6 Instrumented Measures of Sign Production and Perception: Motion Capture, Movement Analysis, Eye-Tracking, and Reaction Times

Martha E. Tyrone

Chapter Overview

One of the central goals of linguistic research on sign language is to elucidate the mechanisms of language perception and production. Instrumented studies of sign language are designed to measure such mechanisms precisely and quantitatively. In many ways, instrumented research on sign language is analogous to acoustic phonetic research: those two subfields of linguistics take a physical, quantitative approach, which complements that of traditional descriptive phonetics. In addition to facilitating phonetic analyses of sign language, instrumented techniques are also useful for the development of systems for automated sign recognition and sign synthesis (which depend on perception and

Research Methods in Sign Language Studies: A Practical Guide, First Edition.
Edited by Eleni Orfanidou, Bencie Woll, and Gary Morgan.

production by machines rather than by human beings). Instrumented techniques and their associated measures can be applied to a broad range of research questions, but they are best suited to studies of adult sign language users. Particular techniques to be reviewed in this chapter include electronic motion capture, data glove systems, video-based motion analysis procedures, eye-tracking systems, and reaction-time paradigms.

Introduction

Linguistics as a field is concerned with discovering how human language (spoken or signed) is produced and perceived. While a variety of techniques are useful for examining theoretical issues in sign language research, experimental research on sign production and sign perception requires the use of instrumentation or automated data analysis. In speech research, instrumental techniques for capturing and analyzing production and perception are well established, but less is available for similar quantitative analyses of sign language.

Many of the techniques that linguists use for recording and transcribing speech cannot be easily applied to a language that uses no sound and has no widely used written form. In sign language research there are no universal conventions for phonetically transcribing sign productions, for representing the physical form of language, or for quantifying the physical correlates of sign structure. By contrast, spoken language research has a widely agreed upon transcription system (the International Phonetic Alphabet, IPA), conventions for representing the physics of speech (e.g., spectrograms and acoustic waveforms), and established physical correlates for the features of speech sounds (e.g., formant frequencies). Thus, describing sign production in precise terms is challenging, due to the comparative lack of specialized equipment, measurement techniques, and quantitative phonetic measurement units.

The limited availability of recording techniques and measurement schemata for sign language creates a challenge for sign language researchers. An additional challenge is the fact that the physical structure of sign language is fundamentally different from that of spoken language. Unlike speech, which uses an auditory–vocal production medium, sign language uses a visual–manual medium, the hands and arms being its primary articulators. Because the sign articulators are much larger than the speech articulators, sign language uses a large articulatory space than speech, which means that sign data must be captured from a broad region of space. In addition, because there are many sign articulators that can act independently of each other, data capture must allow for multiple streams of information to be recorded at once. This can be problematic, for example for automated sign recognition, in which a machine uses visual information to identify and translate signs during real-time production (see Vogler and Metaxas, 2004). Sign language uses a visual medium, which means that the capture system must be able to capture data that are four-dimensional, with dissociable x, y, and z spatial coordinates measured over time. (The one dimension along which sign may be easier to capture than speech is speed: because signs are produced more slowly than spoken words, it is possible to record signing at a much lower sampling rate; see Klima and Bellugi, 1979.)

These issues of data capture technology and sign structure are relevant to the study of sign perception as well as to that of sign production. The complex, multidimensional nature of sign language makes it difficult for researchers to record and present the type of naturalistic language stimuli that are used to assess language perception and processing. Language perception experiments typically involve presenting carefully controlled linguistic stimuli to language users and measuring these users' ability and aptitude for perceiving similarities and differences among the stimuli. For example, in speech research, listeners might be asked to identify the consonant that they hear in multiple recordings of CV syllables, in which acoustic aspects of the consonant vary only slightly (see Lisker and Abramson, 1964). In this way researchers can determine which factors the listeners use to categorize acoustically distinct speech sounds as a single phoneme. This type of study has been challenging for sign language researchers, because it is difficult to manipulate phonological parameters of signs with precision and without one parameter influencing another.

The purpose of instrumented research on sign language is to facilitate our understanding of the perception and production of language in the sign modality. In particular, instrumented techniques allow quantitative, objective measurements that can inform our understanding of sign language structure. For example, researchers have used instrumented techniques to examine co-articulation (Grosvald and Corina, 2012), emphatic stress (Wilbur, 1990), and verb agreement (Thompson, Emmorey, and Kluender, 2006) in the sign modality. The rest of this chapter will review specific techniques for recording and analyzing sign production and perception and will discuss some sign language studies that have used those techniques.

Motion Capture

One technique for collecting precise quantitative information about sign production is motion capture, which is distinct from standard video in that it collects three-dimensional information about movement at a fast sampling rate (~60–100Hz). It also differs from video in that it uses markers placed on the body in order to record the data. This has advantages and disadvantages, which will be outlined below. Motion capture is a technique that is primarily used in laboratory settings, although a few portable systems exist. Portable systems are useful for recording signers who are home-bound or have limited mobility, for example. In general lab-based systems are more reliable, because no measurement error is introduced by the reconfiguration of the cameras. Moreover, a lab-based setting allows for control over issues such as lighting. Motion capture recordings can be used to study a wide range of linguistic and paralinguistic phenomena, including the use of signing space, prosodic aspects of signing, differences in typical and atypical sign production, and the major phonological parameters of signs (handshape, movement, and location).

The markers that are attached to the body in a motion capture recording session send signals to some type of electronic device, which processes the information and relays the data to a personal computer. Various systems use different types of physical

Figure 6.1 Optotrak cameras.

signals (e.g., light, sound, magnetic fields) to track the positions of the markers. Optical systems use a set of cameras in conjunction with a set of light sources to track movement; for example, the Optotrak system (Northern Digital Inc.) uses a set of diodes that emit infrared light. The diodes are attached to participants' articulators (usually the hands), and a set of two or more cameras record pulses of light emitted by the diodes and compute the 3D coordinates of their locations over time (see Figure 6.1). Optical systems can be spatio-temporally very precise, with sampling rates of 100–750 Hz (depending on the software) and spatial resolutions of 0.1–0.15 mm. However, because they use an optical signal, they cannot record data when anything opaque comes between a diode and the cameras. Such situations include instances where the hand changes orientation, so that part of the hand itself comes between the diode and the cameras. This can be problematic for capturing sign production, because the hands change orientation often in the course of signing.

Other types of optical motion capture, such as the Vicon system, use passive markers placed on the body that reflect infrared light emitted by strobes on the camera units (see Figure 6.2 and Figure 6.3). The reflected light is received by the electronic cameras, and the locations of the markers are tracked over time. This type of system has comparable spatial resolution but lower temporal resolution (60–120 Hz), though this capture rate is sufficient for capturing limb and body movement during signing. Both types of optical motion capture systems have been used for research on speech (Ostry, Gribble, and Gracco, 1996), motor control (Lang and Bastian, 2002), and sign language (Brentari, Poizner, and Kegl, 1995; Poizner, Klima, and Bellugi, 1987; Mauk, 2003).

The motion capture systems that have been used most often for sign language research are optical. However, motion capture systems can instead be designed to

Figure 6.2 Vicon camera.

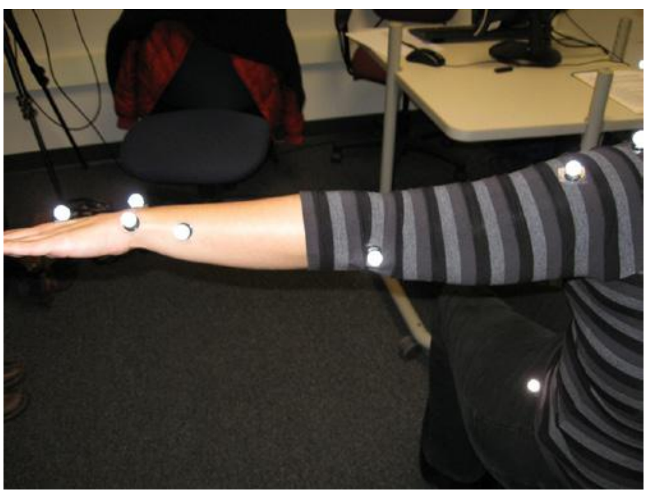

Figure 6.3 Reflective markers from a Vicon system.

detect changes in magnetic fields, ultrasonic sound waves, or rotational inertia and to use these signals to track motion (see Huenerfauth and Lu, 2010; Grosvald and Corina, 2012). These non-optical systems have an advantage over optical ones in that movement data are less likely to be occluded when the markers are oriented away from the device detecting the signal. Like the strobed optical systems, the non-optical systems are slightly less precise temporally (60–100 Hz), but they are nonetheless sufficient to capture most aspects of sign production. Only the most rapid fingerspelling is likely to pose a problem for motion capture recording.

Regarding the physics of the non-optical systems, magnetic motion capture systems employ a large magnet in combination with a group of small magnets: the small magnets are attached as markers on the body and, when they move through space, they disrupt the magnetic field generated by the large magnet, so that their positions can be tracked. Ultrasonic motion capture systems include a set of sound emitters that are placed on the body as markers; and a pressure-sensitive microphone tracks the ultrasonic waves emitted by the markers. Finally, inertial motion capture systems use electronic gyroscopes attached at different points on the body to detect movement rotation. This type of system is unable to track markers' absolute positions in space, but it is able to track markers' relative positions. So it is good for global, multi-articulator measures of signing – such as the timing and coordination of sequences of sign movements – but less well suited to measuring movements in relation to an external target.

As outlined above, with any of the motion capture systems, markers must be attached to a signer's articulators in order to record movements during signing. Strategic marker placement is one of the main challenges of experimentation with motion capture. Correct camera placement is important too, but this is simpler to implement, since cameras can remain in fixed positions from one experimental session to another, while markers are reapplied for each session. Light-emitting or -reflecting markers must be placed so that they are detectable by the capture system most (if not all) of the time. Marker occlusions are most problematic for hand and finger movements, since those movements are faster, smaller, and have many degrees of freedom. The best way to address this problem is either to use a large number of markers on those articulators or to use a data-glove system (see below). In addition, markers must be placed so that they do not interfere with the signer's movements. For example, for signs that require contact with the body, it is important not to place markers in a way that would block this contact. At the same time it is important to place markers so that they give information about where the articulators of interest are located. Thus, in order to measure where the hand is located during a particular sign, it is necessary to attach markers on the hand, but not on parts of the hand that will make contact with the body during signing. In addition, it is necessary to attach markers such that they will not change position in relation to the body part that is being tracked. So, for example, it is not feasible to use a marker on the forehead as an indicator of the chin's position, because the chin can move partially independently of the forehead. Finally, a single marker cannot indicate the orientation of a particular sign articulator. In order to determine orientation, the articulator must be defined as a three-dimensional rigid body, which requires placement of three markers that do not move relative to each other. Figure 6.4 illustrates the marker arrangement for a sign production experiment with a Vicon system. Note that three or more markers are used to represent individual sign articulators, such as the chin or the hand, and these markers remain fixed in relation to each other. (The color scheme in the figure has no special significance and is used simply to facilitate visualization for researchers during data-processing.) If markers are placed very precisely at specific joints or according to specific anatomical landmarks, then the movement data can be analyzed according to established biodynamic models for the movements of particular body parts. However, attaching markers directly on the finger joints can both impede movement and increase the likelihood of marker occlusion. Moreover, the more proximal joints of the arm are often not visible to the experimenter, which makes it impossible to place markers precisely in relation to those joints.

In this type of experiment the signer usually produces utterances from a script, which means that specific signs are produced in a specific order, as indicated by the

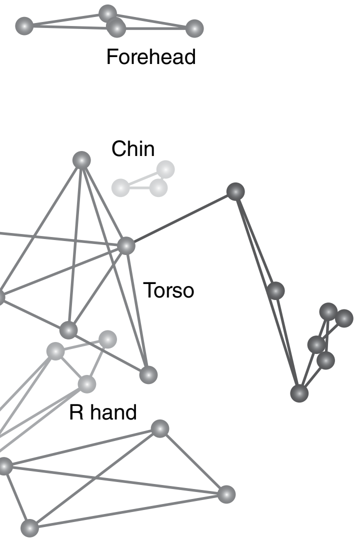

Figure 6.4 Marker schema for a Vicon experiment.

experimenter. There are two main reasons for this. First, motion capture systems collect many streams of data from multiple articulators, and the capture rate is high. As a result, it can be difficult to differentiate meaningful patterns in the data from random noise. On a related point, the other reason for the use of scripts is that motion capture lends itself well to comparisons of minute distinctions in sign production, such as the variation in a sign's position that results from the co-articulatory effects of other signs; hence it is useful to design one's experiments, in order to make sure that specific distinctions are elicited. In experiments that utilize spontaneous data, it is difficult to ensure that the relevant contrasts will emerge in the data set. In addition, it can take an enormous amount of time to identify and extract the data of interest from spontaneously generated productions. That said, studies that are more interested in automatic recognition or synthesis than in measurement are beginning to use motion capture to collect extended periods of signing (Jantunen, Burger, De Weerdt, Seilola, and Wainio, 2012).

Many of the earliest studies that used motion capture to analyze sign production investigated differences between typical signing and signing that was disrupted by neurological disorders such as stroke or Parkinson's disease. Poizner et al. (1987) were among the first sign language researchers to use motion capture. They carried out a series of studies that compared the productions of ASL signers who had aphasia, apraxia, or right hemisphere damage as a result of stroke. The goal of their research was to determine whether aphasia would take a different articulatory form from deficits in the production of meaningful gestures (apraxia) or from visuospatial deficits caused by right hemisphere damage. Poizner, Bellugi, and Klima (1990) extended this line of research to include signers with Parkinson's disease, a disorder that is primarily motoric rather than linguistic in nature. Similarly, Brentari et al. (1995) carried out an Optotrak study to compare an ASL signer with Parkinson's disease and an ASL signer with aphasia. Using motion capture data, they were able to show that the signer with Parkinson's disease preserved linguistic contrasts in production but showed a deficit in the coordination of handshape and movement. By contrast, the errors produced by the signer with aphasia were linguistic rather than motoric in nature.

Some early motion capture studies examined normal variation rather than comparing typical and atypical production. Wilbur (1990) used a WATSMART system to examine the realization of linguistic stress in the sign modality. WATSMART was one of the earliest motion capture system – its name is an acronym for Waterloo Spatial Motion Analysis and Recording Technique. Using this technique, she showed that native signers modified the duration of the movement transition prior to a stressed sign, whereas non-native signers increased the displacement of the sign itself. This distinction between native and non-native stress patterns would likely not be identifiable from descriptive analyses of video. In another study using WATSMART, Wilcox (1992) examined the production of ASL fingerspelling. His findings demonstrated that there was a large amount of co-articulation in fingerspelling and that features from an individual letter in a fingerspelling sequence would carry over into subsequent letters in the sequence. Moreover, he found that the transitions between letters were important for comprehension of fingerspelling in ASL.

More recent studies have applied motion capture methodologies to typical variation in sign production in order to study phenomena such as co-articulation, reduction, and articulatory undershoot and overshoot. Mauk (2003) used a Vicon system to examine articulatory undershoot of handshape and location in ASL (that is, situations when the articulators do not move far enough to achieve handshape or location targets). He found that undershoot occurred in both of these parameters as an effect of signing rate and phonetic environment. Similarly, Tyrone and Mauk (2012) collected a larger data sample using Optotrak in order to investigate phonetic reduction in the realization of location in ASL. Like the earlier study by Mauk (2003), theirs found that phonetic reduction in ASL occurred as an effect of factors that would be predicted from speech research. Their main result was that ASL signs with locations that are high in the signing space tended to be lowered at faster signing rates and when they were preceded or followed by a sign that was low in the signing space.

Grosvald and Corina (2012) used an ultrasonic motion capture system to examine linguistic and non-linguistic co-articulation in sign production, which they compared to co-articulation in acoustic speech data. They examined not only the effects of adjacent signs on the realization of location, but also the effects of signs that precede or follow the target sign at a distance of up to three intervening signs. They found that co-articulatory effects were weaker in the sign modality than in speech – in particular, distant speech segments had a stronger influence on vowel formants than distant signs had on sign location. In addition, they found that, in terms of co-articulation, linguistic co-articulation patterned more like non-linguistic co-articulation than like co-articulation in speech.

Data Glove Systems

A few studies have used electronic data glove systems to collect hand movement data during signing (Huenerfauth and Lu, 2010; Eccarius, Bour, and Scheidt, 2012). These systems have strain gauges (i.e., pressure-sensitive bendable strips), which are embedded in tight-fitting gloves that the signer wears during data collection. As the hand changes position, the strain gauges are bent and convert the mechanical energy from the hands' movements into an electrical signal transmitted to a computer. In this way researchers can measure degrees of flexion for the different joints in the hand.

In a recent study, Eccarius et al. (2012) measured the realization of ASL handshapes by using a set of cybergloves. The goal of their study was to define an articulatory joint space for ASL handshapes that would be analogous to the acoustic–articulatory vowel space from speech research. This was the first study to investigate the physiological limits of handshape structure and to try to determine the distribution of handshapes produced within those limits. Like the vowels in spoken language, the commonly occurring handshapes were maximally distinct from each other and were located near the outer limits of the articulatory space.

Huenerfauth and Lu (2010) used cybergloves in conjunction with an inertial motion capture system to develop a database of naturalistic ASL signing data. They put special emphasis on collecting productions of verbs that included spatial inflections, because it has been challenging for automated sign recognition and synthesis systems to deal with the variety of articulatory forms that occur with spatial verb inflection. Their database served as the basis for an automated sign synthesis program for ASL.

One of the limitations of data gloves is that the calibration process is lengthy and complicated. The gloves have to be calibrated separately for each participant. During this process, the signer is asked to hold a series of different hand configurations, so that the system can record the range of his/her movements and the approximate size and position of the hands' joints. In addition, probably the biggest limitation of data glove systems is that they are not only expensive but also relatively fragile, which means that the calibration has to be carried out slowly and carefully. If an experimental participant puts on or removes the gloves too quickly or too forcefully, the strain gauges can be irreparably damaged. In general it is difficult to collect instrumented data with children, but this is more true of data gloves than of any other type of instrumented movement recording.

In terms of phonological parameters, gloves can be used to measure handshape, while motion capture systems with separate markers are better suited to measuring location and movement. Either of these techniques can be used to examine phonetic correlates of phrase- or discourse-level phenomena such as emphatic stress or prosody. It should be noted that data glove systems vary in the number of motion sensors embedded in each glove, and only the systems with a large number of sensors are capable of distinguishing the full range of phonological handshape contrasts in a sign language.

Video-Based Movement Analysis

In addition to recording movement with a marker-based system, it is possible to record movement during signing by using ordinary video and applying a motion detection algorithm to the video data. The algorithm uses information about color and contrast in the video to detect edges and determine what is a moving figure versus a fixed background. In this way it is possible to track the articulators during sign production. Motion detection is advantageous, because standard video is widely available and video can be collected in a variety of settings, including Deaf[1] clubs or individuals' homes. By contrast, motion capture systems are less accessible and they most often have to be used in a laboratory setting, which is less naturalistic. Another important advantage of video motion detection is that it avoids the problem of markers that might constrain or otherwise affect movement.

Video-based motion detection is particularly useful for automatic sign recognition, in which a sign recognition algorithm searches for an approximate pattern of movement and compares it against a stored template of what a given sign should look like. However, motion detection is less useful for the type of precise measurements required for phonetic analysis. The temporal and spatial resolution of motion detection is only as good as the resolution of the video input itself. As a result, it is sufficient for differentiating signs from one another (i.e., for identifying phonological contrasts), but less good at capturing and quantifying phonetic variation that is non-contrastive but may reveal information about extralinguistic factors such as accent or language experience.

Vogler and Metaxas (2004) used video-based motion detection to develop an automatic recognition system for ASL. They used a movement-hold model to parse the data in a way that it could be processed by their recognition system. In a similar study, Karppa, Jantunen, Koskela, Laaksonen, and Viitaniemi (2011) collected videotaped signing data in Finnish Sign Language and used automated motion detection to analyze head and hand movement during signing. They demonstrated that information from these articulators could be extracted from video data, and resulting visual displays could be integrated into existing sign language annotation programs, such as ELAN. Both of these studies suggest that video data can successfully be used for automatic sign recognition.

Eye-Tracking

Eye-tracking is a technique that uses small cameras to measure eye movements and periods of gaze fixation. The cameras detect infrared light reflected off the cornea or pupil and use it to compute gaze direction. Cameras can be mounted on the head or placed in a fixed location in front of a research participant (see Figure 6.5). (Note that, if the eye tracker is not mounted on the head, then researchers must track head movement in addition to eye movement, or must require participants to hold their head in a fixed position.) This technique is useful for sign language research, because it allows researchers to determine what information a signer is attending to during sign perception or to analyze how eye movements are coordinated with other movements during sign production. Eye-tracking is unique among the techniques discussed here, in that it has been used both as a method for looking at sign perception and as a method for looking at sign production. In the realm of sign perception, eye-tracking can be used to discover what signers attend to while they are perceiving sign language. Further, it can be used to determine how attentional patterns differ with signing skill or language background.

Figure 6.5 A head-mounted eye-tracking system. With permission of William C. Schmidt, SR Research Limited, http://www.sr-research.com/EL_II.html

Muir and Richardson (2005) carried out an eye-tracking study with Deaf users of British Sign Language (BSL). The purpose of the study was to determine which regions of a signer's body carried the most important information for sign perception. Their longer term objective was to optimize standards for videophones, so that they would be well suited to sign language communication. The eye movements of Deaf volunteers were recorded as they watched BSL stories on videotape. To do this, the researchers measured the amount of fixation time at different anatomical regions such as the upper and lower face, the upper and lower body, and the hands. What they found was that signers spent the most time fixating on the face and that most signers looked preferentially at the upper face.

Emmorey, Thompson, and Colvin (2008) compared eye movements during sign perception in native Deaf signers and in hearing beginners. Both groups of signers fixated primarily on the face of the person who was signing. However, the two groups differed in that native signers fixated on the upper face, whereas beginners fixated on the lower face – specifically, the region around the mouth. Both groups made occasional saccades to the signer's hands, but these almost always co-occurred with the signer's fixation on her own hand. In other words, the signer and the interlocutor shared visual attention by fixating on the signer's hands.

Thompson et al. (2006) used an eye-tracking system to examine eye gaze as a component of ASL sign production rather than sign perception. They measured eye gaze during the production of agreeing, non-agreeing, and locative verbs. They found that signers looked toward locative objects and objects of agreeing verbs, but not toward objects of plain verbs. More recently, Thompson, Emmorey, and Kluender (2009) expanded this research to examine eye movements in native signers, non-native beginning signers, and non-native skilled signers during ASL verb production. In this later study they found that beginners did not show a consistent gaze pattern toward one sign type or another and skilled non-native signers showed a gaze pattern similar to that of native signers for locative and agreeing verbs, but not for plain verbs.

The most difficult aspect of using eye-tracking to measure sign production is likely to be the measurement interference created by hand and body movement. For example, a signer might bump into a head-mounted eye tracker while producing a sign located at the forehead. Similarly, a signer's hand might come between an external eye tracker and the eye, such that eye movement data are blocked from being captured. Head motion can also interfere with eye-tracking, if it causes an eye tracker to shift position on the head. This could arise from sign-related movement in a production experiment or from non-signing movement in a perception experiment.

Reaction-Time Studies

Reaction time refers to the time interval between when a stimulus is presented in an experiment and when a participant responds to that stimulus. This is a technique that is used to collect information about sign perception. There is a long history of using reaction times to analyze speech perception (see Studdert-Kennedy, Liberman, and Stevens, 1963), and the technique has also been adopted by researchers who investigate the perception of sign language. In a sign perception experiment participants are typically asked to make judgments about signs by responding to the signed stimuli as quickly as possible. Thus it is not only their responses that are recorded, but also the speed with which they respond. This is thought to reflect the ease or automaticity of the linguistic task. So, for example, if two stimuli are perceived as very similar, it would be expected that the participant takes longer to carry out the judgment task.

The main types of tasks used with reaction-time measures are discrimination tasks, in which a participant has to determine whether two sign tokens are the same or different, and lexical decision tasks, in which a participant has to determine whether a sign token is a real sign or a pseudo-sign. Information about a participant's reaction time is usually collected by means of a button-press on a computer; but the time it takes a participant to initiate a signed response can also be measured. Using this type of paradigm, Dye and Shih (2006) examined phonological priming in BSL. In each experimental trial, participants were shown two signs in sequence and asked to decide whether the second sign was a real sign in BSL. In some cases the first sign was phonologically similar to the second sign, in other cases it was not. In this way the researchers were able to assess the priming effects of the phonological parameters on participants' responses. They found that native signers showed shorter reaction times when the two signs were phonologically similar (specifically, in terms of movement and location). In other words, the phonologically related form primed participants' lexical decision process.

Reaction-time measurements are often interpreted to reflect the ease with which a participant can carry out a perception task; but one difficulty with this is that unusual or unnatural tasks can yield longer reaction times. Thus it is necessary to independently test tasks for naturalness before using reaction time as a measure of phonological processing, for example. Moreover, it is not informative to compare reaction times across studies or across experimental techniques, because different techniques can create variability in the reaction-time values (see Moreno, Stepp, and Turvey, 2011).

Discussion

Like all techniques for studying sign language, instrumentation has its advantages and its disadvantages. The main advantage of instrumented techniques is that they allow precise quantitative measurements. For sign perception, this means that lexical and phonological processing can be examined almost in real time. For sign production, it means that movement trajectories and endpoints for many productions of the same sign can be closely compared in three dimensions, allowing researchers to study, for example, grammatical versus non-grammatical uses of space. Precise measurement by instrumentation can ultimately lead the way to quantitative and objective analyses, which do not rely on language-specific descriptions.

One of the biggest limitations of instrumented measures of sign production and perception currently is that the amount of production and perception data from typical sign language users is very scarce. This is problematic because it impedes researchers' ability to draw inferences about differences related to signing background or language impairment. The data in these areas are so scarce that it is difficult to separate group differences from individual differences. On a related point, for techniques such as motion capture, different signers have different articulators and different anatomical proportions, so it may not be a valid procedure to pool data from different signers, because that might mask real differences or patterns in signing. To date, there have been no attempts to develop an idealized sign articulator set, analogous to the standardized brain tissue maps used by functional MRI studies. As instrumented sign language research evolves, it will be important to develop a reliable normalization procedure for comparing production data across signers. Finally, one point that researchers need to be aware of is the possibility that the equipment used for instrumented sign language research might interfere with normal production and perception processes. This has been a long-standing issue in speech production research, because placing measurement devices inside the mouth can disrupt normal speech movements. Needless to say, the precision that instrumented techniques permit is worth considerably less if they elicit atypical language patterns.

Instrumented measures of sign production are analogous to acoustic phonetics in speech. Descriptive measures are both useful and necessary in sign language research, but quantitative measures of production provide an analytical approach that is less dependent on subjective interpretation. Moreover, quantitative analyses allow researchers to develop measures that are language-independent. As long as sign phonetics relies solely on descriptive measures, it will never be possible to make truly informed cross-linguistic comparisons. Instrumented data capture and analysis can pave the way for universal measures of production and perception, akin to formant frequencies, voice onset time, or fundamental frequency for speech.

Directions for Future Research

Most recent instrumented studies of sign production have focused on phonetics and phonology (Eccarius et al., 2012; Tyrone and Mauk, 2010; Grosvald and Corina, 2012), or on automatic sign recognition (Vogler and Metaxas, 2004; Huenerfauth and Lu, 2010; Karppa et al., 2011.). Instrumented techniques can also be used to examine discourse-level linguistic phenomena such as turn-taking (by recording two signers simultaneously) or the marking of phrase and word boundaries. These phenomena, like phonetic variation, often involve subtle modifications to the timing or size of sign movements, which are most easily measured with instrumentation. Likewise, instrumented techniques can be used to gauge the size and variability of the signing space across individuals or across groups (see Mauk and Tyrone, 2012) – a necessary precursor to a full understanding of the grammatical use of signing space and of how it varies across groups.

Now that instrumented techniques are more widely available, research on sign language would benefit from more studies that combine multiple techniques, such as data gloves and motion capture, or eye-tracking and motion capture. Similarly, it would be useful if there were more direct comparisons of the effectiveness, precision, and reliability of different techniques (Karppa et al., 2011). On a related note, another area to be explored would be the links between the perception and the production of sign language. One could investigate how modifications to production affect perception. For example, does variability in location have a greater effect on perception than variability in handshape? Looking at production and perception together in a systematic way can reveal more about the structure of sign language than looking at the two only in isolation.

Note

1 In this chapter the capitalized form *Deaf* is used to refer to the cultural group of sign language users; the lowercase form *deaf* is used to refer to clinical hearing loss.

Keywords

kinematics; motion capture; sign perception; sign phonetics; sign production

See Also

Chapter 4; Chapter 5

Suggested Readings

Lu, P., and Huenerfauth, M. 2009. Accessible motion-capture glove calibration protocol for recording sign language data from Deaf subjects. In *Proceedings of the 11th international Association for Computing Machinery SIGACCESS conference on computers and accessibility (ASSETS 2009)*. New York: ACM, pp. 83–90.

References

Brentari, D., Poizner, H., and Kegl, J. (1995). Aphasic and Parkinsonian signing: Differences in phonological disruption. *Brain and Language* 48, 69–105.

Dye, M. W. G., and Shih, S.-I. (2006). Phonological priming in British Sign Language. In L. Goldstein, D. H. Whalen, and C. T. Best (Eds.), *Laboratory Phonology 8*. Berlin: Mouton de Gruyter, pp. 241–264.

Eccarius, P., Bour, R., and Scheidt, R. A. (2012). Dataglove measurement of joint angles in sign language handshapes. *Sign Language & Linguistics* 15 (1), 39–72.

Emmorey, K., Thompson, R., and Colvin, R. (2008). Eye gaze during comprehension of American Sign Language by native and beginning signers. *Journal of Deaf Studies and Deaf Education* 14(2), 237–243.

Grosvald, M., and Corina, D. P. (2012). Exploring the movement dynamics of manual and oral articulation: Evidence from coarticulation. *Laboratory Phonology* 3(1), 37–60.

Huenerfauth, M., and Lu, P. (2010). Eliciting spatial reference for a motion-capture corpus of American Sign Language discourse. In P. Dreuw, E. Efthimiou, T. Hanke, T. Johnston, G. Ruiz Martinéz, and A. Schembri (Eds.), *Proceedings of the fourth workshop on the representation and processing of signed languages: Corpora and sign language technologies*. Paris: ELRA, pp. 121–124.

Jantunen, T., Burger, B., De Weerdt, D., Seilola, I., and Wainio, T. (2012). Experiences from collecting motion capture data on continuous signing. In O. Crasborn, E. Efthimiou, E. Fotinea, T. Hanke, J. Kristoffersen, J. Mesch (Eds.), *Proceedings of the fifth workshop on the representation and processing of sign languages: Interactions between corpus and lexicon*. Paris: ELRA, pp. 75–82.

Karppa, M., Jantunen, T., Koskela, M., Laaksonen, J., and Viitaniemi, V. (2011). Method for visualisation and analysis of hand and head movements in sign language video. In C. Kirchhof, Z. Malisz, and P. Wagner (Eds.), *Proceedings of the second gesture and speech in interaction conference*. (GESPIN 2011), Bielefeld, Germany. Accessed September 12, 2014. http://coral2.spectrum.uni-bielefeld.de/gespin2011/final/Jantunen.pdf

Klima, E., and Bellugi, U. (1979). *The signs of language*. Cambridge, MA: Harvard University Press.

Lang, C. E., and Bastian, A. J. (2002). Cerebellar damage impairs automaticity of a recently practiced movement. *Journal of Neurophysiology* 87, 1336–1347.

Lisker, L., and Abramson, A. S. (1964). A cross-language study of voicing in initial stops: Acoustical measurements. *Word* 20(3), 384–422.

Mauk, C. E. (2003). Undershoot in two modalities: Evidence from fast speech and fast signing. PhD thesis, Department of Linguistics, University of Texas at Austin.

Mauk, C. E., and Tyrone, M. E. (2012). Location in ASL: Insights from phonetic variation. *Sign Language & Linguistics* 15(1), 128–146.

Moreno, M. A., Stepp, N., and Turvey, M. T. (2011). Whole body lexical decision. *Neuroscience Letters* 490, 126–129.

Muir, L. J., and Richardson, I. E. G. (2005). Perception of sign language and its application to visual communications for deaf people. *Journal of Deaf Studies and Deaf Education* 10(4), 390–401.

Ostry, D. J., Gribble, P. L., and Gracco, V. L. (1996). Coarticulation of jaw movements in speech production: Is sensitivity in speech kinematics centrally planned? *Journal of Neuroscience* 16(4), 1570–1579.

Poizner, H., Bellugi, U., and Klima, E. S. (1990). Biological foundations of language: Clues from sign language. *Annual Review of Neuroscience* 13, 283–307.

Poizner, H., Klima, E., and Bellugi, U. (1987). *What the hands reveal about the brain*. Cambridge, MA; MIT Press.

Studdert-Kennedy, M., Liberman, A. M., and Stevens, K. N. (1963). Reaction time to synthetic stop consonants and vowels at phoneme centers and at phoneme boundaries. *Journal of the Acoustical Society of America* 35, 1900.

Thompson, R., Emmorey, K., and Kluender, R. (2006). The relationship between eye gaze and verb agreement in American Sign Language: An eye-tracking study. *Natural Language and Linguistic Theory* 24(2), 571–604.

Thompson, R., Emmorey, K., and Kluender, R. (2009). Learning to look: The acquisition of eye-gaze agreement during the production of ASL verbs. *Bilingualism: Cognition and Language* 12(4), 393–409.
Tyrone, M. E., and Mauk, C. E. (2010). Sign lowering and phonetic reduction in American Sign Language. *Journal of Phonetics* 38, 317–328.
Tyrone, M. E., and Mauk, C. E. (2012). Phonetic reduction and variation in American Sign Language: A quantitative study of sign lowering. *Laboratory Phonology* 3, pp. 431–459.
Vogler, C., and Metaxas, D. (2004). Handshapes and movements: Multiple-channel ASL recognition. *Springer Lecture Notes in Artificial Intelligence* 2915, pp. 247–258.
Wilbur, R. B. (1990). An experimental investigation of stressed sign production. *International Journal of Sign Language* 1(1), 41–59.
Wilcox, S. (1992). *The phonetics of fingerspelling*. Philadelphia, PA: John Benjamins.