Sign & Word Recognition: A First Comparison

François Grosjean

Sign Language Studies, Volume 32, Fall 1981, pp. 195-220 (Article)

Published by Gallaudet University Press
DOI: https://doi.org/10.1353/sls.1982.0003

SIGN & WORD RECOGNITION: A FIRST COMPARISON

François Grosjean

Abstract. The results of a word recognition study (Grosjean 1980) are compared to those of a sign recognition study (Grosjean, Teuber, & Lane 1978) in order to determine which aspects of lexical access are comparable in speech and sign and which are specific to each of the two language modalities. The recognition of a lexical item appears to involve two distinct stages in both modalities: the isolation of the item (word or sign) from other candidates; and after some further processing, the acceptance or recognition of the lexical item. Factors such as the frequency of occurrence of a lexical item play a role in both word and sign recognition, but others such as word length in speech and frequency of location in sign are specific to the modality of expression. In both modalities, however, the word recognition process is complex and is strewn with "garden paths." The future of sign recognition research is discussed in light of these findings.

A first comparison. Much of current research in psycholinguistics is aimed at understanding the on-line processing of language—how language is perceived and comprehended while it is being heard by the user of a spoken language or seen by a signer. The basic problem is to describe the various stages of processing that occur from the moment the acoustic signal impinges on the ear of the listener (or the visual

 ISSN 0302-1475 32:195 $2.50

signal impinges on the eye of the observer) to the moment the signal has been interpreted linguistically and is sent to long term memory. The approach that researchers have used is to break down this complex operation into various processing levels, which interact in real time, and to study them individually. Some have concentrated their efforts on the acoustic-to-phonetic coding during speech perception; others have studied the role of syntax and semantics during comprehension; and others have examined the impact of prosody, of discourse structure, and of pragmatics during on-line processing. A level that has received much attention in recent years is word recognition or lexical access (Cole & Jakimik 1979, Foss & Blank 1980, Marslen-Wilson & Welsh 1978, and Swinney 1979). Studies in spoken word recognition have shown how this process interacts with other processing levels (e.g. syntactic, semantic, pragmatic), how it is very rapid (about three words are recognized per second during spontaneous discourse), and how it is usually extremely efficient (only serious external interference such as noise will make it break down).

In the following paper we examine the word recognition process in sign as well as in speech. To do this we will compare two recent studies—one in speech and one in sign—that have used the same experimental paradigm, the gating paradigm. This procedure consists of presenting a word (or sign) repeatedly and increasing its presentation time (presentation from onset) at each successive repetition. Thus, the subject in a speech task may be given the first 30 msecs of the word at the first pass; at the second pass given the first 60 msecs; at the third pass given the first 90 msecs; and so on until the nth pass at which the complete word is presented. The subject's task is to guess the word (or sign) after each pass and to rate his or her confidence in the guess.

This gating paradigm yields three kinds of results: first, it indicates the isolation point of a word (i.e. the point in time at which it is guessed correctly); second, it yields the subject's confidence ratings at both the isolation point and at the last presentation (when the full item is presented); and third, the paradigm gathers the erroneous guesses made, up to the point at which the stimulus item is finally isolated. These errors, which can be analyzed with respect to frequency of occurrence, number of times proposed, number of guess types, and phonotactic configuration, prove to be very useful in accounting for the narrowing-in process that takes place during word recognition. The gating paradigm is just one of many techniques used currently in psycholinguistics to study various aspects of on-line processing of language. Other techniques such as phoneme monitoring (Foss & Blank 1980) and word monitoring (Marslen-Wilson & Komisarjevsky Tyler 1980) are probably better know, but to our knowledge they have not yet been adapted to the signing modality.

To set the stage for our comparison of word recognition and sign recognition, we will summarize a recent study of spoken word recognition (Grosjean 1980a). This study showed that the underlying recognition process is highly complex and that it is probably made up of two stages: the isolation of a word; and after additional monitoring of the acoustic-phonetic input, the acceptance or recognition of the word. We will then present some of the findings from a first sign recognition study (Grosjean, Teuber, & Lane 1978). Further analysis of some of its data and re-interpretation of some of its results will allow us to compare word recognition in the spoken modality and in the signed modality. Finally, we will discuss the future of sign recognition research and how forthcoming results will have to be integrated into a model that will account for the similarities and the differences of sign and word recognition. Although some of the existing models of spoken word recognition may have to be transformed somewhat to account for sign recognition, they will acquire thereby greater predictive power, in that they will then account for two language modalities. Psycholinguistics will then be a step closer to a complete understanding of on-line processing of language—be it spoken or signed.

Spoken word recognition. In the study by Grosjean (1980a) forty-eight nouns varying in length (one, two, and three syllables) and in frequency of occurrence (high and low) were presented orally to subjects in three context conditions: (1) a no context condition (e.g. the word camel by itself); (2) a short context condition (the word camel preceded by "the kids rode on the"); and (3) a long context condition (the word camel preceded by "At the zoo, the kids rode on the..."). To prepare the stimuli, the complete sentences (long contexts) were recorded and entered into a computer. For the no-context condition, only the stimulus word was read out and recorded in presentations of increasing duration (+30 milliseconds for each presentation). Thus, for the word camel the first presentation contained the first 30 msecs, part of the burst of the consonant /k/; the second presentation included the first 60 msecs of the word; and so on until the whole word had been presented. For the short context condition, exactly the same procedure was used, except that each presentation of the stimulus word was preceded by a clause (e.g."the kids rode on the..."); and for the long context condition each presentation was preceded by the full sentence. Different groups of subjects were run on each of the three conditions. Each subject was asked to guess the word being gated after each presentation and also to indicate how confident he or she felt about the guess. Response sheets were examined to determine the point at which each subject correctly guessed the stimulus word and did not subsequently change the guess. The duration of the gate at that point was adopted as the isolation point. The response sheets also yielded the subjects' confidence ratings at both the isolation point and at the last presentation of the word, as well as the erroneous guesses made before the isolation point.

Three main effects were found in the study. First, as the context became more constraining, the mean isolation times of words decreased substantially: in the no context condition, as much as 83% of a word was needed in order to isolate it; in the short context condition, the percentage was reduced to 60%; and in the long context condition only 37% of the word needed to be heard before it was identified. This effect was explained by the fact that context reduces the number of lexical possibilities and hence less "bottom-up" or acoustic-phonetic analysis (processing) is required in order to isolate the word. In the case of the word camel for instance, no "top-down" or context information was given in the no context condition, and so subjects needed to receive almost all of the word in order to isolate it. However, when the word was preceded by "the kids rode on a..." the subjects knew that the next word would probably be a noun and that it would refer to a vehicle or an animal that can be ridden on by children. This information (in addition to prosodic information carried by the short context) greatly enhanced the narrowing-in process. And in the long context condition where the preceding sentence was "At the zoo, the children rode on a ..." the list of possible word candidates was shortened even more so that even less bottom-up information was needed for word isolation.

A main effect was also obtained for word frequency: as the frequency of occurrence of a word increased, isolation times decreased. High frequency words were isolated 61 msecs sooner than low frequency words across all three context conditions. This effect is explained by the listener's bias toward high frequency words: he or she expects the speaker to use words that oftenest occur, and it is those words that will be primary candidates during word recognition. It is only when a search among high frequency words leads nowhere that the listener will examine the lower frequency words. Finally, a word length effect was also found: the longer the word, the more time it took to isolate it (one syllable words were isolated 80 msecs sooner than three syllable words, for instance). This effect is linked to the phonetic configuration of words. Even though two and three syllable words are probably more redundant than one syllable words, they are also longer and hence more time is required to isolate them. It is interesting to note that an interaction was found between word length and context: as context became more constraining, the effect of word length diminished. This is because top-down information gave the listeners some indication about possible word candidates and this made it possible not to wait so long to isolate the longer words. Thus, depending on the frequency and length of a word and the context that precedes it, its isolation from among a list of candidates will take more or less time. Subjects will need about one third of a word when top-down information is rich and the word itself is quite frequent, they will need to hear most of the word when it is without context and is an infrequently occurring word.

Three main results also emerged from the study of subjects' confidence ratings. First, the isolation of a word took place without the listener feeling very confident about the candidate —no subject ever showed better than medium confidence at the isolation point. Second, the subject's confidence in the proposed condidate increased between the point of isolation and the end of the word. And third, even after the whole word had been heard, the listener's confidence in the candidate proposed was rarely perfect and depended on such factors as the frequency of the word, its length, and the amount of information brought to the isolation and recognition processes by the provided context. These results led us to propose the notion of the isolation point in word recognition; i.e. the isolation point is that point at which the listener has isolated a candidate but may still feel quite unsure about it. He or she will continue to monitor the acoustic-phonetic information until some criterion level of confidence is reached and the word is accepted or reçognized.

An analysis of the erroneous guesses made prior to the isolation point allowed us to study more closely the narrowing-in process employed by listeners when isolating words. In the no context condition the isolation process relies only on the acoustic-phonetic information, and the candidates proposed clearly reflected this. However, in the short and long context conditions, we observed the interaction between top-down and bottom-up information, leading to a smaller and often quite different set of candidates. What is especially interesting is that certain erroneous candidates had a rather long life span because the listener had gone down a garden path. Three such garden paths were isolated: the "word from a word", the frequency, and the semantic garden paths. The "word from a word" path is due to a narrowing-in strategy based on the following premise: if one or any number of syllables from the onset of the word can stand as a word, and either the word is presented in isolation or the syntactic and semantic rules are not violated when the word is in context, then that word will be a candidate in the isolation process. Thus strain was proposed by subjects for the stimulus word stranger, pick for pickle, cult for culture, and so on. It is only with further acoustic-phonetic information that subjects realize that they are dealing with a two or a three syllable word and come out of this garden path. It should be noted that preceding context will usually repress the tendency to go down this "word from a word" garden path. A second garden path results from listener bias in favor of high frequency words. Basing themselves on top-down and bottom-up information, they propose high frequency candidates. Thus, for the word trawler, in the context of "Stephen worked on a . . . ", subjects proposed train, then truck, then trolley—all words of higher frequency than trawler in the Boston area—before being forced by the acoustic-phonetic information to the actual, low frequency, stimulus word, trawler.

Finally, the semantic garden path is due to a conflict between top-down and bottom-up information, with top-down information winning over for a time at least. Thus, for the word picture preceded by the context "Before climbing the North Face, Carl examined the . . .", subjects proposed the word pick. This shows that they were attempting to choose a candidate that would fit the context (what it is you examine before climbing a mountain), and that would also be in accord with the acoustic-phonetic information. As pick fits both constraints (and is helped along by the "word from a word" strategy), they proposed it as a candidate. Once again, only additional bottom-up information forced the garden path subjects to backtrack and change their candidate. (It is interesting to note that for mountain climbers, picture would almost certainly have been an early candidate; one never undertakes a long and dangerous north face climb without having carefully studied the route on a diagram or a picture of the face.)

The study of spoken word recognition confirmed, therefore, that the recognition system makes active use of both bottom-up and top-down information (Marslen-Wilson & Welsh 1978). It also showed that word recognition may involve two distinct stages: isolating an appropriate candidate from among a pool of possible candidates; and after monitoring more acoustic-phonetic information, finally accepting or recognizing that candidate. Finally, the study showed that the narrowing-in process in word recognition is highly complex and that, like syntactic processing, word isolation and recognition is strewn with garden paths. The question that can now be asked is whether any of the aspects of spoken word recognition revealed by the gating paradigm can also be found in sign recognition. How much of the sign is needed to allow the observer to isolate it? Which factors account for isolation times? Do confidence ratings show evidence for a two-stage process in sign recognition as well? How do signers narrow in on the sign being presented? Is there any evidence for garden paths in sign recognition? In the following section we will attempt to answer some of these questions.

Sign recognition. In a first exploratory study of sign recognition, Grosjean, Teuber, & Lane (1978) chose to examine thirty-seven signs (see Appendix 1). Of these, twenty-seven were one-handed (e.g. SUMMER, COW, LOUSY, EAT, STUCK), and ten were two-handed signs. The latter either had both hands active (e.g. MAKE, RUN) or one hand active the other passive (but both with the same handshape, as in TRAIN and SHORT). The signs were chosen so that all handshapes (20 as listed in Klima & Bellugi 1979: 166f) were used at least once (with the exception of /bO/ "baby O" and /W/; ibid.). Also the following locations occurred at least twice: full face, forehead, nose, cheek, mouth-chin, neck, trunk, and neutral space. All signs were executed in isolation and in their canonical form: between each two signs, the signer's hands were at waist level resting horizontally on a table top. The signs were performed at normal rate by an adult fluent signer of deaf parents. Each sign recorded was copied 24 times in succession to make up a stimulus set. Each set was then gated by a timing circuit and sent to a monitor placed in front of the subject. (We should note here that gating took place during the experimental sessions, but in the spoken word experiment the tapes were prepared before running the subjects.) The first gate for a sign lasted 28 msecs. Each successive gate, up to the 20th gate, was incremented by an average of 26 msecs; the duration of the last four gates was incremented by 55 msecs. Thus on the 24th and last presentation of a particular sign set, each subject saw 744 msecs of the sign.

Five adult fluent signers of American Sign Language, who had deaf parents and were all congenitally deaf, took part in the experiment. Each subject was instructed to accomplish three tasks after each gate presentation: copy exactly what was presented, guess the sign, and give a confidence rating on the guess. (In what follows we will report only on the guessing and confidence data.) A prelingually deaf signer timed the stimulus signs in their full version with a chronoscope. Onset was defined as the moment the hand(s) appeared on the screen, and offset as the moment the sign was no longer recognizable; i.e. when the hand(s) began to move back down to resting position. The same timer then determined the points at which each of the four formational parameters (handshape, orientation, location, and movement; Klima & Bellugi 1979: 43-66) were used in a guess correctly and without subsequent change by each subject for each sign.

In the following discussion, we will compare the results obtained in the sign recognition task with those obtained in the no context condition in the word recognition task. The two experiments differ in a number of aspects and it is important to point these out before beginning our comparison. First, the speech study used a parametric design to show main effects for word frequency, word length, and context. The sign study, however, was exploratory and variables were left to interact naturally. Thus only post hoc analysis of the data using partial correlation and multiple regression methods allows us to obtain some indication of the variables that play a role in sign recognition. Second, the signs were presented in their canonical forms, but the words in the speech experiment were extracted from the speech stream. Third, the final gate in each word set corresponded to the total duration of the word (in the speech signal), but in the sign experiment the final gate duration was that allowed as maximum by the gating device. This corresponded to 81% of the sign on the average. And finally, subjects in the sign experiment were not forced to guess the sign at the earlier gates; in the speech experiment care was taken to obtain a guess and a confidence rating at each gate. Despite these differences, we feel that a comparison of the two studies is permissible, especially as many of the findings in the sign study have been replicated by Lorene Clark (1981), who used a procedure almost indentical to that employed in our speech study.

The isolation point. In Figure 1 we present data pertaining to the isolation point in sign and in speech recognition. We note that the average duration of the 37 signs is slightly more than twice that of the 48 words. Signs range in duration from 488 msecs (EAST) to 1232 msecs (COFFEE), with an average duration of 817 msecs; words in the study range in duration from 231 msecs (gull) to 669 msecs (fuselage), with an average of 403 msecs.

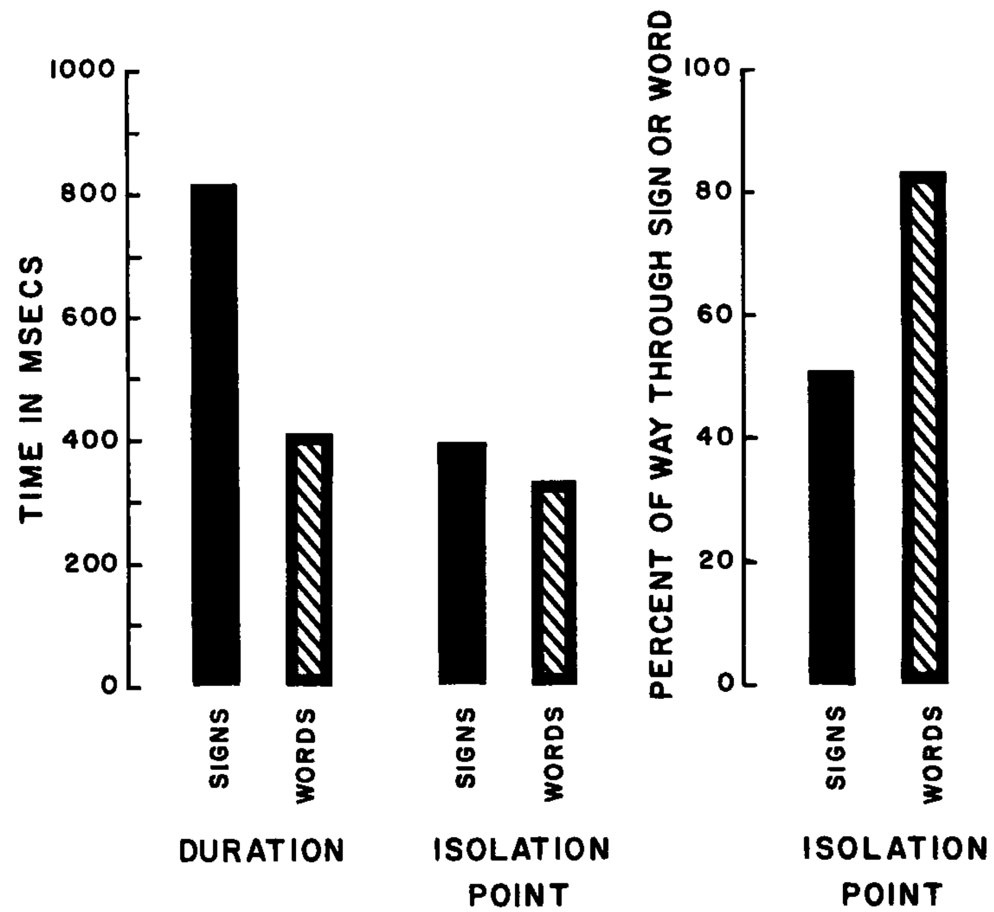

Figure 1. Mean duration and mean isolation times of 37 signs presented in their canonical forms and of 48 words extracted from the speech stream.

These differences in duration confirm other measures of sign and word duration reported by Bellugi and Fischer (1972) and Grosjean (1979). Second, we note that the average isolation time for a sign is slightly longer than that for a word (395 msecs as compared to 333 msecs, a significant difference). The isolation times in sign range from 198 msecs (LIKE) to 585 msecs (SUMMER); in speech they range from 199 msecs (picture) to 563 msecs (fuselage). The longer identification times for signs is not unexpected in itself, as signs are longer than words and therefore will need more time to be isolated. What is very surprising, however, is the amount of a sign needed for isolation compared to the amount of a word (see right hand side of Figure 1). Only 51 % of a sign is needed on the average, but 83 % of a word is needed, for isolation. The range in sign extends from 27 % (COFFEE) to 94 % (SUMMER), and in speech from 55 % (pickle) to 100 % (pharmacist). This finding has been replicated by Clark (1981) with signs taken out of the signing stream, as were the words in the speech study. This difference can probably be explained by the more parallel nature of the "phonetic" structure of the sign [or aspectual structure; see Stokoe 1960]: information is presented in a more simultaneous manner in sign than are the sequential phonemes and syllables of speech. Although there is much debate on the extent of the simultaneity of information in sign and speech (see for instance Studdert-Kennedy & Lane 1980), it is no doubt true that more lexical information is presented at any one time in the sign than in the word. This in turn will help the observer identify the sign when less of it has been presented. We should note, however, that with the exception of one stimulus (LEND), no sign was inflected in our study. It could be that in natural discourse, where nouns and verbs are often inflected by means of movement, the amount of the sign needed for isolation will increase. Only further experimentation will tell us if this will bring the amount of a sign needed for isolation up to 83 %, as it is in speech. We believe that it probably will not.

The question we can now ask is the following: How can one account for a 387 msecs (3:1) range in the isolation time for signs? Although the study was not designed specifically to account for the variables controlling isolation time, a partial correlation and multiple regression analysis allows us to propose a number of possible factors. In our review of these factors, we will report correlation coefficients whenever appropriate. These are usually quite low (on the order of 0.20), but it should be remembered that several factors interacted naturally in the study, thus spreading the variance across several variables. Although a parametric study involving one or two factors and controlling for all others would have produced higher coefficients of correlation, it would not have given us some indication of the importance of the factors that is present here.

A first variable that appears to be important in accounting for isolation times is the frequency of occurrence of signs. As no sign frequency list is currently available, we used as a rough estimate the frequency of occurrence of the most commonly used English gloss of each sign (Kučera & Francis 1967) and obtained a -0.27 Pearson product moment correlation between isolation time and log gloss frequency. Even when the effect of other variables is removed, the partial correlations never fall below -0.20. Thus, as in speech, frequency of occurrence of a sign will most likely prove to be an important variable in accounting for sign recognition, and future studies will be able to show whether the observer's bias for high frequency will lead him or her down the frequency garden path, as in speech. It is interesting to note that we found frequency correlated with sign type ($r = -0.32$); one handed and two handed signs with one hand only moving are often low frequency signs, whereas two handed symmetrical signs with both hands moving are usually high frequency signs. A 0.29 correlation was also obtained between sign frequency and location frequency, indicating that more frequent signs have a tendency to be articulated at a frequent location such as neutral space and trunk.

A second factor that appears to be important in the isolation of signs is the frequency of occurrence of the sign's location. We used a computerized count of the location of one handed and two handed signs as presented in the <u>Dictionary of ASL</u> (Stokoe et al. 1976; see Teuber et al. 1980), and at first found no correlation ($r = -0.06$) between isolation times and frequency of location. But a partial correlation analysis that controlled for distance of the location from rest position showed a 0.20 correlation between the two variables. The less frequent the location, the shorter the isolation time. It would seem from this that observers make active use of location information when setting up an initial set of candidates: the less frequent the location (neck, cheek, e.g.), the smaller the group of candidates and hence the more rapid the final isolation of the stimulus sign. It is interesting to note that a -0.74 correlation was found between sign type and frequency of location, indicating that one handed signs are found more frequently in rare locations, whereas two handed symmetrical signs are found in more frequent location, such as neutral space and trunk (see Frishberg 1975, Siple 1978, among others).

A third factor accounting for isolation times is the number of repetitions involved in the articulation of the sign. A -0.20 correlation (-0.27 after controlling for other factors) was found between the number of times a movement was repeated in the articulation of the sign and the isolation time of the sign: the greater the number of repetitions, the shorter the isolation time. This can be explained by the fact that in uninflected signs a repetition of the movement is usually not necessary for the identification of the sign; part of the first cycle is sufficient for identification. However, such a generalization may not apply to inflected signs, where the nature of the movement and the number of cycles completed may be of utmost importance in correct sign identification. We should also note a 0.48 correlation between number of repetitions of a movement and sign duration; one cycle signs are usually shorter than two or three cycle signs (see Grosjean 1979).

A fourth factor that accounts for sign identification is distance from the starting position of the hands to the sign location. A 0.27 correlation was obtained between these two variables, but interestingly this coefficient was reduced to 0.11 when such factors as frequency, frequency of location, and sign type were controlled for. The reason is that distance and frequency of location were highly correlated ($r = -0.87$) in the study: the further the distance from the rest position (hands on table), the less frequent the location (forehead, nose, cheek, etc.). Despite this artifact, we do not believe that in natural signing the greater the distance from one sign to the next, the longer will be the isolation time for the second. Thus a sign articulated on the forehead preceded by a sign articulated on the trunk will probably be identified more slowly than a sign articulated in neutral space, when it too is preceded by a sign on the trunk. Again we will have to wait for a parametric study that controls for other variables to determine whether distance is truly an important variable in sign recognition.

A fifth factor that appears to play a role in the isolation time of signs is the sign type. We used Battison's classification of signs (1978—as adapted by Clark 1981) to give complexity values to our 37 signs. Two handed signs with similar handshapes and with both hands active were given a value of 1; two handed signs with similar handshapes and with only one hand active were given a value of 2; and one handed signs were given a value of 3. A correlation of 0.26 was obtained between sign type and isolation time (partial correlations were on the order of 0.20). The implication is that the more redundant the information carried by the hands, as in a two handed sign with similar handshapes both moving, the sooner the sign will be identified.

A sixth factor that appears important is the frequency of movement of a sign. We obtained a computerized count of the different movements in one handed and two handed signs (Teuber et al. 1980) and then computed a mean frequency of movement value for each sign by averaging across the movements indicated in the DASL. Much to our surprise, these values correlated -0.12 (about -0.20 after controlling for other factors) with isolation times, indicating thereby that the more frequent the movement, the shorter the identification time. (Note that for frequency of location we found a positive correlation: the more frequent the location, the longer the isolation time.) The negative correlation obtained for movement frequency can be explained by the fact that subjects may have a bias for simple, and therefore frequent, movements (J. Shepard-Kegl, personal communication). As we will see later, certain movement garden paths indicate such a bias. Subjects seem to believe that the sign they are seeing will have a simple vertical, sideways, or horizontal movement and not a bending, opening, closing, or wiggling movement. When one of the latter occurs, subjects have to change their guess, and this affects the isolation times.

We should note that the frequency of occurrence of handshape was not correlated with isolation time ($r = 0.01$), even when other factors were controlled for. In addition, the duration of a sign (in milliseconds) did not seem to be an important factor in explaining sign isolation times ($r = 0.12$). This is in marked contrast with speech, where a correlation of 0.88 was found between word length (also in milliseconds) and identification time. The low correlation in sign may be due to the repetition of the movement in some signs (as in MAKE, TRAIN, EAT) and to the more parallel or simultaneous nature of signs as compared to words.

When the six factors we have discussed in addition to sign duration were inserted into a multiple regression analysis, the multiple R reached 0.60, thus indicating a fairly good prediction of the isolation times. However, much of the variance of isolation times still remains to be accounted for, and future studies may wish to revise the metric we used for each of the factors studied as well as to investigate other factors that may be involved in predicting isolation times. Among these orientation may play a role: the rather long delay in identifying WHO (490 msecs) as compared with a mean of 361 msecs for the other four signs at that location may have been caused by the index finger obscuring the extended thumb touching the chin. Facial expression, head tilt, and body movement that accompany many signs may also influence the isolation time of the sign. For instance, RECENTLY was identified rather quickly for a sign made on the cheek (296 msecs as compared to a mean of 408 msecs for the other three signs at the same location), but it was also the only sign made with a characteristic facial expression and head tilt. Finally, signer idiosyncrasies may always play a role: native informants have told us that the way COW was produced by our signer was characteristic of her idiolect and in our study it was the sign that took the longest time to identify (600 msecs as compared to an average of 461 msecs for the other three signs at the same location). In the end, the number of factors influencing the identification of signs will probably be quite large, although some factors will be more important than others. Only systematic research in this domain will enable us to isolate these factors and to explain how they interact with one another.

The confidence ratings. In Figure 2 we present the mean confidence ratings at the isolation point and at the last presentation of sign or word. We should note that the speech ratings were converted from a 0-100 scale to a 1-5 scale to allow for the comparison between the two modalities, and that the last presentation corresponded to 100% of the words but to only an average of 81% of the signs. Despite these differences, one is struck by the similarity of the data in the two modalities. The mean confidence rating at the isolation point is 3.01 for a sign and 3.25 for a word—a difference that is not significant; and the mean rating for the last presentation is 4.66 for signs and 4.46 for words. Although this last difference is not significant, it probably would have been so had the total duration of the sign been presented. Confidence ratings for signs would probably have reached a perfect 5.0 before the 100%

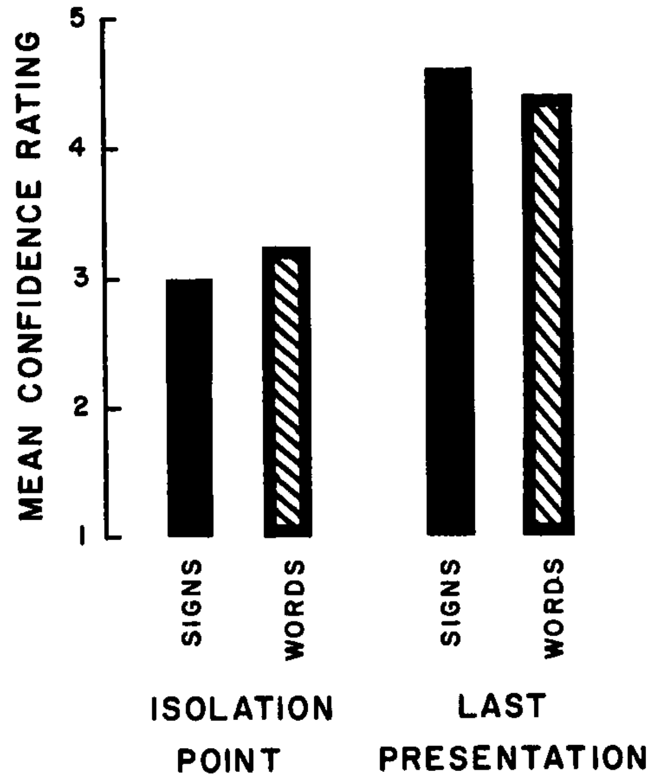

Figure 2. Mean confidence ratings for 37 signs and 48 words at the isolation point and at last presentation (100% of the words and 81% of the signs completed)

mark, and this would have been yet another indication of the more simultaneous nature (parallel encodedness) of signs, at least in their canonical form. Whatever the final value for signs may be, we would like to highlight the difference in confidence ratings that exists between the isolation point of a sign and the last presentation of that sign. First, and as in speech, the isolation of a sign can take place without the observer feeling very confident about the candidate. Second, the subject's confidence in the proposed candidate increases between the point of isolation and the end of a sign. Thus, in sign as in speech, there may well be two stages in the recognition of a lexical item: an isolation stage, at which the observer isolates a candidate but remains quite unsure about it; and an acceptance or recognition stage, which is reached after further monitoring of the bottom-up (visual) information. The one difference between speech and sign may be the point at which acceptance or recognition takes place: in speech, it can vary from being located within the word (for two or three syllable, high frequency words) to being located after the word (for one syllable, low frequency words); but in sign it would seem that the recognition point may always be located within the sign itself. The only exception would be with highly inflected signs in which lexical information is spread throughout the sign.

The sign isolation process. The gating paradigm not only allows us to determine how much of the stimulus sign is needed to be distinguished from other candidates and how confident subjects are about their answers, but it also enables us to have an insight into the sign isolation process. This is done by analyzing the erroneous guesses subjects make before the isolation point of the stimulus sign. In Figure 3 we present the mean isolation time for each of the four formational parameters varied in this experiment; that is, the point in time at which each parameter of the stimulus sign is used correctly in a guess (although of course the guessed sign may not be the stimulus sign). We note first that the location, orientation, and handshape of a sign are isolated at about the same time (307, 309, and 322 msecs respectively); and then some 70 msecs later the movement is isolated; at this point, and ipso facto, the sign is isolated. Movement is therefore the "clincher" parameter that enables the subjects to isolate the sign. An analysis of variance shows a main effect for parameters ($F'(3,57) = 21.79$, $p < 0.01$); but an a posteriori test (Tukey HSD; Kirk 1967) shows that the only significant difference is between movement and each of the three other parameters (these latter not differing significantly from each other). The reason for the 10 msecs difference between the isolation time of movement and the isolation time of the sign is that in some very rare cases another parameter (usually handshape) is the last parameter to be isolated. Overall however, the movement isolation time is an almost perfect predictor of the sign isolation time ($r = 0.98$), but the mean isolation time for location, orientation, and handshape only predicts isolation times with a correlation of 0.62.

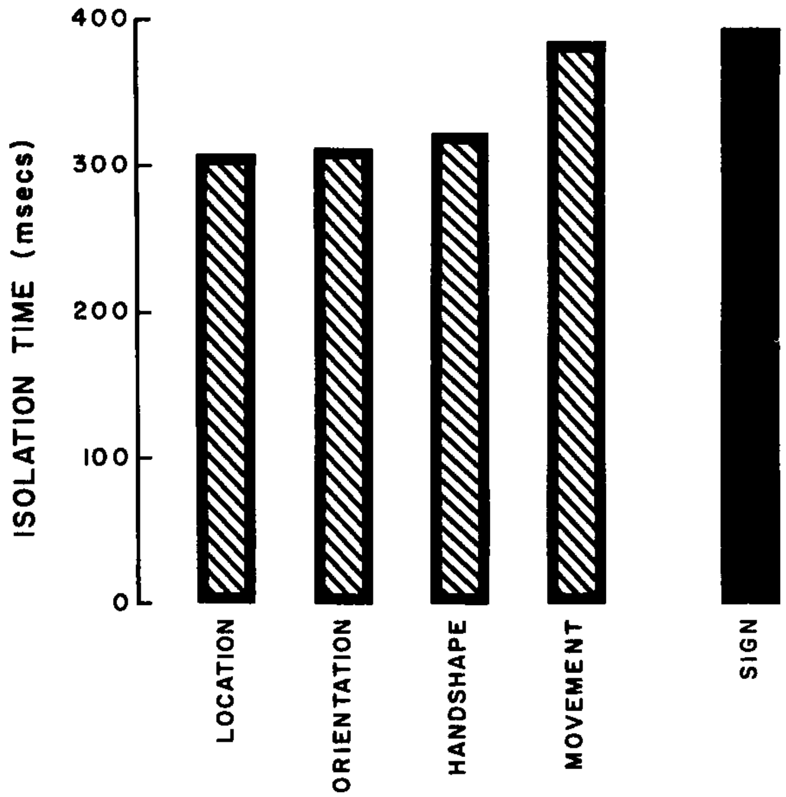

Figure 3. Mean isolation times of each of the four formational parameters of the 37 signs.

We should note that when one compares the mean isolation times of location, orientation, and handshape taken as a group to those of the corresponding movements, the difference between the two is not a stable 70 msecs. In fact it ranges from -37 msecs for FEEL (for which the movement is isolated before the other three parameters) to 278 msecs for LEND (in which movement comes in much later than the other three parameters). This pattern produces a significant interaction between signs and parameters: $F'(108,432) = 3.49$, $p < 0.01$. In the production of some of our signs movement co-occurs with the other parameters (PERFECT, NOW, STUCK, for instance), and hence for each of these the movement is isolated at about the same time as are the other parameters. But for most other signs, the location has to be reached before the movement can be fully realized (e.g. SUMMER, WHO, BUG). For this last category of signs, there appears to be a very strong relationship between the distance from the rest position to the location of the sign and the delay between identifying handshape, location, and orientation and the identification of movement. Thus signs articulated on the forehead, for example, have a mean difference of 84 msecs, but those articulated on the neck have a mean difference of only 15 msecs. In the former case, handshape, orientation, and even location information is available as the hand is moving up to the forehead, but the movement information is not yet available (e.g SUMMER). In the latter case, the four parameters are produced (and hence perceived) at about the same time.

We can conclude from this analysis of the isolation time of the four formational parameters that the narrowing-in on the sign is not an "all or none" operation, in the sense that observers wait to have information about all four parameters before isolating the item with all four together. On the contrary, a two-stage process appears to take place in the isolation of a sign: first, the location, orientation, and handshape of a sign are isolated at about the same time, and then some time later, the movement parameter is isolated; it is this last parameter that "triggers" the isolation of the sign.

An analysis of the erroneous guesses made by subjects during the isolation of each parameter gives some indication of the narrowing-in process that takes place during sign isolation and recognition. However, the picture is not complete, as we did not force subjects to give a guess after each gated presentation. Many opted therefore for a no-guess approach until they felt the information they had was sufficient for a "good" guess. Hence, we only have a few error patterns for handshape, orientation, and location. As concerns orientation, we do note some early confusions of palm orientation (e.g. NINE-O'CLOCK, FRENCH, and PREACH as candidates proposed for PEPPER) and a marked preference for the palm side (or neutral) orientation: this leads a number of subjects to propose SHORT for the sign TRAIN—the reverse is never true. In the handshape data, we note some early misperceptions of /A/ for /O/ (GIRL proposed instead of HOME), /B/ for /Ȣ/ (HAPPY for FEEL), and /A/ for /E/ (GAMBLE for EUROPE). It would be interesting for a future study to examine some of the handshape confusions and to determine how well they are predicted by distinctive feature models, such as the one proposed by Lane, Boyes-Braem, and Bellugi (1976). As for location, we note few early misperceptions; the one exception is the forehead for the cheek in the case of the confusion of FATHER for MOTHER (the fact that all other parameters are identical probably enhanced such a confusion).

As may be expected from the delayed isolation of the movement parameter, most of the errors that we obtained involved movement. An analysis of these shows an interesting garden path. According to J. Shepard-Kegl (personal communication), subjects appear to have a bias for simple movements (vertical, sideways, and horizontal movements, e.g.) and expect signs to contain these movements in preference to twisting, nodding, bending, opening, closing, wiggling, and entering movements. In Figure 4 we present some of these garden paths. For each stimulus sign we give the sign candidates proposed. When candidates are proposed at only one presentation (gate), they are depicted with a dot; when they are proposed over two or more presentations, they are depicted by a continuous line. The number of subjects proposing a particular candidate is represented by the thickness of the line; the more subjects choosing it, the thicker the line.

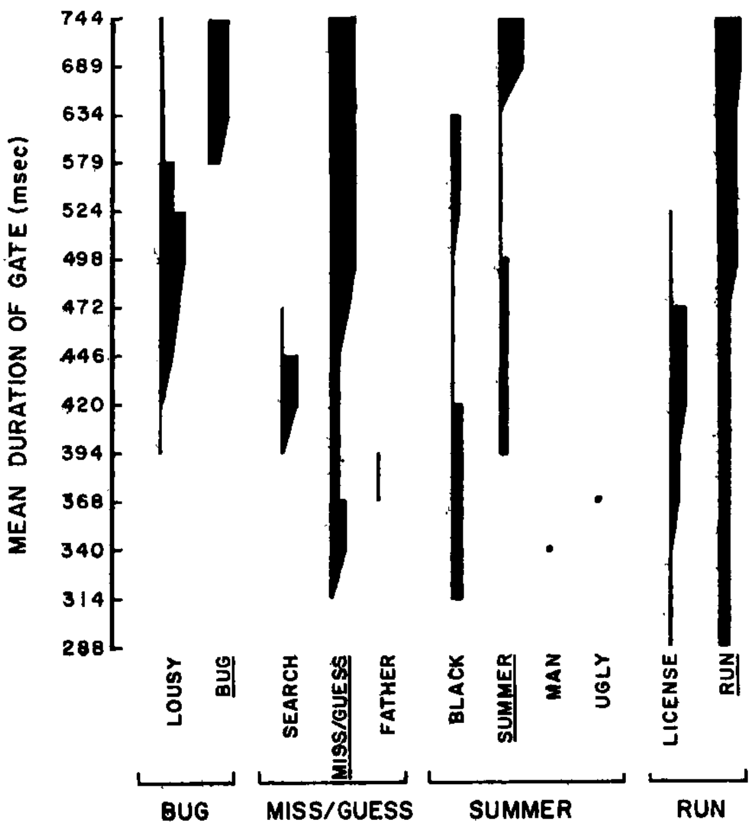

Figure 4. The movement garden path: candidates proposed for 4 stimulus signs. A dot shows those proposed at only one presentation (or gate); continuous lines, proposed twice or more; the thickness of the line varies directly with the number of subjects proposing the sign as candidate.

(Note that we did not include the "no guess" category for any of the stimulus signs). The movement garden path is clearly apparent in each example. Before isolating BUG, for instance, subjects narrow in on LOUSY, a sign that shares the handshape, location, and orientation of BUG but has a simple movement, away and down, instead of the bending action of the index and middle finger for BUG. When LOUSY was presented not one subject proposed BUG as a candidate. (It may be that the preference for LOUSY was enhanced by the higher frequency of the sign; in this case, therefore, two garden paths were combined in one.) In the next example, SEARCH is mistaken for MISS/GUESS, not because of a frequency effect (they have very similar frequencies of occurrence) but because subjects feel that the C-hand will circle instead of close. It is only when they realize that the hand is closing that they all switch to MISS/GUESS. SUMMER is another interesting case. When BLACK is presented, it is perceived as SUMMER by only one subject (and only for two gates), but when SUMMER is presented, many subjects propose BLACK, the sign with a "simpler" movement. It is only when the index starts bending that they switch over to SUMMER. A similar phenomenon takes place when RUN is the stimulus: until the index fingers start bending, subjects propose the sign with the simpler movement, LICENSE (note here that the infrequent occurrence of LICENSE works against the movement garden path). Finally, we should note that when TOMORROW and EVERYDAY were stimuli, the erroneous guesses went both ways (EVERYDAY for TOMORROW and vice versa). This is probably due to the fact that both signs have simple movements, and hence each is open to the movement garden path.

An analysis of a complete set of error data, covering all gate presentations, should be able to confirm this type of garden path, as well as show evidence for other types of biasses linked to the location, orientation, and handshape of the sign as well as to its frequency of occurrence. One may even find the equivalent of the "sign from a sign" garden path. We noted for instance that many subjects proposed NOSE during the presentation of the sign DON'T-MIND; this may qualify as an instance of such a garden path. [This index-hand sign at nose location is glossed 'don't care' as verb and 'indifferent' as adjective in DASL, p. 137.]

The future of sign recognition research. Given the results of the exploratory study we have just presented, future research promises to be as rich and exciting as current research on word recognition. Sign researchers may wish to follow a certain number of directions. First, parametric studies testing one or two variables and controlling for other variables should be undertaken in order to isolate those factors that really do play an important role in sign recognition. Such aspects as the frequency of occurrence of a sign, its configuration in terms of location, handshape, orientation, and movement, as well as in terms of the number of hands involved, the inflectional information attached to the sign, and the non-manual information that accompanies it (facial expression, body shift, eye gaze, and so on) will most probably all play a role in the time it takes to isolate a sign as well as in the narrowing-in process that leads to isolation. Second, sign recognition should be studied with respect to the semantic, syntactic, and prosodic contexts that precede the sign (and follow it). Clark (1981) has undertaken a study in this direction and has found that, as in speech, sign recognition in context is a result of the interaction of bottom-up and top-down information. Not only should context effects be studied and explained, but research should determine whether context interacts with the properties of a sign (frequency of occurrence, sign type, parameter configuration), as it does with the properties of a word (frequency, length, frequency of the first syllable). Third, sign recognition research should develop research paradigms of its own or adapt paradigms from speech research, such as phoneme monitoring (this could become handshape or location monitoring) and word monitoring (sign monitoring). The reason is that it is important to make sure that the results obtained with one paradigm can also be found with other paradigms; in a word, that the results are not paradigm specific.

Finally, sign recognition research will have to work within the framework of a model of lexical access. Among the many models proposed for word recognition, three stand out: Marslen-Wilson and Welsh's (1978) Active Direct Access Model, Forster's (1976) Autonomous Search Model, and Morton's (1969) Logogen Model. Each of these models could be modified and extended to account for results obtained in sign recognition studies. The one that appears the most adaptable to the signing modality is the Active Direct Access Model. The central concept of this model is that a word initial cohort, which is a directly accessed set of potential word candidates, is activated during the earliest phases of the word recognition process. This set is accessed solely on the basis of bottom-up information and consists of the entire set of words in the individual's lexicon that begin with a particular initial sequence. It is this word initial cohort that is assumed to be the basis of word recognition in all contexts. A word is recognized at the point that —going from left to right (or bottom up) through the word—the word in question becomes uniquely distinguishable from any other words in the initial cohort. The role of top-down information is to speed up the process of partitioning the initial cohort of word candidates to isolate a single choice. Just as memory elements will withdraw from the pool of word candidates when they no longer fit the acoustic-phonetic input, so they will also withdraw when they no longer fit specifications of context. This allows for a highly flexible balance between top-down and bottom-up information sources. In fact, the model can make precise predictions about the time course of recognition for individual words when they occur in isolation or in a specified context. Once a single word choice has emerged, a less detailed assessment of the remaining input for that word will be required.

If certain aspects of this model are modified to take into account the word frequency effect and the use of both top-down and bottom-up information in the activation of the initial cohort, and if the monitoring component of the model is strengthened in order to take into account the difference between word isolation and word recognition (see Grosjean 1980 for a complete discussion of these aspects), then the Active Direct Access Model would be a good candidate to account for sign recognition, once it has been adapted to the signing modality. In the long run, we hope that researchers in sign recognition will work in collaboration with researchers in word recognition in elaborating a model that can account for the similarities and the differences between sign and word recognition. As we wrote in a review of the psycholinguistics of sign language:

> Psycholinguistics is addressed, after all, to the perception, understanding, and production of all languages, whether spoken or signed. No model of linguistic performance can be complete unless it describes those aspects of encoding and decoding that are specific to the modality of communication,

oral or visual, and those that are common to all languages, whatever their modality of perception and production (Grosjean 1980b: 34).

. o O o .

Acknowledgment. The preparation of this paper was supported in part by grants from the Department of Health, Education, & Welfare (RR 07143 and NS 14923). The author would especially like to thank Carlos Soares for his help in the statistical analysis, and Lorene Clark, Kerry Green, Mark Mandel, and Judy Shepard-Kegl for their insightful comments on a first draft of this study. Requests for reprints should be sent to François Grosjean, Department of Psychology, Northeastern University, Boston, Massachusetts 02115.

APPENDIX

English glosses of the signs presented in the sign recognition study, duration times, (and isolation times) in milliseconds.

SEARCH 1093 (556)	CURIOUS 987 (407)	COFFEE 1232 (335)
RECENTLY 973 (296)	EAT 1117 (327)	WHO 841 (490)
MAKE 998 (308)	VERY 720 (428)	THINGS 715 (470)
TRAIN 991 (271)	DELICIOUS 918 (367)	LIKE 594 (198)
BORING 676 (265)	MISS/GUESS 559 (377)	EVERYDAY 829 (383)
IDEA 1000 (339)	HOME 735 (385)	DON'T-MIND 600 (363)
PEPPER 905 (564)	BUG 690 (512)	STUCK 553 (320)
SUMMER 623 (585)	MOTHER 750 (287)	EAST 488 (401)
EUROPE 1050 (584)	NOW 534 (228)	RUN 768 (380)
LOUSY 717 (439)	HOUSE 911 (363)	PERFECT 732 (301)
COW 923 (556)	LEND 798 (503)	RESTAURANT 766 (464)
TOMORROW 823 (454)	SHORT 997 (316)	FEEL 839 (342)
BLACK 782 (457)		

REFERENCES

Battison, Robin
1978 Lexical Borrowing in American Sign Language (Silver Spring, MD, Linstok Press).
Bellugi, Ursula, & Susan Fischer
1972 A Comparison of Sign Language and Spoken Language, Cognition 1, 173-200.
Clark, Lorene
1981 Context effects in sign recognition. Working Paper, Department of Psychology, Northeastern University.
Cole, Ron, & Jola Jakimik
1979 A Model of Speech Perception, in Perception & Production of Fluent Speech, Cole ed. (Hillsdale, NJ, Erlbaum Associates).
Forster, Kenneth
1976 Accessing the Mental Lexicon, in New Approaches to Language Mechanisms, Wales & Walker eds. (Amsterdam, North Holland).
Foss, Donald, & Michelle Blank
1980 Identifying the Speech Codes, Cognitive Psychology 12, 1-31.
Frishberg, Nancy
1975 Arbitrariness and Iconicity: Historical Changes in American Sign Language, Language 51, 696-719.
Grosjean, François
1979 A Study of Timing in a Manual and a Spoken Language: American Sign Language and English, Journal of Psycholinguistic Research 8, 379-405.
1980a Spoken Word Recognition Processes and the Gating Paradigm, Perception & Psychophysics 28, 267-283.
1980b Psycholinguistics of Sign Language, in Recent Perspectives on American Sign Language, Lane & Grosjean eds. (Hillsdale, NJ, Erlbaum Associates).
Grosjean, F., Hartmut Teuber, & Harlan Lane
1978 The on-line processing of gated signs in American Sign Language. Working paper, Department of Psychology, Northeastern University.
Kirk, Roger
1967 Experimental Design: Procedures for the Behavioral Sciences (Belmont, CA, Brooks/Cole)
Klima, Edward, & Ursula Bellugi
1979 The Signs of Language (Cambridge, MA, Harvard University Press).

Kučera, Henry, & Nelson Francis
1967 Computational Analysis of Present-Day American English (Providence, RI, Brown University Press).
Lane, Harlan, Penny Boyes-Braem, & Ursula Bellugi
1976 Preliminaries to a Distinctive Feature Analysis of Handshapes in American Sign Language, Cognitive Psychology 8, 263-289.
Marslen-Wilson, William, & Lorraine Komisarjevsky-Tyler
1980 The Temporal Structure of Spoken Language Under-
* standing, Cognition 8, 1-71.
Morton, John
1969 Interaction of Information in Word Recognition, Psychological Review 76, 165-178.
Siple, Patricia
1978 Visual Constraints for Sign Language Communication, Sign Language Studies 19, 95-110.
Stokoe, William
1960 Sign Language Structure= Studies in Linguistics: Occasional Papers, 8; revised 1978 (Silver Spring, MD, Linstok Press).
Stokoe, W., Dorothy Casterline, & Carl Croneberg
1976 A Dictionary of American Sign Language on Linguistic Principles (revised; Silver Spring, MD, Linstok Press).
Studdert-Kennedy, Michael, & Harlan Lane
1980 Clues from the Differences between Signed and Spoken Languages, in Signed & Spoken Language: Biological Constraints on Linguistic Form, Bellugi & Studdert-Kennedy eds. (Deerfield, FL, Verlag Chemie).
Swinney, David
1979 Lexical Access During Sentence Comprehension: (Re) Consideration of Context Effects, Journal of Verbal Learning & Verbal Behavior 18, 645-659.
Teuber, Hartmut, Robbin Battison, Harlan Lane, Joe Heck, & Jim Stungis
1980 A Computerized Lexicon of American Sign Language, The DASL 1965 in Fortran, Sign Language Studies 29, 349-369.

*
Marslen-Wilson, William, & A. Welsh
1978 Processing Interactions & Lexical Access During Word Recognition in Continuous Speech, Cognitive Psychology 10, 29-73.

François Grosjean obtained his Ph.D. and Doctorat d'Etat in psycholinguistics from the University of Paris in 1972 and 1978, respectively. He is currently an Associate Professor in the Department of Psychology at Northeastern University, where he coordinates the linguistics program and teaches courses in psycholinguistics and bilingualism. He also conducts research on the perception and production of spoken and signed languages. He has just edited a book in collaboration with Harlan Lane, entitled Recent Perspectives on American Sign Language (Hillsdale, New Jersey: Erlbaum) and is completing a general introduction to bilingualism for Harvard University Press.