception of discourse is concerned, we are in a state of happy ignorance. A major issue in the study of discourse production is what I have called "linearization" (16): how does a speaker order propositions for expression? For spoken language, linearization is an absolute requirement. Levelt (17) discusses two sets of determinants of linearization: structural determinants, derived from the organization of the information to be expressed and the mutual knowledge involved (e.g., temporal information is preferably expressed in chronological order, etc.), and processing determinants derived from the speaker's and hearer's memory limitations. As far as the experimental results go, the latter limitations have a strikingly uniform effect on discourse production. They predictably maximize the connectivity of discourse, and they minimize backtracking to already expressed information. But where backtracking cannot be avoided because of the multiple connections of that piece of information to other parts of the informational structure, it, again according to predictions, follows a very regular pattern: items to be returned to are treated in a last-in-first-out manner. If, moreoever, an item allows for two or more (connected) continuations, the alternatives get ordered in such a way as to minimize the number and duration of return addresses in memory. The combination of these features leads to maximally right-branching structures in spoken discourse. Levelt (16) shows that such right-branchingness is to the advantage of both speaker and listener. Apart from the experimental evidence which involved the description of spatial patterns, there are anecdotal indications (17) that right-branchingness obtains for other types of discourse as well (such as descriptions of dinners and of kinship structures).

Here we would like to support Grosjean's (7) hypothesis that speech and sign language "probably share some common production mechanism." We would like to suggest that the generation and linearization of propositions in discourse proceeds from the same mechanism in both sign and spoken language. Levelt (17) argues that this mechanism is pre-linguistic and largely unaffected by the formulating operations which cast these propositions in linguistic form. We thus do not share Battison's (1) expectations that "temporal distribution of information in sign utterances, and over several utterances in a discourse, may be different from that of speech," at least not for the discourse level.

SOME REMAINING ENIGMA'S CONCERNING SUCCESSIVENESS AND PARALLELNESS

The gist of the previous paragraphs is that on-line processing is highly similar for spoken and sign language, and that the similarity increases from word/sign to sentence to discourse level. It was furthermore put forward that parallelness of processing is not the exclusive mark of a visual language, but that normal spoken language is highly parallel in nature as well: different sources of knowledge (stimulus information, syntactic, semantic, and other context information) are applied simultaneously and on-line in order to derive an interpretation for the incoming speech.

But this leaves us with some unresolved issues which are in part empirical, psychological ones, and in part biological ones that may never be answerable. The central question is this: <u>if parallel processing is within the range of human capabilities, why then is it not used to a larger degree in sign language?</u>

As the earlier paragraphs have tried to clarify, the parallelness in natural language processing is a very special character. It resides in the simultaneous usage of different sources of information in the interpretation of a given constituent. But for both the spoken and the signed modalities parallelness never seems to involve the simultaneous interpretation (or production) of two propositions. The existing evidence on dichotic listening shows the same pattern. Yates and Thul's (38) short review of the literature and their own experiments show that parallel processing occurs in dichotic listening, but only to the extent of affecting a single word meaning or proposition. There is no evidence so far for the simultaneous computation of two propositions. The same seems to hold for visually presented words. Shaffer and Laberge (21) review the literature and show that there is parallel semantic processing of simultaneously presented words in the visual field. But again the only effect demonstrated is that the processing of the attended-to word is affected by the meaning of adjacent words. There is no evidence for the simultaneous derivation of two word-interpretations in the sense of "conscious availability," let alone for the parallel derivation of two propositions. There is still one more notable form of simultaneous processing of natural language, namely, in simultaneous interpretation, where listening to one language is parallelled with speaking in another language. Karmiloff-Smith (12) shows that there is nothing exotic about this activity; it is based on forms of simultaneous processing which show up on other linguistic activities as well. What is important here, however, is the question of whether the interpreter is working at two different propositions simultaneously. The evidence is equivocal, but if this would turn out to be the case, we must conclude that language perception and production may call on different and somewhat independent conceptual processes, which can run in parallel.

The empirical, psychological question is just how much interpretation of a signer can do in parallel. Consider some cases of increasing complexity: can a native signer recognize two nouns that are simultaneously presented by two hands (of one or of two signers)? If so, is this still possible if iconic storage is experimentally excluded? If the native signer cannot do this, would he or she be able to recognize a simultaneously presented adjective/noun pair which allows for a single interpretation (instead of two independent ones)? If such a capability exists, does it also hold for a noun/verb pair which can be interpreted as a single actor/ action proposition? And what, finally, can a signer do with two simultaneously presented propositions, either independent ones, or related ones (e.g., one being a relative clause)? These and similar questions should be studied, since it is important to know whether the visual modality indeed allows for more than what is in fact used. (But notice that most of these questions have never been studied for the dichotic listening case either.) The observations reported in footnote 1 show that this can be a viable research program.

My conjecture is that the results will not be much different for simultaneous signing and dichotic listening, or to put it differently, that the limitations on the simultaneous processing of propositional information are essentially central in origin.

This, then, brings us to the final biological question: what are the evolutionary causes of these limitations? It has been proposed that natural language evolved from a visual-gestural mode (this is a longstanding tradition in biology, but for recent statements see Hewes (10) and McNeill (26)). However, as Levelt (18) argued in response to McNeill's paper, there is no evidence whatsoever to support this view. If one takes the other stand, that from the beginning, language evolved in the vocal-auditory mode, then one can reason that the limitations of the vocal channel shaped the way in which homo sapiens deals with propositional information. Since it is impossible to articulate two constituents at a time, there was no survival value in developing the capability of parallel processing of propositions; this would have most useless in the niche of our ancestors. If the structure of the vocal tract thus constrained the evolution of propositional processing to an essentially sequential nature, it is indeed not surprising to find that native signers are subject to the same limitations as are native speakers: they have the same genetic endowment.

Acknowledgements. I am grateful to U. Bellugi, A. Karmiloff-Smith, W. Marslen-Wilson, J. Morton, and M. Studdert-Kennedy who made very helpful comments on an earlier version of this paper.

REFERENCES

(1) Battison, R. 1979. On-line processing in English and American Sign Language (Research proposal).

(2) Battison, R.; Lane, H.; and Grosjean, F. 1978. Probed recall latencies reflect both within clause and across clause boundaries in American Sign Language. Working paper, Northeastern University.

(3) Bellugi, U., and Fisher, S. 1972. A comparison of sign language and spoken language. Cognition 1: 173-200.

(4) Fodor, J.A.; Bever, T.G.; and Garrett, M.F. 1974. The Psychology of Language. New York: McGraw-Hill.

(5) Forster, K. 1974. The role of semantic hypotheses in sentence processing. In Problèmes actuels en psycholinguistique, No. 206. Paris: Colloques Internationaux du C.N.R.S.

(6) Foss, D.J., and Jenkins, C.M. 1973. Some effects of context on the comprehension of ambiguous sentences. J. Verb. Learn. Verb. Behav. 12: 577-589.

(7) Grosjean, F. 1978. Cross-linguistic research in the perception and production of English and American Sign Language. Paper for National Symposium on Sign Language Research and Teaching, Coronado.

(8) Grosjean, F. 1980. Spoken word recognition and the gating paradigm. Manuscript, Northeastern University.

(9) Grosjean, F.; Teuber, H.; and Lane, H. 1979. When is a sign a sign? The on-line processing of gated signs in American Sign Language. Working paper, Northeastern University.

(10) Hewes, G.W. 1973. Primate communication and the gestural origin of language. Curr. Anthrop. 14: 5-11.

(11) Hoemann, H., and Florian, V. 1976. Order constraints in American Sign Language. Sign Lang. Stud. 11: 121-132.

(12) Karmiloff-Smith, A. 1978. Adult simultaneous translation: a functional analysis of linguistic categories and a comparison with child development. In Proceedings of the Symposium on Language, Interpretation and Communication, ed. D. Gerver. New York: Plenum.

(13) Klima, E.S., and Bellugi, U. 1975. Perception and production in a visually based language. In Developmental Psycholinguistics and Communication Disorders, eds. D. Aaronson and R.W. Reiber. New York: The New York Academy of Sciences.

(14) Klima, E.S., and Bellugi, U. 1979. The Signs of Language. Cambridge, MA: Harvard University Press.

(15) Levelt, W.J.M. 1978. A survey of studies in sentence perception: 1970-1976. In Studies in the Perception of Language, eds. W.J.M. Levelt and G.B. Flores d'Arcais. New York: Wiley

(16) Levelt, W.J.M. 1979. Linearization in describing spatial networks. Proceedings of Symposium on Formal Semantics and Natural Language Processing, Austin, Texas.

(17) Levelt, W.J.M. 1979. Linearization in discourse. Report, Max-Planck-Projektgruppe für Psycholinguistik, Nijmegen.

(18) Levelt, W.J.M. 1980. The origins of language and language awareness. Comments on papers by McNeill and Foppa. In Human Ethology, eds. M. von Cranach, K. Foppa, W. Lepenies and D. Ploog. Cambridge: Cambridge University Press.

(19) Marslen-Wilson, W.D. 1973. Linguistic structure and speech shadowing at very short latencies. Nature 244: 522-533.

(20) Marslen-Wilson, W.D. 1975. Sentence perception as an interactive parallel process. Science 189: 226-228.

(21) Marslen-Wilson, W.D. 1978. Recognizing spoken words. Working paper, Max-Planck-Projektgruppe für Psycholinguistik, Nijmegen.

(22) Marslen-Wilson, W.D., and Tyler, L.K. 1980. The temporal structure of spoken language understanding. Cognition 8: 1-71.

(23) Marslen-Wilson, W.D.; Tyler, K.L.; and Seidenberg, M. 1978. Sentence processing and the clause boundary. In Studies in the Perception of Language, eds. W.J.M. Levelt and G.B. Flores d'Arcais. New York: Wiley.

(24) Marslen-Wilson, W.D., and Welsh, A. 1978. Processing interactions and lexical access during word-recognition in continuous speech. Cognitive Psychol. 10: 29-63.

(25) McIntire, M., and Yamada, J. 1976. Visual Shadowing: An Experiment in American Sign Language. Philadelphia, PA: Society of America.

(26) McNeill, D. 1980. Language origins. In Human Ethology, eds. M. von Cranach, K. Foppa, W. Lepenies and D. Ploog. Cambridge: Cambridge University Press.

(27) Nakatani, L.H., and Dukes, K.D. 1977. Locus of segmental cues for word juncture. J. Acous. Soc. Am. 62: 714-719.

(28) Reef, S.; Lane, H.; and Battison, R. 1978. Visual persistence in handshapes in American Sign Language. Working paper, Northeastern University.

(29) Shaffer, W.O., and La Berge, D. 1979. Automatic semantic processing of unattended words. J. Verb. Learn. Verb. Behav. 18: 413-426.

(30) Studdert-Kennedy, M.G. 1977. Universals in phonetic structure and their role in linguistic communication. In Recognition of Complex Acoustic Signals, ed. T.H. Bullock. Berlin: Dahlem Konferenzen.

(31) Tanenhaus, M.K.; Leiman, J.M.; and Seidenberg, M.S. 1979. Evidence for multiple stages in the processing of ambiguous words in syntactic contexts. J. Verb. Learn. Verb. Behav. 18: 427-440.

(32) Tervoort, B.T.M. 1953. Structurele analyse van visueel taalgebruik binnen een groep dove kinderen I, II. Amsterdam. (Developmental Features of Visual Communication. Amsterdam: North-Holland, 1975.)

(33) Tweney, R., and Heiman, G. 1977. The effect of sign language grammatical structure on recall. Bull. Psychonomic Soc. 10: 331-334.

(34) Tweney, R.D.; Heiman, G.W.; and Hoemann, H.W. 1979. Psychological processing of sign language: The effects of visual disruption on sign intelligibility. Bowling Green State University.

(35) Tweney, R.D.; Liddell, S.K.; and Bellugi, U. 1978. The perception of grammatical boundaries in ASL. Manuscript, The Salk Institute and Bowling Green University.

(36) Tyler, L.K. 1980. Serial and interactive-parallel theories of sentence processing. In Psycholinguistics 3, eds. J. Morton and J. Marshall. London: Elek Science.

(37) Winograd, T. 1974. Five Lectures on Artificial Intelligence. Stanford.

(38) Yates, J., and Thul, N. 1979. Perceiving surprising words in an unattended auditory channel. Q. J. Exp. Psychol. 31: 281-286.