Coarticulation in Fluent Fingerspelling

Thomas E. Jerde,[1,2] **John F. Soechting,**[1] **and Martha Flanders**[1]

[1]Department of Neuroscience, University of Minnesota, Minneapolis, Minnesota 55455, and [2]Department of Neuroscience, Brown University, Providence, Rhode Island 02912

In speech, the phenomenon of coarticulation (differentiation of phoneme production depending on the preceding or following phonemes) suggests an organization of movement sequences that is not strictly serial. In the skeletal motor system, however, evidence for comparable fluency has been lacking. Thus the present study was designed to quantify coarticulation in the hand movement sequences of sign language interpreters engaged in fingerspelling. Records of 17 measured joint angles were subjected to discriminant and correlation analyses to determine to what extent and in what manner the hand shape for a particular letter was influenced by the hand shapes for the preceding or the following letters. Substantial evidence of coarticulation was found, revealing both forward and reverse influences across letters. These influences could be further categorized as assimilation (tending to reduce the differences between sequential hand shapes) or dissimilation (tending to emphasize the differences between sequential hand shapes). The proximal interphalangeal (PIP) joints of the index and middle fingers tended to show dissimilation, whereas at the same time (i.e., during the spelling of the same letters) the joints of the wrist and thumb tended to show assimilation. The index and middle finger PIP joints have been shown previously to be among the most important joints for computer recognition of the 26 letter shapes, and therefore the dissimilation may have served to enhance visual discrimination. The simultaneous occurrence of dissimilation in some joints and assimilation in others demonstrates an unprecedented level of parallel control of individual joint rotations in an essentially serial task.

Key words: sensorimotor integration; movement sequences; fingerspelling; sign language; hand movement; coarticulation

Introduction

Studies of sensorimotor integration have focused primarily on the various types of eye movement and on reaching and grasping movements of the arm and hand. These studies have yielded a reasonably good understanding of the multidimensional aspects of individual movements (Hess and Angelaki, 1997; Crawford et al., 2000; Santello et al., 2002; Flanders et al., 2003). However, less is known about the coordination of sequences of movements, although particular aspects of sequences (such as shoelace tying or fingerspelling) appear to be selectively impaired in apraxias and diseases of the basal ganglia (Poizner and Soechting, 1992; Tyrone et al., 1999).

Some of the most accomplished movement sequences are those of speech, in which it is well known that certain phonemes are articulated differently depending on which other phonemes will follow (Kent and Minifie, 1977; Fowler and Saltzman, 1993; Matthies et al., 2001). This phenomenon is called coarticulation; it suggests a high level of sophistication in the neural planning, as well as the neuromuscular generation, of speech movement sequences.

A level of fluency similar to that of speech might be expected to govern the control of well practiced hand movement sequences such as those used to type text or play the piano (Rumelhart and Norman, 1982). However, when we recorded and analyzed such movements, we found only a limited amount of coarticulation (Soechting and Flanders, 1992; Engel et al., 1997). In consonance

with earlier studies of drawing movements involving the entire arm (Morasso, 1983; Soechting and Terzuolo, 1987a,b; Pellizzer et al., 1992), we proposed that the limb motor system tends to produce sequences on a segment by segment basis. The relative fluency of spoken sequences could reflect the high level of sophistication of cortical language areas, the lifelong period of speech learning, or differences in the musculoskeletal execution.

The fingerspelling sequences of American sign language (ASL) combine aspects of speech (for planning) and hand movement (for execution). Fingerspelling forms an adjunct to the gestural language of ASL. Although the main component of the language has its own syntax rather than providing a transliteration of English (Bellugi et al., 1989; Poizner et al., 1990), words that have no sign (such as proper names) are spelled, letter by letter, using hand shapes corresponding to the English alphabet. Experts on fingerspelling have tabulated cases in which a particular letter should be spelled slightly differently depending on the preceding or the following letter (Battison, 1978). However it is not entirely clear whether these suggested and observed alterations represent the learning of additional hand shapes for certain letter pairs or, alternatively, whether the hand motor control system is capable of language-like fluency in modulating the segments of movement sequences.

Materials and Methods

Overview

The purpose of this experiment was to search for evidence of manual coarticulation by quantifying the finger, thumb, and wrist movements of professional sign language interpreters. The interpreters were asked to hold the hand shape for each letter of the ASL fingerspelling alphabet and then to spell selected letter strings. As shown in Table 1, the letter strings were either words or non-words, and the non-words were either pronounceable or not. Although these string categories provided an opportunity to test for linguistic influences, our main goal was to determine the

Table 1. Experimental design: letter string categories

I–S–C–(vowel)	N–T–R–(vowel)
Same initial letter	
Words	
DISCARD	INTRACTABLE
DISCERN	INTREPID
DISCIPLE	INTRICATE
DISCOVER	INTRODUCE
DISCUSS	INTRUDER
Non-words (pronounceable)	
BRISCANT	PLONTRALD
BRISCENSE	PLONTREM
BRISCIARD	PLONTRIMER
BRISCOZE	PLONTROTH
BRISCUDGE	PLONTRUDIP
Different initial letter	
Words	
CONFISCATE	ENTRANCE
VISCERAL	ENCHANTRESS
OMNISCIENT	VENTRICLE
PERISCOPE	CONTROLS
BISCUIT	TANTRUMS
Non-words (not pronounceable)	
QWKXISCAVB	SXBCNTRAQF
GDTISCEXI	PTCHNTRERF
XBTKISCIGR	DZQDNTRIKN
HVZSISCOWJ	HMZRNTROTZ
PCLMISCUDF	JMGSNTRUNL

extent of coarticulation. Thus, in each letter string, a fixed sequence (I–S–C or N–T–R) was followed by one of the five vowels (A, E, I, O, or U). The main question was whether the hand shaping for the penultimate letter of the four-letter sequence ("C" or "R") differed depending on which vowel would follow (see Fig. 1).

Experimental procedure
Four female subjects (three right-handed, one ambidextrous), recruited from an interpreter service, participated in the experiment. All were fluent in ASL and had normal hearing. All subjects gave informed consent to procedures approved by the Institutional Review Board of the University of Minnesota.

Subjects completed each trial with the right elbow resting on a flat surface; each trial started and ended with the hand relaxed. Subjects were presented with the target letter or letter string before each trial. On hearing a "go" command, subjects fingerspelled with the right hand. The experiment consisted of two blocks: single letters (static block) and letter strings (dynamic block). For the single letters, subjects held the corresponding hand posture for several seconds. For letter strings, the subjects were instructed to fingerspell at a "normal, conversational rate." The static block was always presented first.

In the static block, the stimuli were the 26 letters of the English alphabet. In the dynamic block, the stimuli consisted of letter strings containing either I–S–C or N–T–R (Table 1), followed by one of the five vowels: A, E, I, O, or U. (These strings will be referred to as "ISC_" and "NTR_.") Within each block, each target was presented 5 times in random order, for a total of 130 trials in the static block and 200 trials in the dynamic block. In half of the strings, the ISC_ or NTR_ was always preceded by the same initial letter (Table 1, top half), whereas in the other half the sequence could be preceded by different initial letters (Table 1, bottom half). As also shown in Table 1, the strings included words and non-words, and the non-words were either pronounceable or not. Thus the string categories were defined as being ISC_ or NTR_, same initial or different initial letter, and word or non-word (pronounceable or not).

Data acquisition
We recorded the subjects' hand postures dynamically using sensors embedded in a right-handed glove (Cyberglove, Virtual Technologies, Palo Alto, CA). The glove fit tightly but was thin and flexible and open at the fingertips. We recorded the motions of 17 df, with an angular resolution <0.5°, at 12 msec intervals. The measured angles were the metacarpal phalangeal (MCP) and proximal interphalangeal (PIP) joint angles for the thumb and four fingers; abduction of the thumb, middle, ring, and little fingers; thumb rotation; wrist pitch; and wrist yaw (Santello et al., 1998).

For the static block, we recorded data for 3 sec and then defined hand posture by averaging the values for each joint angle over the final 720 msec of each trial. For the dynamic block, we recorded data for 9 sec during each trial and then later identified and isolated segments of interest, as described below.

Data analysis
Finding hold times. To isolate distinct letters from the dynamic fingerspelling data, for each subject we used a discriminant analysis based on the static hand postures collected for that same subject. Given a training set of grouped data (in our case, a single letter or letter string for one subject), discriminant analysis maps these data into a multidimensional space (one dimension for each measured variable) and defines axes in this space that best maximize the ratio of between-groups variance to within-groups variance (Santello and Soechting, 1998). For each unknown data vector $\mathbf{y}$ (composed of angle measurements from 17 df), Mahalanobis distances to each group mean vector $\mathbf{u}$ (from the training set) were computed as $d_{ij}^2 = (\mathbf{y}_i - \mathbf{u}_j)' \, \mathbf{A}^{-1} \, (\mathbf{y}_i - \mathbf{u}_j)$, where $\mathbf{A}$ is the pooled covariance matrix. Thus the Mahalanobis distances are defined as being in a space that is normalized by the inverse of the measured variance of individual joints and the correlated motions of pairs of joints.

Taking the static hand postures as a training set composed of 26 groups, we could compute the distances at each point in time, between a dynamic measurement of joint angles and every letter cluster. This is shown in Figure 2, where values of Mahalanobis distance are plotted (in *gray scale*) across time for each letter cluster. To isolate the hold phase from transition postures, we focused on time points coinciding with local minima in the summed angular velocity of all measured joints (see Fig. 2, *top panel*). We then classified the hand posture defined by the joint angle vectors at each of these points as belonging to the letter cluster for which the Mahalanobis distance was smallest.

Using this automated letter recognition procedure, we first recorded the time points corresponding to I–S–C and N–T–R, plus the immediately preceding and following letters. So that we could combine the dynamic data across trials, we then resampled the data to normalize the time scale (for example, see Fig. 3).

Classifying hand postures during transitions. Figure 2 shows that it is also possible to attempt classification of hand postures during the transitions from one hold period to the next. In many cases there was a distinct switch in the classification at the time of peak velocity, resulting in vertical stripes in the gray scale plot (see for example, the time of peak velocity in the transition from N to F). Because the goal of our study was to identify the time course of coarticulation, we made further improvements in this temporal classification procedure.

For this purpose, we evaluated the information content of the letter and transition hand postures using another discriminant analysis. In this case, instead of using the hand postures recorded from the static block as the training set, we used hand postures at various points in time during dynamic fingerspelling. We defined a cluster in discriminant space for each letter string (i.e., each word or non-word) (Table 1). At any point in the normalized time scale, we could then compute the Mahalanobis distance between the joint angle vector of a trial and the clusters composed of vectors from the remaining trials of that same letter string, as well as the other letter strings within that category, at the same normalized time. For example, we could attempt to classify a given trial from the first string category (ISC_, same initial letter, words), as DISCARD, DISCERN, DISCIPLE, DISCOVER, or DISCUSS, on the basis of the angle vectors recorded at any time point. We did not expect correct classification during D–I–S, but we hypothesized that the hand shapes used to spell the C might fall into five distinct categories depending on the word, thus predicting the upcoming vowel.

The results of this analysis can be plotted as confusion matrices (see Fig. 5) in which entries along the diagonal represent correct classification
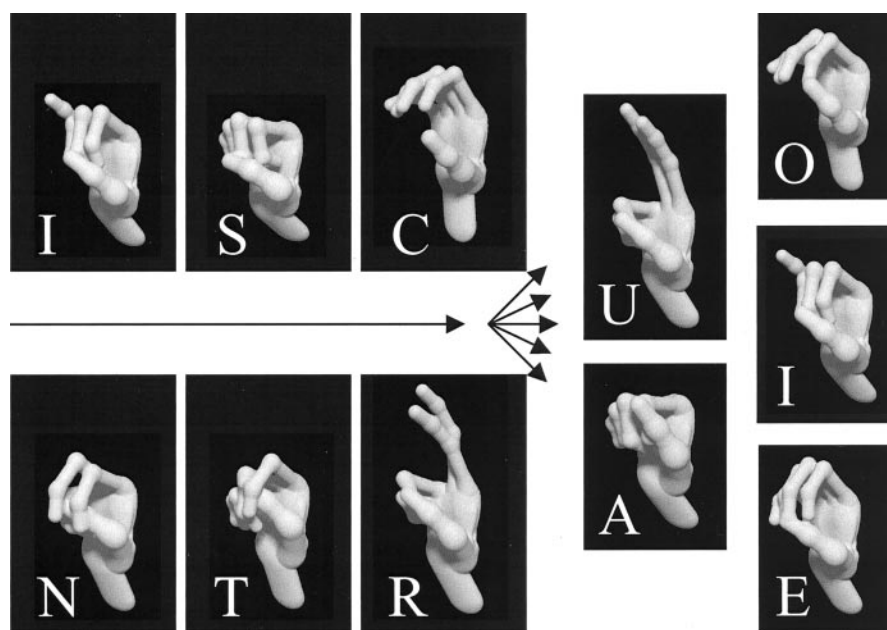
Figure 1. Cartoon images of hand shapes representing our letters of interest in the ASL manual alphabet, rendered with POV-Ray software. The layout illustrates our experimental design: two fixed-letter strings (*I–S–C* and *N–T–R*) were followed by one of the five vowels (*A*, *E*, *I*, *O*, or *U*).

(Sakitt, 1980; Johnson and Phillips, 1981). The correct rate is defined as the number of correct classifications divided by the total number of classifications. We calculated correct rates at intervals of 5% of the normalized time between the first letter (I or N) and the final vowel of each four-letter sequence of interest. This allowed us to examine information content trends across the movement time (see Figs. 6, 7).

To establish upper and lower confidence limits for significant deviation from the chance level in correct rate (1 of 5 or 20%), we used a bootstrapping procedure. For each subject and each category, at every time interval we ran the discriminant analysis 1000 times, each time generating a new training set by assigning trials to clusters randomly with replacement. Statistical significance ($p < 0.05$) was then established as achieving a correct rate higher or lower than 95% of bootstrapped runs (see Figs. 6 and 7, *dotted horizontal lines*).

Identifying the type of coarticulation. Coarticulation in fingerspelling is typically characterized as assimilation (where sequential hand shapes become more similar to one another) or dissimilation (where sequential hand shapes become more different). To quantify this between-letters influence on hand shape and to distinguish between the two types of influence, we performed a linear regression analysis within each string category, for each subject and for each joint angle measured. We correlated the angle at the time of the penultimate letter (C or R) with the angle at the time of the following vowel. A significant positive correlation represents assimilation; a significant negative correlation represents dissimilation.

Graphics. To facilitate both the analysis and the presentation of the results, we sometimes converted the Cyberglove data into a picture of the hand. Images of hand shapes were modeled and rendered using Persistence of Vision Ray Tracer (POV-Ray, copyrighted freeware).

Results

This study sought to identify and quantify instances of coarticulation in dynamic fingerspelling. Thus the main experimental question was whether the hand movements for spelling the C (in I–S–C) or the R (in N–T–R) differed depending on which vowel would follow. Figure 1 illustrates the hand shapes that we focused on, using images of the hand rendered from the Cyberglove data. In the *top row* we show the static hand shapes for the I (little finger extended), the S (a closed fist), and the C (an open but rounded

hand shape, resembling the printed letter). In the *bottom row*, we show the shapes for the N (with the thumb inserted between the ring and middle finger), the T (with the thumb inserted between the middle and index fingers), and the R (with the middle and index fingers extended and crossed). The experimental design was such that the C or R was followed, with equal probability, by each of the five vowels; the vowel shapes are illustrated on the *right side* of Figure 1. The shape for the U resembles the R, except that the fingers are not crossed. The shape for the O resembles the C except that it is closed, with at least one finger touching the thumb. The A is similar to the S (a closed fist) except for the placement of the thumb; the E also resembles the S, but is more open, with the fingertips touching the side of the thumb.

In the following sections, we will show that we could reliably predict which vowel followed the C (or R) by evaluating the Cyberglove data recorded during the I–S–C (or the N–T–R) epoch. We will start by showing that the speed of the S–C (or T–R) transition was a poor predictor, and we will then focus on the time-normalized movements of individual joints and of the entire hand (i.e., all 17 simultaneously recorded joint angles). We will also evaluate the extent to which the coarticulation represents assimilation or dissimilation.

Table 2. Transition times, from hold point to hold point

	S–C transition (msec)	T–R transition (msec)
Same initial letter		
Words		
Subject 10, mean (SE)	224 (06)	246 (10)*
Subject 11	265 (13)	319 (13)
Subject 12	443 (15)	436 (10)
Subject 13	266 (10)	307 (11)
Grand mean	*300*	*327*
Non-words (pronounceable)		
Subject 10	234 (08)	251 (09)
Subject 11	260 (09)	287 (08)*
Subject 12	446 (17)	434 (14)
Subject 13	306 (17)	291 (11)
Grand mean	*312*	*316*
Different initial letter		
Words		
Subject 10	232 (05)	254 (08)
Subject 11	246 (08)	300 (09)
Subject 12	448 (12)	442 (11)
Subject 13	289 (10)	295 (11)
Grand mean	*304*	*323*
Non-words (not pronounceable)		
Subject 10	304 (13)	293 (13)*
Subject 11	297 (12)	319 (08)
Subject 12	498 (16)	471 (13)
Subject 13	306 (12)	310 (15)
Grand mean	*351*	*348*

*Significant difference between movement times for A, E, I, O, and U; ANOVA, $\alpha = 0.0125$.

Velocity profiles and movement times

Words and pronounceable non-words were typically spelled at a rate of three to four letters per second (Table 2). Subjects 10, 11, and 13 spelled at comparable rates, with transition times (from one hold to the next) between a particular letter pair in each word or non-word ranging from 224 msec (for subject 10 spelling the S–C in DISCARD, DISCERN, DISCIPLE, DISCOVER, and DISCUSS) to 319 msec (for subject 11 spelling the T–R in words and nonpronounceable non-words). Subject 12 was the slowest, with transition times ranging from 434 msec for the T–R intervals in pronounceable non-words to almost 500 msec for the S–C intervals in nonpronounceable non-words. Considering the grand means (across subjects) for each letter string category, the nonpronounceable non-words were spelled substantially slower than the other types (the grand mean transition times are given in italics).

We wondered whether the penultimate transition time (i.e., the S–C in ISC_ or the T–R in NTR_) could predict which vowel would follow. Thus in Table 2 we have listed only the movement times for the S–C (left column) and the T–R (right column) transitions, and we have also indicted the results of multiple one-way ANOVAs comparing mean values across the different words (i.e., across the five different vowel cases). After correction for multiple comparisons (Bonferroni $\alpha = 0.0125$), we found only three significant cases in which the movement times for T–R differed depending on which vowel would follow. In each case (and in several others that narrowly missed statistical significance), this was attributable to the relatively slow spelling of the T–R when the R was followed by the U. T–R transition times were ~50–70 msec longer before the U than before the other vowels. Because the R and the U are very similar hand shapes (Fig. 1), this phenomenon may represent a slow-down before an R–U digraph. In addition, the T–R–A, T–R–E, and T–R–I transitions may have been expedited by the fact that after the T, the middle and index finger had to be extended almost into the R position to release the thumb, before full finger flexion for the vowel (see Fig. 1). Thus the R may have been formed somewhat "on the fly" in these cases.

As described in Materials and Methods, we used each subject's data from the static block of trials to automatically classify the hand shapes recorded at each point in time during the conversational spelling of letter strings (in the dynamic block). An example is shown in Figure 2, where Cyberglove data from one subject and one trial are correctly classified as representing the target word CONFISCATE. The speed profile in the *top panel* was computed as the sum of the absolute values of the angular velocities; the local minima correspond to hold times, where the letter should be clearly visible to the fingerspell reader.

In Figure 2, we have dropped lines from these hold points and have circled the darkest stripe, indicating the letter cluster (in
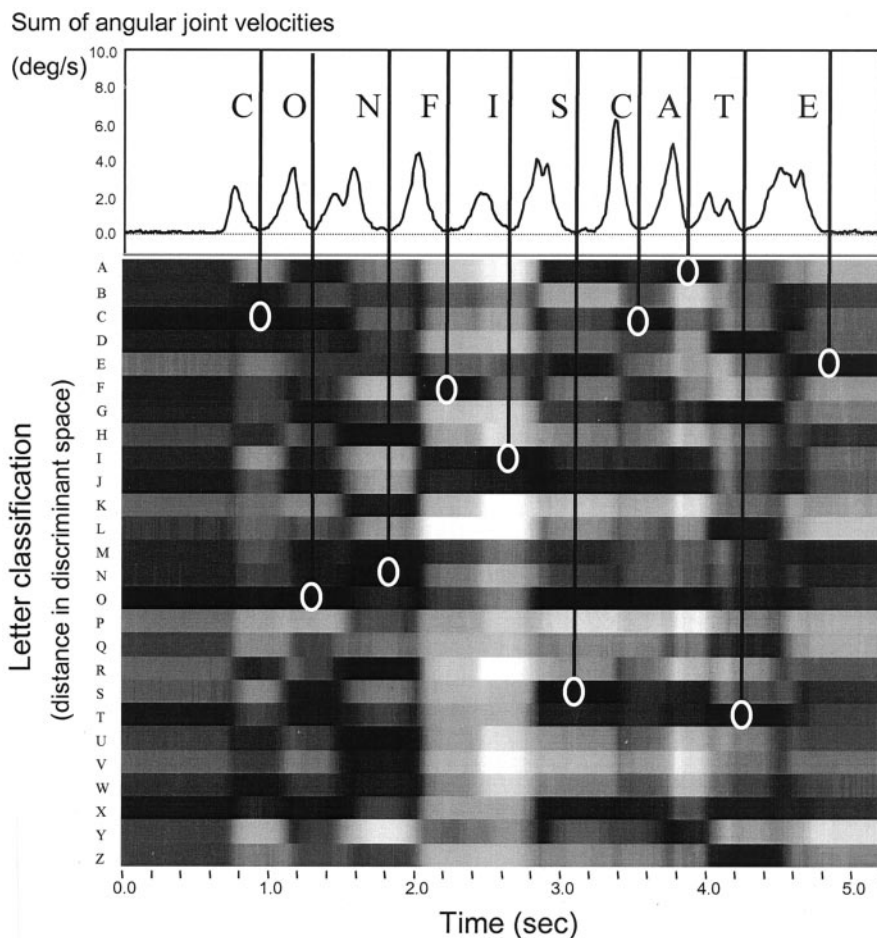

Figure 2. Automatic word recognition to isolate the letters of interest (here, I–S–C plus the preceding and following letters). Data are shown for one trial of subject 12. *Top panel,* Instantaneous angular velocity over time, summed over all 17 measured joints. The *vertical lines* mark local minima, indicating hold times. *Bottom panel,* Distance in discriminant space over time, between the measured joint angle vector and the vectors associated with each letter cluster in the training set. *Darker values* correspond to shorter distances. *White circles* indicate the letter cluster to which the measured vector is closest at each of the isolated hold times.

discriminant space) where the Mahalanobis distance is the smallest (see Materials and Methods). Notice that the other short distances (i.e., the other dark stripes at the same point in time) should represent letters with hand shapes that are similar to the one being classified. For example, the C and the O are similar, N and M are similar, I and J are similar, etc. In Figure 2, one may also notice biphasic speed profiles in cases in which the transitions to and away from a letter involve opening and then closing the fingers to insert or remove the thumb (e.g., N and T).

Examples of coarticulation in individual joints

The hold points identified as shown in Figure 2 (as local minima) were used to normalize the movement epochs for the transitions between hold times. We then plotted angular position across normalized movement time for each subject, joint, and letter string. Examples are shown in Figure 3, using data from subject 10 (*top panels*) and subject 12 (*bottom panels*) spelling DISCARD and DISCUSS (*left column*) and CONFISCATE and BISCUIT (*right column*). Hold points for the I, the S, and the C are marked with thick vertical lines; the traces end with the hold point for the vowel.

The examples in Figure 3 were chosen because they most clearly demonstrate coarticulation. We show the movement of only one joint, the PIP joint of the index finger. The movements
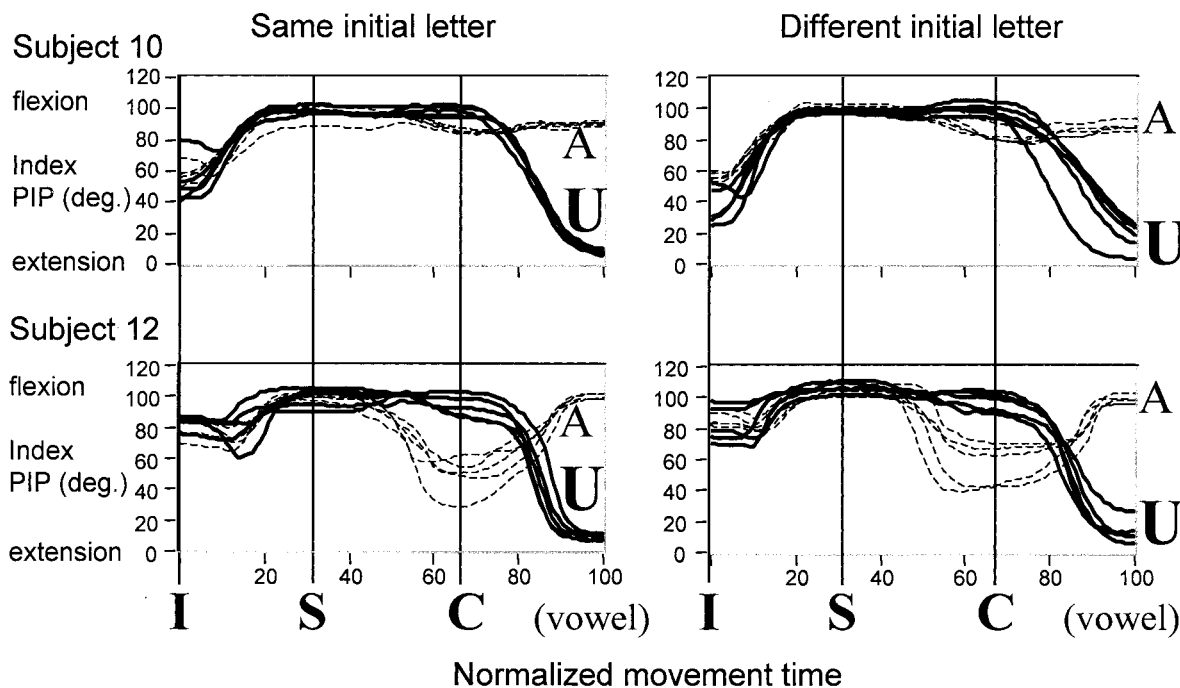
Figure 3. Joint angle data (flexion is positive, full extension is zero) of the index finger proximal interphalangeal (*PIP*) joint for two subjects spelling I–S–C–A (*dashed lines*) or I–S–C–U (*solid lines*), over the normalized time course of the movement. The movement from S to C differed depending on whether the following vowel was A or U. This difference is an example of dissimilation, a negative correlation between the measurements at C and at the vowel.

for DISCARD and CONFISCATE (the "A words") are shown as dashed lines; the movements for DISCUSS and BISCUIT (the "U words") are shown as heavy solid lines. The movement from the S to the C differed depending on whether the following letter was an A or a U.

Figure 3 also shows that there was some variation in the index PIP angle during the I. In the cases in which there was a different letter preceding the I for A words and U words (*right column*), this was a potential source of the variability. For example, in subject 10, the PIP was more flexed in CONFISCATE (the A word) than in BISCUIT (the U word). We will return to this issue below. Figure 3 also shows that at the hold point for the S (a closed fist) (Fig. 1), the index PIP was always tightly flexed.

A relatively wide range of PIP joint angles was observed for the letter C (Fig. 3), perhaps because the letter C can be recognized over a range of hand apertures (Fig. 1). Interestingly, this joint was more extended for the C when it would subsequently be fully flexed for the A (Fig. 3, *dashed lines*), and more flexed for the C when it would be subsequently fully extended for the U (*solid lines*). This is an example of dissimilation, a phenomenon that emphasizes the differences between adjacent letters and therefore may improve the reader's word recognition.

The PIP dissimilation shown in Figure 3 (for subjects 10 and 12) is representative of all four subjects. The index PIP data for subject 13 are displayed in Figure 4 (*third row* from *top, middle column*), along with data for all of the other measured joints. The A word (CONFISCATE) is represented by light blue lines, and the U word (BISCUIT) is represented by red lines. In Figure 4, we have included traces from the other vowels as well, color coded as indicated by the letters in the bottom row of the figure.

Although most joints showed distinct postures at the time of the vowel and some variation in posture at the time of the C, instances of dissimilation and assimilation varied from joint to joint. For example (Fig. 4, *middle column*), the index and middle

PIP joint angles at the C are clearly negatively correlated with the subsequent angles at the vowel (dissimilation). In contrast, the thumb and ring PIP joint angles at the C are positively correlated with the subsequent angles (assimilation), as are the wrist pitch and yaw angles (*top row*). One may also notice an apparent word by word variation at the time of the I in some joints (e.g., little MCP, thumb rotation and abduction).

Quantification of coarticulation using all joints
The next step was to develop a more complete quantification that would lend itself to statistical testing. Thus we developed a discriminant analysis using all recorded joint angles to quantify the time course of coarticulation in the hand as a whole (see Materials and Methods). In this analysis we posed the following question: at what points in time can we reliably classify the Cyberglove data from a particular trial as belonging to a particular word or nonword? We did a separate analysis for each subject and each letter string category (Table 1). Trends across subjects were very similar, so we will first show data from one subject (Fig. 5) and then report these trends as the mean values averaged across all subjects (Figs. 6, 7). Trends differed, however, depending on letter string category, and these results will be presented separately for each category.

In Figure 5, we show the success of discriminant analyses at seven different normalized time points: the hold points for each letter and the midpoints of each transition. Focusing first on the category of ISC_/same initial letter/words (Fig. 5A), the confusion matrix at each time point gives a graphical representation of a gradually increasing success rate. The trials being classified (*vertical scale*) are plotted against the classification result (*horizontal scale*), with the gray scale indicating the number of times that trials were classified as particular words. At the time of the hold period for the vowel (*far right*), the data from this subject were perfectly classified, as indicated by the black shading on the diag-
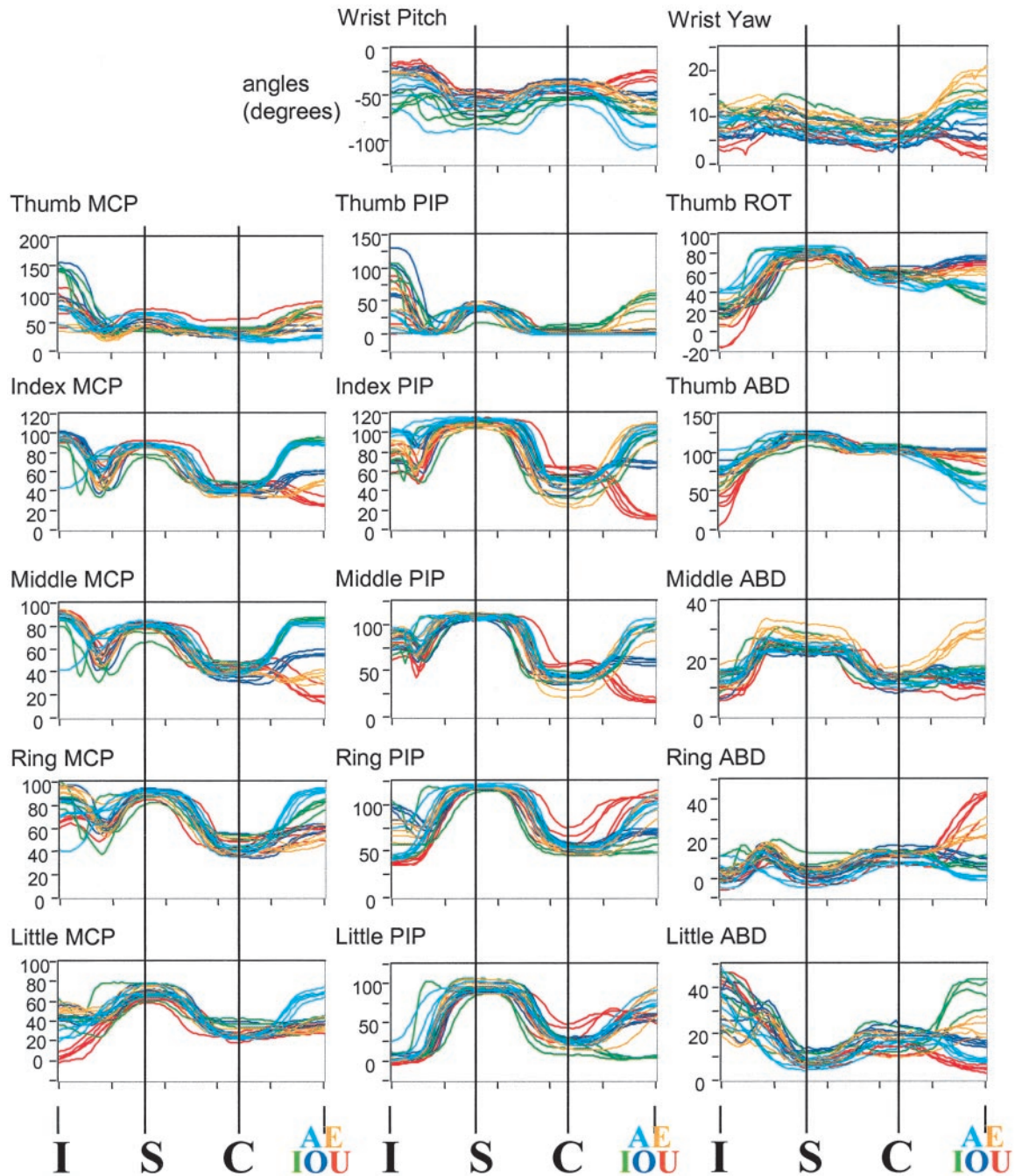
Figure 4.　Joint angle data from subject 13, spelling I–S–C followed by a vowel (different initial letter), for all measured joints (*MCP*, metacarpal phalangeal; *PIP*, proximal interphalangeal; *ABD*, abduction; *ROT*, rotation). For MCP and PIP joints, flexion is positive; for ABD angles, abduction is positive; for wrist pitch, downward is negative. Cases of dissimilation (negative correlation between angles at the C and the final vowel) can be observed in the index and middle PIP joints, and cases of assimilation (positive correlation) are evident in ring PIP joint and in the thumb and wrist.

onal. (As shown below, because of the variability of the joint positions, 100% correct classification was not usually achieved, even at the vowel hold.) Thus, in this case, DICUSS was correctly classified as a U word, DISCOVER was classified as an "O word," etc.

In contrast to the 100% correct classification rate at the time of the vowel, during the hold period for the I (Fig. 5A, *far left*), classification was no better than chance level (20%). For example, reading the top row of the confusion matrix, trials for the word DISCUSS could be classified as an A word (one trial), an "I word" (two trials), or an O word (two trials). Success rates were also near chance during the S, but improved dramatically at the hold period for the C, thus predicting what the upcoming vowel would be.

Figure 6 displays the time course of the success rate, for words and non-words with the same initial letter. For ISC_ words (*top left panel*), the success rate gradually increased during the transition from the S to the C. As emphasized by the arrow, at the time of the hold point for the C (*vertical lines*), the rate of correct classification was well above the chance level. A similar, gradual increase in correct classification was also observed for NTR_ words (*top right panel*) and, to a more limited extent, for the ISC_ and NTR_ pronounceable non-words (*bottom panels*).

The letter strings with different initial letters showed a different trend, in that the amount of information about the target word was already high at the time of the hold period for the I in

A. Same initial letter
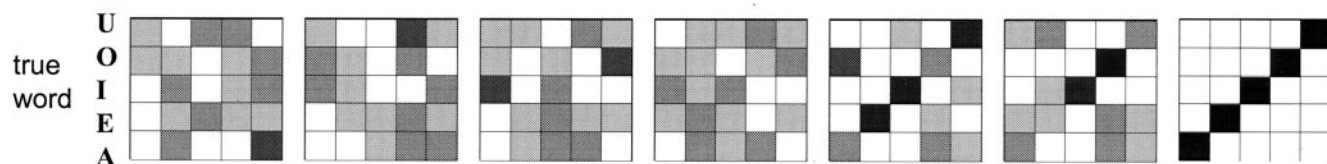
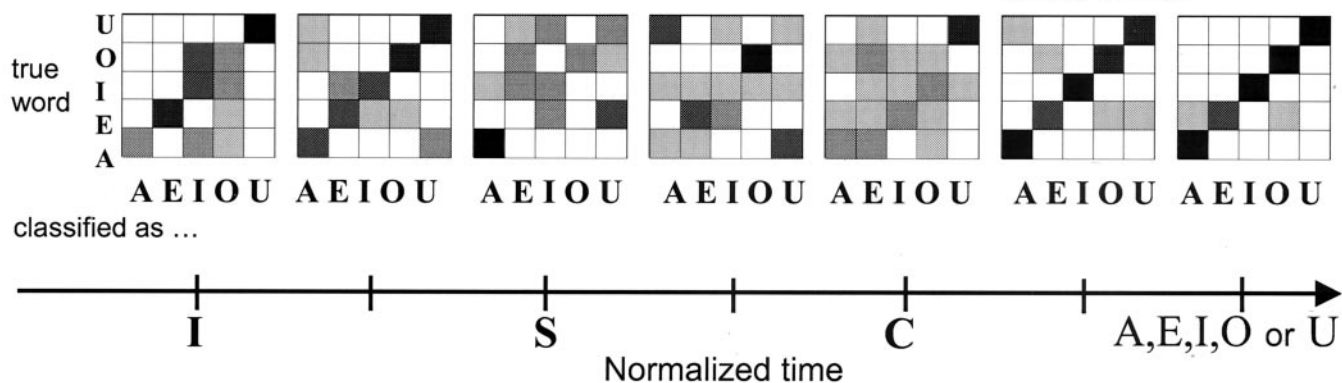## B. Different initial letter

Figure 5. Confusion matrices showing results of discriminant analyses classifying trials by target vowel. Separate analyses were done at various points in time and for trials with the same (*A*) or different (*B*) initial letters. *Darker squares* indicate more trials; *diagonal entries* represent correct classification. In the same initial letter category (*A*), there was a trend of increasing correct rate as the vowel was approached; this trend was already apparent by the time of the C, indicating that there was some information in the C hand shape about the target vowel (a reverse influence of the vowel on the C). In the different initial letter category (*B*), there was information about the target vowel in the I hand shape as well, indicting a forward influence from the different initial letters (which varied with the vowel).

ISC_ letter strings or the N in NTR_ letter strings (Figs. 5B, 7). For example, in Figure 5B, BISCUIT was correctly classified as a U word at the time of the I. Thus in BISCUIT, the B–I transition resulted in an I that was shaped differently than the I in the words containing R–I (PERISCOPE), N–I (OMNISCIENT), V–I (VISCERAL), OR F–I (CONFISCATE). Correct classification was subsequently diminished between the S and the C before it rose again during the spelling of the vowel. In fact, the letter before the I had such a strong "forward influence," on the spelling of the word, that correct rates (declined but) stayed above the chance level throughout the I–S transition. This is quantified for all subjects in Figure 7 (*left panels*). Comparable results are shown for the NTR strings in the right panels (although there were also some differences between the ISC_ and NTR_ categories as discussed below).

At the time of the penultimate letter of the ISC_ (or NTR_) sequence, the high rates of correct classification were attributable to a "reverse influence" of the upcoming vowel on the spelling of the C (or R). Conversely, in strings with different initial letters, the high initial correct rates were attributable to a forward influence of the previous letter on the spelling of the I (or N). Thus, when the previous letter was always the same (Figs. 5A, 6), correct rates started at chance levels, whereas when the previous letter differed from word to word (Figs. 5B, 7), the hand shape at the I (or N) could correctly reflect the word of origin. As illustrated schematically in Figure 8, this suggests an elaborate scheme of temporal blending of sequential hand movements.

Assimilation and dissimilation
In the sections above we have given joint by joint examples of a reverse influence of the vowel on the hand shape for the letter C (Figs. 3, 4), as well as evidence from the analysis of all joints, for

both reverse and forward influences, for both the I–S–C and the N–T–R letter strings (Figs. 5–7). We wondered whether each of these cases represented assimilation (i.e., sequential hand shapes becoming more similar) or dissimilation (i.e., sequential hand shapes becoming more distinct). The example from the index finger PIP joint (Fig. 3) was clearly a case of dissimilation, but considering all joints (Fig. 4), one also finds cases of assimilation (e.g., in the thumb and wrist in Fig. 4), as well as many cases in which there was no apparent correlation between the joint angles at the time of the C and the joint angles at the time of the vowel.

To quantify the extent and type of coarticulation, we first calculated the correlations of joint angles at the time of the C (or R) with the corresponding angles at the time of the subsequent vowel. For example, for the index PIP joint (Fig. 4, *third row, middle column*) of subject 13, there was a significant negative correlation ($r = -0.48$; $p = 0.01$), representing dissimilation.

Correlation coefficients for each of the 17 joints in each subject are displayed in Figure 9 (*circular symbols*). We confined this analysis to the spelling of words, because the fluency seemed to be slightly better (Figs. 6, 7, compare the *top panels* with the *bottom panels*). However, we used the data from the ISC_ and the NTR_ words and from the same and different initial letter categories; thus each joint is represented four times in each histogram. The critical value for a significant correlation coefficient ($n = 25$; $\alpha = 0.05$) was ± 0.39, as indicated by the vertical lines.

Each subject showed a wide range of negative and positive correlations (Fig. 9). Especially in subject 11, there were many cases in which joint angles were completely uncorrelated with the values for the following letter (i.e., the correlation coefficients near zero). However, each subject had at least a few cases of large positive correlations (assimilation) and large negative correlations (dissimilation).

It was not the case that the NTR_ strings tended to show assimilation, whereas the ISC_ strings tended to show dissimilation, or vice versa. Instead, for a given subject spelling a single word, some joints showed assimilation, whereas at the same time other joints showed dissimilation. We noticed that the joints of the thumb and the wrist tended to show assimilation, whereas the index and middle finger PIP joints tended to show dissimilation. In Figure 9, we have therefore marked the symbols representing the thumb and the wrist with an X, and we have filled in with black the symbols representing the index and middle finger PIP joints. Subject 11 may have been somewhat of an exception to this rule, because the largest negative correlations were observed in her ring and little finger PIP joints (*open symbols* at $r = -0.59$). However, in all other subjects the values for the index and middle finger PIP joints were among the largest negative correlations. In all subjects the majority of the significant positive correlations were in the thumb and wrist, possibly suggesting a strategic early "preplacement" of these joints in preparation for the posture of the vowel (Fig. 8).

Our study was designed primarily to focus on the reverse influence of the vowel on the spelling of the preceding letter (C or R). As detailed above, we did find an influence and were able to determine that it could represent both assimilation and dissimilation (in different joints, during spelling of a single letter string) (Fig. 9). By comparing strings with the same or different initial letters, we also found evidence for a forward influence of preceding letter on the shape of the I or the N (compare Figs. 6, 7). This was clearer in the ISC_ words than in the NTR_ words, perhaps because the I could be preceded by five different letters (F, V, N, R, or B), whereas the N was preceded by only three different letters (E, A, or O) (Table 1, NTR_, different initial letter, words).

Because the comparison (Fig. 6 vs Fig. 7) of correct rate trends strongly suggested the presence of coarticulation at the time of the I or N, we sought to further identify it as assimilation or dissimilation. Thus we also computed correlation coefficients for joint angles at the I or N compared with the joint angles at time of the hold for the preceding letter. In this case we found mostly positive correlations, some of which were unexpected and quite strong. For example, because the I is spelled with the little finger, one might have expected dissimilation, for emphasis. However, the correlations represented assimilation, in this case the phenomenon of "leaving behind" a particular joint angle as the others go on to spell the next letter (Fig. 8). All four subjects showed significant positive correlations for the little finger MCP joint, ranging from $+0.86$ in subject 10 to $+0.65$ in subject 11, indicating assimilation rather than dissimilation.

In this analysis of the I or N, the evidence for dissimilation was relatively weak. We found only two, marginally significant negat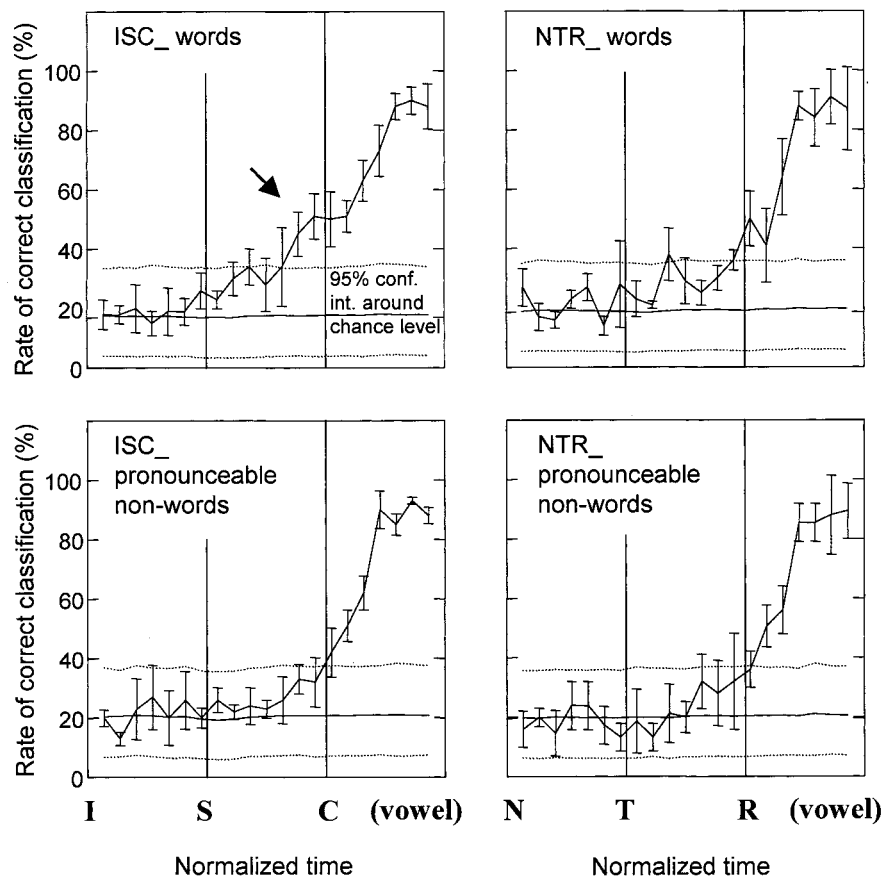ive correlations for the ISC_ words. For the NTR_ words, however, the index finger MCP joint had a substantial negative correlation both in subject 10 (-0.51; $p < 0.01$) and in subject 12 (-0.50; $p < 0.01$). There were only two other significant negative correlations in NTR_ words (both in abduction angles) and many strong positive correlations (mostly in the little finger PIP and in the wrist angles). Of course, a more exhaustive word list could potentially reveal additional cases of forward dissimilation.

Discussion

In this study we addressed the question of how movement sequences are organized, using fingerspelling in American sign language as a model system. This task has several advantageous characteristics. ASL has a strong linguistic component, like speech, but the movements are much easier to measure and characterize. Furthermore, the elements of the movement sequence in fingerspelling are self-evident, corresponding to the letters of the alphabet, with pauses at each letter. In contrast, the criteria to define elements in a sequence for other gestural tasks may not be as clear (Soechting and Terzuolo, 1987a,b).

In fingerspelling, we found substantial evidence of coarticulation, and we characterized the time course and classified the types of parallel control of the 17 joint angles.

Time course of coarticulation

The hand shape for a particular letter could depend on the letter that was to follow, as well as on the preceding letter. Thus there is
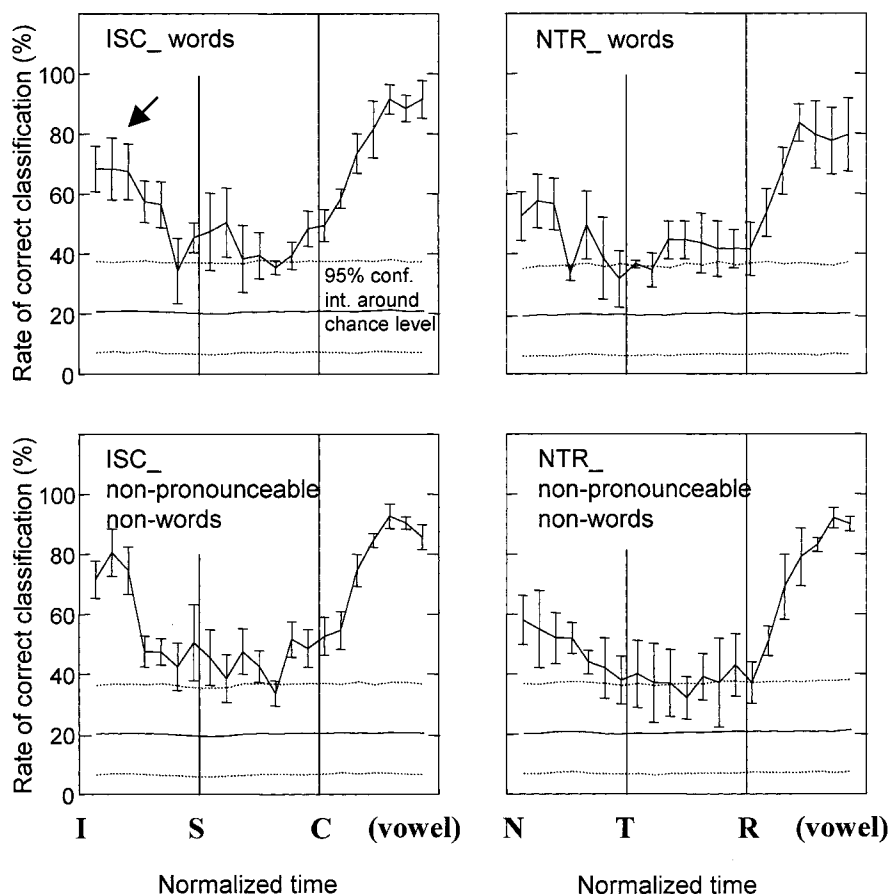

Figure 6. Correct rates over normalized movement time, averaged across the four subjects (grand mean ± SE). These values were obtained from the discriminant analysis for the same initial letter categories (Fig. 5A). Chance level correct rates (1 of 5 or 20%) are plotted along with 95% confidence intervals determined by bootstrapping. Each category showed a trend of increasing correct rate with movement time, which began before the penultimate letter (C or R). In the best case (ISC_ words, top left), the correct rate reached statistical significance well before hold time for the C (arrow), indicating a strong reverse influence on the penultimate hand shape by the final vowel.

Figure 7. Correct rates over normalized movement time, averaged across the four subjects (grand mean ± SE). These values were obtained from the discriminant analysis for the different initial letter categories (Fig. 5*B*). Chance level correct rates (1 of 5 or 20%) are plotted along with 95% confidence intervals determined by bootstrapping. In each category, in addition to a trend for reverse influence from the final vowel, there was a trend representing a forward influence from the different initial letters (*arrow*). Because of the combined effects of the forward and reverse influences, correct rates barely fell below the significance threshold during intermediate letters.

estimate of six phonemes as a maximum for anticipatory coarticulation in speech (Benguerel and Cowan, 1974). At the other extreme, although typing has a linguistic component, we observed a distinct lack of anticipatory coarticulation in this task (Soechting and Flanders, 1992). This may be attributed to the fact that typing differs from fingerspelling in the use of a reference position. Professional touch typists return to the home position after each key press; presumably this helps them to keep track of the spatial relationship between the hand and the keyboard. There is no such requirement in fingerspelling, and instead of returning to a standard posture after each letter, fingerspelling entails a series of transitions between letter shapes.

Although there were some major differences in the speed of the four subjects (with subject 12 being substantially slower than the others) and the variability (with subject 11 being the most variable), the normalized time course (Figs. 6, 7) and the use of particular joints for assimilation and dissimilation (Fig. 9) were very similar across subjects. A slight exception was subject 11 who showed substantial dissimilation with her ring and little finger PIP joints (Fig. 9). However, this subject also showed the normal pattern of dissimilation with distal joints and assimilation with more proximal joints (although the correlation coefficients sometimes failed statistical significance because of the large variability in this subject's performance) (Fig. 9).

a bidirectional flow of information (Fig. 8) defining the kinematic characteristics of each element of the movement sequence. We showed this using discriminant analysis and information theory to determine the extent to which hand shape at a particular instant could predict which word was being spelled. A reverse influence was demonstrated by ascertaining the effect of the vowel on the preceding consonant (Figs. 5*A*, 6). We also showed evidence for a forward influence by considering words in which there was a constant trigraph (I–S–C or N–T–R) but various letters preceding the trigraph. In these cases with "different initial letters" (Figs. 5*B*, 7), the hand shape at the I or N could predict the word of origin at better than chance levels. This result implies a forward influence of the different initial letters on the hand shape at the time of the I or N.

As shown in Figure 7, the effect of the preceding letter was clear at the time of the I or N but was nearly gone by the time of the next letter (the S or the T). Conversely, the reverse influence of the vowel began only during the transition from the S or T to the following letter (the C or R). Thus we can estimate a time course of ~1.5 letters (~0.5 sec) for the time spread of forward and reverse influences. This may be an underestimate because of the relatively closed hand shape of the S and the T (potentially limiting the amount of variability at this time). However, it is useful to compare this 1.5 letter estimate with the more extreme

Concurrent assimilation and dissimilation

Our results showed that the phenomenon of coarticulation could take two different forms: dissimilation, in which the differences in joint angles for the two letters were accentuated, and assimilation, in which they were minimized. Instances of dissimilation involved mainly the PIP joints of the index or middle fingers, whereas instances of assimilation were found primarily for the thumb and wrist joints. Two points should be noted. First, in a previous study (Jerde et al., 2003), we sought an economical means for computer recognition of static hand shapes in fingerspelling. We found that we could correctly classify letters 88% of the time using only four joint angles, including the PIP joints of the index and middle fingers. It is an open question whether the posture at a restricted number of joints conveys privileged information to human observers. However, the fact that we found instances of dissimilation primarily in these two joints is consistent with our previous results and suggests that its function is to aid in letter recognition.

There were instances in which we observed dissimilation at one joint and assimilation at other joints for the same letter combinations. This observation bears on the extent to which motion at the individual finger joints is coordinated. In studies of grasping, we found that two principal components could account for much of the variance in the postures and movements of the many
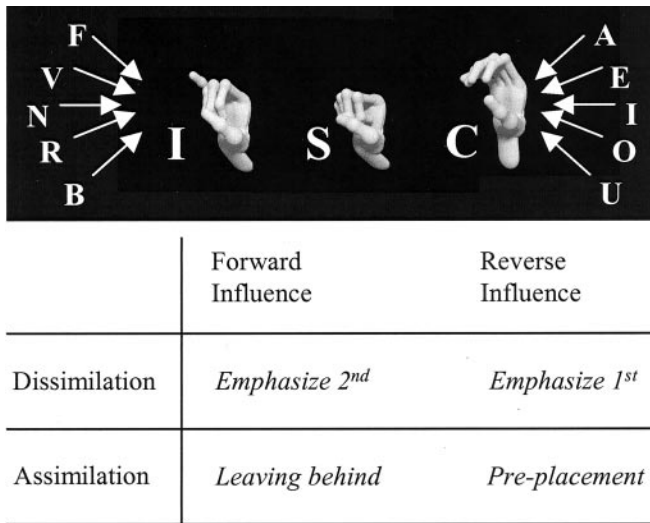
Figure 8. *Top*, A schematic representation of the forward and reverse influences on hand shape in our experiment. In the example letter string I–S–C, the shape of the I is influenced by the preceding letter (*F, V, N, R,* or *B*) (Table 1), whereas the shape of the C is influenced by the following vowel. *Bottom*, A table characterizing the motor and communicative strategies reflected by assimilation and dissimilation in the forward and reverse influences.

Figure 9. For each subject (*10, 11, 12,* and *13*), histograms of correlation coefficients computed between the penultimate letter (C or R) and the final vowel, for each measured joint angle. All data from ISC_ words and NTR_ words are included. *Vertical lines* indicate critical values for the statistical significance of each correlation ($n = 25$; $\alpha = 0.05$). Significant negative correlations were often found in the index and middle finger proximal interphalangeal (*PIP*) joints (*filled symbols*), whereas most of the significant positive correlations were found in data from the thumb and wrist joints (*symbols* marked with *X*).

joints of the hand (Santello et al., 1998, 2002). Likewise, we found a high degree of temporal coordination across joints and fingers during typing (Soechting and Flanders, 1997). These results show synergistic movements involving all (or many) of the mechanical degrees of freedom of the hand, rather than individuation of finger motion (Schieber, 1991, 1995). However, concurrent instances of assimilation and dissimilation argue against synergistic control. A closer inspection of the results from our grasping study also reveals a more complex picture: higher-order principal components, although they were small, did contribute information about the object to be grasped (Santello et al., 1998). Thus, although there is an overriding tendency for a coordination of motion of all fingers, there is a superposed ability for individuated control.

Organization of movement sequences

Our results, as well as a considerable body of previous evidence, indicate that at one level a sequence of movements is organized as a unit. In the present experiments, letter strings embedded in nonpronounceable non-words were executed at a slower pace than the same strings embedded in pronounceable words and non-words. This linguistic effect agrees with observations on typing (Viviani and Terzuolo, 1983). Terzuolo and Viviani (1980) also found that the rhythmic pattern of intervals between key presses showed word-specific characteristics, a phenomenon that may have been echoed here in the slowdown of the T–R transition only when it was followed by an R–U digraph (Table 2). In another study of typing, in a learning paradigm in which the location of two keys was switched, subjects tended to pause at the beginning of a word containing such a switched letter as well as before the letter itself (Gordon et al., 1994).

An organization of movement sequences in their entirety was first suggested by Lashley (1930). He proposed that all of the elements of a sequence would be represented simultaneously, the element with representation that was strongest at any one time being the one that would be executed. Patterns of neural activity consistent with this hypothesis have recently been found by Averbeck et al. (2002), who recorded prefontal cortical activities in
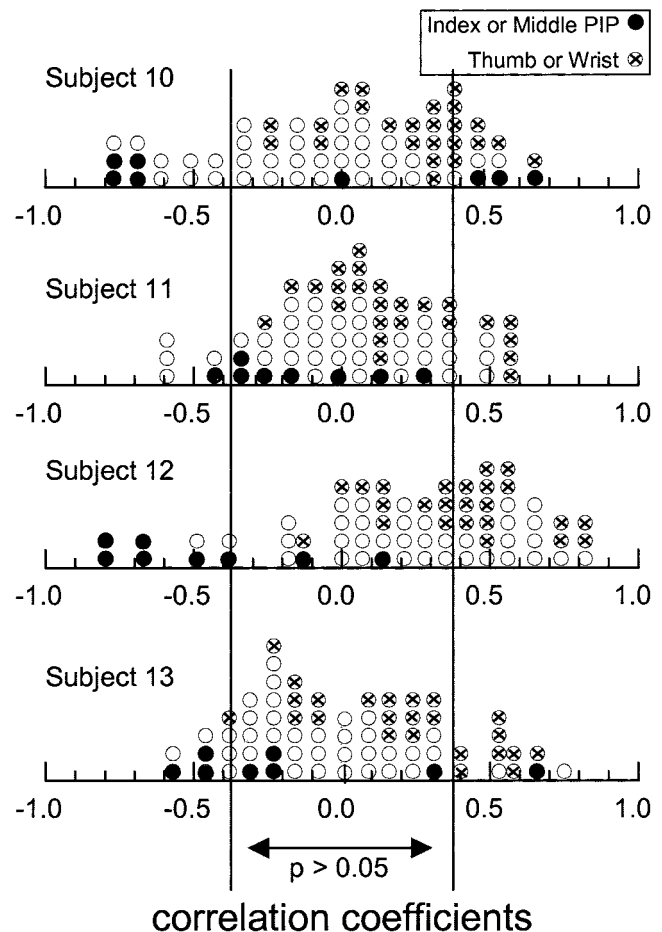
monkeys trained to copy geometric shapes. More generally, neural activity that is specific to (or dependent on) the location of an element in a sequence has been found by several investigators (Carpenter et al., 1999; Tanji, 2001). Note that at the kinematic level, Lashley's hypothesis (1930) is compatible with a strictly serial organization of movements, such as we found in typing (Soechting and Flanders, 1992), or one in which there is an overlap of the elements in the sequence (i.e., the case of assimilation). However, it is not compatible with the phenomenon of dissimilation, in which information flows backward in time to accentuate differences in postural transitions.

Over several decades, the study of speech has revealed many extreme examples of both reverse and forward overlapping of sequential elements (in this case, the articulation of phonemes). However, theoretical models that attempt to explain the organization of this control scenario are still controversial. In a comprehensive review article, Kent and Minifie (1977) favored the development of a somewhat hierarchical model, with the speech rhythm at the upper level and the "pattern of articulatory transitions" at a lower level. Intermediate to these two levels were phonemes (the sounds required for successful communication) as the loosely defined targets of the articulatory transitions. The

results of our fingerspelling study are also compatible with a characterization of sequential behavior as involving transitions between flexible goals. However, it is clear that there are still many open questions regarding the neural organization and implementation of these transitions.

References

Averbeck BB, Chafee MV, Crowe DA, Georgopoulos AP (2002) Parallel processing of serial movements in prefrontal cortex. Proc Natl Acad Sci USA 99:13172–13177.

Battison R (1978) Lexical borrowing in American sign language. Silver Spring, MD: Linstok.

Bellugi U, Poizner H, Klima ES (1989) Language, modality and the brain. Trends Neurosci 12:380–388.

Benguerel A-P, Cowan HA (1974) Coarticulation of upper lip protrusion in French. Phonetica 30:41–55.

Carpenter AF, Georgopoulos AP, Pellizzer G (1999) Motor cortical encoding of serial order in a context-recall task. Science 283:1752–1757.

Crawford JD, Henriques DYP, Vilis T (2000) Curvature of visual space under vertical eye rotation: implications for spatial vision and visuomotor control. J Neurosci 20:2360–2368.

Engel KC, Flanders M, Soechting JF (1997) Anticipatory and sequential motor control in piano playing. Exp Brain Res 113:189–199.

Flanders M, Hondzinski JM, Soechting JF, Jackson JC (2003) Using arm configuration to learn the effects of gyroscopes and other devices. J Neurophysiol 89:450–459.

Fowler C, Saltzman E (1993) Coordination and coarticulation in speech production. Lang Speech 36:171–195.

Gordon AM, Casabona A, Soechting JF (1994) The learning of novel finger movement sequences. J Neurophysiol 72:1596–1610.

Hess BJM, Angelaki DE (1997) Inertial vestibular coding of motion: concepts and evidence. Curr Opin Neurobiol 7:860–866.

Jerde TE, Soechting JF, Flanders M (2003) Biological constraints simplify the recognition of hand shapes. IEEE Trans Biomed Eng, in press.

Johnson KO, Phillips JR (1981) Tactile spatial resolution. I. Two-point discrimination, gap detection, grating recognition. J Neurophysiol 46:1177–1191.

Kent RD, Minifie FD (1977) Coarticulation in recent speech production models. J Phonetics 5:115–133.

Lashley KS (1930) Basic neural mechanisms in behavior. Psychol Rev 37:1–24.

Matthies M, Perrier P, Perkel JS, Zandipour M (2001) Variation in anticipatory coarticulation with changes in clarity and rate. J Speech Lang Hear Res 44:340–353.

Morasso P (1983) Three dimensional arm trajectories. Biol Cybern 48:187–194.

Pellizzer G, Massey JT, Lurito JT, Georgopoulos AP (1992) Three-dimensional drawings in isometric conditions: planar segmentation of force trajectory. Exp Brain Res 92:326–337.

Poizner H, Soechting JF (1992) New strategies for studying higher level motor disorders. In: Cognitive neuropsychology in clinical practice (Margolin D, ed), pp 435–464. New York: Oxford UP.

Poizner H, Bellugi U, Klima ES (1990) Biological foundations of language: clues from sign language. Annu Rev Neurosci 13:283–307.

Rumelhart DE, Norman DA (1982) Simulating a skilled typist: a study of skilled cognitive-motor performance. Cognit Sci 6:1–36.

Sakitt B (1980) Visual-motor efficiency (VME) and the information transmitted in visual-motor tasks. Bull Psychonom Soc 16:329–332.

Santello M, Soechting JF (1998) Gradual molding of the hand to object contours. J Neurophysiol 79:1307–1320.

Santello M, Flanders M, Soechting JF (1998) Postural hand synergies for tool use. J Neurosci 18:10105–10115.

Santello M, Flanders M, Soechting JF (2002) Patterns of hand motion during grasping and the influence of sensory guidance. J Neurosci 22:1426–1435.

Schieber MH (1991) Individuated finger movements of Rhesus monkeys: a means of quantifying the independence of the digits. J Neurophysiol 65:1381–1391.

Schieber MH (1995) Muscular production of individuated finger movements: the roles of extrinsic finger muscles. J Neurosci 15:284–297.

Soechting JF, Flanders M (1992) Organization of sequential typing movements. J Neurophysiol 67:1275–1290.

Soechting JF, Flanders M (1997) Flexibility and repeatability of finger movements during typing: analysis of multi-degree of freedom movements. J Comp Neurosci 41:29–46.

Soechting JF, Terzuolo CA (1987a) Organization of arm movements. Motion is segmented. Neuroscience 23:39–52.

Soechting JF, Terzuolo CA (1987b) Organization of arm movements in three-dimensional space. Wrist motion is piece-wise planar. Neuroscience 23:53–61.

Tanji J (2001) Sequential organization of multiple movements: involvement of cortical motor areas. Annu Rev Neurosci 24:631–651.

Terzuolo CA, Viviani P (1980) Determinants and characteristics of motor patterns used for typing. Neuroscience 5:1085–1103.

Tyrone ME, Kegl J, Poizner H (1999) Interarticular coordination in deaf signers with Parkinson's disease. Neuropsychologia 37:1271–1283.

Viviani P, Terzuolo C (1983) The organization of movement in handwriting and typing. In: Language production, Vol 2 (Butterworth B, ed), pp 103–146. London: Academic.