The Rate of Speaking and Signing

8 For hearing and speaking people, language is produced by modifications of the stream of air that passes through the oral, nasal, and pharyngeal cavities. For deaf signing people, language is produced by modifications of the hands and fingers moving in space. What are the consequences of this difference in production mode? One consequence might be a difference in the rate of articulation for the two languages: clearly the sizes of movements made by the articulatory organs in the case of speech and signing are radically different, and signing, unlike speech, is independent of breathing. If there are in fact differences in the rate of production of units, does this difference affect the rate of producing sentences or underlying propositions in the two language modes? Are differences related to the ways the two kinds of language are structured? How far-reaching are the differences in the production mode for speech and sign?[1]

Measuring the Rate of Language Production

What is known of the rate at which speech is produced? Goldman-Eisler (1968) reports a series of studies of spontaneous speech that are relevant to this question. We tend to think of speech as an even flow, a stream of sound; but Goldman-Eisler notes that spoken language is really very fragmented and that the flow of sound is frequently interrupted by hesitations or pauses. In response to a request to describe picture stories, most of her subjects spent between 40 and 50 percent of their total speaking time in pauses. Thus when investigating the rate

at which language is produced, it is important to separate out the amount of pausing time.

In the Goldman-Eisler studies, rate of articulation was measured as the number of syllables per minute of the time spent in vocal activity (pauses subtracted out). The studies found that although individuals differ in articulation rate, within individuals the rate of articulation is remarkably constant, even in very different types of situations. Goldman-Eisler suggests that what is experienced as a variation in the speed of talking within individuals turns out on careful analysis to be a variation in the amount of pausing: "What is experienced as an increase of speed in talking is therefore due largely to the closing of gaps" (p. 26). The important point here is that within an individual, rate of articulation seems to be a constant of considerable invariance: "Considering that the mechanics of speech production are in the normal adult a skill of high order and stability of output is characteristic of skillful performances, the relative invariance of the rate of articulation is not surprising" (p. 26).

A Comparison Study: Bilingual Subjects

Since there are individual differences in rates of articulation for speech but the rate within an individual remains constant, ideal subjects for a comparative study of rate of production in two languages are people who are highly practiced and fluent in both languages. Fortunately, there is a special group of people who can be fluent in both speech and sign: hearing sons and daughters of deaf parents. If the parents' primary mode of communication with each other is sign language—and that is the usual case—it may also be the primary mode of communication with hearing as well as with deaf children; thus the hearing child may learn sign language as a native language. He or she may learn spoken language from older hearing children in the family, from relatives, or from neighbors and children on the street. The hearing child of deaf parents may from a very early age play a special role as interpreter; he may translate into sign language what hearing people say to his parents and translate into spoken language what his parents sign. He may thus become not only bilingual but an unusually fluent bilingual interpreter.

The subjects for our first study were three young hearing adults who had learned ASL as a native language from deaf parents and who had signed all their lives. All three were presently using ASL as part of their work, their studies, and their living situations. They were therefore extremely fluent in both sign and speech and highly practiced and accomplished in both modes.

Each subject was asked to tell some personal anecdote or story he

knew well. Without specifying at the beginning that the subject would be requested to repeat the story, we asked (in different orders) for three different renditions: one in ASL, one in spoken English, one simultaneously signed and spoken. Each rendition was videotaped and each videotape was carefully transcribed, resulting in four transcriptions for each subject: one of the story in sign language alone; one of the story when it was spoken only; one of the signed part, and one of the spoken part of the simultaneously signed-and-spoken version of the story.

Rates of Words and Signs

To compare the rate of articulation in ASL signing and spoken English, it is necessary to compare words and signs as the units of measurement, since there is no obvious direct analogue in ASL to the syllable in English.[2] In counting lexical units in English, spoken contractions and polymorphemic units were counted as one word (*don't* and *jumped,* for instance). For ASL, an item was counted as a single sign even if it had other information incorporated into it. The root sign INFORM, for instance, can be varied by changing its direction to mean 'you inform me'; it was nevertheless counted as a single sign.[3] Some terms, such as *O-F-F, D-O, B-Y*, were frequently fingerspelled by the hearing signers; in the signed stories, between 2 and 12 percent of the words were fingerspelled. Since they were short (an average of three letters) and often highly practiced, condensed forms, we counted the fingerspelled words as single signs.[4]

Each story was timed from the start of the first utterance to the end of the final utterance. To measure the time spent in pausing, a scorer watched and listened to the videotapes and recorded all durations of measurable pauses. This was done using a telegraph key signal attached to an Oscillomink equipped with a 100 Herz signal. The pauses were measured three times for each condition, and median measured time was used. Signing in all cases was measured at slow motion at the ratio of 3 to 2, and the results were then adjusted to normal speed.

Measuring pauses in signing presents special problems. Although it is easy to distinguish between vocalization and silence, it is less easy to distinguish signing from nonsigning. A signer's hands are always visible; and though nonmovement of resting hands can of course be distinguished from movement of gesturing hands, transitions to and from a sign (or between signs) must be distinguished from movement of the sign itself. Moreover it is sometimes difficult to distinguish a normal final hold of a sign from some sort of extra lengthening of the sign, which is one way of pausing or hesitating. For example, one signer when signing MANY YEARS BEFORE held the sign BEFORE for

nearly one second before moving her hands to the position for the next sign; another signer repeated the bounce contact of SHOE five times at the end of a sentence, where the normal citation form would be made with only one repetition. In this early study, we thought our estimate of pausing time might have been underestimated. However, a more detailed reanalysis of the same data with far finer measurements produced essentially the same results.[5]

For the individually signed and spoken stories the rates of production (excluding pauses) were as follows:

	Mean words per second	Mean signs per second
Subject A	4.0	2.3
Subject B	4.9	2.3
Subject C	5.2	2.5

These data suggest a striking difference between rate of articulation for the two modalities: for each subject the rate of articulation for words is roughly double the rate for signs. Further, the differences across modalities for each subject were considerably greater than the differences between subjects.

Rates in Simultaneous Speaking and Signing

In the individually signed and spoken renditions, there were, as there would be in any two versions of a story, detailed differences in the way each idea was conveyed. For instance, a subject said, *My sister was always a lot bigger than I was and a lot stronger;* in her signed version she signed the equivalent of 'My sister was always much stronger than I, bigger than I was.' The two are roughly paraphrases; they are the same in meaning but do not use precisely the same words or structures.

It is perhaps more accurate to compare the rate of production between two languages if one can ascertain that the propositional contents match. In the case of languages in two different modes, this sort of comparison is facilitated by the possibility of simultaneous production. The subjects in this study were very accomplished at the difficult feat of speaking and signing simultaneously, for it was common experience for them to converse with a mixed group of hearing and deaf people.

For their simultaneously spoken and signed versions of the story we again measured total times, subtracted time spent in pauses, and counted total words and signs. Pause times were measured separately for the spoken and signed versions. More time was spent pausing in speech than in sign, and the percentage of time spent in pausing by

each subject when signing and speaking simultaneously was somewhat greater than that when producing either modality separately. (This may reflect the greater cognitive load involved in producing languages in two modes simultaneously.)

	Pausing (percentage of time)			
	Separate production		Simultaneous production	
	Speaking	Signing	Speaking	Signing
Subject A	29.6%	20.9%	33.6%	28.1%
Subject B	23.6%	10.6%	26.4%	25.0%
Subject C	30.2%	12.4%	34.9%	16.8%

Under the special constraints of producing a narrative in sign and speech simultaneously, the rate of signing remained virtually unchanged, but the rate of speech was somewhat slower than when speaking alone. The rates of production for signed and spoken stories produced simultaneously (excluding pauses) are shown below.

	Item rates: simultaneous production	
	Mean words per second	Mean signs per second
Subject A	3.4	2.2
Subject B	4.4	2.5
Subject C	4.1	2.5

Even when expressing the same propositional content, then, the subjects filled the temporal intervals with different numbers of basic units. The rate of articulation for words was at least one and a half times the rate of articulation for signs.[6]

The study described here used relatively gross measurements and was made before we understood much about the temporal properties of signs and signing. We have redone some of the measurements in a far more refined and detailed way and find our first measurements substantially confirmed; furthermore, Grosjean (1977) has studied the rate of signing and speaking in memorized narratives and reports results comparable to ours, namely, that the mean duration of signs is twice the duration of words.

Temporal Processes Underlying Sentence Production

Until now we have been discussing physiological aspects of rate of production of American Sign Language and spoken English. In observing that ASL signs take longer to produce than words, we are not discussing the potential rate of articulation of the fingertips as compared

duction for propositions in the two modes. The data suggest that there may be a common underlying temporal process governing the rate of production of propositions in language, regardless of the mode.

Our production of language is not determined solely by physiological factors such as the rate at which the basic units—words or signs—are produced. Both speaking and signing involve coordinated sequences of articulated gestures and muscular contractions, whether these are in and around the mouth or in the fingers and hands; clearly, language production requires a plan to direct the order and timing of muscles. Cognitive processes are involved in planning sequences of words or signs and executing them while planning the next; such planning operations are thought to be reflected in hesitation pauses—for example, those that tend to occur at grammatical junctures and constituent boundaries in spoken language (Goldman-Eisler 1968).[9]

McNeill (1974) has proposed that the basic encoding process in speech is the planning operation that produces elementary sentences. From a study of adults and children, he suggests that a constant amount of time is taken to construct underlying elementary sentences and that this is on the average of 1.0 to 2.0 seconds. He hypothesizes that the constant rate for producing underlying sentences may be linked with shifts of attention, which ordinarily occur every one or two seconds. Each new elementary sentence encodes further information into some sort of semantic form and thus perhaps requires a shift of attention; pauses give time for the process of encoding underlying elementary sentences to catch up with the utterance of syllables, words, or surface phrases. The function of such pauses, McNeill argues, is to permit speech to proceed smoothly at the underlying level, even at the cost of interruptions at the surface level.

The mean number of seconds per proposition that we have found in signing alone, speaking alone, and signing and speaking simultaneously is well within the range posited by McNeill for spoken language. This suggests that cognitive processes underlying the production of propositions may not differ in the two modalities, even when comparing deaf and hearing subjects.

What's in a Sign: A Comparison of Signing and Speaking

Since signs take longer to produce than words, how does it come about that propositions in ASL and in spoken English are produced at similar rates? What kinds of information are packaged in the individual units, the words or the signs? In the simultaneous production of a familiar narrative, for instance, the underlying propositions and the general meanings were the same. But on the surface the narratives were expressed in different numbers of basic units; for example, the si-

multaneous story told by one subject has 122 signs as compared with 210 words. Is it really possible that the same message was conveyed?

At the beginning of our research on ASL, we noticed that in the ASL sentences deaf signers gave as translations for English sentences, many English words (in particular, grammatical morphemes) were not directly represented as separate signs. Our impression was that ASL simply used fewer morphemes and that there was some premium on economy of expression. Deaf researchers would point to a number of different morphemes in the English sentence and sign ELIMINATE THAT, THAT, THAT, then condense the whole sentence into as few signs as possible.[10]

For example, the sentence *It is against the law to drive on the left side of the road* was characteristically translated into three signs: ILLE-GAL DRIVE LEFT-SIDE. (Presenting the signs for retranslation yielded *It is illegal to drive on the left side*—three signs as compared with nine words.) The omitted words are primarily grammatical morphemes, noncontent words such as *it, is, to, on, the.* So far, this seems not unlike the way we construct telegrams, keeping the main contentive words and eliminating functors like articles and copulas. Perhaps ASL economizes and saves time by omitting such nonessentials.

Such economizing would account for a few of the differences between the number of words in the spoken version and the number of signs in the ASL signed version of the simultaneously produced stories. One signer, for example, used 23 words while simultaneously producing only 9 signs:

(1) They both looked at me.   TWO-OF-THEM LOOK-AT.[11]
(2) And they looked at each other.   THEN LOOK-AT.
(3) And they started laughing and   START LAUGH.
    laughing.
(4) This made me burst out cry-   MAKE (ME) CRY.
    ing.

Judging from our transcription of the signing, it appears, however, that more may have been omitted than "nonessentials": pronouns such as *me, they, each other;* parts of the description of actions, such as the repetition in the phrase *laughing and laughing;* and *burst out* in the phrase *burst out crying.* From such word-for-sign gloss translations, it appears that our claim of equivalent messages might have been ill-advised. Nonetheless, when we showed the videotaped sign rendition of these sentences to bilinguals, the retranslations essentially matched the spoken versions. Evidently considerably more information is compacted into the signs than appears from our word-for-sign transcriptions.

The practice of transcribing signs of ASL by writing down an En-

glish gloss for each sign is very common and was our practice in the first years of our research. The rule we followed was that one must always use the same English gloss to represent the "same" ASL sign. For some time this method insured that we would ignore any modification in the form of a sign that might signal a change in meaning.

In fact, the signed version of the narrative just quoted contains information omitted in our transcription. Careful attention to the form of each sign reveals that many of the signs in this story context are not made in uninflected form; in various ways they incorporate additional meaning. For example, in sentences (1) and (2), the sign LOOK-AT was made in two ways, both different from the root form (one hand, directed away from the signer). In sentence (1) LOOK-AT was made with two hands oriented toward the signer, in sentence (2) with two hands oriented toward each other. The difference in orientation and direction resulted in the two different translations into English: *they looked at me, they looked at each other*. The signs LAUGH and CRY in sentences (3) and (4) were also different from their root forms. LAUGH was made repeatedly with slow movement, intertranslatable with the English *laughed and laughed;* CRY, which is ordinarily repeated in root form, was made only once and with an intensified movement, which led signers to translate it as *burst out crying*. We were to discover that such modifications—changing the orientation and location of the hands in space, adding hands, adding or deleting repetition, changing the manner of movement—are systematic methods of incorporating additional information into signs (see chapter 12).

One way of investigating how information is expressed in ASL signing is to begin with a signed narrative in ASL—not a translation from English—and then develop a matching English translation, comparing the two in terms of the expression of equivalent messages. We began with a brief, simple ASL signed narrative videotaped by a deaf native signer. Four fluent signers independently transcribed the tape, under instructions to make the English translation match closely with the ASL signs; from this we constructed a best fit in the view of the deaf signers, and retranslation verified that indeed the English and ASL matched in message content. The English translation and the ASL transcription are given side by side.

English	ASL[12]
(1) A man was carefully washing his brand new car.	NEW[+] CAR, VEHICLE-classifier[+]; MAN WASH[+].
(2) Another man and a dog happened by.	DOG MAN COME-OVER[+].

(3) Suddenly a cat came along.	WRONG[+], CAT COME-OVER[+].
(4) The dog and the cat were definitely not friendly.	DOG CAT FRIEND[+].
(5) They snarled and clawed at each other.	SNARL[+], CLAW[+].
(6) The man who was washing the car became angry because they jumped all over it.	MAN WASH[+] ANGRY[+]. WHY[+] JUMP[+].

Compare the renditions of utterance (6) in the two languages. Certainly the ASL version is more economical: five signs compared with fifteen words. Yet the ASL version is not like even a headline or telegram; in fact, it seems cryptic. What did the man wash? Who was angry? Who or what jumped on what?

There are three essential types of clues in the ASL signing of utterance (6) that make it intertranslatable with the English version: the special use of space for pronominal and anaphoric reference, the modulation of meaning by changes in the movement and location of the signs, and the use of facial expression to indicate clausal subordination. Of the five signs, only the first (MAN) is made in its root form; the others all have something added or incorporated.

Space, Time, and Memory

In English the order of words in a surface string is important. *The man washed the car* and *The car washed the man,* though not equally likely, are certainly different in meaning, the difference being signaled by the order in which the words appear in the sentence. Furthermore, in English we insert morphemes even when they could easily be understood from context: *The man washed the car and waxed* is a complete thought but not a complete sentence. In ASL, how is it that the two signs MAN WASH can be intertranslatable with *The man who was washing the car,* and the single sign JUMP can convey the same message as *jumped all over the car*?

Part of the answer lies in an elaborated use of space in ASL signing, which permits compression of information into single sign forms. The sign CAR (putative object of WASH and JUMP) was made only once, in the first sentence of the narrative. It was followed by a classifier sign, a sign that stands for 'vehicles,' which was set in a particular spot to the right of the signer. Thus that spot, that locus in space, was reserved as the car locus from that point on. Other signs made use of other areas and radials in the signing space: the cat, the dog, and the second man

made their entrances and had their arguments without intruding on the locus reserved for the car. Now in (6), fifteen signs later, WASH and JUMP are directed and oriented toward the locus reserved for the car, as if the classifier sign were still in place. Signer and addressee must remember the establishing sign and the location where it was made in signing space; the memory lingers and persists, as if the sign still occupied that spot on a kind of stage in front of the signer.

Modulation of Meaning

Another clue to the incorporation of meaning comes from dynamic changes in the quality of movement of a sign. The sign ANGRY is not made in citation form; there is a qualitative change in its movement. It seems somehow more intense. At first one might guess that it is an expressive rendition meaning 'very angry.' But if we study the movement of the sign under slow motion and make detailed comparisons between the normal form of the sign and its form in this sentence, we can make a more precise description: the movement is slow and heavy at first; then it accelerates to an abrupt stop. In chapter 11, we shall see that this specific change in the dynamic quality of the movement (which is demonstrably different from a stressed form) regularly adds the meaning of 'completed change of state.' Thus, concealed in the manner of signing is a systematic change that leads deaf native signers to translate this as *became angry,* rather than *was angry.*

Facial Signals and Grammar

An unexpected clue to the message conveyed by $\overline{\text{MAN WASH}}$, ANGRY comes from a particular use of facial expression. Rapidly shifting expressions, even grimaces, are common accompaniments to signing communication. Much of this panorama of facial expressions and head movement is lively expressive background to signing, but there are also specific, well-defined, restricted facial signals that in fact serve as signals of clausal embedding (Liddell 1977). The facial expression that accompanies $\overline{\text{MAN WASH}}$ (a head tilt and the tightening of certain facial muscles) provides the final motivation for translating the signs $\overline{\text{MAN WASH}}$, ANGRY as: *The man who washed the car became angry.*[13]

Throughout the narrative, then, signs exhibit special locations, entrances, movement, within the signing space; signs are made in special manners to incorporate modifications of meaning (NEW in sentence (1) is translated as *brand new*); signs are accompanied by facial signals that modify meaning ($\overline{\text{FRIEND}}$ is translated as *definitely not friendly;* the negation is in the facial gesture alone). The structured use of space, the modulation of movement of signs which incorporate additional meanings, the use of facial expression not only to convey nu-

ances of meaning but also to indicate syntactic patterning—these are all widespread devices of the grammar of American Sign Language.

How special these devices are to ASL can be seen by comparing the ASL version of this story with a version presented in Sign English, based on Signing Exact English (Gustason, Pfetzing, and Zawolkow 1972). The Sign English version uses ASL signs and adds affix-markers (loan translation signs for articles, for inflections such as -ly, -ing, -ed, for forms of the copula is, was, were, and so on).[14] In the ASL version there were 21 signs; in the Sign English version there were 51 signs and 11 sign-inflections to match the English affixes. Deaf signers experienced in ASL and S.E.E. rendered both versions several times on videotape. The average durations of units were comparable in the two signed versions, though longer in ASL (1.2 signs per second) than in S.E.E. (1.7 signs/units per second). The proposition rates, however, differed significantly: an average of 1.5 seconds per proposition for ASL, an average of 2.8 seconds for S.E.E.[15]

ASL has developed as a separate language, quite distinct from the spoken language of the community that surrounds it; its favored form of patterning is different from English, showing a preference for compacting information into single sign units. Attempts have been made to combine ASL signs with the grammar of English, beginning with the methodical signing system of the early 1800s. These are excellent ways of presenting English on the hands and are significant educational tools; yet somehow the combination of ASL signs and English grammar, though valuable in the classroom, has apparently not had much influence on the grammatical patterning of ASL, as will become evident in part III. The methodical signs (including signlike affixes) were soon abandoned as unwieldy and cumbersome (see chapter 3). ASL so far has resisted absorbing such imposed mechanisms from the spoken language. Perhaps the requirement to produce (and process) a greater number of sign units in order to sign a proposition—and the consequent increased duration of the proposition itself—contributed to the sense that the combined mechanism was unwieldy.

English, Sign English, and ASL sentences may convey the same propositional message, but they differ greatly in the number of lexical units required to convey that message and in the ways those units are elaborated. It is the special linguistic devices of American Sign Language that make possible a proposition rate for ASL identical to that of a spoken language in which the unit articulation rate is double its own.

In sum, what are the effects on language production of differences in language mode? Words are articulated by "speech organs" within the mouth and vocal tract; signs are articulated by the hands moving in

space. Given the radically different sizes of the movements made by the articulatory organs, it is not surprising that there is a difference in the rate at which signs and words are produced. What is intriguing is that this difference apparently has no consequences for the rate of producing propositions in the two languages: though signs are produced at half the rate of words, the rate of producing propositions does not differ in the two modes. ASL economizes by doing without the kinds of grammatical morphemes that English uses; ASL has special ways of compacting linguistic information which are very different from those of a spoken language like English.[16] (1) The structured use of space, (2) the superimposed modulations of the movement of signs, and (3) the simultaneous use of facial expression for grammatical purposes permit compacting of information without significantly increasing production time.

It is possible that the tendency toward compacting linguistic information in signs may be a response to temporal pressure on language production. Cognitive processes underlying language might well create an optimal production rate for propositions, regardless of language mode. Under such temporal pressure, a relatively slowly articulated language of signs might well exploit the possibilities of simultaneous elaboration of meaning which exist in the visual-spatial mode.

Opposite: Movement trajectories of grammatical processes (photographer, Frank A. Paul).