Fall 1-31-2007

A kinematic analysis of sign language

Chemuttaai C. Koech
New Jersey Institute of Technology

ABSTRACT

A KINEMATIC ANALYSIS OF SIGN LANGUAGE

Chemuttaai Koech

Signed languages develop among deaf populations and employ manual communication instead of voiced communication. *Stokoe* attributes classify individual signs in American Sign Language (ASL) and include handshape, hand location, movement, orientation, and facial expression. Signed and oral languages are not mutually understood, and many deaf individuals live in linguistic isolation. This research addresses computer translation between signing and speech, investigating sign duration in sentence context versus in isolation and identifying kinematic sign markers. To date, there has been little study of continuous signing kinematics; it was previously unknown if kinematic markers existed.

Kinematic data were collected from a proficient signer with electromagnetic Flock of Birds® sensors (position/ orientation of both wrists) and CyberGloves® (18 joint angles/ hand). The data were collected for each sign in isolation and in sentences.

Mean sign duration decreased in sentence context due to coarticulation. There was evidence of finger joint and wrist velocity coordination, synchronicity and hand preshaping. Angular velocity maxima and minima indicated differentiation between handshapes. Minima in the wrists' tangential velocity signified *Stokoe* locations, and maxima indicated movement (sign midpoints or transition midpoints), which can serve as anchors in the segmentation process. These segmented locations and movements can be combined with handshape and wrist orientation to identify likely signs based on the kinematic database developed at NJIT.

A KINEMATIC ANALYSIS OF SIGN LANGUAGE

by
Chemuttaai C. Koech

A Thesis
Submitted to the Faculty of
New Jersey Institute Technology
in Partial Fulfillment of the Requirements for the Degree of
Master of Science in Biomedical Engineering

Department of Biomedical Engineering

January 2007

APPROVAL PAGE

A KINEMATIC ANALYSIS OF SIGN LANGUAGE

Chemuttaai Koech

Dr. Richard Foulds, Thesis Advisor — Date
Associate Professor of Biomedical Engineering, NJIT

Dr. Sergei Adamovich, Committee Member — Date
Assistant Professor of Biomedical Engineering, NJIT

Dr. Bruno Mantilla, Committee Member — Date
Special Lecturer, NJIT

BIOGRAPHICAL SKETCH

Author: Chemuttaai C. Koech

Degree: Master of Science

Date: January 2007

Education:

- Master of Science, Biomedical Engineering
  New Jersey Institute of Technology, Newark, NJ, 2007

- Bachelor of Arts, Concentration in Mechanical Engineering and Assistive Technology
  Hampshire College, Amherst, MA, 2005

Major: Biomedical Engineering

Tamu ya mua kifundo.

-Swahili Saying-

"Sugarcane is sweetest at the joint."

meaning

Fruits of hard labor are enjoyed the most.

Dedicated to the entire Koech and Lang'at families

who believe education is invaluable.

ACKNOWLEDGMENT

I would like to express my deepest appreciation to Dr. Richard Foulds, who not only served as my advisor, providing valuable and countless resources and insight, but who also constantly gave me support, enthusiasm and encouragement. I would like to thank Dr. Sergei Adamovich and Dr. Bruno Mantilla (Rafiki) for their guidance, support and for actively participating in my committee.

I am grateful to the Pfeiffer Research Foundation for supporting this study, which was conducted at the Neuromuscular Engineering Laboratory, NJIT.

I would like to extend my gratitude to my peers in the Neuromuscular Engineering Laboratory (Kate, Sally, Diego, Qinyin, Darnell, Kai, Rico, Olga, Luz and Juan, Matt, Hamid and Nishant) whose academic insight, friendship and humor made this experience immense. I would also like to thank our Lab Coordinator, Amanda Irving for all her assistance and for her friendship.

My sincerest appreciation to all my teachers and friends throughout my years at Kabarak High School, Hampshire College, UMASS (Amherst) and NJIT for providing me with a strong background and a wholesome educational experience.

I wish to express my deepest gratitude to my family for their never-ending support. To my parents, Prof. and Dr. Koech for all the sacrifices they have made and for their encouragement. To my sister, Chela, who is the light of my life, and to my brother, Charlo, for giving me a reason to work hard. I am most grateful to my husband, Bena, who is a constant source of love and encouragement.

-All glory to GOD, who makes all things possible-

TABLE OF CONTENTS

TABLE OF CONTENTS
(Continued)

TABLE OF CONTENTS
(Continued)

Chapter **Page**

LIST OF TABLES

LIST OF FIGURES

LIST OF FIGURES
(Continued)

CHAPTER 1

INTRODUCTION

1.1 Objective

Sign language entails swift, dynamic and complicated movements of the upper extremities, specifically of the hands and fingers. Successful sign language recognition requires an understanding of the linguistics, spatial-temporal factors (visual spatial patterns with respect to time), and the biomechanics of sign movements. The core purpose of this thesis was to study the kinematic behavior of the wrists, hands and fingers of a proficient signer during signing. This was done in an attempt to identify the kinematic markers, if any, that denote the boundaries (beginning and end) of signs.

For many years, it was common belief that signs had no internal structure and were regarded as unanalyzable wholes (Wilcox, 1992). Previous research has attempted to segment and recognize signs based on vision perception strategies (such as eye tracking studies), pattern recognition, and linguistics. Vision based recognition systems use position trackers on the hands and cameras to record movement. These systems have the disadvantage of requiring large amounts of computation just to extract the hand position before performing any segmentation analysis (Monhandes et al. 2004). Segmentation based on linguistics involves investigation of the syntactic structure of the sign language, the syntax, semantics, and stress pattern in signs. Linguistics allows for the comparison of the intonation and rhythm mechanism in sign language, but it does not address the biomechanical principals of the signing movements (Wilbur, 2000).

The variables used in signing are of kinematic nature: displacement, duration, velocity, and acceleration. Not much research has been done on the kinematics of sign language and it is unknown whether there are any kinematic markers that distinguish sign boundaries and sign transitions (brief shift between adjacent signs). Many assumptions have been made regarding potential positions of sign boundaries, but the issues of whether such markers are present, their locations, and if they are of kinematic or linguistic nature, have not been resolved.

This thesis was a preliminary study of signs' kinematic nature and the characteristics of sign boundaries. The goal was to identify kinematic parameters present in signed sentences that suggest segmentation, with the aim of investigating the possibility of continuous sign recognition based on sign kinematics. This would enable signers to interact with non-signers in practical settings, such as the workplace and classrooms. So far, progress in the recognition of sign language as a whole has been limited.

Thus, the study focused on two analyses: a comparison of the duration of signs in isolation versus in sentence context, and a sign segmentation study based on kinematics. This investigation was compounded by the presence of coarticulation (kinematic blending of adjacent signs in the sentences); a common occurrence in fluent signing, which exacerbates the segmentation process.

The research hypothesis was that an analysis of the position and velocity profiles of the wrists and finger joints would show coordination, patterns and synchronicity that would assist in the segmentation of signs from signed stream.

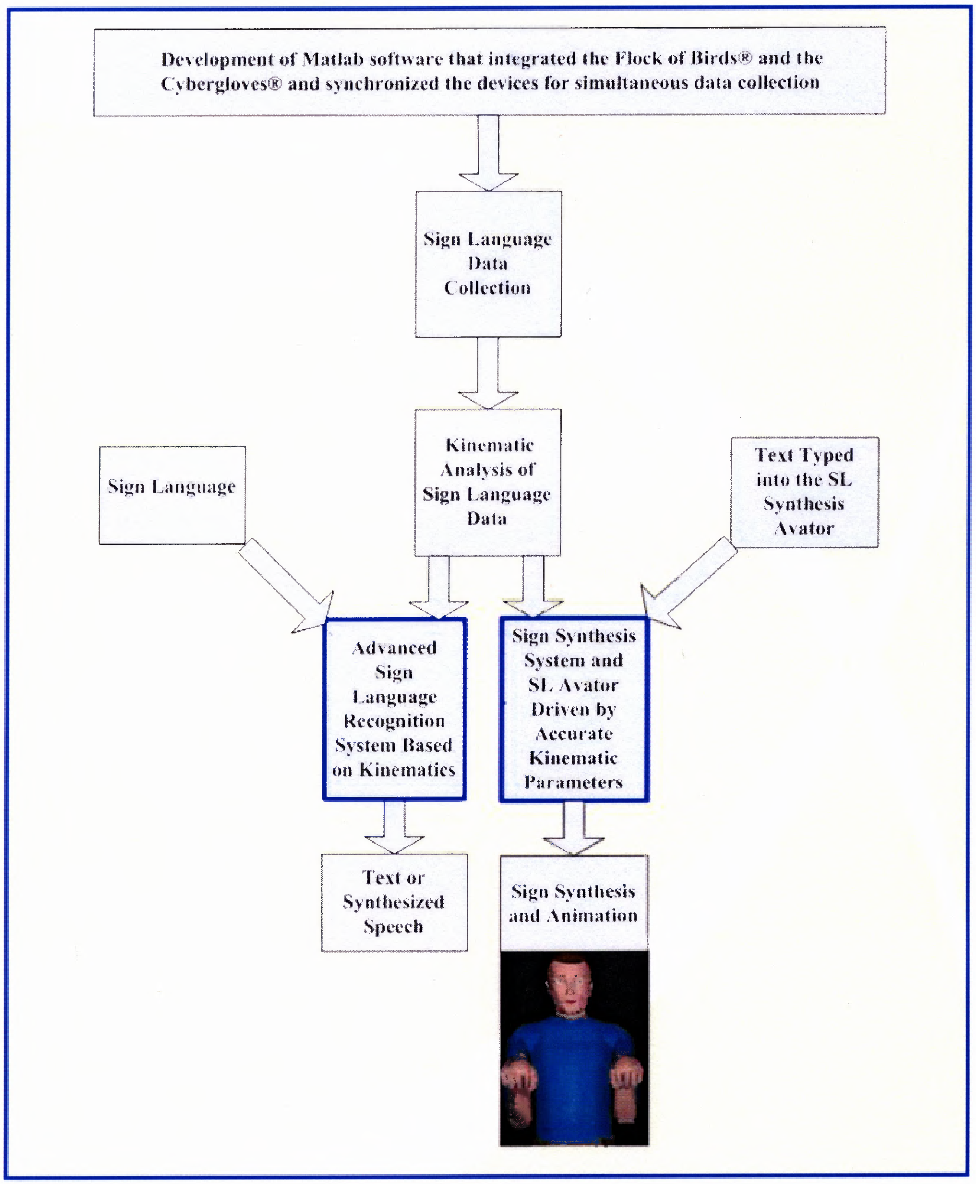

Figure 1.1 Basic structural design of kinematic analysis thesis.

1.2 Deafness and Sign Language

1.2.1 Deafness

Deafness is defined as the partial or complete loss of the ability to perceive auditory information, which could lead to a permanent disability. Deafness is classified into two main categories: sensori-neural and conductive hearing loss. Sensori-neural deafness, sometimes referred to as nerve deafness, is a permanent condition. It occurs when the auditory nerve fails to transmit the sound signal from the cochlea of the ear to the brain and is mostly caused by damage to the fluid-filled cochlea. This type of hearing loss is typically age-related and can increase with exposure to loud sounds, such as loud music, construction site noise, rifle shots, explosions and airbag deployment.

Conductive deafness could be temporary or it could deteriorate into permanent hearing loss. It occurs when sound vibrations are prevented from reaching the inner ear. This could be due to a wax build up in the ear canal, an infection that affects the ear drum, fluid build up in the middle ear or stiffening of the small, auditory ossicles (*incus*, *malleus*, and *stapes* bones which connect the cochlea and the ear drum) in the inner ear.

Deafness varies from mild to severe and can be the result of diverse events such as aging, diseases (for example, rubella and German measles), injuries, problems during pregnancy (due to alcohol and drug abuse, contraction of venereal diseases), and genetic dispositions. Deafness at birth is known as congenital deafness and an onset of hearing loss after birth is called adventitious deafness. More than one quarter of the people afflicted with hearing loss have adventitious deafness.

A person's ability to communicate is affected greatly by when in their life they lost their hearing. Deafness can create a problem for communication with people who do not know sign language and the deaf may encounter physical and social isolation due to their hearing impairment. Safety is also a critical concern for them since they are unlikely to be aware of warnings designed for hearing individuals such as sirens and alarms.

1.2.2 Sign Language

Philosophies on language evolution have influenced the communication means and the education of deaf children. It was believed that in order for one to successfully integrate into the world, which was predominantly hearing, the development of speech, listening (through hearing assistive technology), and lip-reading were the fundamental skills to acquire (Stewart & Akamatsu, 1988). This feat is very difficult for most deaf individuals who have severe hearing loss and as a result are at a disadvantage and are further impaired. Most do not have the capability to lip-read others and create intelligible sounds (Furth, 1966). In consequence, a visual mode (signing) became the preferred means of communication among many deaf people.

Throughout the world, deaf communities have developed visual, sign (or signed) languages. Sign languages utilize manual communication to convey information and combine hand shape, movement and orientation of the fingers, hands, arms, shoulders as well as facial expressions and gestures to express messages and thoughts fluently. Analogous to spoken language, sign language comprises phonology (the distribution and patterning of signs), semantics (relationship between signs and what they denote), syntax (patterns of formation of sentences and phrases from signs), pragmatics (connection between sentence context and interpretation) and morphology (sign structure). Signs offer more possibilities in their visual-gestural modality and have iconic ('pictorial') origins that are culturally determined and use visual space. This enables spatial mapping of persons and places in narrative for clarity of reference (Swisher, 1988).

Individual signs can be divided into five primary components (*Stokoe*, 1976). *Stokoe*'s parameters have been incorporated into various sign languages, including American Sign Language, and they form the categorical foundation of signs. They are:

1. **Handshape**- formation of the joints in the hand, excluding wrist angles
2. **Location**- position of the hands in space
3. **Orientation**- direction the palm is facing ('pointing' direction as if the handshape was a closed fist with the index finger extended
4. **Movement**- change in handshape or location during a sign (including linear, circular, arced and repeated trajectories)
5. **Facial expression**- any expression of the face used to describe or reinforce concept being conveyed

1.2.3 American Sign Language

In the United States, approximately 28 million Americans are reported to have severe to profound hearing loss. In response to the practical need for a communication system specialized for the American and Canadian deaf community, a signed language emerged with grammatical structure dissimilar to that of spoken language. American Sign Language (ASL) is that native signed language. It is a natural and autonomous language, used by more than half a million people in the US. It is object-oriented and therefore differs from both written and spoken English. ASL signs are also used in various visual communications methods, also known as Manually Coded English.

1.2.4 Manually Coded English

Manually Coded English (MCE) symbolizes the structure of the English language and consists of signs that are a visual code for spoken English. The various types of MCE were created artificially and predominantly observe English grammar and syntax. They are completely manual and incorporate less facial expression and body positioning. Forms of MCE can be used with simultaneous communication, since MCE is signed English language. In North America, the most common varieties of Manually Coded English include:

- **Signed Exact English (SEE)**
  This MCE adopts ASL signs and special supplementary signs or inflections that enable English to be manually reproduced in the same order that it is spoken. The supplementary signs include pronouns (examples: the words "he" or "she"), the verb "to be", possessions and plurals.

- **Fingerspelling**
  Each of the 26 letters in the English alphabet has its own respective sign in ASL. Fingerspelling represents the written form of a spoken language and is integrated into ASL where English words are spelled out one letter at a time. Fingerspelling consists of about 8.7% of casual signing in ASL (Patterson & Morton, 2003). This manual alphabet is typically used when words have no sign equivalent, for emphasis, clarification, or during teaching or learning a sign language. Fingerspelling is not a very efficient method of communicating since it is time-consuming, with a typical rate of 3-4 letters per second for fluent signers.

- **Signed English**
  Signed English is a much simpler MCE system than SEE and was originally used for teaching young children. ASL signs are used in English word order with fewer grammatical markers. It is common to speak English as one simultaneously signs in Signed English.

1.3 Assistive Technologies for the Deaf

Various hearing assistive technologies are available for the deaf, including hearing aids and cochlear implants. Hearing aids are electronic, battery-operated devices that amplify and alter sound for improved hearing. Cochlear implants are surgically implanted devices designed to provide some measure of hearing to the individual with profound neural-sensural hearing loss. Hearing aids require that the person has some degree of hearing, while cochlear implants have been shown to be successful with most hearing loss types, including children with some or no hearing, and adults who have been deaf all their lives. Although cochlear implants do not restore normal hearing they can provide a useful representation of sounds in the environment and help the deaf person understand speech.

Assistive technologies available for communication include digital mobile devices such as cell phones (for example the *Sidekick*) and pagers, widely used for their short message service (sms). These text messages are typed and sent between mobile phone and other handheld devices. Tele-typewriters (TTYs) are devices that allow deaf persons, hard of hearing, and speech impaired individuals to communicate with each other and with hearing people through the telephone system. TTYs transmit and receive signals that are converted into text and printed to the TTY screen. Some TTYs are connected to computers through modems and the non-hearing person can either call the hearing person via the computer or through a Relay Service (service that allows hearing callers to communicate with TTY users and vice versa through specially trained personnel). There are other variations of relay services: Internet Protocol (IP) Relay (accessible with an Internet-capable computer and specially trained personnel who relay messages between the caller and receiver) and Video Relay Service (caller signs to a video interpreter through a computer, who speaks to the voice user and the video interpreter signs the voice user's communication back to the signer). Even though Relay Services are very important for communication between the non-hearing and the hearing, they have a few drawbacks. There is a lack of privacy during the conversation and the caller has no control over what information is relayed by the Relay personnel, and in what manner it is relayed.

1.3.1 Need for Assistive Technology for Communication

Communication is the process of transmitting and receiving information, and is commonly done verbally. In that way, the profoundly deaf, deaf-blind, and speech-impaired people are segregated from the hearing and speaking people, incapable of communicating easily with the mainstream world. This can be partially overcome by replacing spoken word with text or sign, or by using a system that converts from one mode of communication to another.

Currently, the universal method of communication between the deaf and the hearing is through interpreter. Sign language interpreting is the translation of a spoken language into sign language and vice versa. This process relies on the presence of a sign language interpreter for information to be relayed and received, which infringes upon conversation confidentiality. In a classroom situation, a stenographer performs a job similar to an interpreter, typing words into a portable computer that simultaneously appear on a computer screen provided to the student.

The most successful communication technology available for the deaf is computer-human interface systems. These systems integrate computer technology with sign language and facilitate communication between hearing people and signers.

An example of a successful computer-human interface system is the sign language synthesis system developed by J. Allen, A. Irving and R. Foulds at the Neuromuscular Rehabilitation Laboratory in the Biomedical Engineering Department, NJIT. This sign synthesis system allows relatively easy creation of new signs through the manipulation of a commercial avatar that is available in UGS's Jack software (Irving and Foulds, 2005). Jack is an ergonomics and human factors software product that enables users to position biomechanically-accurate digital humans of various sizes in virtual environments. At Jack's core is the powerful JackScript Toolbox which is written in Python computer language, which allows access to movement control of the avatar (Phillips and Badler, 1988). A sign editor has been developed, that utilizes a graphical user interface to aid in the formation of signs. Currently, an initial electronic collection of 5000 signs (defined in Random House Webster's Concise American Sign Language Dictionary) is being created. The sign synthesis allows for coding of signs with kinematic parameters that can be used to define the animation of signed languages. It also acts as a supplement to human interpreting services that are often in limited availability (Irving and Foulds, 2005). The system consists of a sign editor that utilizes a graphical user interface (GUI) providing visual and text feedback to the user. The sign editor saves the information as an ASCII text file, storing handshapes, end-effectors, target locations and orientations, special via points, and the English, gloss and definition the sign. These files are immediately available to control the avatar. A set of hand configurations has been set up for ASL as well as fingerspelling. Data from the ASCII file are used to calculate fundamental positions for each sign and the path of the sign is constrained by end-effector locations and orientations. Primary movements to change position and orientation are calculated and attained via Jack's inverse kinematics. Secondary movements involve time-dependent handshape transitions. Text is entered real-time through a text file, the keyboard or a speech recognizer and is compared with the available words in the dictionary. If the word is found, the respective file is executed and if it is not, it is fingerspelled. Coarticulation between signs is achieved by the use of the final end-effector's position and orientation of the previous sign as the starting positions of the subsequent sign.

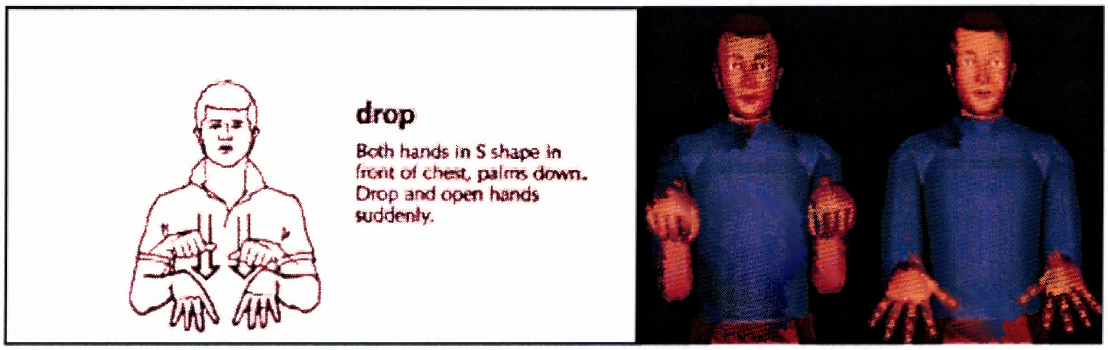

Figure 1.2 The sign DROP and still images of Jack signing DROP.
(Sources: Gallaudet Survival Guide to Signing and Irving & Foulds, 2005)

The analysis of ASL signs and signed sentences done in this thesis will provide essential kinematic information that will be used to improve the avatar's coarticulation and increase the intelligibility of the animated signs ensuring that messages appear more human-like.

This system allows for the translation of text/ speech into real time animated sign, but not for the reverse process. There still exists great need for technology that is efficient at recognizing and translating sign language back into text/ speech. The systems available either cannot accurately recognize signs from signed stream, are video or camera based, are not real time, or require an interpreter.

1.3.2 Current Status of Machine Sign Language Recognition

Sign recognition can be divided into two categories: static sign recognition and continuous sign recognition. In view of the fact that signing involves a series of smooth connected movements, and is not static in nature, continuous sign recognition has more practical applications. Sign recognition has to adeptly segment the signed stream in order to accommodate for the coarticulation problem. The baseline system for continuous sign recognition consists of data collection from gesture input devices, followed by segmentation, feature extraction, and a matching module. The matching module compares and matches signs with those in a sign database/ dictionary.

Continuous sign recognition research has been done by Murakami & Taguchi (1991), who used a Dataglove system and recurrent neural networks. Neural networks are a computing paradigm modeled after cortical structures of the brain. Murakami & Taguchi carried out static sign recognition and a continuous sign recognition experiment, in which they were able to recognize 10 different signs in Japanese Sign Language, with user dependent accuracies up to 96% in constrained situations.

Ma et al (2000) designed a multi-modal based dialog system using speech, sign language and 3-dimensional virtual technologies. Their integrated system had a database of 220 words and 80 sentences (2-15 words each). For continuous sign recognition, their correct word rate was 91.4% for 200 sentences. They found a few errors with the speech and sign recognition, which they said could be corrected by repeating the wrongly recognized sentence.

An automatic Australian Sign Language recognition system using 2-dimensional rotations, signing speed and Hidden Markov Models (HMM) was developed (Holden et al. 2005). An HMM is a statistical model that establishes the hidden or unknown parameters from observable parameters and is mostly used for pattern recognition. The recognition system tracked multiple target objects, on the face and hands, throughout an image sequence and extracted important features. Experiments were conducted using 163 test sign phrases each with varying grammatical formations. The system achieved over 97% recognition rate on a sentence level and 99% success rate at a word level. One constraint was that sign representation dealt mainly with global motion in space, and provided limited information on local motion of the hands. This made it difficult to differentiate signs with the same trajectory. The use of a prediction tracking algorithm based on spatio-temporal velocity did not handle direction changes of the hands very well, and caused discontinued velocity at these points. The segmentation algorithm was only tested with foreground and did not deal with background, which is important in occlusions of the hands where both foreground and background are used in the sign.

The limitations of the three mentioned recognition systems are that they fail to integrate atypical sign procedure and are still restricted by coarticulation. Temporal segmentation and interpersonal variance are additional difficulties that have to be accounted for. This thesis examines kinematics as the key to sign segmentation, contingent on *Stokoe*'s parameters of location, movement, handshape and orientation. This is done with the intention of creating an advanced sign recognition system, unaffected by interpersonal variance that accurately distinguishes individual ASL or MCE signs from a continuous stream of signs.

1.3.3 Significance of Thesis

This thesis will enhance the sign language synthesis system developed by J. Allen, A. Irving and R. Foulds, NJIT. The analysis of ASL signs done in this thesis will provide vital kinematic information that will be used to improve the avatar's coarticulation and increase the intelligibility and appearance of the animated signs.

In addition, through the application of kinematics as the key to sign segmentation, this thesis suggests to the broader scientific community a way forward in the creation of an advanced sign recognition system. The thesis documents the effectiveness of the kinematic approach taken in segmenting individual ASL signs from a continuous stream of signs while controlling for coarticulation, based on movements, timing and location. The thesis also provides further evidence in support of *Stokoe*'s parameters of location, movement, handshape and orientation.

CHAPTER 2

KINEMATICS OF UPPER EXTREMITY MOVEMENT

2.1 Coordination of Movement

Movement is the constant change in the position of a body relative to a reference point and it plays an elementary role in the existence of humans. Human movement is made up of simple and complex, static and dynamic motions and can be regarded as a network of coupled oscillators operating in multiple degrees of freedom. Various parts of the body play important functions in the physiology of human motion, including the brain, the nervous system, muscles, tendons and bones. Coordination of movement occurs because of synchronization and desynchronisation of the actuators (muscles), which shorten and lengthen to provide torque at the joints and move the bones. The cerebellum in the brain coordinates opposing muscle groups resulting in smooth and refined movements. Motor coordination control is the fundamental element of the movement of multiple body segments. Although the relationship between neural activity, muscular activity and movement is ambiguous, a theory known as coordination dynamics addresses the coordination of human movement.

2.1.1 Coordination Dynamics

Coordination dynamics endeavors to discover the broad pattern formations and self-organization in human movements. This is done using the dynamical system theory and dynamics concepts to gain insight in interlimb coordination and limb movement coordination with the environment. The concept of a dynamical system, originating from

Newtonian mechanics, is the affiliation with the state of the system only a short time into the future. Particular concepts of dynamics are very important in studies of inter and intralimb human movements, and they include coordination patterns, stability, timing and symmetry.

> *"The new science of coordination is called coordination dynamics. The basic idea is that the coordination of life is an emergent, self-organized process in which all the many and diverse parts of the system come together and form coherent patterns of behavior. Coordination is dynamic: it evolves and changes in space and in time. These patterns of coordination are always meaningful patterns, but they are transient and fleeting, never staying still for too long, making them difficult to comprehend."* (Kelso & Engstrøm. Coordination Dynamics, excerpt from The Complementary Nature, 2006)

A study by Jeka and Kelso (1992) demonstrated the significance of symmetry as a conceptual tool in the differentiation of coordination between components with similar and different anatomical properties. The results showed that the addition of weight to the arm or leg minimized and enhanced coordinative asymmetry, respectively, and that the response to a perturbation slowed as movement frequency increased, replicating the underlying coordinative asymmetry. These coordinative effects suggest the influence of the central nervous system and assert the important role symmetry plays in the comprehension of coordination in systems.

This thesis anticipated results similar to Jeka and Kelso's, demonstrating the role that coordination dynamics plays in sign language, through symmetry and asymmetry of the articulators, as well as the coordination of the fingers and wrists during signing.

2.1.2 Reaching and Grasping

Reaching and grasping is comparable to signing because the many joints in the hand have to coordinate and position accurately into a task-specific or goal-oriented action. Reaching and grasping of an item also entails positioning of several potential degrees of freedom depending on the size, shape and location of the object. There is a progressive opening of the hand and straightening of the fingers, followed by gradual closure of the grip until it matches the target size (Figure 2.1.a). This is a skilful action taking into account the body segments (arm, hand and fingers) utilized and the different hand shape configurations, such as the precision grip and power grip. Coordination of both hand shape (grasp) and arm movement (arm and/or trunk transport) conveys temporal constraints in the form of relative timing or phasing relationship, which is characteristic of various multi-articulator actions (Kelso et al. 1994). This temporal coupling theory is also suggested by Gentilucci et al. (1992), who argued that the two components (grasp and arm transport) are not independent of each other and are coordinated by a time-dependent mechanism to preserve optimal performance.

Phase, peak velocity, peak aperture, preshaping, timing and coordination patterns are the common parameters that are investigated in reaching and grasping studies. In the present study, most of these parameters were examined in the kinematic analysis of sign language.

In the reaching and grasping studies, finger angles and finger velocity profiles show preshaping of the hand, whereas the peak velocities exhibit the speed of movement. Coordination patterns symbolize synchronicity of the movement, while timing represents the control of the movement. Peak aperture is studied in relation to the size, shape and manipulation of objects. The maximum peak aperture occurs within 60-70% of the time duration (msecs) of the reach and correlates to the size of the target (Figure 2.1.b.). Figure 2.1.c.demonstrates the difference in grip aperture with a change in object width.

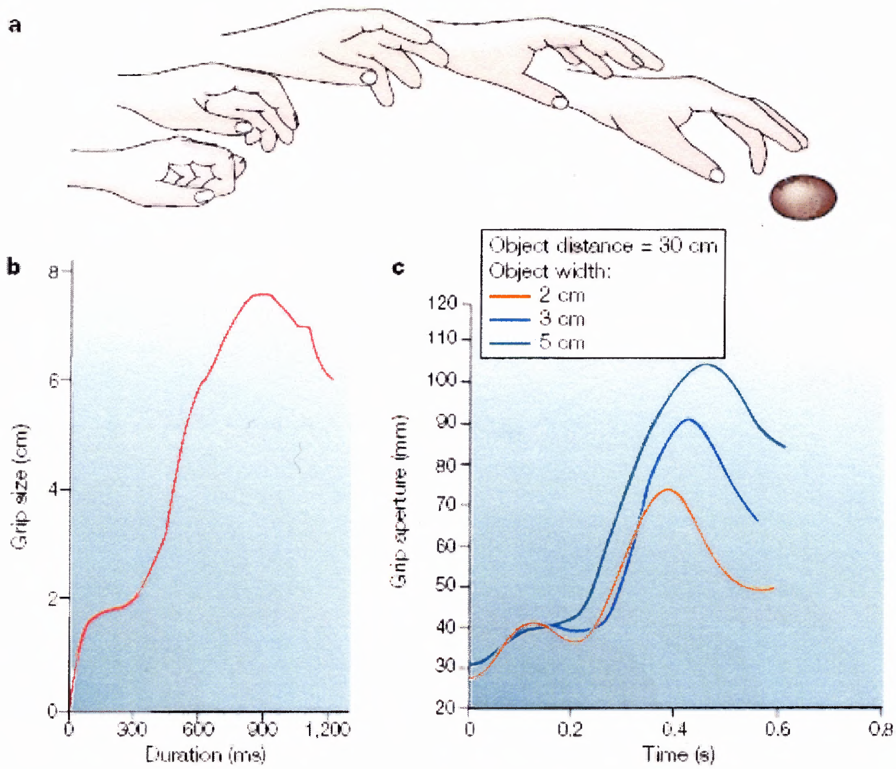

Figure 2.1 Kinematics of grasping.
a. Hand preshaping during its journey to the target.
b. Maximum grip aperture occurring within 70% of movement completion.
c. Traces demonstrating the scaling of maximum grip aperture with respect to object size.
(Source: Castiello, 2005)

In a study of coordination of the body segments involved during reach-to-grasp movements, the temporal and spatial kinematic parameters related to grasping were invariant regardless of whether the had was transported to the target by the arm, trunk or both (Wang and Stelmach, 1998). The tasks were achieved with differing coordination patterns across conditions, while keeping the peak aperture, time to peak aperture and closing distance of the hand constant. This important feature of preshaping showed that the hands' opening distance systematically increased as the hand-transport distance increased to reach the target. The authors hypothesized that the opening/ closing distances increasing as a function of movement time reflected that the spatial features of aperture formation are dependent on the temporal mechanism.

A study by Soetching and Flanders (1997) investigated the behavior of the fingers and joints during a typing task, and the results demonstrated a high degree of temporal coordination across the joints and fingers. This showed synergistic movement involving the many degrees of freedom of the hand more than individuation of finger motion. On the other hand, instances of simultaneous assimilation and dissimilation argue against synergistic control. In their studies of grasping, they looked into the components that could explain the discrepancy in the postures and movements of the joints of the hand (Santello, Soetching and Flanders, 1998, 2002). The authors proposed that there are small higher-order principal components that contributed information about the object to be grasped. Therefore, they suggested that even if there is a dominant tendency for coordination of motion of all fingers, there is a superposed ability for individuated control. This would explain the coordinated and individuated control of the fingers during typing and signing.

2.1.3 Coarticulation

Coarticulation of signs is the assimilation of the point of articulation of one sign to that of an adjacent sign. It is done in such a way that the movements needed for adjacent signs are produced almost simultaneously. Soetching et al (2003) designed a study that quantified coarticulation in the hand movement sequences of sign language interpreters while fingerspelling. Evidence of coarticulation was found with display of both forward and reverse influences across letters. The authors categorized these influences into assimilation (tendency to reduce the differences between sequential hand shapes) or dissimilation (tendency to emphasize the differences between sequential hand shapes). The proximal interphalangeal (PIP) joints of the index and middle fingers mostly demonstrated dissimilation, while the joints of the wrist and thumb showed assimilation during the spelling of the same letters. The dissimilation that occurred may have assisted in improving visual discrimination since the PIP joints of the index and middle fingers are known to be among the most important joints for computer recognition of the 26 letters in fingerspelling.

One objective of this thesis was to demonstrate the coarticulation between adjacent signs, and more importantly, to find out if the coarticulation can be kinematically identified when the signs are produced in sentence context. For an analysis of coarticulation, the kinematic data from individual signs and those signs in sentences was compared. Similar to the study by Soetching, the tendency for assimilation or dissimilation between adjacent signs was checked.

Exploring the theory of coordination dynamics and the kinematic parameters of reaching and grasping movements impacted the formulation of this research. Important concepts including preshaping, coordination and synchronicity of the articulators, coarticulation, and duration and velocity analyses of movements, developed into the basis of this kinematic analysis of sign language. This chapter looked into human movement and coordination as a means to understand the biomechanics of signing. The next chapter analyzed the structure of sign language in order to explore segmentation of individual signs from signed stream.

CHAPTER 3

LANGUAGE SEGMENTATION

3.1 Internal Structure of Signed Languages

The phonological structure of signed language for the deaf has been shown to have an internal structure, similar to the words of spoken languages. This conception was brought forward by in the publication of the classic monograph, "Sign Language Structure" (*Stokoe*, 1960 & 1965). *Stokoe* established that signs must be described in terms of three parameters or sublexical units (*chereme*): *dez* (hand configuration), *tab* (location), and *sig* (movement). Researchers later suggested that *Stokoe* was illustrating the phonological structure of ASL and renamed them phonemes. A fourth parameter, orientation, was brought forward by Battison in 1973.

Additional phonological research on ASL has been introduced on sign structure indicating that signs are analyzable sequentially (Liddell 1984, 1990; Wilbur, 1990). This was implied by Liddell, who argued that signs are sequentially segmentable on the basis of movement sequences. The segment types are divided into two broad categories: movements (M) and holds (H). M refers to the motion of the hand/s along a path and H to the brief moment (50-100 msecs) when the hand/s remain/s stationary. Liddell refers to the *Stokoe* parameter of location as a hold.

An example Liddell used is the sign THINK, which traditionally was viewed as a simple sign consisting of a single handshape (articulator), a single location for contact (place of articulation), and one motion (articulation).

Liddell claimed:

> *"THINK requires two activities to be carried out in sequence. First, the hand must move towards the forehead (movement). Second, it must come to a brief stop (hold). The motion without the stop is not sufficient for the sign to be well-formed; but these two activities cannot possibly be regarded as simultaneous."* (Liddell, 1984, p.372)

A different autosegmental model of ASL was proposed by Sandler (1986) in which the segment types are movements and locations, with handshape on an independent autosegmental course. This model provides better clarification for the handshape spreading that occurs in ASL signs. Additionally, she suggested that holds may occur phonetically (list rhythm), phonologically (at utterance boundaries), morphologically (aspectual inflection on ASL verbs insert holds), and pragmatically (at the ends of conversational turns).

However, this is contrary to the incidence of holds in the underlying structure of ASL. Wilbur (1982 &1987 & 1990) argued for a model of ASL phonology that included syllables and later came up with a model that incorporated Liddell's holds and movements segments. Wilbur clarified the relation between location (holds and target position) and movement, and she specified that path movement is actually a change in location, thereby defining path movement as the dynamic correlate of location and likewise for handshape, and orientation.

Table 3.1 Phonologically Related Static, Full dynamic, and Reduced Dynamic Forms

Static	**Full Dynamic**	**Reduced Dynamic**
Handshape	Handshape change	Flutter/ repeat bend
Location	Location change	Tremored contact
Orientation	Orientation change	Tremor (wrist/ elbow)

(Source: Wilbur, 1987)

3.2 Sign Language Segmentation

There is not much literature available on the segmentation of sign language, and in particular on the location of sign boundaries in American Sign Language. Visual information has been shown to provide significant sign boundary characteristics of either the grammatical structure of ASL or as phonological parameters of the language. These visual cues could originate from the hand shape of a sign, its location, the lexical movement of a sign, facial expression, eye blinks, and head nods. All these cues are known to be significant to the grammatical structure of ASL (*Stokoe* et al. 1965, 1976; Battison, 1978; Liddell, 1977; Baker and Padden, 1978).

Green (1984) suggested that an analysis of the transitions (movement connecting contiguous signs) between signs could assist in the marking of the end of one sign and the beginning of the next. Beginning and end holds were also proposed as marks of sign boundaries (Newkirk, 1977; Klima and Bellugi, 1979).

Green conducted a research experiment to investigate sign boundaries in American Sign Language. The research addressed the following questions, "Do deaf signers agree on the location of sign boundaries in American Sign Language and where in time are the boundaries located?" The study was carried out with six deaf subjects, all of whom were fluent at ASL. Thirty five videotaped ASL sign sequences were edited out of sentence contexts and were slowed down to a tenth the normal speed. The videotape of sign sequences had been used in an earlier sign recognition study (Clark and Grosjean, 1982). The sentences of signs were constructed such that each of the test signs promptly followed the sign for SISTER (SISTER was chosen because it provides a well defined anchor point: both hands move together in front of the body, creating a clearly visible contact). The subjects' role was to observe the tapes and determine the beginning and end of a test sign in the sign sequences. The deaf subjects judged the sign sequences first, and thereafter the hearing subjects judged. All the subjects agreed on the location of sign boundaries and relied on changes in facial expression, hand configuration, or manual movement differences as visual cues for the boundaries.

Using the analysis videotape, the experimenter and an independent judge scrutinized each sign and ascertained whether a visual cue occurred within that window. The experimenter and observer agreed 91% of the time. The deaf observers responded with about equal consistency for the beginnings and endings of the signs. The mean beginning point for each sign occurred before the start of the transition movement, immediately following the contact break (end of the contact). This did not depend on the length of the sign and was true even for signs with relatively short transition movement, such as TRAIN. Likewise, the end mean value for each sign appeared before the start of the transition movement of the sign or at the very beginning (first 50-75 msecs) of motion.

These results suggest that the 'transition movement' between two signs is considered by deaf signers to be part of the second sign. Green proposed that the start of the transition movement could be one cue used by deaf observers when judging sign boundaries. According to the results, the deaf subjects primarily used two types of visual cues to determine the sign's beginning: the contact break of the previous sign's end and facial expression of the signer. Each test sign began after the contact break at the end of SISTER and it is possible that the subjects relied on this predictable cue for their beginning judgments. Their judgments might have differed had the experiment consisted of variable test-sign context. The results from the end boundaries showed that all the proposed cues were used to some amount in the location of the end-boundaries, with movement change, facial expression change, and hand shape change being the most critical.

A second study was carried out by Green (1984) hypothesizing that the deaf subjects were using their linguistic knowledge of ASL to identify the sign boundaries. To assess this interference, he conducted an experiment in which six hearing observer-judges performed the same task that had been performed by the deaf ASL signing subjects. They did not know any sign language and therefore had no linguistic knowledge of any sign language that could influence their findings. There was a fairly good agreement on the judgment values amongst the hearing subjects. However, the boundary locations selected by the hearing subjects differed from those chosen by the deaf subjects. The beginning boundaries occurred much later, usually in the middle or towards the end of the transition movement. The hearing subjects also chose different visual cues, such as the point where the hand shape was first well-formed, beginning holds, and lexical contact.

One interesting incident was that even though the contact break of SISTER (the end of the sign before the test sign) was pointed out to the subjects as a reference point, the hearing observers still did not consider that contact break as important in identifying the beginning of the subsequent sign. The contact break was the highest ranked cue for the hearing subjects for the end of an event.

In conclusion, the results from the second study indicate that the deaf subjects, in some part, utilized their linguistic and grammatical experience with ASL in the location of the sign boundaries. The deaf subjects also proposed the evaluation that a sign begins almost immediately after the sign preceding it has ended or that sign boundaries overlap in time, just as the acoustic information for words in speech.

This is an important finding that applies to this current research and in the analysis of the segmentation of signs this thesis studied the overlapping of signs and the functional role of transitions between signs and their occurrence.

3.3 Speech Segmentation

The acoustic waveform of speech is continuous in nature, yet humans can process a language and identify the boundaries between words, syllables or phonemes so as to hear individual, distinct words and afterward comprehend the sentence. Speech segmentation occurs on a phonetic and a lexical level, and is mostly of linguistic characteristic.

Sign language segmentation is also considered to occur on a lexical level and this is a similarity that could relate segmentation characteristics of speech with sign. The segmentation of sign language is further supported by the theory that speech evolved from manual gestures and hand movements. Gentilucci et al (2006) argue that spoken language developed from gestures and not vocalizations because manual actions provide a more prominent iconic link to objects and actions in the physical world. It is well acknowledged that signs are essentially different from gestures; nonetheless, there is an analog element of sign suggesting a link to a more iconic mode of communication. In the course of development, conventionalisation (the transition of iconic gesture to arbitrary symbols over time) may have occurred. Pantomimes of actions, referring to the telling of a story without words, may have incorporated gestures that are analog representations of objects and actions.

However, the merging of bordering words and sounds poses a serious problem in the recognition of speech. The co-articulation of speech sounds blend and modify the adjacent sounds and this can occur between words or within a word. Isolation and identification of individual words is very difficult, and context, grammar and semantics of the words must be considered. This problem might arise in sign language as well, with the merging of contiguous signs in sentences, unless the boundaries of signs have kinematic markers. Coarticulation was expected to be the principal problem in this research with the segmentation process from signed stream.

CHAPTER 4

METHODOLOGY

4.1 Overview of Sign Language Experiment

The primary goals of this research were a) to contrast sign duration in isolation and in sentence context, b) to carry out a spectral analysis of the frequency bandwidth of sign language and c) to identify the kinematic markers of sign boundaries.

Matlab computer programs were developed for integration and synchronization of the input devices and for data collection, filtering and analysis. Kinematic data were simultaneously collected from two electromagnetic Flock of Birds® sensors (position and orientation data) and two CyberGloves® (joint angle data). The data were collected from a proficient signer for a list of isolated signs and signed sentences. Wrist tangential and angular velocity and finger joint angular velocity were derived from filtered position and angular data and computed with a central difference program.

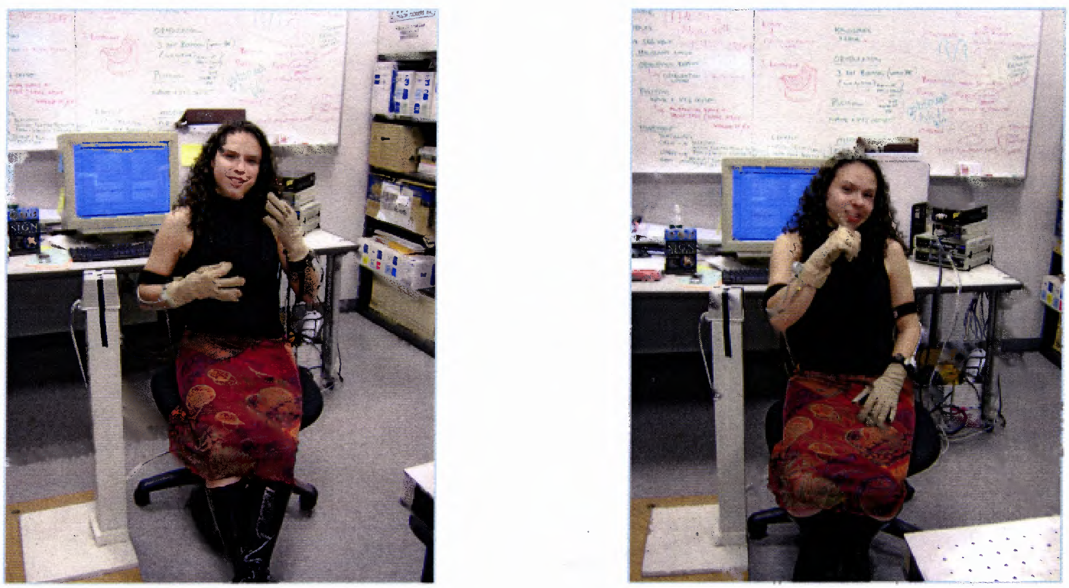

Figure 4.1 Experimental setup showing signer wearing input devices (Flock of Birds® sensors and CyberGloves®) and signing JUSTICE and GIRL.

The first analysis of the data included a statistical analysis of the duration of signs in isolation and in sentence context. Additionally, the tangential velocity profiles of the wrists and the angular velocity profiles of the finger joints were analyzed in relation to the position and orientation graphs, with respect to time. A video of the data collection was used as a reference for the actions of the signer.

4.2 Experimental Task

A member of the laboratory staff who is a proficient sign language interpreter, assisted in the generation and collection of the data. She is female, right-handed, with normal hearing, and not a native ASL signer. During data collection the signer was presented with a visual representation of a sign or sentence. She was asked to sign at a comfortable, moderate similar to normal conversational speed. Each trial began with a beep sound and a visual prompt that alerted the signer to start. The signer began each trial tapping her lap twice with both her hands, and ended with her hands returning to her lap. In between, the signer signed the appropriate sign/s.

Ten declarative sentences were signed four times each. ASL signs were used with the grammatical structure of Manually Coded English because of MCE's declarative nature in signed sentences. The ten sentences consisted of at least five ASL signs each and were divided into two groups:

- **Sentence 1-5**: Had no repeated signs
- **Sentence 6-10**: Shared one sign with corresponding sentence from 1-5 (repeated)

Table 4.1 a Signed Sentences (Non-Repeated Signs) and their English Translations

SIGNED SENTENCES (ASL Signs with MCE grammatical structure)	**ENGLISH TRANSLATION**
1. THAT MAN HIS DAUGHTER GO SCHOOL.	That man's daughter goes to school.
2. HUSBAND WIFE THEY HAPPY WITH BABY.	The husband and wife are happy with their baby.
3. YOUR PANTS BLACK SHIRT GREEN.	Your pants are black and your shirt is green.
4. EVERY NIGHT MY FRIEND DRIVE GYM.	Every night my friend drives to the gym.
5. MY GRANDMOTHER THIRSTY SHE NEEDS WATER.	My mother is thirsty she needs water.

Table 4.1 b Signed Sentences (with Repeated Signs) and their English Translations

SIGNED SENTENCES (ASL Signs with MCE grammatical structure)	**ENGLISH TRANSLATION**
6. MONDAY MORNING I GO DENTIST.	On Monday morning I am going to the dentist.
7. MY SISTER HER BABY HAS MANY TOYS.	My sister's baby has many toys.
8. WOMAN DRINKING HOT BLACK COFFEE.	The woman is drinking hot black coffee.
9. MY SPANISH TEACHER DRIVES CLASS.	My Spanish teacher drives to class.
10. I HUNGRY I WANT PIZZA AND WATER.	I am hungry I want pizza and water.

Furthermore, the ten sentences were broken up into the 46 individual signs that comprised them and each sign was recorded in isolation. There were three data collection trials for each sign, following the same experimental procedure as mentioned above for the ten signed sentences.

4.3 Calibration

Calibration of the measurement devices was necessary prior to the collection of data due to the range of variability in hand sizes and range of motion of the fingers.

The CyberGloves® were calibrated using Immersion Technology's *Device Configuration Utility* (DCU) software. Figure 4.2, demonstrates the two-stage calibration procedure, in which the hand was held still in two positions, and the DCU created and saved a unique calibration configuration file based on specific hand structure.

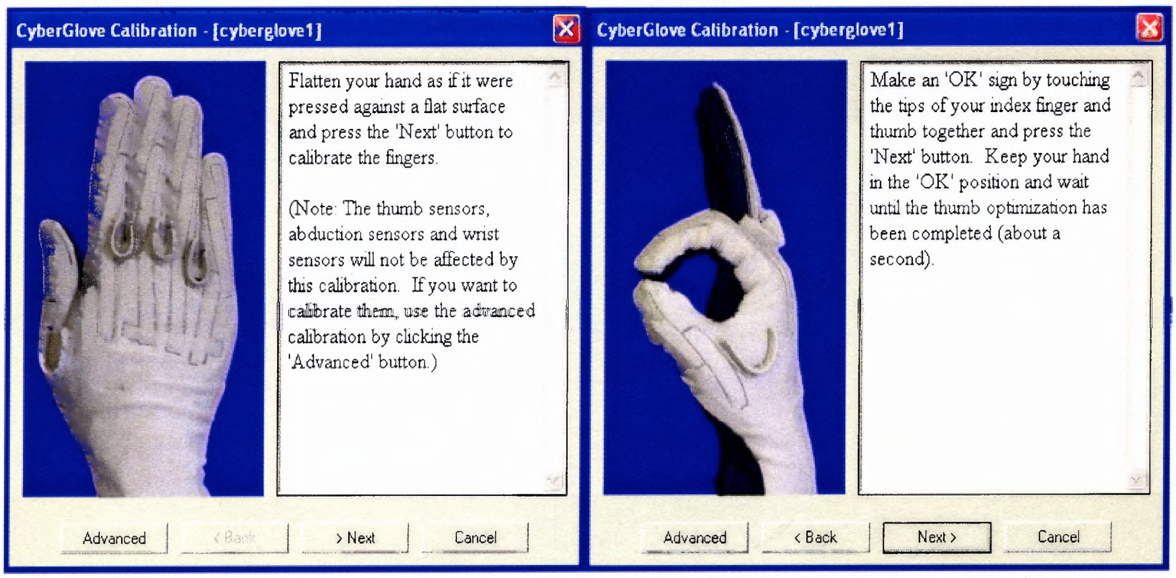

Figure 4.2 Calibration procedures for the CyberGloves®.
(Source: Virtual Technologies, 1998)

The following equation converted raw analog to digital sensor value into appropriate joint angles, where A/D is the analog to digital CG sensor value:

$$Angle = Gain * (A/D - Offset)$$

Gain affected the range of motion of the joint angles and the offset refers to the difference between the analog to digital values and the default hand-geometry position. An advanced calibration allowed the user to manipulate the gain and offset parameters for individual sensors.

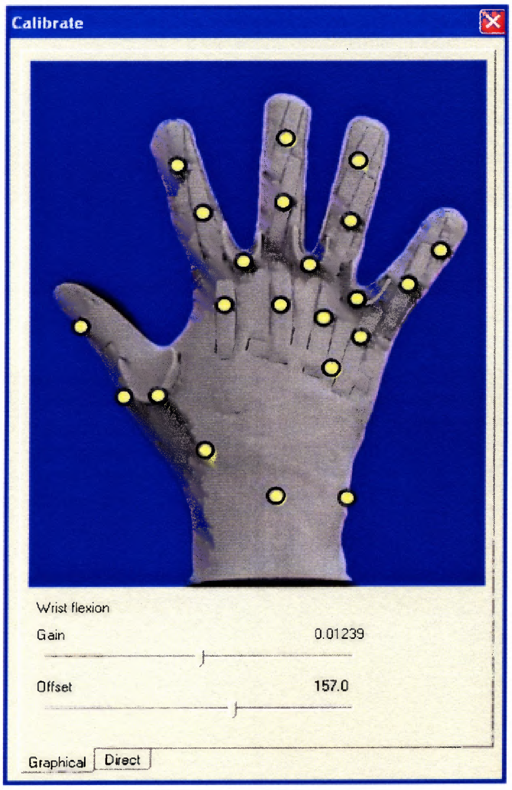

Figure 4.3 Advanced calibration of gain and offset.
(Source: Virtual Technologies, 1998)

Accuracy of the Flock of Birds system was checked by running the *WinBird* software from Ascension Technology and comparing the position and angle values against data collected from our Matlab programs. Physical measurements using a tape measure were also taken and confirmed against the collected data.

4.4 Data Collection

4.4.1 Instrumentation

In addition to the Matlab programs, the two devices that were used in the sign language data collection were two Flock of Birds® (FOB) sensors and two CyberGloves®.

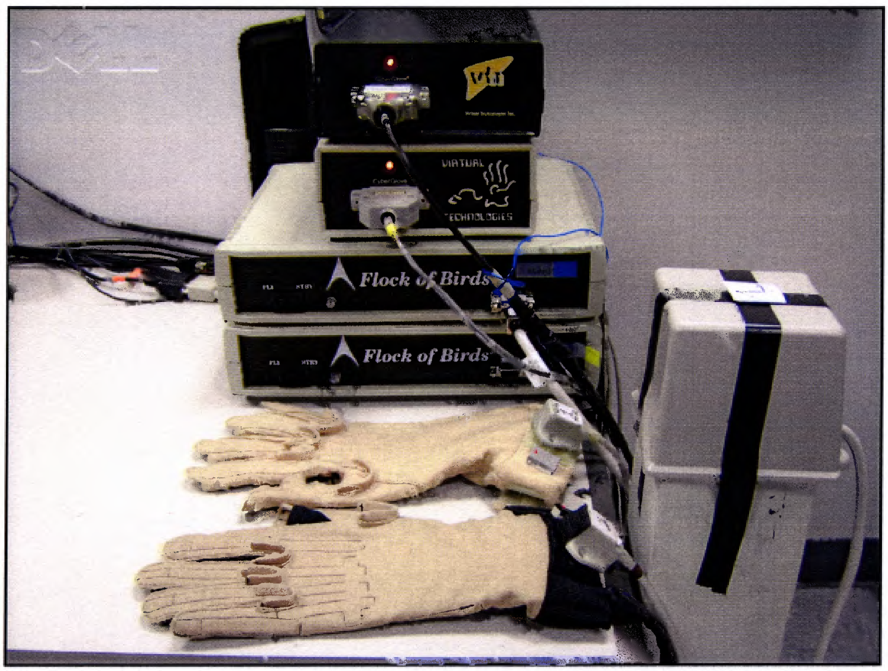

Figure 4.4 Flock of Birds® sensors and transmitter and CyberGlove® system.

Ascension Technology's Flock of Birds® (FOB) is a six degrees-of-freedom measuring device that can be configured to simultaneously track position and orientation. This system is used for various purposes including 3-dimensional graphics control, biomechanical/ human factors analyses, instrument-body tracking and rehabilitative feedback and assessment. The FOB consists of a transmitter that sends a pulsed DC magnetic field that is simultaneously measured by sensors. Two sensors were strategically placed on the subjects' wrists. The FOB tracked the position and orientation of two sensors via an RS-232 interface to the host computer and a Fast Bird Bus (FBB) cable between them.

The CyberGlove® (CG) system by Immersion Corporation (formerly Virtual Technology Inc.) consists of one right-handed and one left-handed glove. The CG provides a high-accuracy output proportional to the angle between the bones via the embedded piezoelectric bend and abduction sensors. The 18 flexible bend sensors measure the angles of the metacarpophalangeal (MP) and interphalangeal (IP) joints, while the abduction sensors measure the amount of corresponding finger movement to the lateral plane of the palm. There are also two wrist sensors that measure wrist pitch and wrist yaw.

Joint angle data collected from the CyberGloves® consisted of digitized output ranging from 0 to 255 (counts) which could be converted into angular units (degrees or radians). Using counts as the unit of bend angle instead of converting into angular units ensured that computed angular velocities were not distorted. The conversion of counts into degrees involves the approximation of joint angle measurements in relation to the initial conditions set during calibration. As a result of this approximation, there could be inaccuracies that would be amplified in the velocity computation. For that reason, counts were used as an alternative to degree conversion, removing the conversion ambiguity.

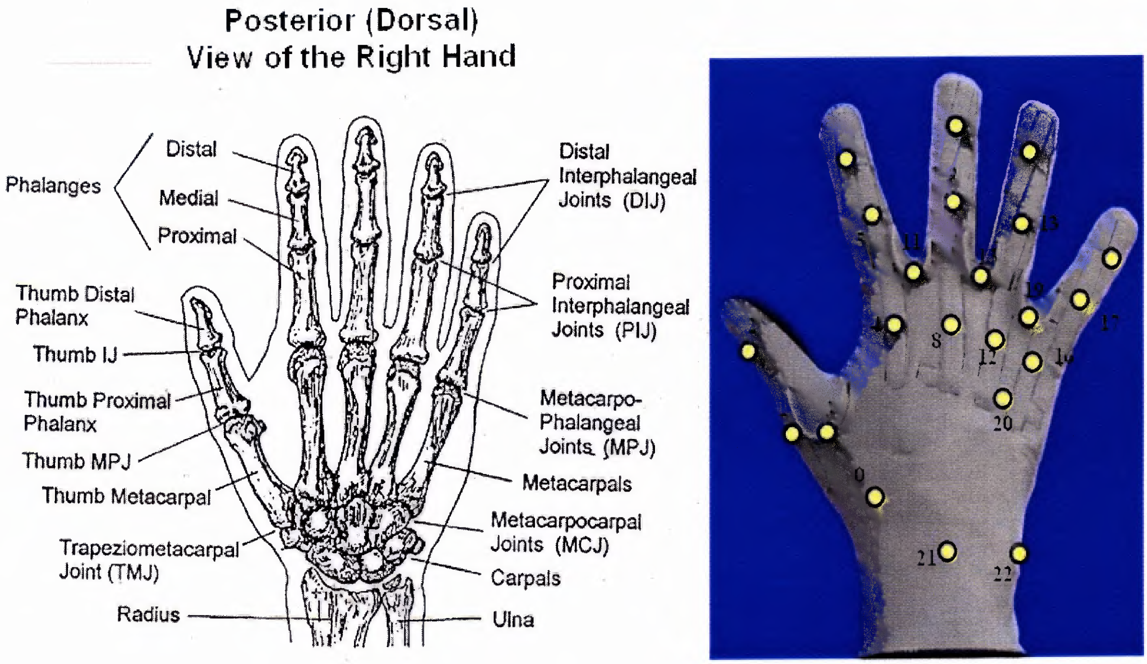

Figure 4.5 Posterior view of the hand (Left) displaying the finger joints and CyberGlove® with sensor locations (Right).
(Sources: Kapit W. & Elson L. M, 1993 and Virtual Technologies, 1998)

Table 4.2 CyberGlove® Sensor Descriptions

CG Sensor	Finger and Joint Description
0	THUMB- Rotation
1	THUMB- Metacarpo- Phalangeal Joint
2	THUMB- Interphalangeal Joint
3	THUMB- Abduction
4	INDEX- Metacarpo- Phalangeal Joint
5	INDEX- Proximal Interphalangeal Joint
8	MIDDLE- Metacarpo- Phalangeal Joint
9	MIDDLE – Proximal Interphalangeal Joint
11	MIDDLE-INDEX- Abduction
12	RING- Metacarpo- Phalangeal Joint
13	RING- Proximal Interphalangeal Joint
15	RING-MIDDLE- Abduction
16	PINKIE- Metacarpo- Phalangeal Joint
17	PINKIE- Proximal Interphalangeal Joint
19	PINKIE-RING- Abduction
20	PALM- Arch
21	WRIST- Pitch
22	WRIST-Yaw

4.4.2 Software

Data were collected via computer programs that had been created to synchronize and integrate the input devices. Matlab programs were developed at the NJIT Neuromuscular Engineering Laboratory. The programs incorporated and synchronized the Flock of Birds® and CyberGloves® for data collection, analysis and graphing purposes.

A GUI (Graphical User interface) was created to assist in data collection utilizing the developed Matlab programs. Push button control allowed for convenient data collection and a beep sound (in program) indicated the start and stop of each 10 second trial.

The developed Matlab programs included:

1. **Data Collection**

   - Opened serial ports (FOB, CG1, and CG2)
   - Set and checked all device measurement rates for 100 Hertz
   - Simultaneously collected data for 10 seconds from FOB sensors, CG1 and CG2
   - Saved position and angle data and closes serial ports

2. **Unpack**- Converted binary to decimal (units in millimeter for FOB and counts for CG) and plotted position (FOB) and joint angles (CG) versus time

3. **Filter**- 5th order Butterworth filter (6 Hz f_c) increased signal to noise ratio and smoothed the data

4. **Velocity**- Tangential velocity (FOB) and angular velocity (CG) were computed and plotted using the central difference formula

5. **Image Acquisition**- Acquisition of video data of the signed trials from web cam collected at 10 Hertz

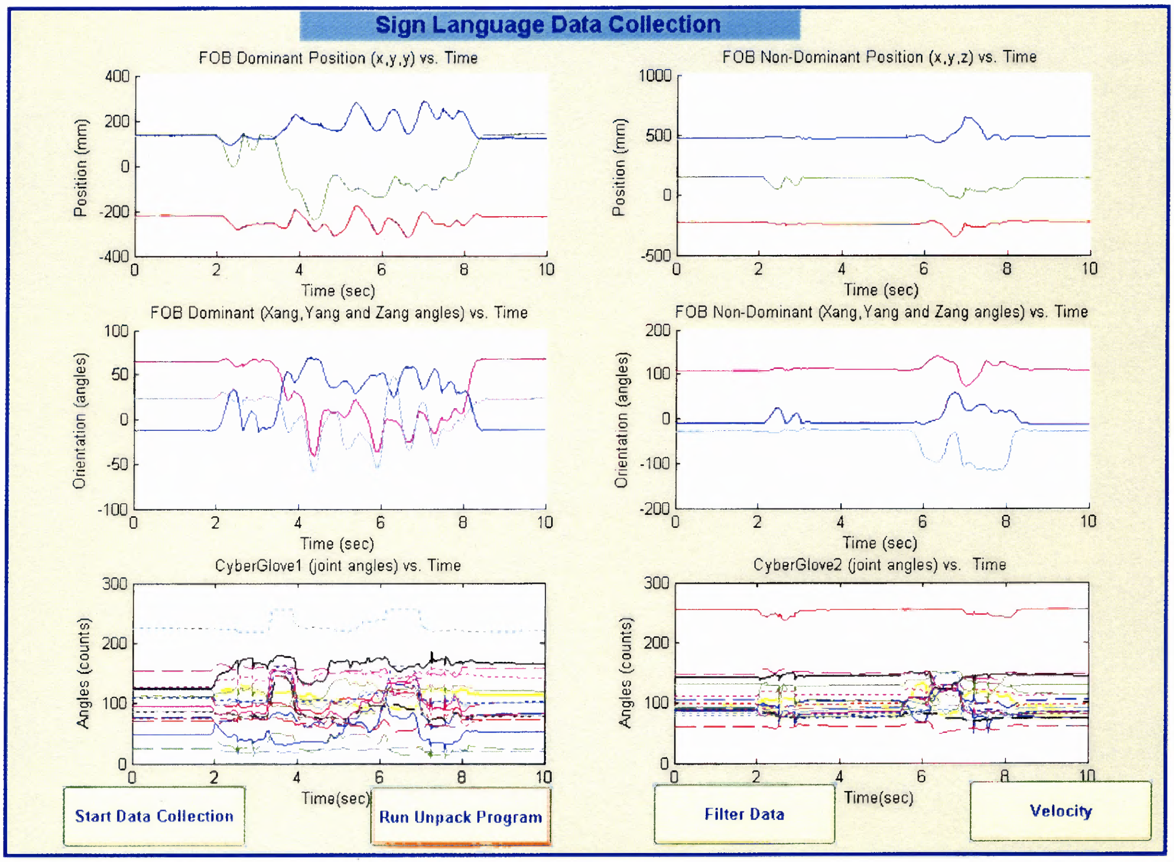

Figure 4.6 Data collection and analysis Matlab GUI.

4.5 Data Processing

Prior to the analysis of the raw data, some data pre-processing was required. A smoothing operation was required to increase the signal to noise ratio. Data from human movement experiments could contain additive noise from many sources, such as vibration, high frequencies, environmental interferences (60 Hz electrical lines), metal interference (FOB electromagnetic sensors) and error (human and measurement). Therefore, it is essential for the raw data to be smoothened and yet retain the useful information.

4.5.1 Filtering the Data

In biomechanical experiments, the most important frequencies are much lower than the sampling rate (f_s). Normal human movements have maximum frequencies in the range of 5 to 20 Hz. The raw, unprocessed sign language data were filtered with a 5th order Butterworth filter, at a cut-off frequency (f_c) of 6 Hz. Winter (2005) suggests a low pass filter with f_c of 6 Hz for biomechanical data to attenuate all unwanted high frequencies. The 6 Hz f_c also followed Nyquist's Sampling Theorem as it is less than half of the sampling frequency. Signals above 49 Hz would be unreliable and would likely be the result of aliasing (distortion due to high frequency signals). A 5th order, zero-lag filter smoothed the data and canceled the phase lead and lag that can occur in digital filtering.

Butterworth filters are digital, recursive filters that provide a superior linear representation of amplitudes of low frequencies. Matlab's function ***butter*** was used to compute the Butterworth coefficients; the filter order and the variable proportional to $2f_c/f_s$.

4.5.2 Velocity Derivation

Velocity was computed by using a central difference algorithm, which is a numerical differentiation of displacement (angular and X, Y and Z position) over time. Assuming the data are spaced equally and *i* represents the sample, the central difference equation is:

$$\text{Central Difference} = (\text{Position}_{(i+1)} - \text{Position}_{(i-1)}) / (\text{Time}_{(i+1)} - \text{Time}_{(i-1)})$$

This calculation of velocity represents the velocity at a point in time midway between two adjacent samples.

The X, Y and Z directional velocities of the FOB sensors were combined into one resultant known as the tangential velocity ($velocity_{tan}$). The tangential velocity combines three dimensions into one with direction tangent to the path of the sensor. It is computed by taking the square root of the sum of the squares:

$$velocity_{tan} = \sqrt{((velocity_X)^2 + (velocity_Y)^2 + (velocity_Z)^2)}$$

The units of the $velocity_{tan}$ of the FOB sensors were in millimeters per second and the angular velocities of the finger joints were in counts per second.

CHAPTER 5

DATA ANALYSIS

5.1 Velocity Analysis of Reaching and Grasping

A preliminary experiment of the tangential velocity of the wrists and the angular velocity of the finger joints was carried out for a reaching and grasping task. A right-handed, male was asked to reach and grasp an object (a cell phone) that was located approximately 15 inches from his hand and then return his hand to the starting point. This task was performed repeatedly over a period of ten seconds.

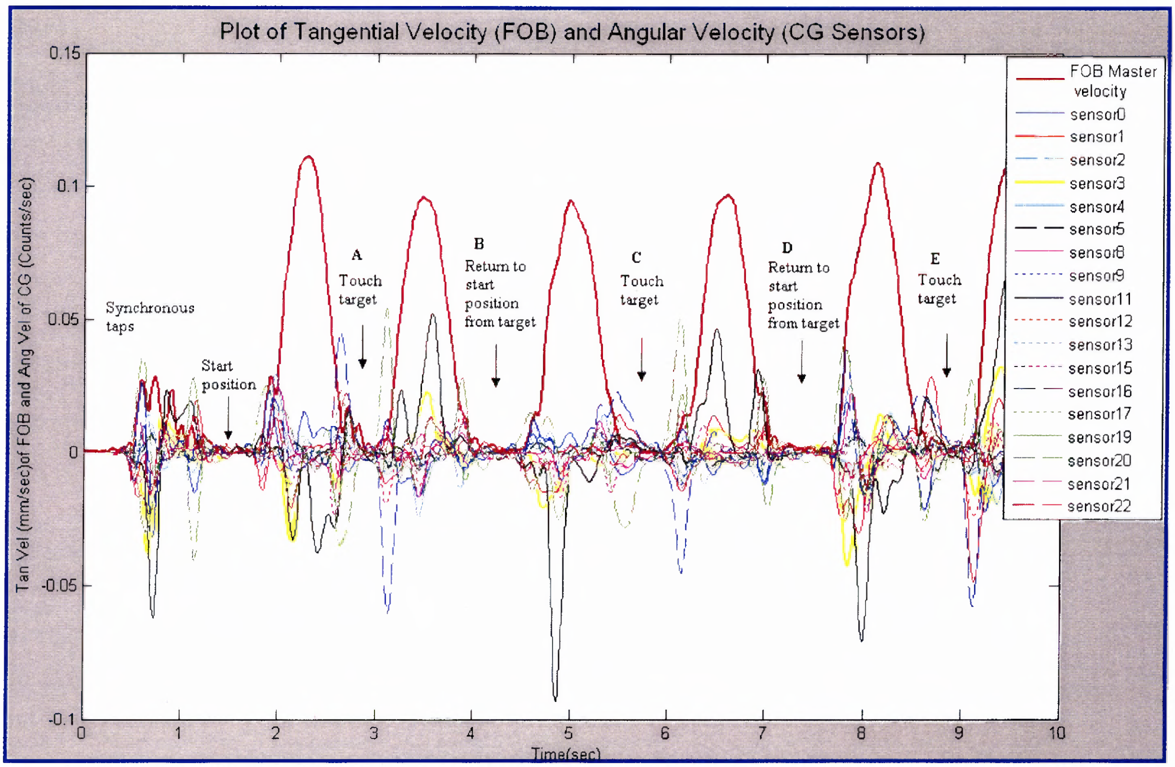

Figure 5.1 Tangential velocity of the FOB and angular velocities of the CG sensors in a reaching and grasping task.

FOB velocity $_{tan}$ troughs matched up with points in time when the CG velocities$_{ang}$ approach velocity of zero (points A, B, C, D and E). These velocity$_{tan}$ minima indicated the position when the hand had arrived at the target or returned to the starting point. The FOB velocity$_{tan}$ peaks were the points at which the wrist was in maximum velocity, between the target and the starting position, suggesting movement.

There was visible evidence of preshaping of the hand before it arrived at the target, shown in the CG angular velocities (Figure 5.1). Peaks in the velocity$_{ang}$ of the finger joints signified finger flexion (closed handshape), while troughs in the velocity$_{ang}$ of the finger joints indicated finger extension (open handshape). Most of the velocity$_{ang}$ of the CG sensors showed a reciprocal pattern in the graph. For example, the arch of palm extended prior to target arrival, opening the hand in preparation for the grasp (sensor 20- green dashed line). It then flexed preceding the movement back to the start position. The velocity$_{ang}$ of the MIDDLE-INDEX abduction sensor (sensor 11- black line) peaked during movement from the target to the start position and troughs during movement to the target. This was a display of preshaping, where the fingers spread out as the subject's hand moves away from the target and the fingers come together before the target was grasped.

Figure 5.1 also demonstrates the repeatability and coordination of the hands and fingers during the motor task. The spatial variability of the movement was small, a finding that conforms to the observation of Laquaniti and Soetching (1982). The synchronous and temporal behavior of finger joints and the subject's wrists is invariant throughout the task.

5.2 Kinematic Bandwidth of Sign Language

The kinematic bandwidth of ASL signs and fingerspelled words has been graphically shown to have a spectral energy concentration in the 0-3 Hz frequency range (Foulds, 2004). Foulds analyzed fingertip motion in the x, y and z directions during a fingerspelling task. An analysis similar to that of Foulds was carried out on the FOB displacement variables (x, y and z) and orientation variables (pitch, roll and yaw) and CG joint angle data to evaluate their frequency range. The data was filtered at 49 Hz cut-off frequency and a power spectrum was plotted for each variable.

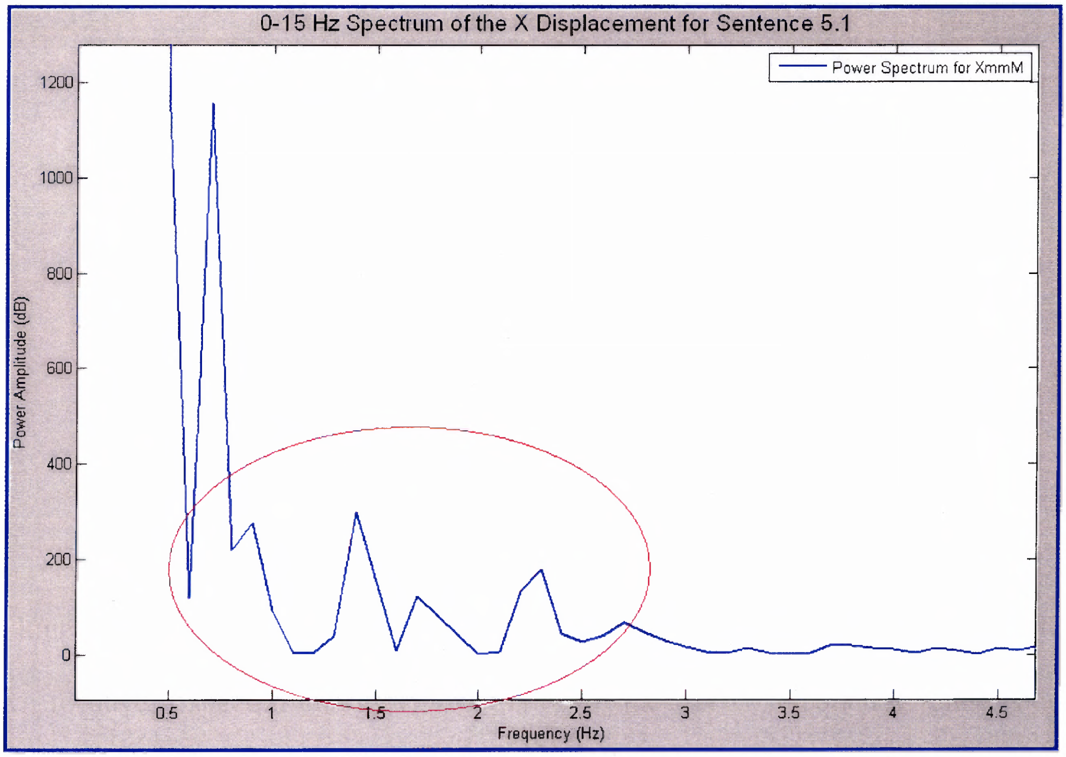

Figure 5.2 Spectrum of the X-dimension movement of the wrist in sentence 5.1.

A 0-3 Hz biomechanical bandwidth was found for the x displacement variable of sentence 5.1 and is shown in Figure 5.2. Similar results (0-3 Hz biomechanical bandwidth) were found for the other displacement and orientation variables of sentence

5.1, variables of other signed sentences, as well as for finger joint angle data (CG). Therefore, this frequency range accurately represents movements important to sign language and in that way, supports the cut-off frequency of 6 Hz for filtering the sign language data.

5.3 Transformation of the Flock of Birds® Sensor Position

To avoid distortion in the FOB velocity profiles, such as that shown in Figure 5.4, the transformation of the FOB sensor position from the top of the distal end of the forearm to the center axis of the wrist was computed. The distortions were present in the FOB tangential velocity due to the location of the sensor, approximately 40mm from the central axis of the wrist. During signing, rotation of the wrist should result in no tangential velocity. However, displacement of the sensor by 40mm (see Figure 5.3) translated the position of the sensor, and provided incorrect velocity calculations.

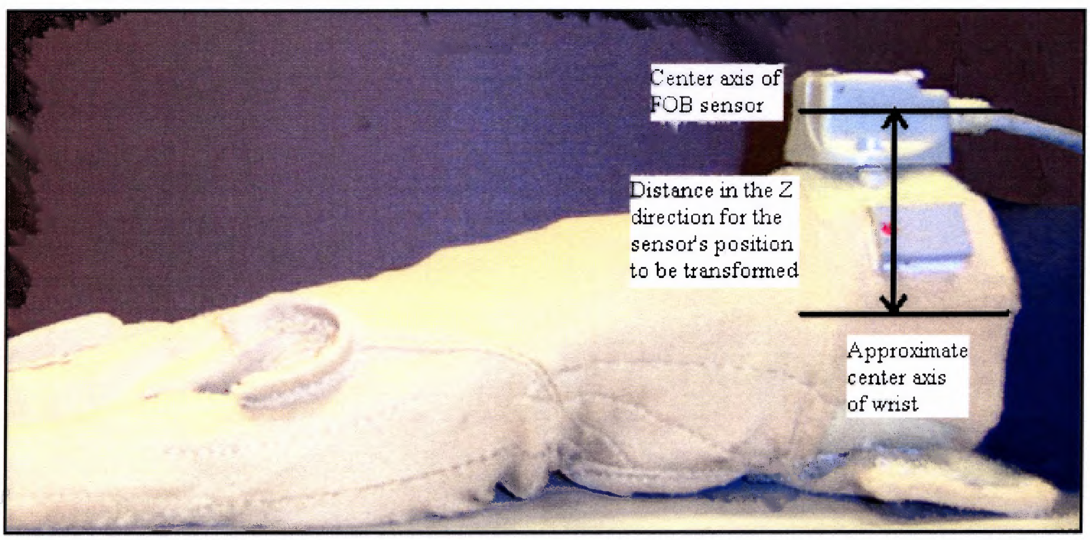

Figure 5.3 Displacement of the position of FOB sensor from the center axis of the wrist.

This problem was corrected by mathematically translating the senor to the central axis of the forearm. The new position was computed with a program that utilized the *Matlab Robotics Toolbox.* Using a function, ***rpy2tr***, the roll (rotation in the X-axis), pitch (rotation in the Y-axis), and yaw (rotation in the Z-axis) angles were converted into a rotation matrix, and another function, ***transl,*** converted the X, Y, and Z positions in mm into a translation vector. The rotation matrix and translation vector were used to form the new transformation matrix. The transformation matrix $^{0}T_{b}$, represents the position and orientation matrix of the FOB sensor atop the wrist, and can be found using the measured FOB sensor's position and rotations.

$$^{0}T_{b} = \begin{bmatrix} r_{11} & r_{12} & r_{13} & x_b \\ r_{21} & r_{22} & r_{23} & y_b \\ r_{31} & r_{32} & r_{33} & z_b \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

Next, the dorsal width of the subject's forearm (beneath the FOB sensor) was measured and added to half of the thickness (Z-direction) of the FOB sensor. A value of 40 mm was found. The transformation matrix $^{b}T_{c}$, was set up and comprised an identity rotation matrix and a translation vector with a z translation value of 40 mm (z_{new}).

$$^{b}T_{c} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & z_{new} \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

Since the translation is in the positive Z direction, the new transformation matrix, $^{0}T_{c}$, was computed to represent a position on the center-axis of the wrist transformed from the sensor's original location.

$$^{0}T_{c} = {}^{0}T_{b} * {}^{b}T_{c} = \begin{bmatrix} r_{11} & r_{12} & r_{13} & x_b \\ r_{21} & r_{22} & r_{23} & y_b \\ r_{31} & r_{32} & r_{33} & z_b \\ 0 & 0 & 0 & 1 \end{bmatrix} * \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & z_{new} \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

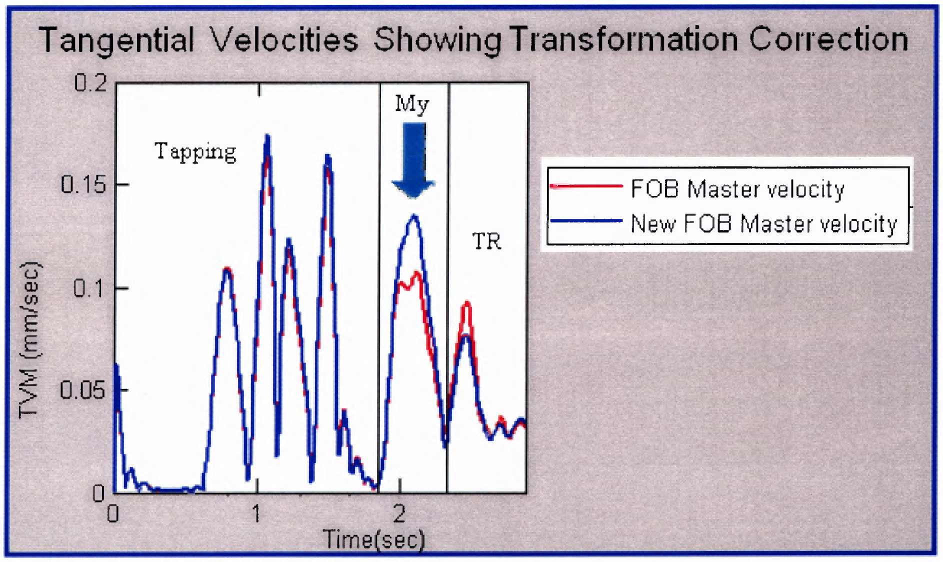

Figure 5.4 Graph showing the velocity profile of the sign MY before (red) and after (blue) the transformation correction.

The velocity profile showed improvement following the transformation of the FOB sensor to the center axis of the wrist. There was still a slight depression in the peak of the tangential velocity since the center axis of the wrist is not fixed and will rotate slightly, even with the position transformation. There was also an increase in the tangential velocity of the sign MY, due to the transformation in the positive Z-direction, which incremented the value of the displacement and consequently increased the tangential velocity of the sensor.

5.4 Sign Duration

The duration of a sign in isolation was retrieved from the graphs of wrist position versus time, using the Matlab function, ***ginput,*** to approximate the length (in seconds) per sign or signed sentence. Throughout the data collection procedure, the signer began each trial by tapping her lap twice, with both her hands. This movement was clearly visible on the graphs as four positive peaks, which aided in the identification of the start of the signed trial. The end of each sign was also explicit as the signers' hands returned to her lap immediately she finished signing, shown as the last tangential velocity peak. This was done for each trial of the ten sentences and the forty six individual signs. The sign GYM (Figure 5.5) involves a handshape of formed fists, followed by pumping the arms and wrists up and down twice. This is recognized as four positive velocity peaks.

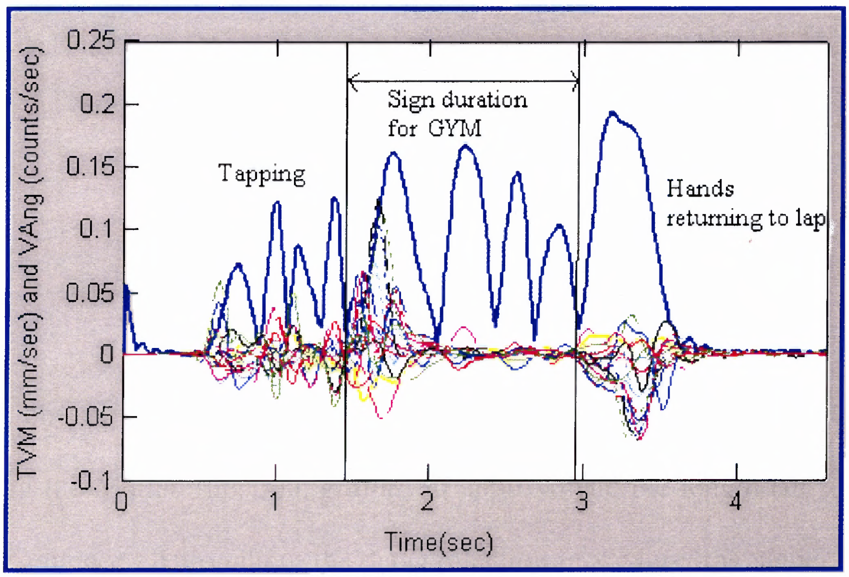

Figure 5.5 Demonstration of the sign duration of the sign GYM from velocity profile.

5.5 Measurement Error

Errors in measurements and graphs constitute two components, the systematic error/ bias (the same for every measurement) and the random error (varies from measurement to measurement and averages out to zero over time). The bias was trivial in this experiment since the measurements were collected by one experimenter, and the data analysis was across data collected from one subject, using the same equipment. To ensure that the measured value of the durations of the signs were estimated accurately from the graphs, a calculation of the measurement error was carried out to check for uncertainty in the measurement and data processing methods. It is important to note that the data were collected, processed and captured from the graphs by the author of this thesis. Therefore, although information was acquired by estimation of the ***ginput*** function there was less random error from the measurements because they were collected by the same person.

To confirm accuracy of the process, a precision analysis of the measurements was performed on the graphs of five randomly chosen signs. Precision refers to the degree to which repeated measurements of the same quantity are likely to agree with each other (Navidi, 2006). Position, joint angle and velocity graphs of the wrists and fingers versus time were plotted for the signs: AND, GREEN, WIFE, GO, and TOYS. Sign duration was retrieved from the graphs with the ***ginput*** function five times for each sign.

Table 5.1 Statistical Accuracy of Sign Duration from Position vs. Time Graphs

	AND	GREEN	WIFE	GO	TOYS
Mean duration (second)	0.5	0.48	0.68	0.44	0.54
Std. Dev (δ)	0.002849	0.001753	0.003371	0.002281	0.001789

Table 5.1 shows the sign duration measurements, their calculated means and standard deviations. The precision is determined by the standard deviation, δ, of the measurements. The smaller the value of the standard deviation, the higher the accuracy of the measurement process. The repeated measurements were approximately the same each time, demonstrating high precision of the measurement procedure. In Figure 5.5, comparative boxplots of the five values from each sign were plotted. The interquartile range (IQR) is the difference between the first and third quartiles of the data of each sign. There is no outlier (points that is more than 1.5 IQR above the third quartile or below the first quartile). The values are densely crowded implying high reliability of the measurement procedure.

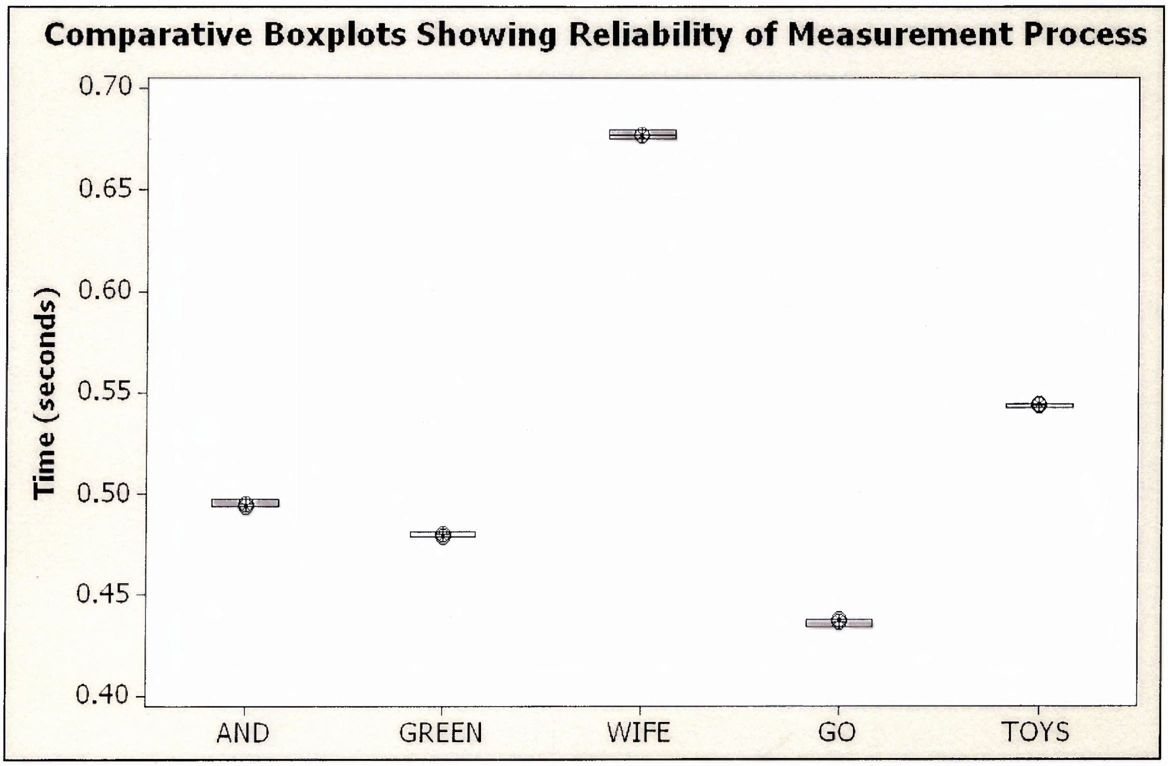

Figure 5.6 Comparative boxplots of accuracy of the measurement process for five signs.

All the data were collected at a measurement rate of 100 samples/ second. As a result, the data are actually only accurate to one hundredth of a second or to two decimal places. Data returned from Matlab have up to four decimal places and therefore the data were rounded off to two decimal places, respectively.

CHAPTER 6

RESULTS AND DISCUSSION

6.1 Sign Duration

After measuring sign durations from the position and velocity graphs, boxplots (Figure 6.1) of each sign were plotted to show the mean, median, highest value and lowest value for each sign and across all the signs.

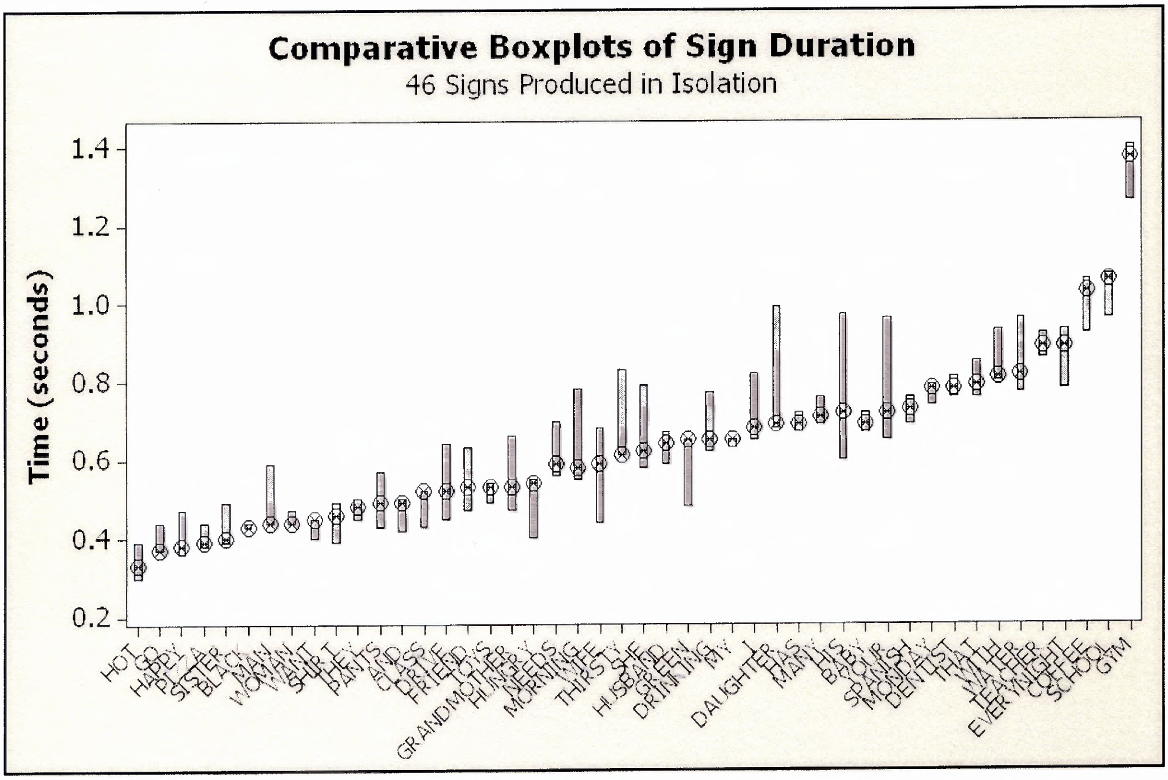

Figure 6.1 Boxplots of the sign duration of all 46 signs signed in isolation.

Signs were shown to vary in length, with a range of less than half a second to over a second, for the signs chosen for this experiment. There was also some observable variation in duration within the three samples of each sign. This disparity may be due to the fact that the signer is not a native signer and is not fluent at ASL.

The same measurement procedure, as mentioned above, was used to find the duration of the signed sentences and boxplots of their duration are shown in Figure 6.2.

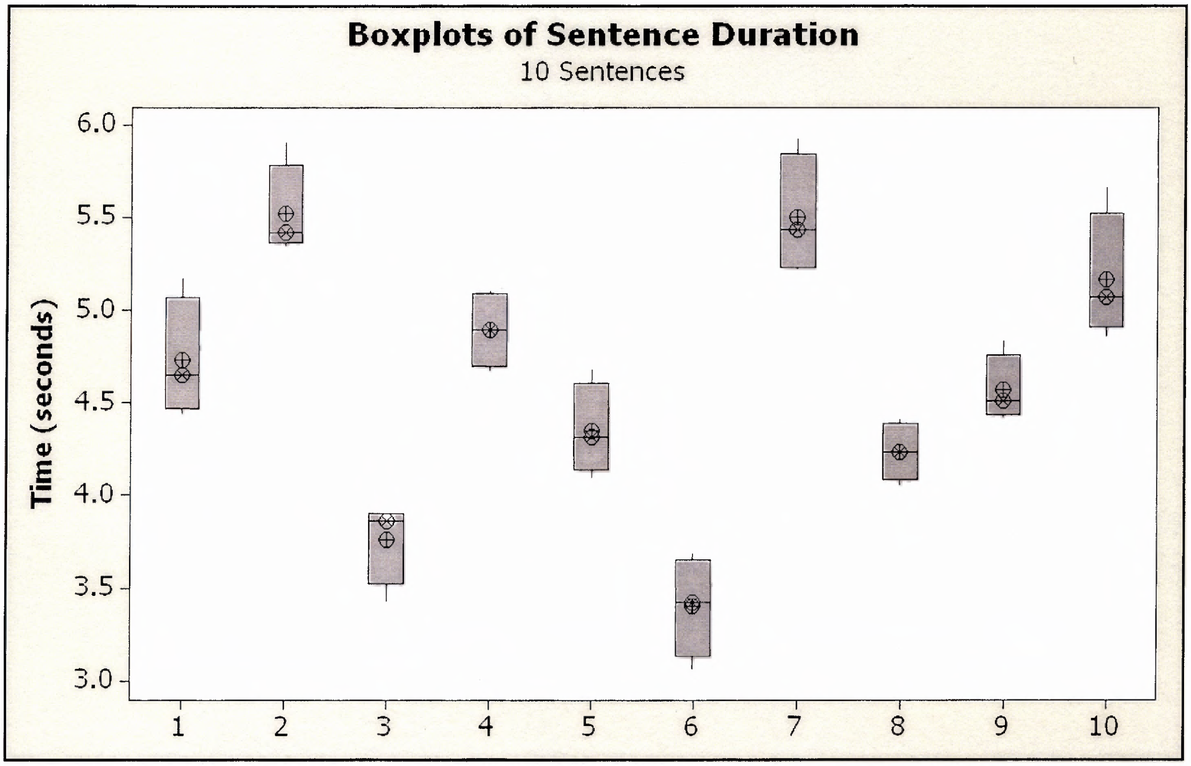

Figure 6.2 Boxplots of the 10 signed sentences.

In order to evaluate sign duration in isolation versus in sentence context, sign durations (from isolation and from segmented sentences) were measured from the position and velocity profiles. For the signs in sentence context, *Stokoe* locations, based on time, were determined where individual signs began and ended. This segmentation is described in detail in Section 6.4. After segmenting the sentences into individual signs, their duration was noted and compared with the same signs that had been produced in isolation.

Comparison of the mean durations of signs produced in isolation and within sentences confirmed that sign duration lessens in sentences. This is similar to the situation in spoken languages, where the length (in time) of a word is shorter in sentence context than when produced separately. Figure 6.3 demonstrates this decrease in time for the signs that comprised sentence 5.1. This decrease in duration is due to coarticulation of the signs, a phenomenon analogous to the coarticulation that occurs in speech.

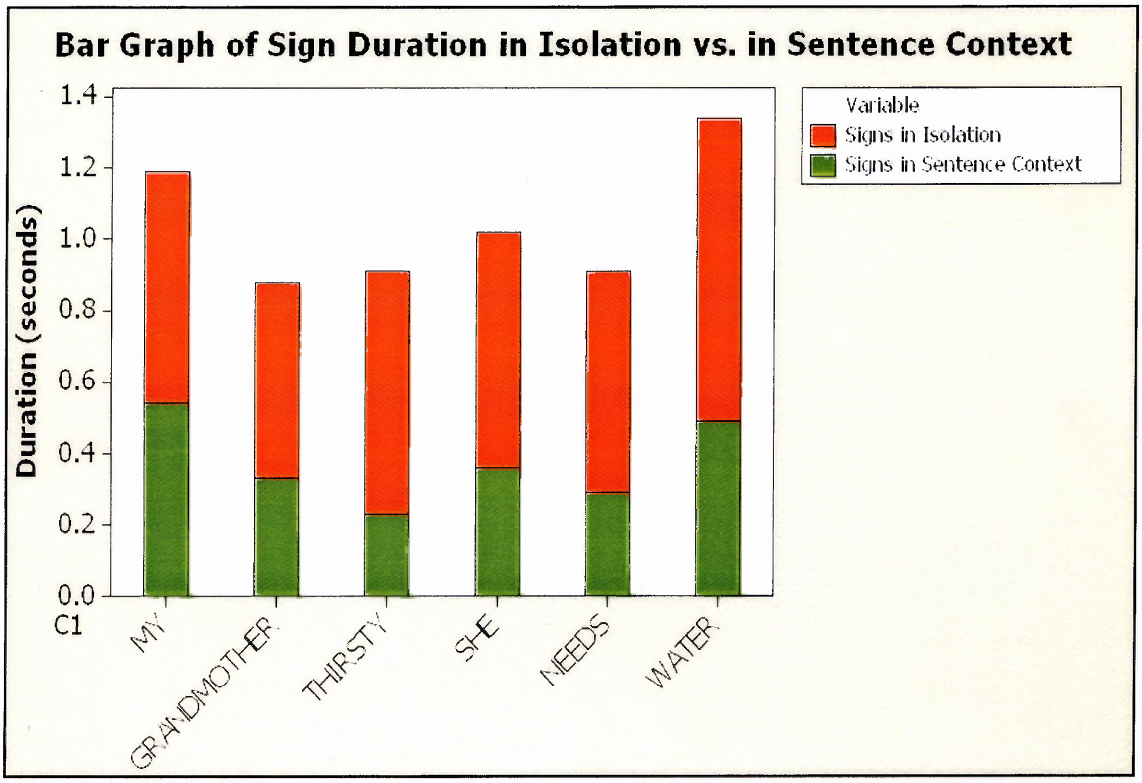

Figure 6.3 Bar graph of the durations of signs in isolation and in a signed sentence.

6.1.1 T-Tests

A paired t-test was carried out on the mean durations of the signs comprising Sentence 5 (both in isolation and in sentence context). The t-test was done to test the mean difference between paired observations, to compute a confidence interval and perform a hypothesis test of the mean difference between paired observations in a population. The paired observations were the signs that were produced in isolation and the same signs produced in a sentence. A paired t-test matches responses that are dependent or related in a pairwise manner accounting for variability between the pairs which usually results in a smaller error term, thus increasing the sensitivity of the hypothesis test or confidence interval. This test provided information on the significance of difference of sign duration in sentence versus in isolation.

The average sign duration for each trial was computed for the Sentence 5 signs (in isolation and in sentence context) and they are shown below:

	Isolation (secs)	**Sentence Context (secs)**
Sentence 5		
MY	0.65	0.54
GRANDMOTHER	0.55	0.33
THIRSTY	0.68	0.23
SHE	0.66	0.36
NEEDS	0.62	0.29
WATER	0.85	0.49

The null hypothesis was that the means of sign duration in isolation would be equal to the means of sign duration in sentence context, and there would be no difference between them. The results of the t-test gave a t-value = -6.14, and a p-value = 0.002.

	Samples	**Mean**	**Standard Dev.**	**SE Mean**
Sentence	6	0.373333	0.119108	0.048626
Isolation	6	0.668333	0.099883	0.040777
Difference	6	-0.295000	0.117771	0.048080

Since the p-value (0.002) is less than 0.05, the null hypothesis was rejected. The mean sign duration of signs in isolation was significantly different when compared to the mean sign duration of signs in sentence context.

A second t-test (two sample t-test) was also carried out to investigate the sign durations of repeated signs in different sentences. The two samples were the sign durations of BLACK in the Sentence 3 trials, versus the sign durations of the same sign in Sentence 8 trials. This refers back to Section 4.2, which described how specific signs in Sentences 1-5 were repeated in corresponding Sentences 6-10.

Table 6.1 Comparision of the Sign Duration of the Sign BLACK in Different Sentences

Sentence	**3.1**	**3.2**	**3.3**	**8.1**	**8.2**	**8.4**
Duration (sec)	0.41	0.29	0.29	0.29	0.33	0.37

The null hypothesis: [H_0: $\mu_3 = \mu_8$] hypothesized that the mean durations in the trials of Sentence 3 would to be equal to the mean durations in Sentence 8 trials.

Difference = μ (Sent3) - μ (Sent8)
T-Test of difference = 0 (vs not =):
T-Value = 0.04 **P-Value** = 0.974 **DOF** = 3

The results gave a p-value = 0.974, therefore the null hypothesis is not rejected. Sign durations of BLACK in different sentences are shown to be statistically insignificant. This demonstrates the temporal relation of sign duration of a sign produced in different sentences.

6.1.2 ANOVAs

Two two-way Analyses of Variance (ANOVA) were performed to compare sign duration of signs in sentences and across trials of the same sentences. This analyzed the equality of population means when classification of an observation is done by two fixed factors.

The first ANOVA was carried out to evaluate the mean duration of signs in and across three trials of the same sentence. The fixed factors were:

- Identical signs across the sentences

-Signs in each sentence

Signs were segmented from Sentence 5 (5.1, 5.2, 5.4) and 8 (8.1, 8.2, 8.3) and their durations (seconds) were measured.

	5.1	**5.2**	**5.4**
MY	0.48	0.68	0.45
GRANDMOTHER	0.35	0.28	0.35
THIRSTY	0.20	0.21	0.29
SHE	0.50	0.32	0.25
NEEDS	0.22	0.31	0.33
WATER	0.43	0.55	0.49

	8.1	**8.2**	**8.3**
WOMAN	0.88	0.89	0.97
DRINKS	0.36	0.38	0.40
HOT	0.21	0.19	0.21
BLACK	0.29	0.33	0.37
COFFEE	0.92	0.79	0.87

The results for individual signs comprising Sentence 5, showed a significant value ($p<0.05$) for the difference in sign duration in each sentence. The comparison of identical signs across the three sentences (5.1, 5.2, 5.4) gave a p value that was not significant ($p=0.193$).

Source	**DOF**	**SS**	**MS**	**F**	**Probability**
ROW	5	0.209517	0.0419033	5.04	0.014
COLUMN	2	0.003633	0.0018167	0.22	0.807
Error	10	0.083100	0.0083100		
Total	17	0.296250			

Results from Sentence 8 followed the same trend, with p=0.014 for the assessment of signs in each sentence and p=0.807 for identical signs across the sentences (8.1, 8.2, 8.3).

Source	DOF	SS	MS	F	P
ROW	4	0.102307	0.0255767	2.61	0.116
COLUMN	2	0.002520	0.0012600	0.13	0.881
Error	8	0.078413	0.0098017		
Total	14	0.183240			

These results reveal that signs in sentences have dissimilar durations, but they confirm the repeatability of the signer, displaying no statistical significance in the duration difference of the same signs signed in different sentences. This analysis further supports the segmentation of signs on a basis of kinematics with a temporal relation.

6.2 Synchronicity

Synchronicity in the angular (finger joints) and tangential velocities (wrists) was apparent in all the velocity profiles. It was present more so in the angular velocities of the finger joints, where there was overlapping of the velocity peaks and troughs. This is an indication of coordination dynamics and the temporal synchronicity of the individual articulators.

The signer is not a native ASL signer and she may have exhibited some variation in synchronicity due to lack of experience (Wilcox, 1992). Wilcox demonstrated variation in synchronicity of articulator peak velocities for a fluent and non- fluent fingerspeller (Figure 6.4).

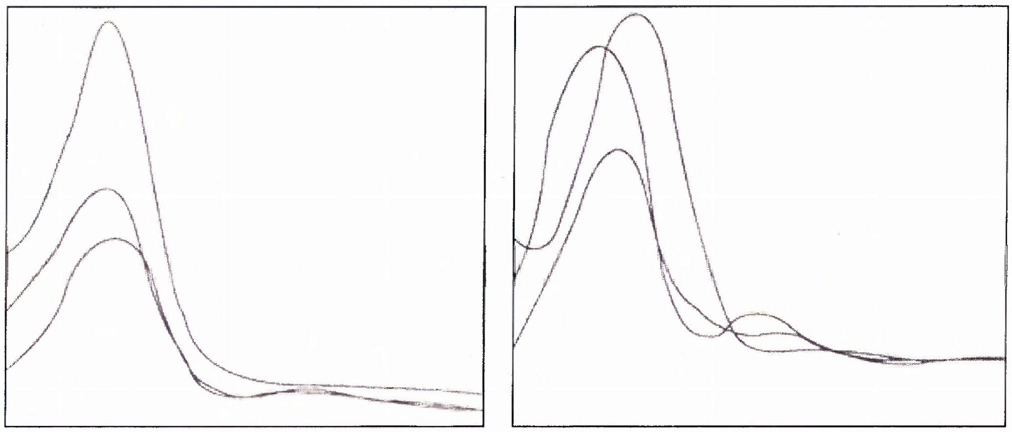

Figure 6.4 Demonstration of synchronicity with fluent fingerspelling (Left) and non-fluent fingerspelling (Right).
(Source: Wilcox, 1992)
(No axes are shown in original)

6.3 Preshaping

Evidence of preshaping of the fingers was observable as the hand approaches locations important in signs. This was shown graphically (Figure 6.5) as a coordinated clustering of finger velocities.

Preshaping was present in the transition between the signs PANTS and BLACK. During the sign PANTS, both hands were held open and were drawn up along the thighs. BLACK was signed by moving the index finger along the eyebrow. Between the two signs, the signer created the handshape for the second sign, shown in the transition in the midst of PANTS and BLACK. The sensors in the thumb, middle, ring and pinkie fingers (sensor 1, 2, 3, 8, 9, 12, 13, 17) were shown as local maxima, representing the flexing (closing) of the fingers, while the velocity index proximal interphalangeal sensor (sensor 5) stayed constant, remaining extended from the previous sign, PANTS. The metacarpo-phalangeal joint in the index finger bent slightly inwards for the sign BLACK and was represented by a peak. The velocities of all the finger joints then remained constant, signifying the handshape being held during the sign, BLACK.

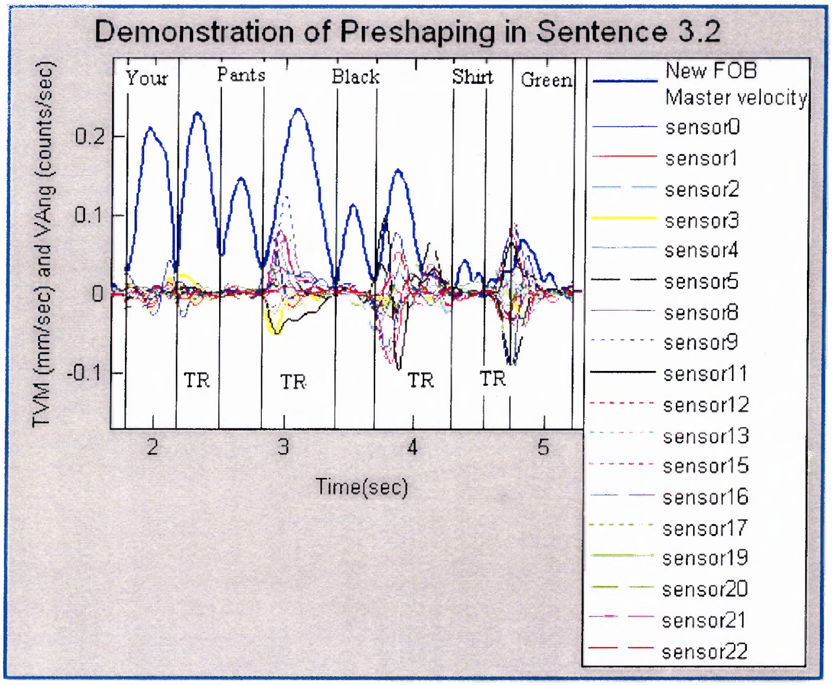

Figure 6.5 Sentence 3.2 demonstrating preshaping of the fingers(between PANTS and BLACK and between BLACK and SHIRT).
(TR-Transition)

6.4 Segmentation of Signs

Sign recognition entails separating signed stream into individual signs. *Stokoe's* notation classifies signs into handshape, location, orientation and movement.

Examples:

-MY has handshape, orientation, location and no movement.

-GRANDMOTHER has handshape, orientation, two locations and movement between the two locations.

This research expected to find kinematic markers that could identify one or more of *Stokoe's* notations. Using the three dimensional position of the wrists, their orientations and the behavior of the fingers, the signed sentences were divided into (approximate) signs, according to time. The velocity profiles of the sentences corroborated these times and segmented the sentences into locations and movements. Locations were categorized as local minima (troughs) in the tangential velocity, places in signs where the signer makes contact or changes position. At these locations, the angular velocities of the finger joints also approach zero velocity, signifying the shaping of the signers hand. Midpoints in movements were identified as peaks in the tangential velocity, and could be part of the signs or transitions between signs. These locations and movements are two of *Stokoe's* parameters for sign classification. Transitions linked certain signs, such as between MY and GRANDMOTHER, where the right hand moved from the chest (MY) to the chin (beginning of GRANDMOTHER). Other signs, for example THIRSTY, SHE, and NEEDS do not have any transition between them, visible in the velocity profile. One explanation is that during these signs the right hand remains in generally the same location and most of the movement is made by changing the handshape and the finger positions. These patterns of velocity changes therefore serve as cues for *Stokoe* location and movement, either in the sign or in sign transitions. Figure 6.5 exhibits the segmentation of Sentence 5.1, and illustrates the tangential velocity troughs and peaks as well as the angular velocities of the roll, pitch and yaw of the wrist.

Figure 6.6 shows the segmentation of Sentence 3.2, following the technique used to segment Sentence 5.1. Velocity profiles of each sign produced in isolation and orientation vs. time graphs were used as verification for segmentation. The sign MY involved the use of one hand, with a flat palm handshape and one location on the signer's chest. Both hands then maintained the same handshape and the sign PANTS (consisting of two locations-upper thigh and lower thigh, with a movement between the locations) was signed. The transition between the two signs was a result of the movement of the hands from the location of MY to the first location in PANTS. There was a transition following PANTS, in which the right had formed the new handshape (index finger extended, all other fingers closed). The angular velocities of the finger joints were kept constant as the sign BLACK was signed across the signer's forehead. This point in time occurred when the hand was at the most negative z-position. (FOB z-axis is negative in the upward direction). There was then a transition when the right hand preshaped (thumb, index and middle fingers flexing into a pinch shape), forming the handshape as it arrived at the location for SHIRT. SHIRT was signed by pinching the shirt while moving hand up and down. The right hand transitioned into GREEN (comprised the sign of the letter 'g' and shaking the hand from the wrist). The sentence ended with the signer's hands returning to her lap.

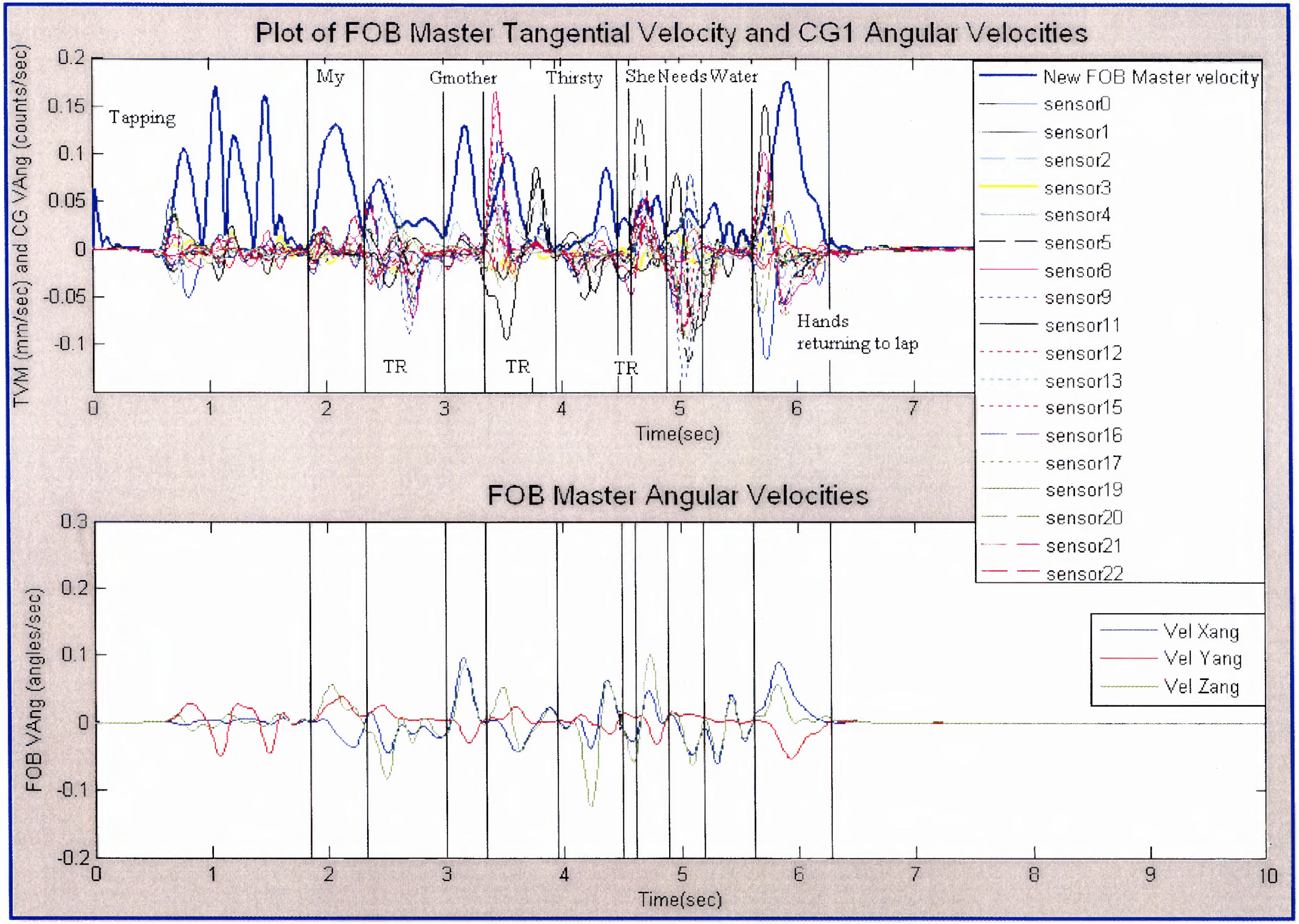

Figure 6.6 Segmentation of signs from Sentence 5.1, based on time, *Stokoe* locations and movement. (TR- transition)

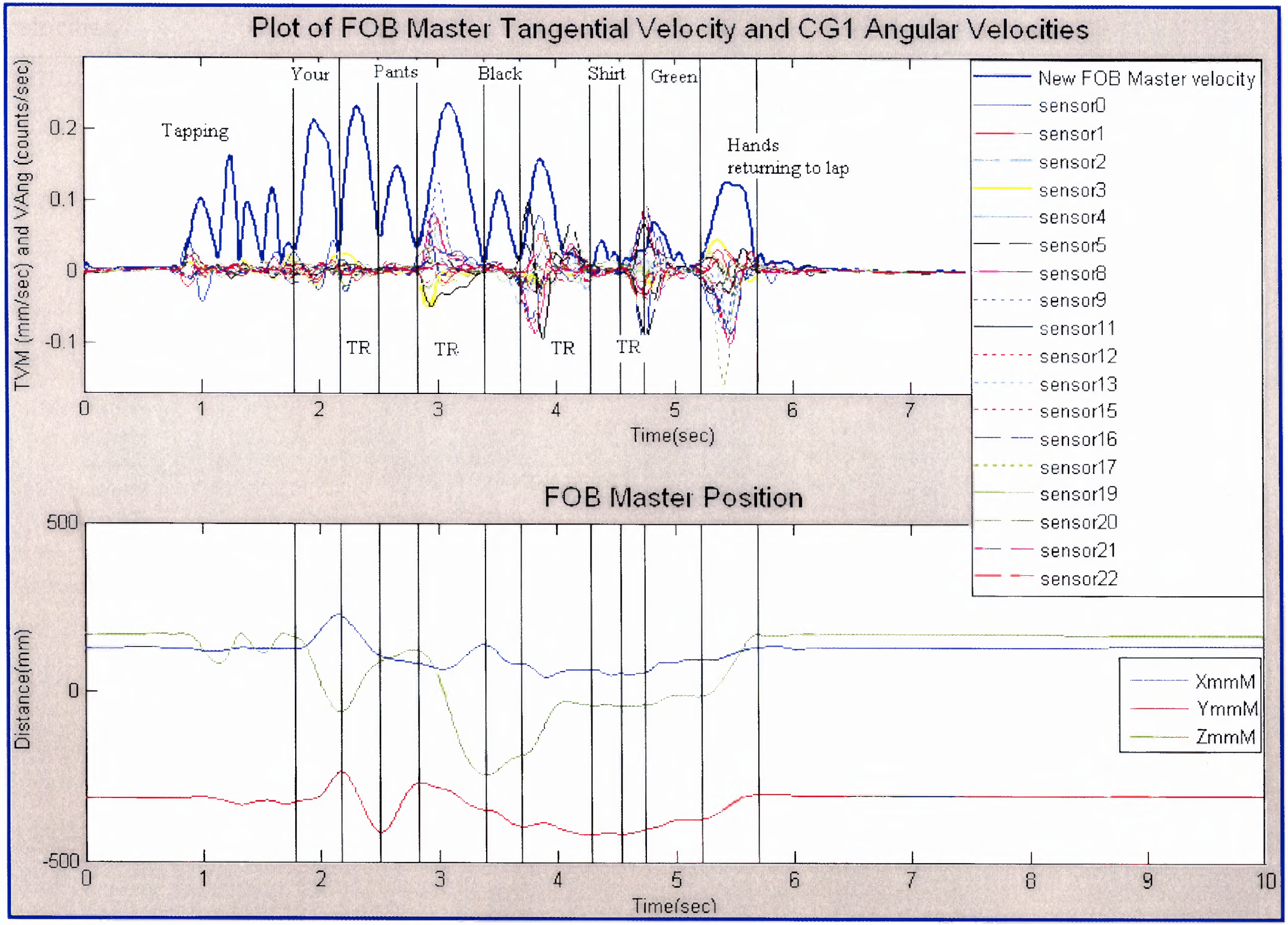

Figure 6.7 Segmentation of signs from Sentence 3.2, based on time, *Stokoe* locations and movement. (TR- transition)

CHAPTER 7

CONCLUSION

This analysis of sign language has provided a foundation for ongoing work in sign recognition in the context of signed sentences. Recognition research has been successful in the recognition of signs in isolation, but has had little success in continuous sign recognition. This thesis connects various aspects of kinematics: a spectral analysis, timing and duration of signs, and the identification of kinematic markers of *Stokoe* locations and movements, with respect to time.

The spectral analysis carried out confirmed results of the kinematic bandwidth utilized during signing. Signs and signed sentences have been shown to have a spectral power concentration in the 0-3 Hz range. This trend applied to the position and angular data of the wrist (from the FOB), and the joint angle data of the fingers (CG), irrespective of the sign or signed sentence analyzed.

Comparison of the mean durations of signs produced in isolation and within sentences showed that sign duration decreases in sentence context. This is due to coarticulation that occurring between signs when produced in stream. Sign synthesis and recognition systems will need to account for this variation of sign duration in sentences. It is important for sign synthesis to appear natural and human-like. Within sentences, sign durations of individual signs have been shown to be quite consistent; indicating that duration may be an additional kinematic parameter that can be used in recognition.

Analyses of the displacement, angular and velocity profiles of the FOB sensors, with respect to time, have shown important kinematic markers of handshape, location and movement. Peaks in the tangential velocity of the wrists clearly indicate the midpoints of movements within signs or the midpoints of transitions between signs. Troughs (velocity minima) in the tangential velocity signify *Stokoe* locations, which can serve as markers in the segmentation process. Time anchors have been identified for moments when these *Stokoe* locations or movement are achieved. Handshape and wrist orientation can be examined at these points, which will prove to be very useful for sign identification and segmentation. Positive peaks in the angular velocity of the fingers indicate the opening (extension) of the fingers and the negative troughs show the closing (flexion) of the fingers. The CG abduction sensors between the fingers also provide information on the spread of the fingers, denoting features of the handshape. A hold in a sign (fingers held in specific handshape) is shown graphically when the angular velocities of the fingers approach and remain constant around 0 counts/ sec.

This kinematic analysis does not yet segment sentences into individual signs, but it does segment the signed sentences into movements and *Stokoe* locations, by identification of respective timing of those locations and movements. In future work, these segmented locations and movements will be combined with handshape, wrist orientation, as well as the actual location of the hand with respect to the signer's body, to recognize likely signs based on the kinematic database of signs developed at NJIT. The signs contained in this database are stored with the kinematic parameters of orientation, handshape, location and movement. Identifying *Stokoe* locations and movements in signs and transitions and improves potentiality of sign language recognition based on kinematics.

CHAPTER 8

APPLICATIONS AND FUTURE WORK

Current techniques of translating spoken language into sign language include sequenced video clips, video animation and sign synthesis using parametrically driven animation. Many of these techniques lack the kinematic parameters required to produce lucid, human-like signs. This thesis provides useful information regarding the kinematics of continuous sign and sign segmentation. This information can be incorporated into the NJIT sign language synthesis system which presents efficient sign translation from English text input and the production of signed sentences. The integration of kinematics will produce more accurate, intelligible signs and signed messages.

The results of the kinematic analysis of this thesis will not only prove to be constructive for sign language translation and animation, but could be applied to numerous other signed languages that utilize *Stokoe* parameters. This study provides kinematic information on sign duration, continuous signing and segmentation, and does not examine the linguistics of the sign language. Therefore, the results of this thesis can be implemented in the recognition of various sign languages around the world.

Most importantly, sign language recognition, based on sign kinematics, will provide an accurate and efficient method of identifying individual signs from streams of signs. Through this kinematic analysis, signs can be isolated from sentences on the basis of time-dependent location and movement parameters. Continuous sign recognition will breach the communication gap between signers and non-signers.

Transitions comprise a significant part of the phonetic structure of sign language and further kinematic analysis of the sign transitions will be required for precise segmentation. Additional work on the classification of handshapes in continuous signing will complement the results of identification of *Stokoe* locations and movement parameters and advance the segmentation process.

Pattern matching techniques are useful for the identification of signs during sign recognition. In conjunction with the kinematic database of signs created at NJIT, an investigation into handshape pattern classification (artificial neural networks, fuzzy logic, statistical classification) will be carried out. Such pattern classification will provide a vaguely defined method of distinguishing signs based on kinematic parameters. This is useful given that people do not sign at the same speed or in the same manner. Handshape pattern classification will allow for the identification of individual signs from stream, regardless of the signer.

REFERENCES

1. Allen J., Foulds R. "An approach to animating sign language: A spoken English to sign English translator system." *Bioengineering Conference, Proceedings of the IEEE 30th Annual NorthEast*, 2004.

2. Battison R. "Phonology in American Sign Language: 3-D and digit vision." *California Linguistic Association Conference*, Stanford, CA, 1973.

3. Castiello U. "The neuroscience of grasping." *NatureReviews*, 2005; 6 (99): 726-736.

4. Clark L., Grosjean F. "Sign recognition process in ASL: The effect of context", 1982.

5. Costello E. "Random House Webster's Concise American Sign Language Dictionary", 2002.

6. Foulds R. A. "Biomechanical and perceptual constraints on the bandwidth requirements of sign language." *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 2004; 12(1):65-72.

7. Furth H. G. "A comparison of reading test norms of deaf and hearing children." *Am Ann Deaf,* 1966; 111(2):461-2.

8. Gentilucci M., Corballis M. "From manual gesture to speech: A gradual transition." *Neuroscience and Behavioral Reviews*, 2006; 30: 949-960.

9. Green K. "Sign boundaries in American Sign Language: Review." Linstok Press, Inc, 1984: 65-91.

10. Holden E. J., Lee G., Owens R. "Australian Sign Language recognition." *Machine Vision and Applications*, 2005; 16 (5): 312-320.

11. Irving A., Foulds R. "A parametric approach to Sign Language synthesis." *Proceedings of the 7th International ACM SIGACCESS Conference on Computers and Accessibility,* Baltimore, MD, 2005.

12. Jeka J. J, Kelso J. A. S, "Manipulating symmetry in the coordination dynamics of human movement." *Journal of Experimental Psychology and Human Perception Performance*, 1995; 21 (2): 360-74.

13. Jerde T., Soetching J., Flanders M. "Coarticulation in Fluent Fingerspelling." *Journal of Neuroscience,* 2003; 23 (6): 2383-2393.

14. Kapit W. and Elson L. M. "The Anatomy Coloring Book." Harper Collins 2nd Edition, p. 27, 1993.

15. Kelso J. A. S., Buchanan J. J., Murata T. "Multifunctionality and the switching in the coordination dynamics of reaching and grasping." *Human Movement Science*, 1994; 13: 63-94.

16. Kelso J. A. S., Buchanan J., Fuchs A. "Coordination dynamics of trajectory formation." *Biological Cybernectics*, 1996; 74: 41-54.

17. Kelso & Engstrøm. "Coordination Dynamics." *Excerpt from the Complementary Nature.* [Online Document], [cited 2006 October 17], Available HTTP: http://www.thecomplementarynature.com/TCN%20CD.php

18. Klima E., Bellugi U. "The Signs of Language." Cambridge, MA: Harvard University Press, 1979.

19. Lane L. G. "The Gallaudet Survival Guide to Signing." Washington, D.C: Gallaudet Universal Press, 1990.

20. Liddell S., K. "An investigation into the syntactic structure of ASL." *Unpublished Doctoral Dissertation*, University of California, San Diego, 1977.

21. Liddell S., K. "THINK and BELIEVE: Sequentiality in American Sign Language. Language", 1984a; 64 (2): 372-399.

22. Liddell S., K. "Structures for representing handshape and local movement at the phonemic level." In S. D. Fischer & P. Siple (Eds.), *Theoretica Issues in Sign Language Research.* Chicago, IL: University of Chicago Press, 1990.

23. Ma J., Gao W., Chen X., Shanl S., Wei Z., Yan J., Zhang H., Wu J., Wu F., Wang C. "MDS: A Mulitmodal-based Dialog System." *Proceedings of the 8th ACM International Conference on Multimedia*, 2000.

24. Murakami K. and Taguchi H. "Gesture recognition using recurrent neural networks." *CHI '91 Conference Proceeding*, 1991: 237-241.

25. Neidle C., Kegl J., MacLaughlin D., Bahan B., Lee R. "The syntax of American Sign Language." Cambridge, MA: MIT Press, 2000.

26. Newkirk D. "Surface level handshape assimilation in ASL." *Unpublished Manuscript*, The Salk Institute, San Diego, 1977.

27. Patterson Jill, Morford. "Frequency characteristics of American Sign Language." *Sign Language Studies,* 2003; 3 (2): 213-225.

28. Phillips C., Badler N. "JACK: A toolkit for manipulating articulated figures." *Proceedings of 1st Annual ACM SIGGRAPH Symposium on User Interface Software*, Banff, Canada, 1988: 221-229.

29. Santello M., Flanders M., Soechting J. F. "Postural hand synergies for tool use." *Journal of Neuroscience,* 1998; 18:10105-10115.

30. Santello M., Flanders M., Soechting J. F. "Patterns of hand motion during grasping and the influence of sensory guidance." *Journal of Neuroscience*, 2002; 22:1426-1435.

31. Schettino L. F, Adamovich S. V, Poizner H. "Effects of object shape and visual feedback on hand configuration during grasping." *Experimental Brain Research*, 2003; 151 (2): 158-166.

32. Sherry G. and Foulds R. "Pattern recognition considerations for continuous Sign Language recognition." *Bioengineering Conference, Proceedings of the IEEE 29th Annual NorthEast*, 2003.

33. Soechting J. F, Flanders M. "Flexibility and repeatability of finger movements during typing: analysis of multi-degree of freedom movements." *Journal of Computational Neuroscience*, 1997; 41: 29-46.

34. Stewart D. A, Akamatsu C. T. "The coming of age of American Sign Language." *Anthropology & Education Quarterly,* 1988; 19, (3): 235-252.

35. Stokoe W. C. "Sign Language Structure: An outline of the communicative systems of the American deaf." Silver Spring, MD: Linstock Press, 1960.

36. Stokoe W. C. "A Dictionary of American Sign Language on Linguistic Principles." Silver Spring, MD: Linstock Press, 1976.

37. Swisher M. V. "Similarities and differences between spoken languages and natural Sign Languages." *Applied Linguistics*, 1988; 9(4):343-356.

38. Virtual Technologies, Inc. "*CyberGlove®* Reference Manual." Palo Alto. CA, 1998

39. Wang J., Stelmach G. "Coordination among the body segments during reach-to-grasp action involving the trunk." *Experimental Brain Research*, 1998; 123: 346-350.

40. Wilbur R. B. "Stress in- ASL: Empirical evidence and linguistic issues." *Language and Speech*, 1999; 42 (2-3): 229-250.

41. Wilbur R. B. "Physical correlates of prosodic structure in American Sign Language." *Chicago Linguistic Society*, 2000; 38.

42. Wilbur R. B. “A multi-tiered theory of syllable structure of American Sign Language.” *Annual Meeting, Linguistic Society of America*, San Diego, CA, 1982.

43. Wilbur R. B. “American Sign Language: Linguistic and Applied Dimensions.” San Diego, CA: Little, Brown and Co., 1987.

44. Wilbur R. B. “Why syllables? What the notion means for American Sign Language research.” *Theoretical issues in sign language research.* Chicago, IL: University of Chicago Press, 1990.

45. Wilcox Sherman. “The Phonetics of Fingerspelling.” Amsterdam, Philadelphia: John Benjamins Publishing Company, 1992.

46. Winter David. “Biomechanics and Motor Control of Human Movement.” Hoboken, NJ: John Wiley and Sons Inc., 2005.