FACILITATING COMMUNICATION FOR DEAF INDIVIDUALS WITH MOBILE TECHNOLOGIES

A Thesis
Presented to
The Academic Faculty

by

Valerie Henderson Summet

In Partial Fulfillment
of the Requirements for the Degree
Doctor of Philosophy in
Computer Science

School of Interactive Computing
Georgia Institute of Technology
May 2010

FACILITATING COMMUNICATION FOR DEAF INDIVIDUALS WITH MOBILE TECHNOLOGIES

Approved by:

Dr. Thad E. Starner, Advisor
School of Interactive Computing
Georgia Institute of Technology

Dr. Gregory Abowd
School of Interactive Computing
Georgia Institute of Technology

Dr. Jennie Carroll
School of Property, Construction, and Project Management
Royal Melbourne Institute of Technology

Dr. Rebecca E. Grinter
School of Interactive Computing
Georgia Institute of Technology

Dr. Mark Guzdial
School of Interactive Computing
Georgia Institute of Technology

Date Approved: March 28, 2010

To my mom,

Susan Amorette Vogt,

who never told me I couldn't or wouldn't and only occasionally told me I shouldn't.

ACKNOWLEDGEMENTS

I wish to thank:

- My advisor, Thad. For his vision and encouragement.
- My committee, Gregory, Beki, Mark, and Jennie. For their comments and critiques which were instrumental in the creation of this work.
- My lab, the CCG. For support above and beyond the call of duty, especially Tracy, James, and Helene.
- My husband, Jay. Words cannot express how thankful I am.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

SUMMARY

Communication between deaf individuals and hearing individuals can be very difficult. For people who are born deaf, English is often a second language with the first language being American Sign Language (ASL). Very few hearing people in the United States sign or are aware of Deafness, Deaf culture, or how to appropriately communicate with people with hearing loss.

In this thesis, I concentrate on the role that mobile technologies can play in ameliorating some of these issues. In formative work with Deaf teenagers in the metro-Atlanta area, I investigate the role that communication technologies play in the lives of many Deaf individuals and examine how these devices have effected their communication patterns and social circles. Specifically, the teens identified problems communicating with hearing individuals such as close friends and family in face-to-face situations.

Having identified sign language use at home as one of the earliest interventions for Deaf children, I investigated the use of mobile phones for learning survival-level ASL. I created a prototype software application which presented short ASL lessons via either a mobile phone or desktop web-browser. The software presented the lessons via one of two different scheduling methods designed to take advantage of the spacing effect during learning. I designed and conducted a study of forty individuals with no prior ASL knowledge which compared the effects of both scheduling algorithm and platform. My results show that individuals who used a mobile phone platform and received a group of lessons at one time performed better on post-test receptive and generative ASL metrics than did participants in the three other conditions.

CHAPTER 1

INTRODUCTION

Mobile communication technologies have revolutionized communication practices in several countries including the United States. Their basic affordance – mobility – caused many new practices and routines [86]. Individuals adapted their mental models of traditional, “landline” phones to mobile phones and began to pioneer new uses. As computing technologies have advanced, the mobile phone has become a mobile computing platform. Instant messaging, text messaging, mobile web browsing, calendaring, scheduling functionality, and media capabilities (such as mobile video viewing, video messaging, and music playing) are just a few of the possible uses of these devices. Mobile device adoption has grown rapidly and surpassed 2.8 billion subscribers worldwide during the first quarter of 2007 [111]. In many developed countries, the devices are becoming more of a necessity and less of a luxury. The ubiquity of mobile devices presents interesting opportunities to researchers.

The deaf community has enthusiastically adopted mobile devices [93, 92, 107]. While this may seem counterintuitive as deaf individuals typically have no need for the voice telephony capabilities, the mobile computing platform has opened many alternate communication channels that were not previously available. In this dissertation, I will explore some of the other capabilities of mobile devices and how they can affect communication for deaf individuals.

Communication between deaf individuals and hearing individuals can be very difficult. For people who are born deaf, English is often a second language with the first language being American Sign Language (ASL). Very few hearing people in the United States sign or are aware of Deafness, Deaf culture, or how to appropriately communicate with people with hearing loss. Difficulty in cross-cultural communication can affect everyday interactions such as making a purchase at a store or meeting new neighbors. Furthermore, cross-cultural communication difficulties can have adverse effects on employment [94], legal representation [73], and medical care [40, 97]. Improving cross cultural communication between Deaf and hearing individuals is a significant problem, and many aspects of the problem continue to be addressed by the education, linguistic, and medical research communities.

In this dissertation, I concentrate on the role that mobile technologies can play in ameliorating some of these issues. In Chapter 3, I investigate the role that communication technologies play in the lives of many Deaf individuals and examine how these devices have changed their communication patterns and social circles. Working with Deaf teenagers in the metro-Atlanta area highlighted many of the difficulties they encounter in everyday communication. In particular the teens identified problems communicating with hearing individuals such as close friends and family in face-to-face situations.

As discussed in Chapter 2, many deaf children are born to hearing parents [78]. These parents are then faced with an overwhelming number of choices and decisions to make for their child. One decision to be made is how to communicate with their child. The decision to utilize medical interventions such as a cochlear implant (CI) and attempt to communicate with English must be made at a very young age to give children the optimal chance to succeed using the CI. However, children begin responding to language earlier than 12 months of age, the age at which CIs have been approved for use by the US Food and Drug Administration [80]. As a result, many deaf children are not immersed in language until they are beyond the critical period for language acquisition. This lack of exposure can impair English fluency later in life.

In order to address this issue it is crucial to supplement early language learning from the source of the majority of a child's language: his or her parents. In Chapters 4 and 5, I present a technological intervention which leverages the ubiquity and video output capabilities of mobile phones to assist parents in learning survival level American Sign Language and a study to show the effectiveness of this technology. This study tests the hypothesis that utilizing the inherent affordances of the mobile phone is a feasible means to enhance second language learning of American Sign Language for hearing individuals. I also present a study to evaluate whether the mobility and ubiquity afforded by mobile phones increased the language learning over a traditional, desktop or laptop based platform. By leveraging existing mobile devices and proactive software for language learning, I show the advantages and disadvantages of using mobile devices and instruction to improve adults' command of ASL. In Chapter 6, I present some final recommendations and implications for design of future mobile learning systems.

1.1 Thesis

I hypothesize that participants who learn American Sign Language vocabulary:

1. *utilizing mobile phone platforms as a content delivery mechanism will demonstrate better receptive and generative language abilities than participants who learn using a traditional desktop-based platform as measured by post-intervention tests.*

2. *using the spacing effect, which presents the material on a distributed schedule instead of a massed schedule, will demonstrate better receptive and generative language abilities as measured by post-intervention tests.*

1.2 Contributions

The exploration of this thesis will yield the following contributions:

1. Research into the role that mobile communication technologies currently play in the social lives of Deaf teenagers and with whom they communicate (Chapter 3).

2. An investigation of algorithms for ASL vocabulary learning (Chapter 4).

3. Comparison of mobile devices and traditional desktop/laptop computers as platforms for ASL vocabulary learning (Chapter 5).

4. Assessment of different scheduling methods for ASL vocabulary learning (Chapter 5)

1.3 Overview of Dissertation

First, in Chapter 2, I present background on American Sign Language and the developmental delays which can result from being born with significant hearing loss. I also discuss forms of current communication technology and how they are utilized by both deaf and hearing individuals. I explore work related to learning technologies and, in particular, mobile learning platforms. I present an algorithm to maximize learning in a set amount of time and discuss how this algorithm was derived.

In Chapter 3, I discuss a study conducted at Atlanta Area School for the Deaf and show how a lack of language and geographical distance can combine to isolate Deaf children. I further discuss mobile communication use among Deaf teenagers and how these technologies affect their social networks. In particular, I focus on cross cultural communication between the Deaf and hearing communities.

In Chapter 4, I discuss different scheduling algorithms for computer-aided second language learning. I present an algorithm developed by Atkinson [4, 3] which leverages the intrinsic difficulty of certain vocabulary words to create an adaptive learning schedule of word presentations. I also discuss a data gathering study conducted to establish the relative difficulty of 80 age-appropriate American Sign Language vocabulary words selected from the MacArthur-Bates Communicative Development

Inventory [37]. Twenty participants were enrolled in this study and the results of their post-tests were compared to the adaptive algorithm scheduling study presented in Chapter 5.

I discuss communication between parents or caregivers and deaf children in Chapters 4 and 5. I then present my web-based, mobile phone application to assist learning American Sign Language and discuss the benefits and drawbacks. Chapter 5 discusses the evaluation of this application and the comparison to other, more traditional methods of content delivery. In Chapter 6 I outline future work which investigates questions and issues raised by the work presented in this dissertation. Finally, in Chapter 7, I present a discussion of the larger implications of my dissertation work.

CHAPTER 2

RELATED WORK

In this chapter, I discuss work related to several areas that are relevant to this dissertation. In particular, I examine background relating to American Sign Language (ASL) and the Deaf culture as they relate to this dissertation, work related to software for learning and mobile learning platforms in particular, and work related to how individuals learn.

2.1 American Sign Language and Deaf Culture

There are several issues surrounding American Sign Language and Deafness which relate to this dissertation. These issues must be understood in order to make appropriate decisions when developing and assessing technologies for this community.

For many people who are born deaf, their native language is American Sign Language (ASL) instead of English. Additionally, unlike English, ASL does not have an analogous written language form in common use. ASL, the dominant sign language of North America, is a visual-spatial language which uses different hand, face, and body gestures to communicate. ASL's grammar is different from English, and it uses a spatial structure for many linguistic constructs. For a full discussion of ASL linguistics, readers are referred to Klima and Bellugi [63] and Valli and Lucas [108].

Language acquisition problems can particularly affect deaf children born to hearing parents. There are different varieties of sign languages in the United States, and it is important to understand how ASL is different from other types of signing. The history of communication technologies for the deaf and how communication has historically been conducted between the deaf and hearing communities provide examples of how imperfect technologies can still affect great change.

2.1.1 Language Acquisition

There are many linguistic issues for individuals who are born deaf. Linguists have identified the existence of a "critical period" for language development – a period during which a child must be exposed to and immersed in a language to avoid delays in memory and linguistic development. In 1988, children born with a hearing loss were identified at an average age of 2.5–3 years old with many not being identified until age 5 or 6 [27], which is well beyond the usual age of language acquisition. Although originally thought to exist only for spoken languages, research has shown that this critical period also applies to ASL acquisition [75, 83]. Mayberry showed that even after 20 consecutive years of signing ASL, normative (i.e., deaf children of hearing parents) signers performed worse on generative tasks than did native (i.e., deaf children of Deaf parents) signers. In fact, performance on the tasks declined as a linear function of the age at which ASL was acquired [75]. This fact indicates a critical need for early immersion in ASL, and the effects of delaying language acquisition impact a deaf child throughout his or her life. For those whose primary mode of communication is sign language, this delayed language acquisition can lead to a lifetime of communication difficulties.

2.1.2 Deaf Children Born to Hearing Parents

Children who are born deaf to hearing parents are often not immersed in language in the same way hearing children of hearing parents (or deaf children of deaf parents) are and miss the critical period for language acquisition. Strong and Prinz [106] have shown that there is a strong relationship between early ASL proficiency and later English literacy for Deaf children. The relationship means that English is often an ineffective means of communication between deaf and hearing individuals unless ASL proficiency was achieved early in life.

Ninety percent of deaf children are born to hearing parents who may not know sign language [42, 78]. Often these children's only exposure to language is signing at school, whereas hearing children are immersed in spoken English from birth. By two years of age, hearing children learning a spoken language are combining words in expressive communication [110]. By one and a half years, deaf children of Deaf parents are also combining signs to communicate [7]. A third group, deaf children of hearing parents, develop language in the same sequence as the first two groups but at a much slower pace.

Parents play a critical role in their child's language acquisition. Social interactions with the child allow the child to develop language and make sense of the world around them [87]. Parents are often confused by the choices, both medical and social, they must make for their child, and this confusion can lead to a delay in language. As Bailes, Erting, Erting, and Thumann-Prezioso point out:

> Many parents, because they are not fully aware of what is at stake and thus accept the long-standing but erroneous claim that acquiring a signed language will impede speech development, seek resources devoted to supporting spoken language learning. If they decide they want to use a signed language, they soon realize they are unprepared to fully meet their child's immediate linguistic and cognitive needs. Meanwhile, every day that goes by is another day these deaf children live without the opportunity to acquire language because they are not exposed to a natural language that is fully accessible to them through their eyes. [7]

The slower linguistic development of deaf children born to hearing parents has been attributed to incomplete language models and lack of daily interaction using a language [104, 51]. Studies have linked delayed language acquisition with delayed short term memory development (now more commonly referred to as working memory) [50]. Due to this delayed linguistic development, the fact that English is a second language to many deaf students, and a myriad of other factors, the average 17–18 year-old deaf student reads English at a fourth grade level [58].

2.1.3 ASL, MCE, and Home Sign

There are many different gesture based communication systems in use in the United States today. In order to focus the scope of this work, it is useful to explain the differences and distinguish among the systems.

Manually Coded English (MCE) and its subset, Signed Exact English (SEE), were invented as a way to make English accessible to Deaf children [35]. Usually, MCE is based on the English language and translates a sentence word–for–word using signs. MCE uses gestures similar to ASL, but often replaces the hand shape with the ASL hand shape of the first letter of the English word thus "initializing" the word. Additionally, MCE often adds suffixes such as *-ed* or *-ing* to verbs. It also introduces signs for articles such as *a*, *an*, and *the*. While common in English, none of these linguistic constructs are found in ASL [35]. MCE generates phrases that require more time to utter than either spoken English or ASL. Thus, it can be quite slow and cumbersome for conversational use [63].

Home sign is the term used for gesture based communication systems which are invented by individual families. Families may spontaneously develop a series of signs for use with their children and family. Home sign is usually family specific and can be understood only by members of a single family. [43, 41].

Neither home sign nor MCE are complete languages as defined by linguists. According to Yule [112], languages, as defined by linguists, share common features:

1. **Intentionality** – Language is used for intentional communication. For example, sneezing or coughing are not communication and are not part of a language.

2. **Displacement** – We can communicate about past and future as well as present time. We can also talk about other places, abstract ideas, or people who are

not present.

3. **Arbitrariness** – On the whole, the symbols we combine and use to communicate are not connected to the things they represent. For example, the letters "s," "u," and "n" combine to form the word "sun" which represents a fiery ball of gas millions of miles away.

4. **Creativity** – Language can be used to create new communication. The production of a language is infinite.

5. **Cultural Transmission** – Language is transmitted from person to person. It is not inherited.

6. **Discreteness** – Languages can be broken up into smaller units. The units (e.g., sounds or signs) used in a language lead to a difference in meaning.

7. **Duality** – Languages are organized on two levels: the physical and the meaningful. The physical level of a language refers to the structure, syntax and rules of a language, while the meaningful refers to the ideas and common meaning a language conveys.

ASL is linguistically complete, as are most signed languages around the world. As a linguistically complete language, ASL includes both linguistic concepts such as those listed above and linguistic constructs such as rules of syntax, rules of semantic interpretation, phonology, and a lexicon [29].

Conversational rates (both spoken and signed) range from between 180-200 wpm. Research has shown that hearing people speak about 4 to 5 words per second and deaf individuals sign about 2 to 3 signs per second. However, the information conveyed by both is about 1 to 2 "propositions" or conceptual units per second and is therefore equivalent [13, 38].

ASL is more grammatically similar to French than English, as it was brought to the US by a graduate of a French deaf school, Laurent Clerc. Along with Thomas Gallaudet, he founded the American School for the Deaf (originally known as the American Asylum for Deaf-Mutes). Although Clerc originally instructed in French Sign Language, his students incorporated and morphed signs. As time passed, this collection of signs became ASL, a distinct entity from French Sign Language [66, 35].

2.1.4 Deaf vs. deaf

At this point, it is worth a slight digression to explain the usage of the word "deaf" in this dissertation. Deaf can have several meanings, but here I concentrate on the medical and cultural definitions. Medical deafness focuses on the severity and cause of a hearing loss. This classification is often used in legal and medical terminology and is denoted with a lowercase 'd,' "deaf." The cultural definition of Deaf, with an uppercase 'D,' is a voluntary classification and refers to the community formed by individuals whose primary method of communication is ASL. Using a language other than English differentiates these individuals from the larger hearing-based population in the US and allows them to form strong bonds of community. Where appropriate in this dissertation, I have used "deaf" to refer to individuals who are medically deaf and "Deaf" to refer to individuals who identify themselves as associating with the Deaf community. At the Atlanta Area School for the Deaf, where some of my research occurred, most students consider themselves Deaf, and thus I also use the term "Deaf" to refer to a majority of participants in the collective sense.

The identification of "Deaf" is a complex issue. The Deaf community is a very strong cultural community which often grows out of ties formed at residential schools. In the early 1800's schools for the deaf were seen as a place to save deaf individuals in the evangelical Protestant tradition. Thus, intellectual pursuits were seen as secondary to religious ones. Gallaudet's profession as a minister affected his views on deaf individuals, and the language used to teach religious instruction was seen as a secondary concern. If sign language could be used to save deaf individual's souls, then sign language would be used to communicate in Gallaudet's school [12].

After the US Civil War, sign language fell out of favor. It was seen as important for all individuals to participate in the common context of life as an American, and having a different, non-English language was seen as a deficiency in deaf people.

During the late 1800's and early 1900's, sign language was actively forbidden by many schools which focused on educating deaf children. Alexander Graham Bell was prominently involved in the education of deaf individuals during this time period. Bell's wife was deaf, and it is assumed by many that his inventions of the microphone and telephone were designed to assist her and the deaf community [59]. Bell was a proponent of removing sign languages entirely and teaching deaf children to speak and lip read so as to function in an oral environment. This education method became known as "oralism" in opposition to "manualism" or signing for communication [59]. This divide persisted for many years. Deaf individuals were not taught to sign, nor were deaf teachers placed in classrooms. Individuals who signed were seen as foreign or peculiar and thus, isolated from the larger hearing community. There was also an effort by Alexander Graham Bell and others to remove deaf students from residential schools as these were seen as places where signing could thrive. By arguing that deaf children should be with their parents and not in residential schools, students were effectively removed from the deaf culture and kept in an oral education [12].

In spite of an education climate that was outright hostile to ASL and signing, many deaf individuals learned some sign through family connections or deaf clubs. Deaf social clubs were a crucial factor in keeping ASL alive and growing as a language [85, 59]. Deaf social clubs hosted performances, discussion, and debates in sign. The activities allowed deaf individuals to socialize and form tightly knit bonds to one another. These clubs formed the basis of the Deaf community.

In the 1950's, William Stokoe was hired as a professor in the English department at Gallaudet. In keeping with the prevailing pedagogical bias against ASL, he was hearing. However, he was fascinated with ASL and began to study it as a linguist. Stokoe's seminal work [105] began to turn the tide away from oralism by arguing that ASL was a full and complete language similar to English, French, or any other natural language.

As Stokoe's research became accepted by the hearing world and ASL became accepted by the hearing world as a language worthy of study and recognition, Deaf culture (which revolved around the use of ASL) could once again flourish. Deaf students became free to use ASL in face–to–face conversations without fear of education reprisals [59].

2.1.5 Communication in the Deaf Community

When communicating with a hearing person remotely, deaf individuals have traditionally been required to use a Teletypewriter (TTY), also known as a Telecommunication Device for the Deaf (TDD). If each person has a TTY, they can type messages back and forth using TTYs connected to standard telephone lines. To communicate with someone who does not own a TTY requires a third party relay operator who transcodes the messages from voice to electronic transmission and back. Figure 1 shows the flow of a conversation conducted via relay.

However, the development of a TTY which could be used with the phone system was in and of itself a triumph for the Deaf community. As the popularity of the telephone increased, deaf people were shut out of this means of communicating with one another. In 1964, several California engineers, two of whom were deaf, began to experiment with TTY but were met with resistance by AT&T which had a complete monopoly of the phone system at that time. As a result, they had trouble obtaining equipment and were forced to use an acoustic coupler rather than a direct connection

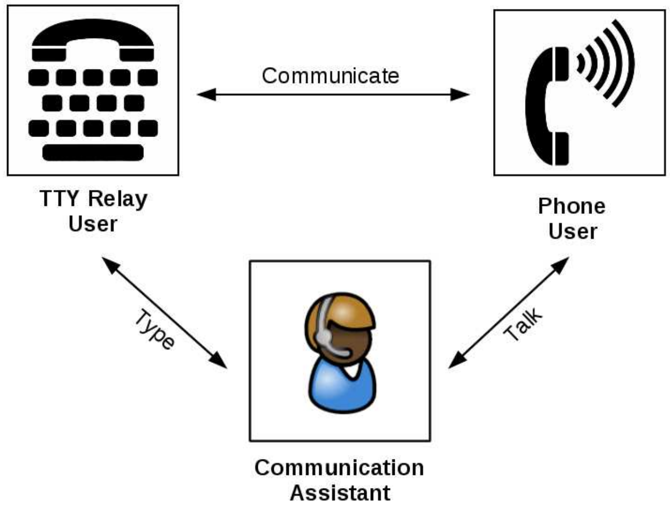


Figure 1: Conversation Between a Deaf and Hearing Individual via a Relay Operator

to the phone line [64]. Additionally, some in the Deaf community were against TTY use because it went against the currently prevailing oralism tradition. Using typing to communicate rather than speaking or lipreading was seen as an inferior way to communicate. In the late 1960's the Federal Communication Commission (FCC) forced AT&T to make their network accessible, and by the mid 1970's TTY's had been made portable. This ruling led to widespread adoption in the deaf community and the advent of a national relay service [64, 59].

Lang notes that developments and advances in technology often come with a price for the deaf.

> "Advances" in voice telephony led to a ninety-year delay in access to the telephone for deaf people. "Advances" in adding the sound track to silent movies led to more than forty years of lost access to films [65].

As mentioned earlier, deaf social clubs provided a means for Deaf individuals to socialize and communicate with one another. As phone use became more prevalent with the widespread adoption of TTYs and a national relay service, many in the deaf community worried that these communication technologies were eroding Deaf culture

[85]. However, the ease of communication and the ability to communicate remotely without a hearing intermediary was a powerful tool to many Deaf individuals.

Recent technologies have been developed which can replace the TTY device. However, these new technologies still require the use of a relay operator when communicating with a hearing individual. For example, IP Relay allows deaf individuals to utilize newer technologies such as computers, Instant Messaging, or wireless devices instead of TTYs. Video relays are also becoming more mainstream and offer a video link between the deaf signer and an operator who is also fluent in sign. Video provides a more expressive means of communication and allows for the inclusion of subtle communication nuances which cannot be communicated via text. While relay operators are governed by a strict code of ethics, many deaf individuals still dislike using an operator and the need to involve a third party in a private conversation.

Despite many of the challenges and limitations of technology, the deaf community have been fast adopters of telecommunication technologies [107]. The quick adoption and evolution of new technologies in the deaf community provide a rich testing ground for mobile applications which could enhance or facilitate communication for deaf individuals.

A 2002 survey by the National Association of the Deaf [15] found that 75% of their respondents used IM at home and 35% reported using it at work. Additionally, 97% used email at home and 79% used email at work. However, it is worth noting that the demographics of respondents to this study were considerably skewed as only 9% of their 884 respondents were under 25 years old and 66% were college graduates.

2.1.6 Electronic Communication

In 1982 Barbara Wagreich, a deaf–blind computer professional wrote an article about the possibilities of a new technology, email, and how it might prove beneficial for people with disabilities, particularly the deaf [109]. Her study, conducted from

1978–1981 involved distributing terminals and access to electronic message servers to both deaf and hearing individuals in the greater Boston area, facilitating easy and asynchronous communication. She found that email was not only a formal medium for business meetings and communications, but also an informal tool for maintaining friendships and furthering acquaintances. The participants were overwhelmingly enthusiastic about the new communication medium, and Wagreich concluded that email was a tool that many deaf individuals could use to communicate with their hearing friends, families, and acquaintances.

Many years later, email is complemented by the newer technologies of Instant Messaging (IM) and Short Message Service (SMS) or text messaging. These technologies have been widely adopted by the hearing population [79, 60, 68], in particular by teenagers [96, 47, 46]. It was unclear how these new methods of communication are being used by the Deaf community. While many sources report that the Deaf are often early adopters of technology [93, 52, 8], there is very little formalized work studying the use of communication technologies.

In their work on SMS, Barkhuus and Vallgårda found that SMS was used to communicate mainly among friends and significant others, but IM was used for a wider range of conversational partners [9]. Grinter and Eldgridge found that teens adopted mobile messaging for a variety of reasons, prominently, to coordinate conversations via another medium [46, 47]. In a market report on mobile technologies, Blinkoff and Barranca found three central themes to users' desires: manage relationships, experience the unexpected, and avoid mobile stress [14]. In a study on teenage communication preferences, Schiano *et al.* found that home phones were the most common communication medium [96]. Much of this work points to electronic communication from mobile devices (usually SMS) as being a transition medium which is used to coordinate voice communications via mobile or "landline" phones. This usage pattern points to a reliance on the voice telephony capabilities which many deaf individuals do not use, and this situation presents a significant difference between hearing and deaf teenagers. While it might seem that work surveying the use of electronic communication methods among hearing teenagers [46, 47, 96] would be sufficient to understand Deaf teens' communication, there are distinct linguistic differences between hearing and Deaf teens which warrant further exploration.

Unlike English, ASL does not have a commonly used written form. It is languages's written form which enables electronic communication such as SMS, IM, email, etc. to exist. Deaf reliance on a medium which requires the use of a foreign language seems improbable. However, Bakken [8] found that Deaf teens in Norway relied on SMS messaging for building social networks, maintaining those networks, and for keeping abreast of trends and gossip. However, Bakken's work may not generalize to the United States population due to the differences between Europe and the US in SMS use and mobile device adoption.

A study by Power *et al.* in 2007, investigated deaf Germans' use of technology and found that 96% of their survey respondents had access to a mobile phone and used text messaging. Most respondents reported that they used their mobile to send an SMS message every day [92]. Only 23% of responders reported having a TTY, and only 69% reported using a computer.

Most previous studies have a population bias in that the participants who responded tended to be highly educated deaf individuals, which could affect the comfort level and technological preferences of the participants. Likewise, most studies focussed on adults or deaf individuals who were out of school. Several early studies of the use of technology in the deaf community focussed on deaf communities in other countries such as Australia [93], Germany [92], and Norway [8]. While these studies are interesting and applicable, differences in social services, early intervention practices, telephony infrastructure, and telecommunication providers introduce large variances in the findings. Given the different characteristics of hearing and Deaf teenagers, we designed a research study to investigate the different ways in which electronic communication technologies were used in the Deaf community. This study is detailed in Chapter 3.

2.2 Second Language Learning

As discussed earlier in this chapter (sections 2.1.1 and 2.1.2), many deaf children born to hearing parents are not immersed in language until they are beyond the critical period for language acquisition. This delay can impair English fluency later in life. In order to address this issue, it is crucial to supplement and support early language learning from the source of the majority of a child's language: his or her parents.

In order to address early sign language acquisition, we must consider the two sides to the problem. One strategy to combat this problem is to enhance young children's signing skills. The other strategy is to increase their immersion in ASL by teaching their parents ASL. There are many projects which address ASL learning in children (e.g., [54, 17, 33, 34]). For example, the CopyCat project [54, 67, 17] focusses on improving a child's command of ASL by allowing them to control a game using sign language recognition. However, there are no projects which focus on a technological intervention for hearing parents of deaf children. Thus, I choose to concentrate on the task of teaching parents basic ASL vocabulary.

Seventy-five percent of hearing parents of deaf children say they use some sign with their child. Yet, a large number never become fluent signers. Based on the National Parent Project, of the 75% of hearing parents that sign, only one-third report that both parents have either good or excellent signing skills [76]. Thus, only about 25% of deaf children with hearing parents have what may be considered "good" language models and communicative partners in their own family.

While vocabulary knowledge alone does not constitute fluency or command of ASL, Marchman and Bates have shown that a knowledge of approximately 150 words is enough to increase the rate at which new words and grammatical skills are acquired [74]. Further, research has shown that some level of ASL exposure for deaf children is vastly superior to no language exposure [101]. Thus, even "survival level" signing is a worthwhile endeavor for families with deaf children.

Caccamise and Newell [21, 82] developed the Sign Language Proficieny Interview (originally known as the Sign Communication Proficiency Interview), or SLPI, as a way to assess ASL skills in a conversational format. This test defines survival level ASL as,

> Able to satisfy basic survival needs in social and/or work situations. Can ask and answer basic questions and has some skills in creating sign utterances based on learned/memorized sign vocabulary. Can get into, through, and out of simple survival situations.

In a 2002 study, Freeman *et al.* studied hearing parents and caregivers of deaf children. They examined how the parents and caregivers obtained early intervention services for their children and helped their children learn language. Freeman notes, "parents play a major role in the child's acquisition of language given that development of communication and language peaks between ages 1 and 4 years" [39]. They make 10 recommendations to help parents facilitate language learning in their deaf or hard-of-hearing child. Those recommendations include:

1. Using signs, gestures, facial expression, and voice to help convey what you are trying to tell your child.

2. Getting down to your child's eye level when you can.

3. Reducing background noise (e.g., television, radio).

4. Signing and talking about things your child is doing (e.g., "Oh, you're playing with your cars").

5. Signing and talking about things your child sees and hears (e.g., "That's a fire truck").

6. Keeping language short, simple, and direct.

7. Being a speech and language model. The parent should get the child's attention before beginning to communicate, by tapping the child on the shoulder, or signing and saying, "Look at me."

8. Taking up sign language instruction courses as needed.

9. Seeking out accurate, nonjudgmental information regarding deafness, deaf education, and communication methods and options.

10. Finding a peer-parent – a parent of an older child who has been there.

The system I developed (described in Chapter 5) helps give parents the vocabulary necessary to "sign and voice daily events, create stories, and describe what is happening during daily routines" [39].

Many parents of children with disabilities face additional difficulties beyond their child's disability. Research has shown that the cost of treating children with disabilities greatly exceeds the healthcare costs of children without disabilities over the course of their lifetimes [84]. Even with insurance, some of this cost is born by the parents in terms of rehabilitation costs, lost time at work, or technological interventions. Additionally, children with disabilities are statistically more likely to be raised by a single, female caregiver [26]. These statistics point to a lack of resources which can make traditional ASL classes difficult for the parent. Traditional classes often take significant time which can take time away from hearing siblings, family, or work.

Mobile devices provide a unique opportunity to address this problem. Mobile devices are ubiquitous and often within arms reach [90]. Additionally, mobile phones are offering increased video capabilities which are well suited to a visual language such as ASL. Given the limited amounts of time that parents of deaf children may have, it is crucial to deliver content in an optimal manner as to maximize the learning over time.

2.3 Related Work on Learning and the Spacing Effect

It is a well accepted fact in psychology that learning over time is more effective than massed practice. That is, massed content delivery (an item which is repeated several times in succession) is less effective than content which is spaced over time or interspersed with other items [31]. This phenomena, known as the spacing effect, was first noted by Ebbinghaus in 1885 [32] and has been the subject of over 300 formal, published studies [19]. Through teaching himself nonsense syllables and testing his retention rate, Ebbinghaus proved that the relationship between the amount learned and the recall rate the next day was roughly linear. Simplistically, learning can be a function of the number of presentations of the material. This theory is known as the *total time hypothesis*.

However, there has been much work which shows that distribution of the presentation of material affects learning rates. *Distributed practice* during which presentation of material is spaced over a time period can be contrasted with *massed practice* in which an item is presented multiple times in a row.

Baddeley studied the effects of distributed and massed practice and summarizes his findings with the statement, “As far as learning is concerned, ’little and often’ is an excellent precept” [5]. He proved this concept studying postal workers learning new keyboard layouts and typing skills. He divided subjects into three conditions receiving either one, two, or four hours of instruction per day. Condition 1 received one one-hour session, condition 2 received either two one-hour sessions or one two-hour session, and condition 3 received two two-hour sessions.

> "The postmen who worked for only one hour a day learned the keyboard on fewer hours of training and improved their performance more rapidly than those who trained for two hours a day and they in turn learned more rapidly than those who trained for four hours per day" [5].

Much of the psychological literature is focussed on performing experiments to validate a particular heuristic of learning. Very few works propose models which can be used, manipulated, or applied by other researchers. One exception to this generalization is the work of Atkinson [3]. Atkinson proposed a three state model of learning that classifies objects in one of three memory states: Long Term Memory (LTM), Short Term Memory (STM), and a Unknown/Forgotten state (U/F). Atkinson's model is for paired-associate presentation of second language vocabulary. For example, subjects were presented pairs of vocabulary words consisting of one word in the language they were attempting to learn and the corresponding word in their native language (e.g., "casa" and "house").

Table 1: Atkinson Model of Presented Item

Prior \ Post	LTM	STM	U/F
LTM	1	0	0
STM	x	1-x	0
U/F	y	z	1-y-z

Table 2: Atkinson Model of Item to be Learned

Prior \ Post	LTM	STM	U/F
LTM	1	0	0
STM	0	1-f	f
U/F	0	0	1

Tables 1 and 2 summarize the model. In both tables, the row header represents the state before presentation and the column header represents the state after presentation. The parameters x, y, and z are "parameters that govern the acquisition process" and f is a parameter that governs forgetting [3] for a particular item. Table 1 is applied when an item is presented. For example, if an item is in STM, then the probability of it being transfered to state LTM is x. Table 2 is applied when a different item is presented and simulates interference forgetting, the forgetting that happens when newly learned material blocks older material. In his 1964 work [4], Atkinson described how to conduct a pilot test using a randomly ordered list of paired vocabulary words to establish the x, y, z, and f parameters.

Atkinson tested his model in two different configurations. The first configuration assumed all items in the list are of an equal level of difficulty. The second configuration used an experimentally determined level of difficulty per word. He compared these two configurations against a random order and against a self-selection condition in which subjects were allowed to determine what pair they wished to study. Participants were presented lists in the varying orders to study and then tested after 7-8 days. After the 7-8 days during which participants were not trained on the words, Atkinson found that participants using his model with individual difficulties recalled significantly more translations than all other strategies (80% recall). Participants using the self-selection condition (54% recall) and the Atkinson model with equal difficulties (58% recall) recalled an approximately equal number of translations. Subjects presented with vocabulary in a random order scored the lowest, recalling only 38%.

2.4 m-learning

Learning on mobile devices, or "m-learning" is becoming a more prominent research area as mobile devices become more prevalent, faster, and capable of many functions. As an example of the quickly expanding mobile device capabilities, a study by NetMarketShare estimates that as of January 2010, mobile web browsing accounts for

1.48% of all web browsing and has more than doubled over the past 12 months [81].

Below, I survey several distinct areas in m-learning research and explain how my research differs.

2.4.1 m-learning in Traditional Environments

Many applications involving m-learning focus on traditional educational environments such as classrooms or field trips. For example, Explore! [28] is an m-learning system that uses a game to “help middle school students to acquire historical notions while visiting archaeological parks.” This system compared two groups of students playing a game, one with a mobile device and one without. They found that the sequential nature of actions using a mobile device hindered the game play and the researchers reached the conclusion that mobile games require significant flexibility. They also found no significant differences between the two user populations, indicating that mobile learning systems did not distract the students from the overall pedagogical purpose of the game. However, the group using the technology was more motivated than the one without.

Likewise, Sànchez and Salinas studied using a mobile gaming system to help eighth graders learn science concepts. They used a trivia software program on a pocketPC during a trip to a science museum. They found no statistical differences between the group using their system and a control group [95].

Lindquist *et al.* explored using mobile phones in active learning exercises in CS courses. Students used mobile phones to send answers to short questions via SMS or MMS during class. They found that MMS photo answers were easier to use than text entry SMS answers, but students had concerns about the recurring costs of using messaging services for class exercises. They also provided some recommendations for designers of mobile learning systems including flexibility in input and being cognizant of continually changing mobile phone input and output specifications [69].

Sharples *et al.* [100] developed requirements to support Contextual Life-long Learning (CoLL) technologies. They hold that, "learning is not confined to pre-specified times or places, but happens whenever there is a break in the flow of routine daily performance" and "formal education cannot provide people with all the knowledge and skills they need to prosper throughout a lifetime." The requirements to support life-long learning include technologies that are highly portable, individual (adapting to abilities), unobtrusive, available anywhere, adaptable (to context), persistent, useful, and easy to use. In Sharples's work, children learn in the traditional context of school and interface with a web client to organize the material. However, in many ways, mobile and portable technologies have great potential to support life-long learning.

2.4.2 Learning via Mobile Phones

When discussing the "mobile" aspect of m-learning, Sharples *et al.* point out that there are many different aspects of mobility in learning [99]. These include:

1. Mobility in physical space: People are on-the-go and location "may be relevant to learning, or merely a backdrop."

2. Mobility of technology: learning platforms are portable and lightweight, and they are meant to be transported from place to place.

3. Mobility in conceptual space: a learner's attention quickly moves from topic to topic driven by a variety of topics.

4. Mobility in social space: learners "perform within various social groups" which may include "family, office, or classroom contexts."

5. Learning dispersed over time: learning is cumulative and requires multiple presentations across both formal and informal contexts.

The research proposed in this dissertation takes advantage of many of these differing definitions of mobility. In particular, it is designed to investigate items 2 and 5: mobility of technology and learning dispersed over time. The study described in Chapter 5 investigates these two variables within the context of learning ASL.

Paredes *et al.* used handheld computers in a field trip exercise assigned by teachers to assist in learning Japanese [88]. The exchange students went on a scavenger hunt during which they interacted with native Japanese speakers. The handheld devices allowed them to record their experiences and language interactions for further reflection during class time. However, this study did not evaluate language learning of the participants; the researchers evaluated the platform and the software implementation via Likert scale questionnaires and interviews.

Kam *et al.* used cell phones and mobile gaming to work with students in India. Their focus was on English as a Second Language (ESL) learning, and they conducted five field studies over 4 years. They noted that students did better when the researchers distinguished between gaming for fun or pleasure and gaming for educational purposes. Much of their game time was situated in a classroom to reinforce the educational nature of the activities and encourage the children to take them seriously. Their participants exhibited improvements in their language abilities as measured via a post-test [62].

Liu *et al.* conducted a survey of m-learning in China. While the number of participants was too low to draw broad conclusions, they nonetheless had several interesting findings. In particular, they found that the ability to participate in m-learning services influenced a quarter of their respondents when deciding what phone to purchase. Fifty percent of their respondents said that they had already used m-learning services on their mobile phone, and 85% said they would like to use such services in the future. They also found that most of their respondents used the m-learning features and services when in a "stable environment" rather then when mobile. Additionally, after 7pm was the time that most respondents chose for conducting m-learning activities [71]. However, it is important to note that their survey consisted of only 65 students at a specific Chinese university. These students may have been particularly motivated to use m-learning services due to the emphasis on learning English and other school activities.

As discussed in the conceptual framework proposed by Parsons et al. [5] the most important features in a mobile environment are: 1) mobility itself, 2) user profiles and roles, 3) device capabilities, and 4) communication support.

Parsons *et al.* defined an m-learning design framework [89]. They list the most important features of an m-learning system as:

- the inherent mobility of a device
- user profiles and roles
- the capabilities of the device
- communication support

The framework can be applied to the design of mobile systems in order to understand whether the objectives of the system are helped or hindered by the design choices made. This framework was used to help narrow potential mobile platforms in the study described in Chapter 5. Moreover, the m-learning experience was designed with these factors in mind.

Thornton and Houser have specifically addressed the issue of learning languages via mobile platforms. They used mobile phones to deliver English vocabulary to Japanese students. They compared performance of students who received short messages ($<$ 100 words) and longer, multi-sentence definitions. They found no statistical advantage to the longer definitions. For evaluation purposes, they compared their mobile content delivery against a PC based implementation and paper handouts with the same content. In statistically significant measurements, students learned more when using the mobile based system when compared to either mobile web-based applications (40% more words learned using mobile platform) or traditional, paper-based learning materials (24% more words learned using mobile platform). 71% of students preferred the mobile system to the PC based system.

More generally, Groot's research on CAVOCA (Computer Assisted VOCabulary Acquisition) [49] details learning 2nd languages on the computer, although not in a mobile situation. This system presented a word and several definitions. The system then used the word in a variety of sentences to allow the user to learn via context. A series of trials using the word in a sentence were presented for the user to determine whether the word was used correctly or not. The user was then presented with several paragraphs from national news sources using the word. In the last section of the program, the user must generate the word from a fill–in–the–blank question.

2.5 Conclusion

Given the many language and communication issues that deaf individuals may face, this dissertation discusses several studies which address different aspects of the communication barrier. The three studies detailed in Chapters 3–5 show different ways of investigating the communication barriers and utilizing mobile phones' properties to reduce these barriers. In the next chapter, I discuss communication patterns between Deaf teenagers, their friends, and their families along with the role that electronic and mobile communications play in these relationships. This study highlights the isolation that occurs when current communication practices are insufficient or break down due to distance or technological limitations. I also discuss communication barriers between the hearing individuals with whom the deaf participants wanted to communicate and the available communication modalities.

CHAPTER 3

DEAF TEENS' COMMUNICATION AND ISOLATION

In the Spring of 2005, I conducted a study at the Atlanta Area School for the Deaf (AASD). This study investigated the role different communication methods played in Deaf teenagers' lives. It also allowed examination of the stakeholders in the teens' lives and the different communication methods used when communicating with these different groups of people. This chapter details the study methodology and participants (Section 3.1) and identifies some of the major findings. In Section 3.2, I discuss further in-depth analysis of the teens' communication patterns and stakeholders (3.2.3) to identify communication breakdowns.

3.1 Study Design and Methodology

I recruited twelve participants from the Atlanta Area School for the Deaf (AASD). AASD is a publicly funded school for students who are deaf, and its enrollment area covers the majority of North Georgia, including the Atlanta metro area.

This school was selected as the researchers have an existing relationship with the school, and it is one of only two schools in the state of Georgia with an exclusively deaf population. As all the students who attend AASD are deaf, recruiting them allowed the researcher to observe intra-cultural communication (i.e., communication between deaf individuals) as well as inter-cultural communication (i.e., communication between deaf and hearing individuals).

The participants were all recruited from a specific grade at AASD. These students

Parts published in [55]: V. Henderson-Summet *et al.* , "Electronic Communication: Themes from a Case Study of the Deaf Community," in *Proceedings of INTERACT, 2007* knew each other and had long established ties to each other. AASD has small classes, so these students had often progressed through multiple grades together. These long-standing ties meant that the researchers could study established communication patterns rather than new, emerging relationships.

The students were offered financial incentive to participate and complete the entire study. Students were paid a total of $25 to participate; however, I offered $5 for the first part of the study, $10 for completing the second part, and $10 for completing the third and final interview.

3.1.1 Demographic Data

The twelve participants ranged in age from 14–17 with an average age of 16. There were six females and six males, although one male student left the study after the first activity. The students were also asked if they preferred English or ASL as a language for communication. Table 3 summarizes the participants' demographic data.

Table 3: Participant Data

Participant	**Gender**	**Language Preference**
P1*	F	Both
P2	M	ASL
P3	M	ASL
P4*	F	English
P5*	M	ASL
P6	F	ASL
P7	F	ASL
P8	F	ASL
P9*	M	English
P10	M	Both
P11	F	English
P12	M	Both

All the participants' parents or caregivers were hearing. Several of the students had some residual hearing, but not enough to make a school with auditory instruction feasible. Four students (one of whom left the study) had enough residual hearing as to have some speech and to use oral communication with some degree of success. These students are marked with an "*" in Table 3.

On average, the students lived almost 30 miles from AASD. The maximum distance between students was almost 100 miles. This distance often precluded the teenagers from associating with their social community when not attending school. The students had a strong community at school; however, these students often could not communicate easily with the hearing teens in their neighborhood or local hang-out spots. Because of this, they often felt isolated. Figure 2 shows a map with the locations of each student (red markers) and the location of the school (yellow marker).

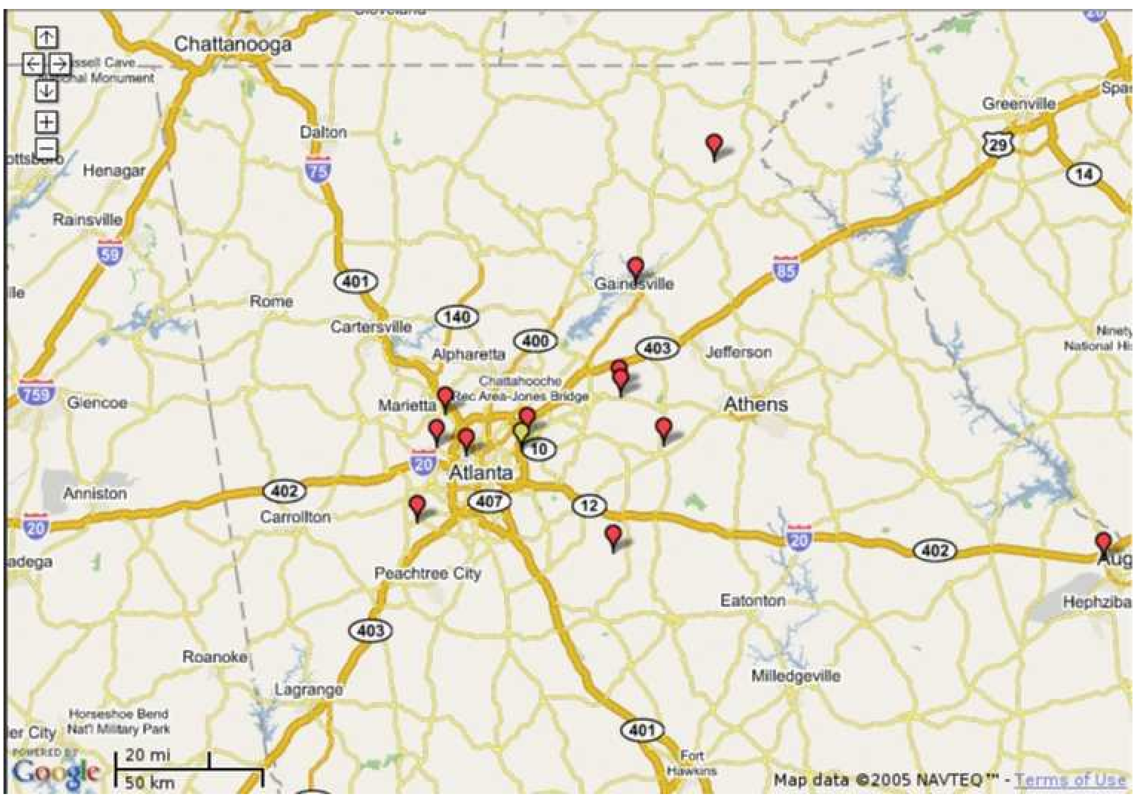


Figure 2: Map of North Georgia Showing Geographic Distribution of Participants

3.1.2 Methodology Development

The study had three phases designed to survey different aspects of the role of electronic communication in Deaf teenager's lives. The three phases were a social mapping activity, a diary study, and discussion groups, and these are described in more detail below.

3.1.2.1 Social Mapping Activity

The social mapping activity was designed to elicit the teenagers' social networks and delineate the boundaries of their communication patterns. Knowing these boundaries helps highlight the choices the participants chose to make regarding communication. Specifically, this activity studied with whom they communicated, whether those people were hearing or deaf, and the main techniques for communicating with those people.

Figure 3: Example of a Social Network

While this study was based upon the work of Smith, Rodgers, and Brady [102], it was more structured to make it easier for the teens to understand. During pilot testing with teachers at AASD, they expressed concern that the activity was too loosely structured and nebulous for their students. Thus the original "draw your social network" activity was redesigned into a series of concrete steps for the students to follow. These steps built upon each other and eventually a complete picture of the networks would emerge.

The teens were given a large sheet of paper (easel size) and pencils. They were first asked to list everyone with whom they felt it was important to communicate. The definition of "important" was largely left up the participants. However, I asked them to consider if they would be upset or unhappy if they could not communicate with a person. They were also told that they could think about their list overnight and add people to it the next day if they felt they had forgotten someone. Second, the students were asked to label each person on their list as “hearing” or “deaf.” After that, they were asked to go through their list and write down how they would communicate with that person if they were face-to-face with that person. In a similar fashion, they were asked to write down how they would communicate if the person was not in the same room or was somewhere far away. They were asked to write N/A or “Nothing” if they did not communicate with a contact. Lastly, the students were given packs of colored markers and asked to group and categorize their contacts in any way they felt appropriate, for example “Family” and “School Friends” were common categories. Figure 3 shows an example of a social network similar to those generated by the students, with the contacts and their characteristics listed in the top half of the figure and the categories listed on the bottom.

3.1.2.2 Diary Study

The diary study phase involved students recording when and where they used electronic communication. Based on the work of Grinter and Eldridge [47, 46], it was designed to give a clear picture of the teenagers’ lives on a daily basis and how electronic communication fit into it. Again, the activity had to be slightly restructured to give the students a concrete perception of what to record and when. Whereas Grinter and Eldridge’s participants noted incoming and outgoing messages, I provided the students with a more calendar-like template to facilitate the recording of activities. The students were told that if they were in doubt about an activity to record it.

This form began at 5am with 12 hours on the front of the page and 12 hours on the back of the page, thus allowing the students to record one day per page. In the middle of the page is a list of times. Each hour is subdivided into four 15-minute blocks. To the left of the list of times is a space for the students to record their physical location. In a space to the right of the list of times they were told to record their electronic communication. Figure 4 gives an example of the form the students used. The full form is included in Appendix B.

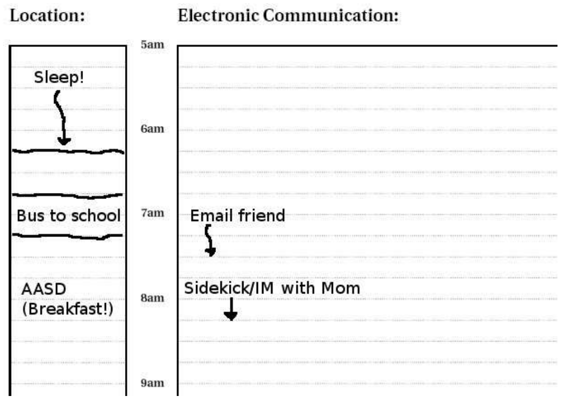


Figure 4: Example of Daily Diary Form

The students were given one sheet during a class period and completed a log form for the previous day with assistance from the researcher. This process allowed the students to ask questions and receive immediate feedback rather attempting to work with just an explanation and an example. After this in–class example, they were given a booklet containing 7 forms, one for each day for a week. They were instructed to fill out a form each night for a week and return the booklets to their teachers at the end of the week.

3.1.2.3 Semi-structured Interviews

After the social mapping activity and week-long diary activity, the researcher facilitated a group discussion. The discussions were conducted in two groups based on the students' class section. Thus, they were with people they already knew and with whom they were comfortable talking. The questions were largely determined using data from the social maps and weekly diaries. I also allowed the students to highlight other topics they felt relevant. Because an interpreter was voicing for the students, audio recording was insufficient for identifying the speaker. Thus, I chose to video record all the interviews. The two group discussions lasted approximately 45 minutes to 1 hour each.

After collecting this data I coded it according to several different schemes. I first coded the interview data by device used and looked for patterns and themes that were specific to a particular technology. I then recoded the data by more general categories such as “Who, what, when, where, why?” and looked for themes which emerged regardless of the device or technology used. The data from the social maps and journals was then used to help support or discount the emerging themes. The preliminary findings were reviewed by external researchers for further validation.

3.1.3 Methodological Issues and Implications to Study Design

This study presented several interesting logistical issues. The most prominent issue was the language barrier between the students and researcher. Additionally, this study relied on a visual language, ASL, and involved the participants recording study data in their second language of written English. Moreover, the delayed linguistic development of many of these students required instructions to be short, with concrete examples and subtasks to ensure completion.

3.1.3.1 Working with Second Language Issues

The first language of most of the participants was ASL, not English. This barrier meant that someone fluent in ASL and English needed to be present at all times to facilitate communication between the researcher and students. In most cases, this person was the classroom teacher from AASD; however, in one case, a certified interpreter helped with the interview sessions. This language barrier did not seem to hinder rapport. While not fluent, the researcher is conversational in ASL. The teens seemed to see these conversational attempts as goodwill gestures and tried to communicate with the researcher via ASL when possible.

3.1.3.2 Task Analysis and Decomposition

As mentioned earlier, many of the students have linguistic difficulties relating to their deafness. The linguistic complexity of the three tasks had to be carefully considered so as not to present methodological problems. Many of the students might have problems with complex, written English instructions or tasks which were structured as open-ended activities due to short term memory issues and problems with understanding language. Before presenting an activity to the students, the researcher performed a task decomposition on it. This breakdown helped the students comprehend the activity and gave them measurable progress towards completing each activity.

Task decomposition is a standard technique used in teaching assistive technologies in a special education setting [20]. For example a high level goal was for each student to make a social map. However, the teacher felt this task would be overwhelming to the students when presented as an open-ended, free-form project. Instead, we defined the five subtasks detailed earlier, each with concrete and measurable goals. I found this practice invaluable. It not only helped the students complete the activity, but it also helped me define exactly what elements of the social map and network I was interested in learning more about and why.

3.1.3.3 Importance of Visual Attention

The importance of visual attention in ASL also presented a problem for the researcher during the interview portion of the study. The researcher tried to take notes, but abandoned the effort after the students perceived it as rude. ASL is a highly visual language. As the children could not hear the interviewer making acknowledgments such as "um-hum" or "yeah", which Brennan and Clarke have stated are crucial for developing a common ground and a mutual understanding [25], the students interpreted the interviewer looking down to take notes as a sign of not paying attention and became self-conscious about their answers. Given the difficulty in balancing rapport with exhaustive note taking, I chose rapport to be more important and relied on the videotapes of the interviews.

3.2 *Analysis*

Originally, the study set out to examine the overall space of electronic communication use including desktop computers, laptops, PDAs, mobile phones and any other electronic communication platform or device. However, it soon became clear that the majority of the teens used some kind of mobile device and the research was narrowed to reflect this preference.

In this section, I discuss the preferred methods of communication used by the teens. I then discuss communication infrastructure in the United States at the time of this study and how it affects the teens' choices. In Section 3.3, I then present five central themes of communication that I extracted from the data and present the teens' communication within the framework of those themes.

3.2.1 Electronic Communication Preferences

Communication technology for the deaf has evolved from telephones and relay or TTY to email to mobile platforms. These technologies level the playing field by lowering the barrier of participation between the deaf and hearing populations and allow the deaf to maintain their friendships in ways that are convenient and simple. When students needed to communicate with someone they knew to be nearby, they would usually seek them out in person. However, when unsure of a person's location or when the person was far away, the students usually turned to an electronic method of communication rather than the more traditional method of relay or TTY. TTYs (also called teletypewriters or TDDs) have been an accepted way for deaf individuals to communicate. If each person has a TTY, they can type messages back and forth using TTYs connected to standard telephone lines. To communicate with someone who does not own a TTY requires a third party relay operator.

The teens I interviewed overwhelmingly favored establishing electronic means of communication after meeting new people. Most expressed a preference for exchanging email addresses or IM screen names. One student noted that many hearing people asked for an email address or a phone number for text messaging. Since he could never remember his phone number, he simply gave them his email address instead.

3.2.2 Devices and Infrastructures

Mobile text messaging has been increasing in popularity in the United States, but had not achieved the widespread acceptance seen in Europe or parts of Asia [68] at the time of this study. Few teenagers pay for their own phone usage or use "pay as you go" plans. Text messages are not included in standard mobile service plans, and providers in the US often charge both the sender and the receiver of text messages and voice calls. Additionally, most providers provide "free night and weekend" plans which allow free voice calls after 9pm on weekdays and all day on weekends. This payment structure leads many hearing teens to wait until free calling periods and not use text messaging.

The most prized mobile device among the Deaf teenagers was clearly the T-Mobile Sidekick. The Sidekick is a device marketed in the US by the service provider T–Mobile. It is designed as an out–of–the–box Internet platform with a mini-QWERTY keyboard. It includes software for web browsing, instant messaging, email, address book, and SMS. T–Mobile also offers unlimited, data-only service for this device, making it attractive to deaf students who do not need the voice capabilities. Many of the teens already owned this device, and some expressed a desire to upgrade. Those that did not have one expressed a desire to own one. However, none of them paid for it themselves. The Sidekick has become so ingrained in the teens' lives that it has a unique sign in ASL which mimics the screen popping up on the device. Only one student with verbal abilities used a mobile phone for voice telephony.

Some of the students in the study did not have a mobile device but did have a family computer. Several students with mobile devices commented that they disliked computers. Two specific complaints were that "you're stuck in one place" and that computers had "a lot of things going on." Interestingly, one student noted that a legacy technology kept him tied to the computer. When he first began using an instant messenger client, he set up a screen name and began building up a list of contacts using that identity. However, when he acquired a mobile device, he discovered that his instant messenger client was not supported on the mobile device. He established a new screen name on a supported client on his mobile device and relegated the older screen name to his desktop computer. Thus, he had two screen names which were used on different devices in different situations. For new contacts, he always asked for buddy names which were compatible with the screen name on his mobile device, but he maintained the older list as well.

The use of the computer for communication was drastically different from the use of mobile devices. Students reported only occasional, not constant, communication via IM or email from a computer. Like Schiano's findings [96], the teens did not mode switch to email once they were online and chatting via IM. When they did use email, they responded to emails as soon as they received them but complained that email was much slower than IM and "it takes a whole day to get it maybe."

3.2.3 Stakeholder Analysis

During the social mapping activity, the students sorted their contacts into self-defined groups. Based on these groupings, I was able to discern the major stakeholders in the Deaf students' lives and look for patterns in that information.

The students' contacts largely fell into five groups: family, school, online, church,

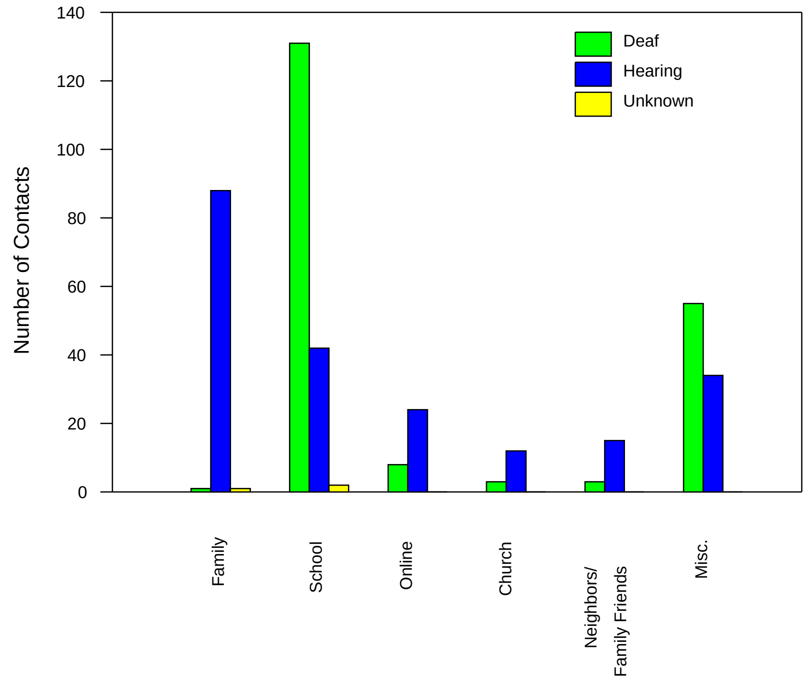


Figure 5: Stakeholders and Hearing Status in Deaf Teens' Lives

and neighbors or family friends. There were a significant number of contacts which I could not place into any category. An example is a student who labeled many of her contacts into two groups: "Buddy girls" and "Buddy boys." When questioned, she explained that these were the boys and girls on her buddy list. While these individuals may have been family or school friends, it was impossible to tell from her classification scheme and are thus included in the "Misc" category in Figure 5.

As can be seen from Figure 5, the vast majority of the teens' families are hearing while the majority of school contacts are Deaf. It is useful to further break down these two main groups. Figure 6 shows how the Deaf teens communicate with people in each of these groups in face-to-face communication. Figure 7 shows the breakdown in remote communication for school and family contacts.

From these figures, it can be seen that students generally choose to communicate with the two major stakeholder groups via sign language when face-to-face, although

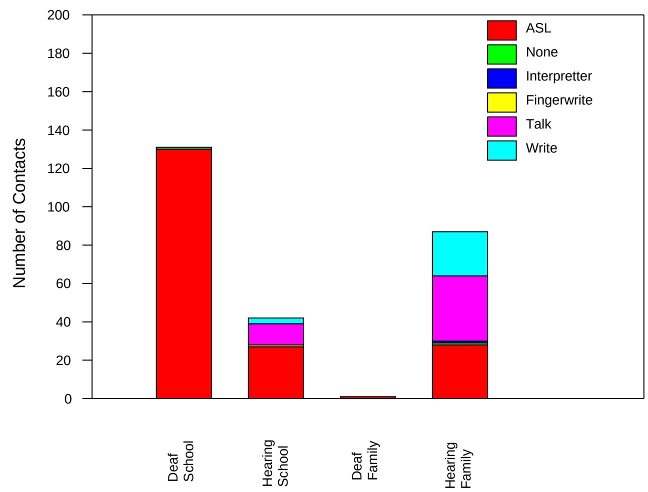


Figure 6: Face-to-face Communication Preferences for School and Family Contacts

they must employ a variety of other methods for the non-signers.

3.3 Analysis of Group Discussions: Themes of Communication

I coded the group discussion portions of the study using open or inductive coding [72]. Combining this data with analysis of the social maps and diary studies, five clear themes emerged. To the teenagers, communication is: **Identity**, **Connection**, **Control**, **Tension**, and **Convenient**. These themes are pervasive in the teens' communication. They are useful to designers because they emphasize the importance of the central purpose rather than the technological specifics of devices, communication methods, protocols, and other issues that often influence design decisions.

3.3.1 Communication is Identity

The teenagers viewed their electronic communication as a vital piece of their identity. They also manipulated and managed the identity they created online. In certain circumstances the teens used their communication to rebel by communicating in times or places when it was forbidden. In some cases, while unable to detect the noise it

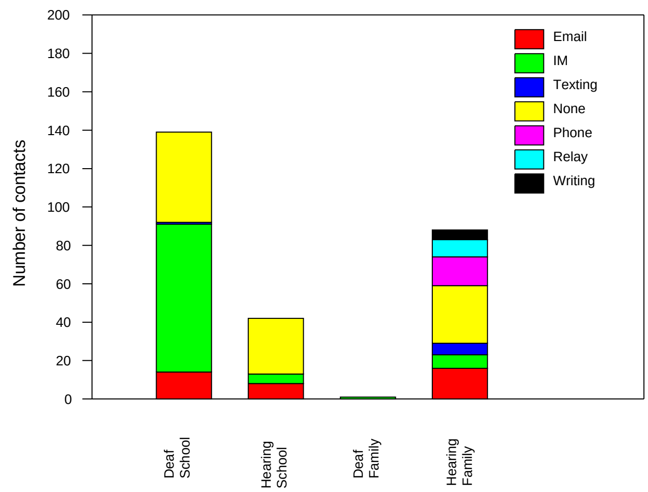


Figure 7: Remote Communication Preferences for School and Family Contacts

made the teens were aware that they needed to mute their device to avoid detection. The vibrate setting was the de facto alert mechanism for most teens, but they were aware that even that could occasionally be detected. When asked why they went to such lengths to avoid detection, one student summed up her feelings as, “I don’t want everyone looking at me.” The style of communication allowed her to preserve her privacy. They were aware of how the communication affected others around them and might reflect on them in public.

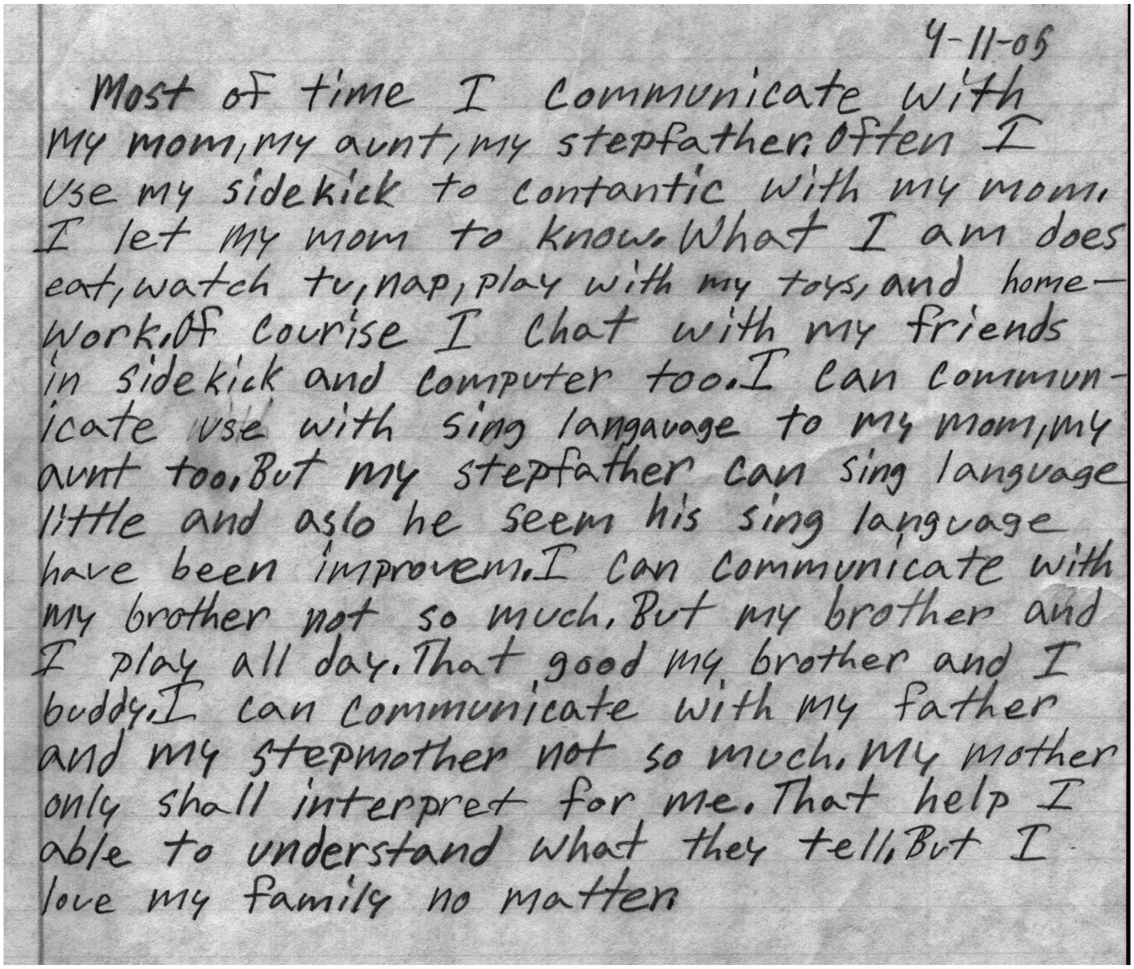
Students had no qualms about IMing someone not co-present even while physically with someone else. The teens didn’t consider it rude if someone they were with also messaged other people. Their diaries often showed that they were collocated with members of their social circle (at dinner with family, for example), but IMing with friends via the mobile device. However, the teens felt that messaging should be conducted during breaks or lulls in the conversation. Being kept waiting during face-to-face conversation by someone messaging was “wasting my time.” Messaging while collocated was seen as something to fill time when their conversational partner was distracted by talking or driving. One participant noted, “You know sometimes, like with hearing people, they'll be talking to someone, and I feel left out. So I IM my sweetheart." Her communication usage allowed her to feel included even when she was with people who excluded her.

Somewhat surprisingly, given their difficulty with written English, most students expressed only minor worries about grammar or spelling mistakes. These mistakes were considered inconsequential for the most part, particularly among friends. One student said, "If I don't know how to spell it, I just make it up" while another noted, "Sometimes, if I get the grammar wrong or whatever, I'll just send [the message] anyway." One participant said he would generally ask his mother for help, but several others relied on the Sidekick, noting that it had built-in spell check and grammar help, for example adding an apostrophe to a student's spelling of "Ive" instead of "I've."

3.3.2 Communication is Connection

All the students considered communication with hearing friends and relatives to be an important component of their lives. Some students saw a mobile device as a means to enable that connection. The method of communication was less important than the ability to convey meaning and establish connections. In the words of one student, "The important thing is that people understand what I'm saying."

Like Barkhuus and Nardi [9, 79], we found that the primary recipients of IMs were the Deaf teens' friends, and they valued this ability to communicate very highly. They mostly reported messaging people who were not co-present, as they preferred to communicate directly with collocated people. One notable exception was a student who told me she used her Sidekick with a hearing person, passing the device back and forth. "When I can't hear [people], but they don't know how to sign, it's the only way we can actually get the point across." Her electronic communication allowed her to easily establish a connection with people she would not have been able to otherwise.



4-11-05

Most of time I communicate with my mom, my aunt, my stepfather. Often I use my sidekick to contantic with my mom. I let my mom to know. What I am does eat, watch tv, nap, play with my toys, and home-work. Of courise I chat with my friends in sidekick and computer too. I can communicate use with sing langauage to my mom, my aunt too. But my stepfather can sing language little and aslo he seem his sing language have been improvem. I can communicate with my brother not so much. But my brother and I play all day. That good my brother and I buddy. I can communicate with my father and my stepmother not so much. My mother only shall interpret for me. That help I able to understand what they tell. But I love my family no matten

Figure 8: Description of Communication Problems by a 16 Year-Old Student

Figure 8 gives a striking example of the communication barriers that many students face and why their mobile devices provide an indispensable lifeline for communication. This description was written by a 16 year-old Deaf student. All the individuals described in his writing are hearing. He has problems communicating with his father, stepmother, and brother. His mother signs, and his stepfather's sign is improving. However, it also becomes clear that there is a language barrier between him and his family. His first language is ASL, while theirs is English. In Chapter 5, I will explore platforms to aid in the acquisition of survival level ASL which parents could use to communicate with young children.

Siblings were frequent targets for messages, but parents were not. The exception was when something was required of the parents such as locating them in a store or needing a pickup from a friend's house. They used IM on their Sidekick for talking to their friends about school, plans, how they were doing, or just to find out 'what's happening?'. One student noted that only after he was done messaging his friends would he consider talking to his parents, clearly viewing such interaction as a last resort for socialization. When unable to visit friends, the teens maintained connections to peers with their communication.

The students also reported IMing from the Sidekick to maintain connections with groups of their friends in large, multi-user sessions. Some of the students clearly enjoyed the large conversations with "lots of chatting going on." However, some students were indifferent or clearly blasé about group conversations. Several students said it depended on what was going on. Another student noted that it could be "kinda annoying." Yet another characterized group IM conversations as "Blah, blah, blah."

One student described using away messages as a social activity, maintaining a large buddy list and reading away messages because she was curious what people were doing. Grinter *et al.* and Nardi found a similar use of IM for awareness in their studies of IM [48, 79]. While not a turn-taking form of communication, the student was still maintaining connections with her peers and awareness of their activities.

The data from the social maps also demonstrates how dependent the teens have become on staying in touch via electronic methods. When analyzing this data, I found it interesting to examine the difference in how the teenagers communicated with their hearing and Deaf friends, both in face-to-face communication and when not collocated. Figure 9 shows a graph of the data obtained from the teenagers' social maps. The top node shows the total of 419 contacts the teens listed in their maps. These were broken into three main categories based on the students' contacts' hearing ability: Hearing Contacts, Deaf Contacts, and Hard of Hearing or Unknown. Each of these categories were further split into methods of communication: Face-to-Face and Remote. The methods of communication were then listed in order of preference. For example, Figure 9 shows that 201 of the 419 contacts were deaf. For 190 of those 201 contacts, ASL was the preferred method for face-to-face communication, and for 124 of 201 contacts, IM was the preferred method of remote communication.

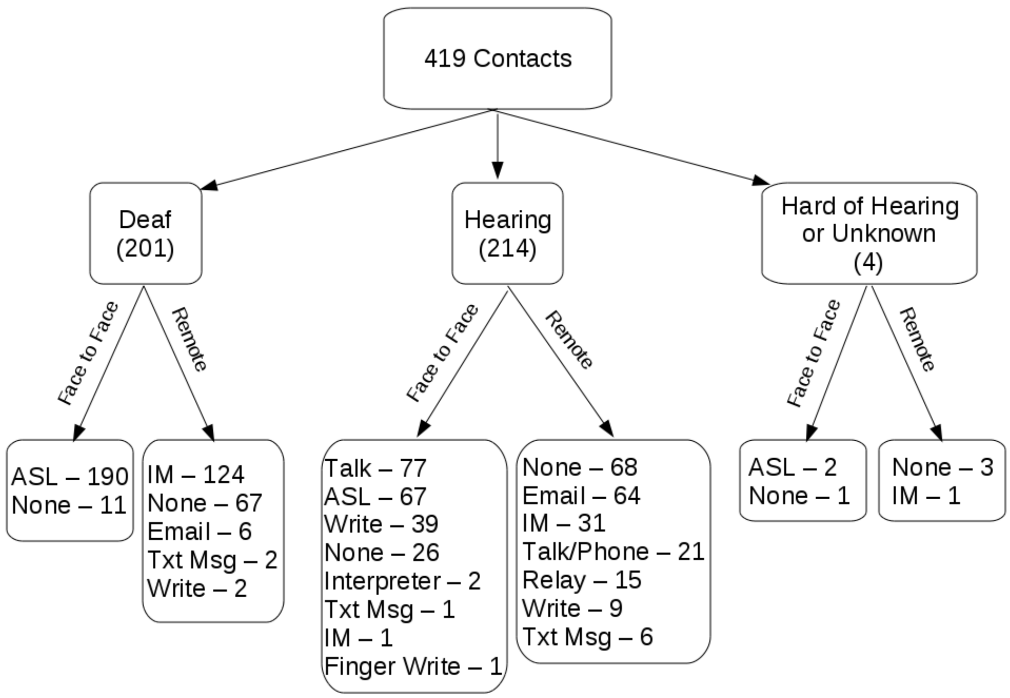


Figure 9: Results for Social Networking Study

While this data should not be generalized due to the limited number of participants, several interesting things can be seen from this chart. An interesting trend is the differences in remote communication methods between the students' deaf contacts and hearing contacts. IM is the preferred communication method between the Deaf teenagers and other deaf people, used for 124 of 201 contacts. However, there is no clear preference for their hearing contacts who are not collocated, with the Deaf teens having no remote contact with 31% (68 out of 214) of hearing contacts, and email being the preferred method of only 32% (64 out of 214).

The category of friends which the students do not maintain face–to–face relationships with (i.e., the "None" classification under the four face–to–face communication paths in Figure 9) also bears closer examination. There were 39 contacts that the students listed that they did not have any face–to–face communication with regardless of hearing status. However, the students communicated with 36 out of those 39 people remotely via the electronic methods of email and IM. (Of the other three people, two listed no communication either face-to-face or remotely and were obvious aberrant data points; the other wrote letters.) Before electronic communication existed and was widely available, remote communication with a person you had not met face–to–face would have taken the form of written letters (i.e., "pen-pals"). However, the teens today are using email and IM to do the same thing but with faster and more synchronous communication.

3.3.3 Communication is Control

The teens' usage of communication also showed how they used it to feel more in control of their lives. An "easy" way to communicate clearly made the teens feel safer which appeared to be a key component of feeling in control.

The students controlled their contacts in a variety of ways, including blocking and multiple screen names. Blocking is a standard feature of most IM clients and allows users to block messages from other users. One student noted she usually only blocked advertisements and spam. Another volunteered that he never blocked anyone. Several would immediately block someone they didn't know, but one student would try to talk with people before blocking them, explaining, "Maybe it's a new friend. I wouldn't mind trying to talk to them." However, that student went on to state she would block them if she did not feel comfortable with them. Students would also block people who simply annoyed them either through the content of messages or the volume of messages sent. They would block friends if they were having a fight or disagreement. Although some students reported that acquaintances they blocked were quite upset, they still utilized the feature. One student noted that she would unblock acquaintances after some time to see if they still bothered her.

While not a specific medium of communication, away messages filled an important aspect in the teenager's communication spectrum, just as Baron found in her work on college students' use of away messages [11]. The teens used a variety of away messages to control the flow of communication and indicate availability. Many students left a time estimating when they would return in their away messages before going out or being involved in other activities. They went to great lengths to maintain an accurate away message, including minute by minute updates. "When I get up in the morning, I immediately have to change [my away message] to my 'Hi, I'm at school now' message. I don't want people to think I'm sleeping then!" The away message removed the obligation of an immediate response.

3.3.4 Communication is Tension

Communication usage also raised tension in the teens' lives. It provided many positive benefits but also added negative artifacts to their lives.

Unlike other studies of hearing teenagers [70], very few of the students reported that their parents used the device as a tool to enforce discipline. Most students reported that their parents put few or no restrictions on the use of the device at home, including use during mealtimes or curfews. Six students reported that their parents did not restrict their use at all, while another two reported some restrictions such as not using the device after a certain time on school nights or putting it away at mealtimes. However, one teen admitted she sometimes turned it on after her curfew anyway. One student's parents had taken away the Sidekick as punishment for a week, but this method of punishment was a rarity. One student pointed out, "Mom is fine because she knows, like, I'm Deaf and I want to communicate with people.

And she knows it's OK. It's not wasting our time." One teen's parents would not allow him to have one due to the expense and possibility of losing it but used the device as incentive, hinting he might get one if he kept his grades up.

The school, however, banned the use of the devices. Students were not allowed to use the device during class hours and instead were limited to using the device during breakfast or lunch times. The rule had recently been strengthened, requiring parents to come and retrieve the device if it was confiscated. Since these rules had been implemented, several students reported they had gotten in trouble for using the device at school. In fact, many told me that they now left their Sidekicks at home to avoid the temptation. However, the students who regularly carried their devices noted that any time without the device felt strange.

Tension also arose from the ease and prevalence of communication availability. Like the hearing teens in the study conducted by Smith *et al.* [102], the Deaf teenagers sometimes felt overwhelmed by the sheer numbers of contacts and the social energy expended maintaining them. One student reported that having more than one screen name was "too much to keep up with." In contrast, another told me that she would simply make up a new screen name when her current one hit the maximum number of contacts allowed (around 200 contacts by her estimation). Other students reported 20, 89, and 72 people on their buddy lists. One student noted that having many people on her buddy list led to people contacting her constantly which "gets kinda silly at times."

Another source of tension was the amount of time the students would spend communicating with friends. A few other students told me that in the past they used their Sidekick every day, but now had grown somewhat tired of it, and "once in awhile" they didn't use it at all. Two students in particular stressed that they liked IMing with their friends but also liked doing other things such as spending time with their families or reading. Another student simply noted, "Sometimes I'm doing things. I'm busy." Another student noted that "sometimes my thumbs get worn out." One student who did not own a Sidekick offered his perspective on the constant chatting of his classmates by saying, "My eyes would be falling out of my head if I did that!"

3.3.5 Communication is Convenient

The students viewed their communication as highly convenient. It was clear that the device preferences and communication modes arose due to the convenience it afforded the users. The freedom of a personal platform was greatly valued, and this freedom was exercised in a variety of ways.

The students used communication via mobile devices heavily. Most students who had a device reported using it daily with a majority of the use occurring during the free time after school or on the weekends. In their daily diaries, students reported several hour-long blocks of IM without interruptions. When questioned, the students assured me that was correct and that they were constantly chatting. Several of the participants gave the following responses:

> P5: Constantly. Constantly chatting. Constantly chatting.
>
> P8: Constantly. I talk a lot. I talk a lot. Even in my sleep I do it. Constantly.
>
> P2: Sometimes I chat and then I'll do other things. Something else like work or play or whatever. Then I'll come back and chat again. I go back and forth, back and forth. Some other times, some other days, I'll chat with my friends maybe 3 hours. Just constantly chatting. It kinda varies.
>
> ...
>
> P10: Me. Yeah. All day. I do. Kinda lazy I guess.

In addition to just chatting, IM was viewed as an optimal, convenient way to schedule activities with friends. This mirrored findings by Nardi *et al.* and Grinter and Palen [79, 48] that the immediacy of IM was useful for coordination and scheduling. Scheduling events was of great importance to the teens given their lack of transportation and distance from friends. The teens could refine their plans on-the-fly, but plans were usually made in advance due to the logistics of meeting. Only if a friend was offline was email employed as a scheduling tool. Even then, email was usually used to establish a time when both would be available to IM and finalize the details.

The mobility of a Sidekick provided great freedom to the teens, and they used them in variety of places. Some would use it in stores to locate their family if they became separated. Another volunteered that he would use it on vacation to get messages right away. The teens would even use their devices in what could be considered socially inappropriate places, such as during church or during a family dinner. They had developed techniques to avoid detection, such as pretending to be asleep and pulling up the hood of a jacket to disguise their gaze direction and hiding the device in their laps. The students were particularly aware that their eyes could be a giveaway as evidenced by this quote, "...if people are looking at me and my eyes are going up and down... You have to be careful with your eyes or they'll figure it out that way."

The students choice of electronic communication medium also reflected the teens' desire for convenience. For example, most students preferred IM and saw text messaging as a backup communication medium (to be used only as a last resort). The Sidekick was capable of many different communication modes, including IM, text messaging, and email. Thus, the students made decisions based on factors other than the availability of a specific communication technology to them.

Text messaging was something to be used only if the other person lacked the capability to use a more convenient technology such as IM. The students might use text messaging if someone didn't have IM or email, if a friend wasn't online, or if a friend wasn't in a position to check email. Some students realized that others might have less capable communication platforms. One participant identified text messaging as a feature central to all mobile phones and said that she used text messaging because most other people had a cell phone thus increasing the number of people with whom she could communicate. Another posited that he might use text messaging only in an emergency situation.

Cost was not a deciding factor, unlike Grinter's study [46] which found that teenagers made a determination based on cost 27% of the time. This lack of differentiation based on cost can be explained by the fact that none of the teenagers in this study paid for their own costs and the unlimited data plan that many of them had.

3.4 Discussion

Several points bear further discussion: first, the teens' *use-centric* view of mobile computing; secondly the *social acceptability* of their chosen device and the tension arising from that choice; and thirdly, the role technology can play in overcoming communication limitations.

3.4.1 "Use-Centric" Perspective

The students' use and understanding of electronic communication was very "use-centric." While not understanding the specifics of the communication mode, they were still able to use the methods. For example, very few students could articulate the differences between text messaging and instant messaging. Instead, they characterized them based on the reply or response time. When asked what text messaging was, many of them responded from a use-centric perspective. Text messaging is "pretty fast" and "you're kinda talking." "...You kinda write a long thing out and then you send it and then you wait a minute and it comes back." This use-centric perspective extended beyond just the differences between text-messaging and instant messaging.

When asked about a specific method of communication, the students responded first with the characteristics of its usage or with a list of whom they could contact by a specific method. However, while they may not have understood the underlying elements of the device, they had no trouble using the device in a variety of different ways. During the course of the study, the teens cited using the Sidekick for voice telephony, email, IM, Internet search, relay, and grammar/spelling checks.

In many ways, the teenagers' confusion of technological specifics mimics Palen's findings about new users of mobile phones and the distinction of the hardware and software components of service–based technologies [86]. Palen found that users of mobile phone had to understand many different aspects of a service, including phone hardware and software, "netware" (e.g., analog or digital service), and "bizware" (e.g., service plans offered by the provider) before achieving mastery of their communication devices. For example, the bizware layer imposed by the service providers clashed with the users' mental models of telephony service and created confusion.

This study highlights that how technologists classify and distinguish technologies is different from mainstream, public use. Technologies that are drastically different can be unproblematic in use for end–users. This clarity in usage models is key for usability. In the study, the teens viewed the device as something to be used for communication and seemed to inherently know what service to use and when to use that particular service.

3.4.2 Social Acceptability

As discussed earlier, the teens drew from a wide geographical area which often limited their contact with each other. The Sidekick may have succeeded in this community because it helps reduce this distance in a socially acceptable way. This device, unlike relay or a TTY, is practical to the entire population, not just the deaf or those trying to communicate with them. The Sidekick is accepted by both hearing and Deaf teens and allows the Deaf teens to be similar to their hearing peers. It helps establish a communication link using that similarity.

However, this similarity comes with a price. In some venues educators and linguists have expressed reservations about instant and text messaging. They have highlighted how the English used in computer–mediated communication differs from that used in more normative language [10, 30]. This difference is a particularly interesting issue given that this population may have problems with their second language of English, especially with structure and grammar. While the students reported using some acronyms and abbreviations, they were more concerned with whether or not the recipients could understand them. However, the teens admitted that others sometimes used acronyms or slang that they didn't understand, with one participant hypothesizing, "Sometimes, they make stuff up." It is worth noting that many slang terms popular in text and instant messaging are phonetically based (e.g., "c u l8r") which would present inherent problems to the Deaf teens. By using informal language, the teens are practicing written English, something their teachers usually encourage.

3.4.3 Isolation

The isolation the teens felt must also be noted. They identified several instances when they were not able to communicate with co-located friends, neighbors, and family. As the writing in Figure 8 shows, the teens recognized and understood this gap. Some of the students used mobile technologies to ameliorate this feeling of isolation, but others did not. This barrier points to a large language barrier when communicating with ASL. The students have adapted both socially and technologically in communicating with the hearing world. However, the hearing world has not necessarily adapted to them. Parents, caregivers and friends of deaf students need support in learning languages (in this case, ASL) to communicate with Deaf individuals in their lives. Deaf children of hearing parents are more likely to have language acquisition problems for both first languages (ASL) and second languages (usually English) [106].

Public education often begins to address the language barrier issue from two sides of the problem. One strategy to combat this problem is to enhance young children's signing skills. The other strategy is to increase their immersion in ASL by teaching their parents ASL. While there are many projects which address ASL learning in children (e.g., [54, 17, 33, 34]), few focus on the communication gulf between parents and children. Thus, I choose to concentrate on the task of teaching parents survival level ASL vocabulary.

CHAPTER 4

DATA GATHERING STUDY AND MODEL DEVELOPMENT

In Chapter 3, I discussed a formative study which was designed to highlight the issues Deaf teenagers face when communicating with other Deaf individuals as well as hearing individuals. This study gave concrete examples which show that many Deaf teenagers are beginning to rely on electronic and mobile communication technologies. However, it also became clear from this study that the teenagers faced more difficulties communicating with hearing friends, family, and acquaintances than when communicating with other Deaf people.

As discussed earlier in Chapter 2, Section 2.1.1, many deaf children born to hearing parents are not immersed in language until they are beyond the critical period for language acquisition. This delay can impair not only the Deaf child's ASL fluency, but also his or her English fluency later in life. The teenagers who participated in the study outlined in Chapter 3 face difficulties communicating with the hearing community in English and hearing individuals could rarely communicate with Deaf students using sign language. This problem has two parts: Deaf individuals can have difficulty learning English due to late exposure to language, and hearing individuals can have difficulty learning sign language. Deaf students are exposed to English daily and English fluency is a skill which most Deaf students are expected to master, despite their difficulty in doing so. However, very few hearing individuals have the

Parts published in [56]: V. Henderson-Summet *et al.* , "American Sign Language vocabulary: computer aided instruction for non-signers," in *Proceedings of ASSETS, 2008* same infrastructure for learning ASL if they need to do so. This chapter outlines the beginning of a mobile language learning platform aimed at assisting hearing individuals in learning sign language. Such a platform would allow hearing parents with Deaf children to begin to expose their children to language, specifically ASL. Since parents are one of the most important sources for a young child's exposure to language, this learning platform would allow parents to begin communicating with their Deaf child at any early age, which is critical for memory development and language acquisition.

Many parents of children with disabilities face additional difficulties beyond their child's disability. Research has shown that the cost of treating children with disabilities greatly exceeds the healthcare costs of children without disabilities over the course of their lifetimes [84]. Even with insurance, some of this cost is born by the parents in terms of rehabilitation costs, lost time at work, or technological interventions. Additionally, children with disabilities are statistically more likely to be raised by a single, female caregiver [26]. These statistics point to a lack of resources which can make traditional ASL classes difficult. Traditional second language classes often take significant time away from hearing siblings, family, or work.

Mobile devices provide a unique opportunity to address this problem. Mobile devices are ubiquitous and often within arm's reach [90]. Additionally mobile phones are offering increased video and movie capabilities which are well suited to a visual language such as ASL. Given the limited amounts of time that parents of deaf children may have to learn ASL during the critical period of their child's language exposure and the fact that the parents may not have much leisure time during which to take classes, it is crucial to deliver content in an optimal manner as to maximize the learning over time.

Previous work has shown that adaptive presentation in second language learning can be highly beneficial for the learner [3]. In this manner, more time is spent on the signs which are hard for participants to learn, and less time is spent on the signs which have already been learned. By using an adaptive algorithm, I hope to be able to help learners make more efficient use of the time they spend learning.

Much of the psychological literature is focussed on performing experiments to validate a particular heuristic of learning. Very few works propose models which can be used, manipulated, or applied by other researchers. One exception is the work of Atkinson [3]. Atkinson proposed a three state model of learning that classifies objects in one of three memory states: Long Term Memory (LTM), Short Term Memory (STM), and a Forgotten/Unknown state (U/F). This model is further detailed in Section 4.2. Atkinson's model is for paired-associate presentation of second language vocabulary. For example, subjects were presented pairs of vocabulary words consisting of one word in the language they were attempting to learn and the corresponding word in their native language (e.g., "casa" and "house").

Using this model requires a pilot study which uses a test–and–train algorithm to establish parameters (x, y, z in Equation 2) related to the relative difficulties of the words under consideration for learning. Section 4.1 details this data gathering pilot study. Section 4.2 describes the development, fine-turning, and validation of the Atkinson Model for the selected subset of ASL vocabulary. For ease of reference, I refer to the study outlined in this chapter as the Data Gathering Study or DGS.

Atkinson's original model was formulated as:

$$\begin{array}{c|ccc|} & L_{n+1} & S_{n+1} & UF_{n+1} \\ \hline L_n & 1 & 0 & 0 \\ S_n & a & 1-a & 0 \\ UF_n & ca & c(1-a) & 1-c \end{array} \tag{1}$$

For ease of reference, this matrix was simplified to

$$
\begin{array}{c|ccc|}
 & L_{n+1} & S_{n+1} & UF_{n+1} \\
\hline
L_n & 1 & 0 & 0 \\
S_n & x & 1-x & 0 \\
UF_n & y & z & 1-y-z
\end{array}
\tag{2}
$$

The goals of the study outlined in this chapter were to establish the per-word parameters necessary to use the Atkinson algorithm (detailed previously in Section 2.2) for American Sign Language. Additionally, I added a post test evaluation after data collection. This post test allowed for some comparisons to be drawn between the naive presentation algorithm and the use of the Atkinson algorithm as detailed in Chapter 5.

4.1 Data Gathering Study

In the Spring of 2008, we conducted a data gathering study (DGS) necessary to determine the parameters for each ASL sign to be used in conjunction with the Atkinson Model.

4.1.1 Methodology

I designed an experiment to evaluate test–and–train methods of instruction with ASL. In a test–and–train paradigm, individuals are presented with a series of "tests" (often multiple choice or short answer) and learn through their correct or incorrect responses to the test. This instruction method was chosen due to its portability and flexibility. Words can be treated singularly, and instruction can be delivered according to a number of different algorithms which might maximize total time spent on instruction, time spent on different themes (e.g., questions, the home, or food), or repetitive testing of problematic signs. During April 2008, we recruited 20 participants, ranging in age from 19-28 with a mean age of 23.15 (σ = 3.36). Our population included seven females and 13 males. Participants were screened to eliminate any participants who had knowledge of ASL beyond fingerspelling (which is sometimes taught in US elementary schools).

4.1.1.1 Experimental Procedure

Participants completed two sessions which occurred approximately one week (between six and eight days) apart. The goal of the sessions was to learn 80 basic ASL signs. During Session 1, the participants completed a set of training trials. During Session 2, they completed two tests, an expressive test and a receptive test, which measured the overall mastery of some ASL vocabulary.

We devised a set of paired associate vocabulary consisting of one ASL sign and one English word as a translation. Vocabulary words and signs were chosen from the MacArthur-Bates Communicative Development Inventory (MCDI) [37] with the help of an ASL linguist. We chose an 80 words subset suitable for communication with a child from infancy. The MCDI is a test designed to evaluate the vocabulary and language development of young children and infants. It has been validated in 44 languages, including American Sign Language, and is a standard assessment metric used in many deaf schools. A full list of the selected vocabulary for this study can be found in Table 18.

4.1.1.2 Session 1

In order to utilize the test–and–train method of learning, we developed a Flash-based web interface using Adobe Flex. Data from each trial, including timing of responses and answer chosen, was saved in an XML hierarchy. In the first screen of the web application, the participant was asked to provide demographic information including handedness. Depending on the handedness of the participant, the video was mirrored appropriately. After completing the demographic information, participants were able to select a button to start the trial.

In Session 1, subjects completed five trials. Each trial consisted of videos of

Figure 10: Progression of Experimental Software

each of the 80 ASL vocabulary words presented in random order. After each video, the participant was presented with multiple choice answer consisting of the correct answer, three incorrect answers randomly selected from the other 79 words, and an "I don't know" option. After participants selected an answer, they would be told whether their answer was correct or incorrect. If they were incorrect, the interface would highlight the correct answer. Figure 10 shows a progression as the subject watches a video (Fig 10A), sees a set of answers and answers incorrectly (Fig 10B), and receives feedback with the correct answer (Fig 10C). In this way we implemented a basic testing and training method of instruction. For this study, we chose to use multiple choice questions because they are capable of being fully automated, are easy to program, cost effective, and focus on learner outcomes [2, 98].

After finishing a trial in which all 80 words had been presented, the participant would see a screen showing how much time had been spent on the trial as well as a pie chart displaying total correct responses, incorrect responses, and responses of "I don't know" for the trial. A timer counting down from five minutes was also displayed. The participant had to wait a minimum of five minutes before starting the next trial. When five minutes had elapsed, a button was enabled which allowed the participant to start the next trial over the 80 words in a different, randomly selected order. After the fifth trial, a message indicating that all trials had been completed was displayed instead of the countdown timer.

Subjects completed five trials in all. Subjects averaged 51.68 minutes (σ = 6.2 minutes) spent on language instruction (not including breaks). The fastest participant spent 44.72 minutes on instruction and the slowest spent 64.58 minutes.

4.1.1.3 Session 2

After completing the five test–and–train trials in Session 1 described above, the participants were scheduled for a second session approximately a week later. While a week long retention interval is not sufficient to claim the subject has mastered the material, research has shown that a week without instruction is a beginning point for measuring whether an item has entered into long-term, permanent memory [6]. If an item is not recalled after a week retention interval, then the item is most likely lost and will not be recalled during subsequent tests.

During Session 2, participants were given two different tests: a receptive language test and a expressive language test. During the receptive language test, participants were shown videos of 40 words and asked to write the English translations. During the expressive test, the researcher called out 40 English words to the participant, and the participant attempted to sign the words to the best of their abilities. The sets of words for the expressive and receptive tests did not overlap. In both conditions, the participants were given 10 seconds to generate either the English word or correct sign. The words for both the receptive and expressive tests were randomly selected from the original 80 word vocabulary list the participants had learned the previous week. Half the participants performed the expressive test first, and the other half did the receptive test first.

The expressive test was analyzed for correctness using a three point scale. A sign is composed of several parts [105] which are:

1. Handshape

2. Palm Orientation

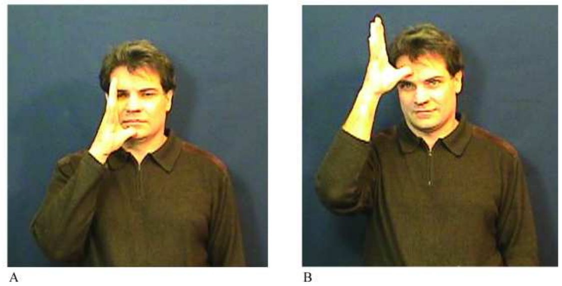


Figure 11: Signs for (a) "mother" and (b) "father"

3. Location

4. Movement

5. Facial Expression

The five parts were used to assess correctness of a sign. A score of '1' indicated the subject made an attempt at the sign but was incorrect or the subject did not know the sign. A score of '2' indicated that some aspects of the sign were correct. For example, a participant might have made the correct motion in the correct location and sign-space, but with an incorrect handshape. However, if the error resulted in a change in the semantic meaning of the sign, the participant's response was coded as a 1. An example of this is shown in Figure 11. In this figure, the signs for "mother" and "father" which are identical in handshape and motion and differ in only the sign-space. A score of '3' indicated an entirely correct sign comprehensible to someone familiar or fluent in ASL. Each sign was scored by two raters and compared against the original video that the participants saw while learning. The raters were required to come to an agreement on what the score of a sign was.

4.1.2 Results

During Session 1, we collected 100 correct/incorrect data points for each of the 80 words (20 participants x five trials). During Session 2, we collected 40 expressive data points and 40 receptive data points per participant. Because the 80 words were randomly divided between expressive and receptive, some words have slightly more receptive data points then expressive, or vice versa.

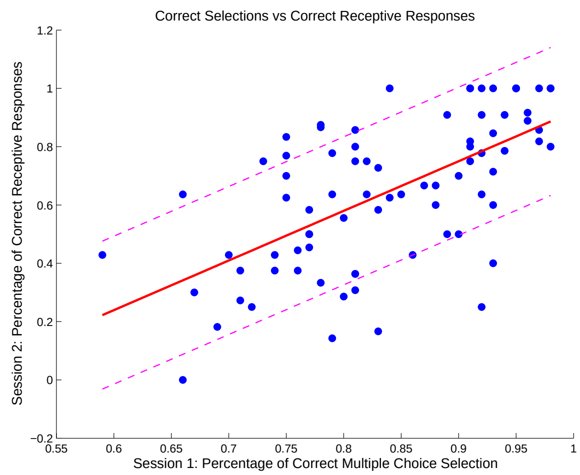


Figure 12: Test–and–train Data Related to Receptive Percentage

To compare the predictive power of the correct answers in Session 1 to the observed performance in Session 2, we summed the results of the multiple choice trials 1–5 in Session 1. For example, the aggregate data points for the correct choice of the English word "school" when presented with the ASL sign were T_1=2, T_2=11, T_3=16, T_4=18, and T_5=19. Thus, T_{total}=66. In this case, two participants answered correctly on Trial 1, 11 answered correctly on Trial 2, and so forth.

For each word, we computed the percentage of correct responses during the receptive test for each word. On a per word basis, we plotted the (x, y) pair where x equaled the percentage of all correct responses in Session 1, and y equaled the percentage of correct responses during the Session 2 receptive test. The resulting scatter plot is shown in Figure 12. (Note: the y-axis on these graphs have been extended in the negative direction for easier data viewing.)

As can be seen, there is a positive correlation (ρ=.6223) between the two variables along the x and y axes. The center (red) line shows the best fit, while the two outer

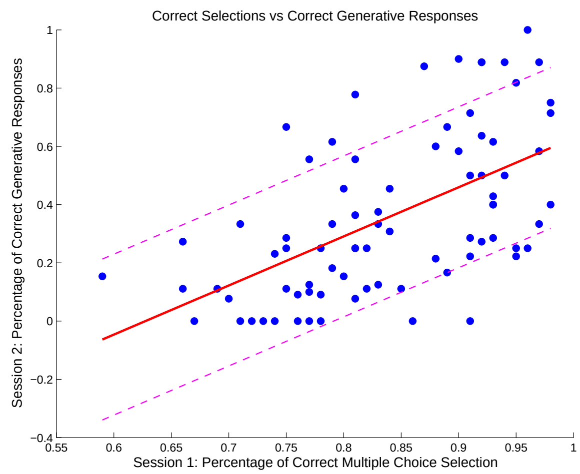


Figure 13: Test–and–train Data Related to Expressive Percentage (Signs Rated "3")

dashed (pink) lines show data within one standard deviation ($\sigma = 0.2538$).

For the expressive test during Session 2, we calculated the percentage of signs which were correct and intelligible signs (i.e., those rated as a 3) on a per word basis. Figure 13 shows the resulting scatter plot of 80 data points. Again, this figure shows a positive correlation ($\rho = .5663$, $\sigma = .2762$) between the Session 1 results and the expressive test results in Session 2.

A corresponding set of data can be seen in Figure 14 which shows the number of the participants' expressive signs which were rated as a '1' (indicating either an incorrect sign or an "I don't know/don't remember" response from the participant) plotted against Session 1 data. This graph shows a negatively sloped best-fit ($\rho = -0.6034$, $\sigma = 0.2815$) line due to the inversion of the y-axis. This result indicates that we can predict the difficulty of words for the participant to generate based upon their performance on our test–and–train questions.

Table 18 in Appendix A summarizes much of this data. In this table, we give equal weight to the expressive percentage correct (column 2), the receptive percentage correct (column 3), and the percentage of correct responses for T_1–T_5 (column 4). We then ordered the list by the average of these three values and presented it in column

5 as a rough estimate of how easy or difficult it is to learn a word. A high measure of "learnability" indicates the word is easier to learn then words with a lower measure. The more difficult words are at the beginning of the table, and the easiest words are listed at the end.

Collectively, our participants correctly generated 278 words and recalled 524 words after less then twenty hours of non-human (i.e., computer based) instruction. As noted earlier, due to the relative inexpensiveness of computer-based training materials, a learning system could be especially advantageous for parents of deaf children who may have limited time or money to spend on instructional methods.

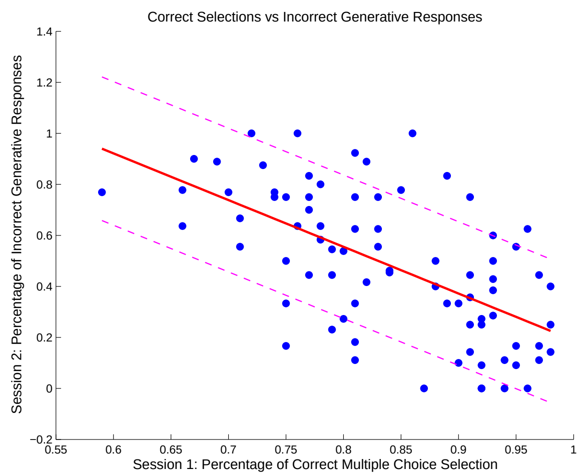


Figure 14: Test–and–train Data Related to Expressive Percentage (Signs Rated "1")

4.2 Model Development

The data collected during the DGS also allowed us to experimentally determine parameters for use with the Atkinson model. The model was generated from the observed data collected in T_2-T_5 which correlate to performance on the expressive and receptive tests as described in Section 4.1.2 and Figures 12 and 13.

In the following sections, I detail Atkinson and Crothers's method outlined in their 1964 paper. I do this to clarify the model and well as make implementation easier for future researchers. I also note changes to methods which have been made possible by more powerful, inexpensive computers.

4.2.1 Equations and Parameter Generation

During the DGS as discussed above, we collected data from the trials and noted correct and incorrect responses. In Equation 3, we follow Atkinson's notation in denoting the 16 possible outcomes (denoted as $O_{outcome,trial\ number}$) from trials n to $n+3$. c_n represents a correct response on trial n while e_n represents an incorrect response on trial n in Equation 3. For example, in Equation 3, $O_{9,n} = e_n c_{n+1} c_{n+2} c_{n+3}$ with $n = 1$ would indicate a incorrect response on trial 1 (e_1) followed by correct responses in trials 2, 3, and 4 ($c_2 c_3 c_4$). These equations cover the space of all 16 possible combinations of correct or erroneous responses for a given vocabulary word over four trials.

$$
\begin{aligned}
O_{1,n} &= c_n c_{n+1} c_{n+2} c_{n+3} & O_{9,n} &= e_n c_{n+1} c_{n+2} c_{n+3} \\
O_{2,n} &= c_n c_{n+1} c_{n+2} e_{n+3} & O_{10,n} &= e_n c_{n+1} c_{n+2} e_{n+3} \\
O_{3,n} &= c_n c_{n+1} e_{n+2} c_{n+3} & O_{11,n} &= e_n c_{n+1} e_{n+2} c_{n+3} \\
O_{4,n} &= c_n c_{n+1} e_{n+2} e_{n+3} & O_{12,n} &= e_n c_{n+1} e_{n+2} e_{n+3} \\
O_{5,n} &= c_n e_{n+1} c_{n+2} c_{n+3} & O_{13,n} &= e_n e_{n+1} c_{n+2} c_{n+3} \\
O_{6,n} &= c_n e_{n+1} c_{n+2} e_{n+3} & O_{14,n} &= e_n e_{n+1} c_{n+2} e_{n+3} \\
O_{7,n} &= c_n e_{n+1} e_{n+2} c_{n+3} & O_{15,n} &= e_n e_{n+1} e_{n+2} c_{n+3} \\
O_{8,n} &= c_n e_{n+1} e_{n+2} e_{n+3} & O_{16,n} &= e_n e_{n+1} e_{n+2} e_{n+3}
\end{aligned} \tag{3}
$$

Using the above series of equations, Atkinson and Crother's derived the equations for $\Pr(O_{i,n})$. $\Pr(O_{i,n})$ is the probability of a sequence of correct/incorrect responses over four trials for a given word. Equation 4 (reproduced from Equation 19 in the original work [4]) gives the resulting derivations as follows:

$$
\begin{aligned}
Pr(O_{1,n}) &= (1 - s_n - f_n - u_n) + (s_n + gf_n)(a + xA_1) + gu_n[c(a + xA_1) + gB_1] \\
Pr(O_{2,n}) &= (s_n + gf_n)xA_2 + gu_n[cxA_2 + gB_2] \\
Pr(O_{3,n}) &= (s_n + gf_n)xA_3 + gu_n[cxA_3 + gB_3] \\
Pr(O_{4,n}) &= (s_n + gf_n)xA_4 + gu_n[cxA_4 + gB_4] \\
Pr(O_{5,n}) &= (s_n + gf_n)yA_1 + gu_n[cyA_1 + (1 - g)B_1] \\
Pr(O_{6,n}) &= (s_n + gf_n)yA_2 + gu_n[cyA_2 + (1 - g)B_2] \\
Pr(O_{7,n}) &= (s_n + gf_n)yA_3 + gu_n[cyA_3 + (1 - g)B_3] \\
Pr(O_{8,n}) &= (s_n + gf_n)yA_4 + gu_n[cyA_4 + (1 - g)B_4] \\
Pr(O_{9,n}) &= (1 - g)f_n(a + xA_1) + (1 - g)u_n[c(a + xA_1) + gB_1] \qquad (4) \\
Pr(O_{10,n}) &= (1 - g)f_nxA_2 + (1 - g)u_n[cxA_2 + gB_2] \\
Pr(O_{11,n}) &= (1 - g)f_nxA_3 + (1 - g)u_n[cxA_3 + gB_3] \\
Pr(O_{12,n}) &= (1 - g)f_nxA_4 + (1 - g)u_n[cxA_4 + gB_4] \\
Pr(O_{13,n}) &= (1 - g)f_nyA_1 + (1 - g)u_n[cyA_1 + (1 - g)B_1] \\
Pr(O_{14,n}) &= (1 - g)f_nyA_2 + (1 - g)u_n[cyA_2 + (1 - g)B_2] \\
Pr(O_{15,n}) &= (1 - g)f_nyA_3 + (1 - g)u_n[cyA_3 + (1 - g)B_3] \\
Pr(O_{16,n}) &= (1 - g)f_nyA_4 + (1 - g)u_n[cyA_4 + (1 - g)B_4]
\end{aligned}
$$

where

$$
\begin{aligned}
x &= (1 - a)(1 - f + fg) \\
y &= (1 - a)(1 - g)f
\end{aligned}
$$

and

$$
\begin{aligned}
A_1 &= a + x(1 - y) \\
A_2 &= y(1 - y) \\
A_3 &= xy \\
A_4 &= y^2 \\
B_1 &= (1 - c)\{ac + cx(1 - y) + g(1 - c)[c(1 - y) + g(1 - c)]\} \\
B_2 &= (1 - c)\{cxy + g(1 - c)[1 - c(1 - y) - g(1 - c)]\} \\
B_3 &= (1 - c)\{cy(1 - y) + (1 - g)(1 - c)[c(1 - y) + g(1 - c)]\} \\
B_4 &= (1 - c)\{cy^2 + (1 - g)(1 - c)[1 - c(1 - y) - g(1 - c)]\}
\end{aligned}
$$

In Equation 4, f_n and s_n are both used to denote the probability of being in a state of short-term memory on trial n and u_n denotes the probability of an item being in an unknown/forgotten state on trial n. Equation 5 corresponds to equations 7a, 9a, and 9b in the original paper by Atkinson and Crothers [4].

$$
\begin{aligned}
u_n &= (1-c)^{n-1} \\
t_n &= \begin{cases} c(n-1)(1-a)^{n-1} & \text{if } c = a, \\ \frac{c(1-a)}{c-a}[(1-a)^{n-1} - (1-c)^{n-1}] & \text{if } c \neq a. \end{cases} \\
f_n &= f t_n \\
s_n &= (1-f)t_n
\end{aligned} \tag{5}
$$

For Equations 4 and 5, the parameters a, c, and f must be calculated. Atkinson minimized the χ^2 error (Equation 7) between the model and the O_i events[1]. In other words, he optimized the a,c, and f parameters of his model so there would be minimal error between the subjects' performance as predicted by the model and the actual performance as recorded in his study. To show this calculation, we follow Atkinson's nomenclature and allow $Pr(O_{i,n}; a, c, f)$ to represent the possibility of event $O_{i,n}$ where a, c, and f are parameters of the expression. Also, we let $N(O_{i,n})$ be the experimentally observed frequency of items which display outcome O_i over trials n to $n + 3$. In this way, the total number of experimentally observed events can be represented by Equation 6.

$$
T = N(0_{1,n}) + N(O_{2,n}) + \ldots + N(O_{16,n}) \tag{6}
$$

These notations allowed Atkinson to define the function

$$
\chi^2(a, f, c) = \sum_{i=1}^{16} \frac{[T * Pr(O_{i,n}; a, f, c) - N(O_{i,n})]^2}{T * Pr(O_{i,n}; a, f, c)} \tag{7}
$$

In their original work, Atkinson and Crothers used a computer to carry out a brute force search to estimate a, c, and f to three significant decimal places and calculate the χ^2 value.

[1] For a full discussion of the chi-squared method and its relevancy, see Atkinson and Crothers [4]

We followed Atkinson and Crothers's original method. However, even running on modern computers, this brute force computation to .001 accuracy required almost three days of compute time to estimate the χ^2 value for one word. Running on a 16 computer cluster, parameter estimation for 80 words required almost a week of runtime.

To overcome this time limitation, we used standard machine learning techniques of hill climbing with gradient descent and simulated annealing with random restarts to optimize the calculation of the Atkinson model parameters. This code picks a predefined number of points in the three dimensional space (defined by axes a, c, and f). It then calculates the χ^2 metric for points along vectors in each cardinal direction from the initial sample point in search of a local minima. If no local minima is computed, the vectors are resized and calculation continues. The code stores the minimum chi-square value computed. Once a χ^2 value with .0001 accuracy or less has been reached, the program terminates and reports the minimum chi-square value found along with the three parameters a, c, and f.

Using this method, we determined the minimum chi-square value and parameters a, c, and f for each word to .0001 accuracy. The hill climbing code implementation allowed for all parameters for all words to be calculated in under 15 minutes. For all words, the minimum chi-square value was as good or better than the original brute force calculations. The final calculations of parameters and chi-square values for each word are included in Table 19 in Appendix A.

4.2.2 Validating the Model

After the parameters were calculated, I wrote several programs to validate the model against the DGS data. The first was a simple program to calculate the probability that a word would transition to LTM from the U/F state. This calculation allows for words to be evenly distributed according to level of difficulty during the next phase of testing (see Section 5.1.1.3 for a complete description). The words and the probability of transitioning from the U/F state to LTM is shown in Table 20 (in Appendix A) in sorted order.

This program also allowed me to simulate a complete learning run (i.e., repeated presentations of words until the model predicted that all words had been learned). After each presentation of a word, the program would randomly guess if the word had been responded to correctly or not with specific odds of a correct response. Over 10,000 trials, if run with a specified 50-50 percentage (that is, the participant guesses randomly and answers correctly 50% of the time and incorrectly 50% of the time), the model predicted that it would take 654.89 word presentations ($\sigma = 90.02$). Note that the 654.89 presentations would not be equally distributed over the 80 words since the model takes into account the difficulty of the words. However, given the experimental data in the DGS, a 50% chance of answering correctly is extremely low.

However, across the five trials, the experimentally observed percentage of correct responses was 83.9%. (Table 4 contains the percentages broken down on a trial-by-trial basis). I re-ran the simulation with 83.9% correct response rate in accordance with the DGS experimental data. In this scenario, the model predicts that it would require 395.82 word presentations (σ = 58.07) for all words to transition from an Unknown/Forgotten state to Long Term Memory on average across 10,000 runs. Because of the predicted number of word presentations, I choose to have a total of 400 word presentations during the final learning study. This number has the additional advantage of being the same as the number of presentations of total words during Session 1 of the DGS (80 words presented five times each). This experimental similarity allows for some interesting comparisons between this DGS and the final study results detailed in Chapter 5.

Once I determined the experimental parameters, I could test the model's predictions against the results of both the receptive and expressive testing of the DGS

Table 4: Percentage Response by Trial

Trial	**Correct (%)**	**I Don't Know (%)**	**Incorrect (%)**
T_1	54.25	26.375	19.375
T_2	83.375	3.8125	12.8125
T_3	91.125	1.25	7.625
T_4	94.75	0.0625	5.1875
T_5	96.0	0.0625	3.9374
Total	83.9	6.3125	9.7875

(described in Section 4.1.1.3). Based on the calculations, the model predicted which of the three states of memory the words would be in after the first test-and-train session (T_2). However, two issues had to be overcome first. The first was how to model a word which was in STM, and the second was how to score expressive sign data.

4.2.2.1 Short Term Memory Modeling

The Atkinson model can predict that any word is in one of three memory states: Long Term Memory (LTM), Short Term Memory (STM), or Unknown/Forgotten (U/F). If a word is in LTM, a question regarding it should be answered correctly. Similarly, if a word is in U/F, a question regarding it should be answered incorrectly. However, if a word is in STM, over time it could regress to U/F or be committed to LTM. When validating the model, how should we treat words which were predicted to be in STM after Session 1? Would a participant correctly identify or generate the word during Session 2? Since psychological literature suggests that short term memory is on the order of seconds [5], we chose to assume that an item in STM after Session 1 would not be recalled during Session 2. This decision may lead to a slight underestimation, but I feel that an underestimation is preferable to an overestimation of the model's performance.

4.2.2.2 Expressive Data Scoring

As mentioned previously when the participants were asked to sign words, we scored their responses into three levels: completely correct, somewhat correct (e.g., handshape or motion correct, but other elements incorrect), or completely incorrect/unknown. If the model predicted a word would be in the U/F state and the participant scored a 0 when generating that sign, we counted this as a correct prediction. Likewise, if the model predicted a word should be in LTM and the participant signed it completely correctly, we also counted this as a correct prediction. However, the words which were signed partially correctly (i.e., a score of "2") were more problematic. In order not to overstate the learning abilities of participants or the correctness of our model, we decided to treat these partially correct signs as incorrect for purposes of validating the model. Therefore, the percentages listed in Table 5 are a slight underestimation of the correctness of the model.

4.2.2.3 Validation Statistics

In order to determine how well the model fit the data in a number of given situations, I compared the experimental data from Session 2 against three different ways of predicting the participants' answers: the Atkinson model formulated as discussed in Section 4.2.1, a random guess of correct or incorrect, and guessing against the most common class. Guessing against the most common class is a common testing method of validation from the field of machine learning. This method of verification tests the efficacy of a model (in this case, the Atkinson model's predictions) against an outcome which occurs most frequently. For example, if result A occurs 30% of the time in an experimental setup, result B occurs 60% of the time, and result C occurs 10% of the time, a commmon class guess would select the most commonly occurring result (in this example, B) and compare this against the model's predictions. If 60% accuracy can be achieved using a simplistic guess and a model only achieves 61%

accuracy consistently, an argument must be made that the 1% increase in accuracy justifies using a more complex model. In the case of the data from Session 2, the participants were more often correct than incorrect and the most common class of answers was that the participants would answer correctly.

Testing against a random "coin-flip" allows us to compare our model's performance against an algorithm which randomly guesses whether the participant would answer correctly or incorrectly. Similarly, testing the experimental data against the most common class ensures that the model performs better than the assumption that the participant is always correct (or incorrect). Both of these tests allow us to justify using the model. For example, if we can achieve an acceptable level of accuracy using either random guessing or always guessing the participant's response as correct (or incorrect), it becomes difficult to justify the use of a computationally complex model. To further the example, if the experimental data shows that we can achieve an accuracy level of 75% by simply always guessing that the participant would always answer correctly, it is difficult to justify using a model which only gives 76% accuracy.

We obtained 1599 data points from Session 2. The 20 participants generated or recognized 80 signs each. The single word discrepancy (1600 theoretical data points vs. 1599 collected data points) is because the experimenter missed a single data point accidentally. Table 5 shows the experimental data from Session 2 against these three statistical tests. These statistics were compiled by averaging the result of 10,000 runs of the model.

Table 5: Validation Statistics of Model, Random Guessing and Common Class Guessing against Experimental Data

	Atkinson Model Prediction	**Random Guess**	**Common Class Guess**
Correct (%)	59.0997	49.9991	50.2189
Incorrect (%)	40.9903	50.0009	49.7811
StDev (%)	1.0241	1.2464	N/A

With a 9% gain, the Atkinson model improves our predictive capabilities by 18% over both random guessing and common class guessing. A breakdown of the correct and incorrect predictions based on true and false positives and negatives is shown in Table 6. This data corroborates the findings shown more generally in Table 5.

Table 6: Error Analysis of Validation Metrics

	Error Type	**Model**	**Random Guess**	**Common Class Guess**
Correct	True Positive (%)	36.6608	32.0007	50.2189
	True Negative (%)	22.4389	17.9986	N/A
Incorrect (%)	False Positive (%)	27.3422	32.0023	49.7811
	False Negative (%)	13.5581	17.9984	N/A

4.3 Discussion and Implications

While 80 signs is a very small subset of ASL vocabulary, our results are promising. A positive correlation for both expressive and receptive language skills based on a naive, random presentation of multiple choice answers shows that we can predict to some degree what words will be most difficult for participants to learn. This data allows us to take advantage of the Atkinson algorithm [4, 3] to optimize the time spent on instruction. In Chapter 5, adaptive instruction was implemented using this algorithm.

4.3.1 Adaptive Instruction

Our data and results make a strong case for adaptive instruction. In the case of a sign such as "banana", by T_5, the correct response percentage was 100%. It was correctly recalled 100% of the time and generated correctly 98% of the time by all participants. This data indicates that the sign was, for some reason, "easy" to learn for the participants. There are a number of reasons certain signs could be easier for participants. Some signs, such as "grandmother," "grandfather," "mother," and “father” are highly related in handshape, sign-space, and motion. Participants who remember one may use it as a trigger to remember others.

Another explanation may be the fact that some signs are highly iconic, such as “book” which mimics opening a book with two hands or “banana” which mimics the peeling of a banana. However, ASL as a whole is not iconic. A sign’s iconicity “refers to the visual similarity between a sign and its referent (object, feeling, or idea) and the extent to which the pictorial nature of a sign provides clues to its meaning” [77]. While sign language researchers disagree on the importance of a sign’s iconicity in the recognition of that sign, in this study, the iconicity may provide a mental mnemonic that allows some people to better remember it.

Whatever the cause, our participants learned certain signs much more effectively then others. The rank orderings we devised in Tables 18 and 20 serve to illustrate this point. In both tables, signs at the bottom of the table were the easiest for participants, while the signs at the top were the most difficult. The differences in the “learnability” of the words can be taken advantage of for improving instruction and (theoretically) learning. While we hope to improve on this ranking metric, it serves as a useful starting place for devising an adaptive instruction algorithm to make learning more optimal.

In the next chapter, I turn to the actual use of the Atkinson Model in a system for delivering ASL vocabulary lessons via mobile devices.

CHAPTER 5

LANGUAGE LEARNING VIA MOBILE DEVICES

In this chapter, I describe an experiment conducted using the per word coefficients of difficulty determined in Chapter 4. The goal of this study was to evaluate whether participants who learn American Sign Language (ASL) vocabulary:

1. utilizing mobile phone platforms as a content delivery mechanism will demonstrate better receptive and generative language abilities than participants who learn using a traditional desktop-based platform as measured by post-intervention tests.

2. using the spacing effect, which presents the material on a distributed schedule instead of a massed schedule, will demonstrate better receptive and generative language abilities as measured by post-intervention tests.

In Section 5.1, I describe the experimental design of a study to test these two hypotheses, the experimental setup, types of data collected, and participant demographics. Then in Section 5.2, I describe the quantitative results and interview data obtained via post-lesson testing and participant interviews. I also discuss some possible explanations for some of the results seen in this experiment.

5.1 Experimental Method

5.1.1 Study Design

This study was designed to evaluate spacing and retention intervals for ASL and a mobile ASL content delivery system. This study was a 2x2 between subjects design and included a comparison to a standard desktop interface. The four participant conditions were:

- Phone delivery with Distributed lessons. (PD)
- Phone delivery with Massed lessons. (PM)
- Computer delivery with Distributed lessons. (CD)
- Computer delivery with Massed lessons. (CM)

These conditions will be described in more detail in Section 5.1.1.1.

The software system was designed around the idea of delivering short ASL video lessons to participants throughout their day. Each lesson consisted of five video/multiple choice pairs and required approximately one and a half to two minutes to complete. The ASL lessons were accessed via the web on a participant's personal device (mobile phone or computer). Alerts were sent via email or text message to notify participants that a new lesson was available to them. After the participant finished a lesson, they would be given a brief status update consisting of the total money they had earned in the study so far and a message instructing them to either continue on to the next lesson or close their browser window and wait for the next lesson.

The study had three parts:

- **A half-hour training session:** During this part, I verified the participant's eligibility for the study and explained the purpose of the study. I also collected brief demographic data including age, native language, other languages known, and handedness. They were then presented with an example lesson on either their iPhone (if they were in a mobile content delivery condition) or a desktop computer (if they were in a desktop content delivery condition) to demonstrate the manner in which content would be delivered to them.
- **A week of ASL lessons:** Participants completed a week of ASL lessons consisting of 80 lessons spread across seven consecutive days (400 word presentations). These lessons were delivered on different platforms and under different

spacing conditions.

Participants were compensated in such a manner as to give them an incentive to finish all 80 lessons. Participants received \$0.625 per lesson they completed; thus, if a participant completed all 80 lessons, they would earn \$50 for that portion of the study. Participants were also compensated \$8/hour for the time they spent in the training and testing sessions with the researcher. Those sessions averaged 1.5 hours in total.

- **Followup testing:** Each participant returned to be tested a week after his/her lessons ended. Each participant completed the NASA Task Load Index (NASA-TLX) [53] workload assessment, receptive and expressive tests, and participated in a brief interview. An example of a participant during the expressive test is shown in Figure 15. During the receptive test, a participant watched a video of 40 randomly selected signs from the original list of 80 words and translated them to English. During the expressive test, the participants attempted to sign the other 40 words. The participants were not told that the expressive and receptive tests were mutually exclusive.

5.1.1.1 Experimental Conditions

The study was designed with two independent variables: the method of delivering content (mobile phones vs. desktop/laptop computers) and the timing of the content delivery (Massed vs. Distributed practice). This design led to four conditions in the experiment:

- Phone delivery with Distributed lessons. (PD)
- Phone delivery with Massed lessons. (PM)
- Computer delivery with Distributed lessons. (CD)

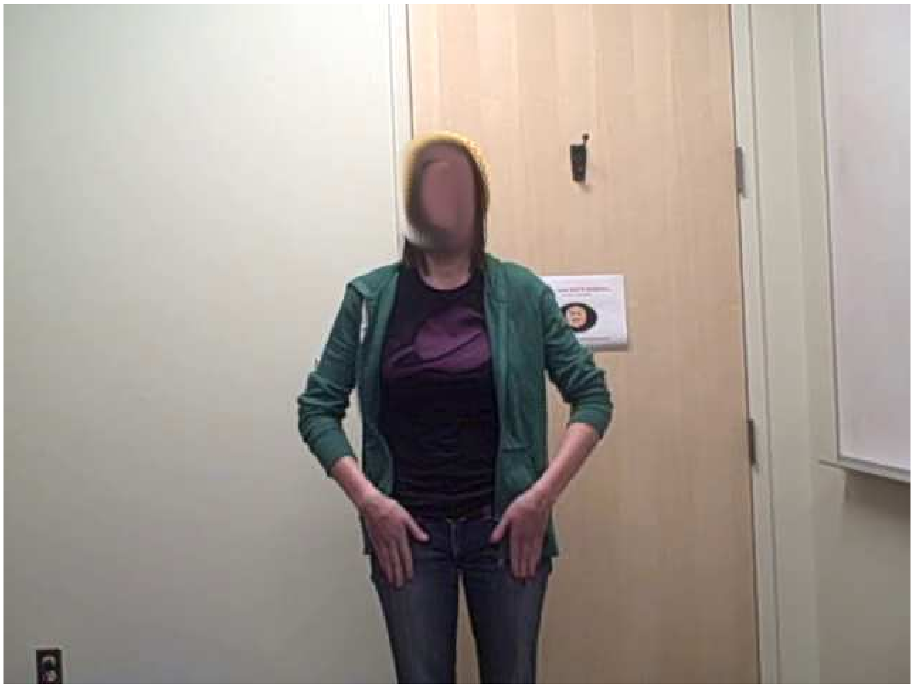


Figure 15: Participant Signing ‘pants’ During the Expressive Post-test

- Computer delivery with Massed lessons. (CM)

In the “Phone” conditions, the participants used their iPhone and accompanying data plan, while in the “Computer” condition they used a computer of their own choosing. During the study, participants reported using their laptops, home computers, or computers available at labs around campus. iPhones were chosen as an easy platform on which to standardize screen real estate, supported video codecs, and included data plans. This choice reduced the inconsistency of mobile phones platforms. Additionally, as of December 2009, iPhones account for 40% of smartphone handsets [1] in use in the United States.

In the “Massed” conditions, participants were given the option to complete all their lessons for a given day at once. In the Distributed condition, lessons are evenly spaced throughout the day, and a minimum time of four minutes between lessons was enforced.

5.1.1.2 Alert Scheduling

Participants where allowed to specify when they would like to receive alerts about lessons by setting starting and ending times. While participants had to choose a minimum five hour block of time, they were allowed to specify as large a window as they wished. This scheduling was done to allow the participants some control over the hours alerts would arrive, as there were external constraints on participant's time such as sleeping habits and work or school schedules. In all conditions, an alert would never be sent outside of the participants' designated times. Additionally, if an alert had been sent and the corresponding lesson was not completed, a reminder alert would be sent an hour later. This timing also mitigated the danger of a participant accidentally deleting a text message or email.

When participants finished a lesson, the software would generate another lesson for them with a scheduled alert time. This time might be immediate (in the "Massed" conditions) or another time during the day more than four minutes away (in the "Distributed" conditions.)

In the "Massed" conditions, participants were scheduled to receive a single alert, usually at the beginning of their self-designated window. If the participants continued through the lessons immediately, no other alerts were sent. However, if the user paused for more than four minutes, an alert was sent prompting them to pick up with the lessons where they had left off.

In the "Distributed" condition, participants were scheduled to receive the alerts evenly distributed throughout their chosen window of time. If a participant delayed several lessons, the system would compensate and adjust the alert times of subsequent lessons in order to try to keep the participant on pace to complete 80 lessons over the seven consecutive days. Thus, the system could give the impression of "speeding up" if a participant in the "Distributed" condition did not respond to several alerts. This scheduling ensured that all participants had the opportunity to complete all lessons if they responded to the alerts promptly. However, many participants did not respond immediately to the alerts and fell behind schedule. The analysis of the number of lessons participants completed and the time to respond to an alert are further discussed in Section 5.2.3.

5.1.1.3 Software design

The system was designed as a web application coded with PHP and a mySQL database backend. The database stored times of alerts, the five words to be presented in a given lesson, the correct or incorrect answers generated by the participants, the number of times an alert was sent before a participant responded, and the beginning and ending times of each lesson.

Figures 16-20 show screen shots of the iPhone web application while Figures 21-25 show the desktop interface. The following interaction sequence summarizes the steps necessary for a user to complete a lesson. In several steps, two figures are referenced. The first figure references a user in the Phone condition using an iPhone, and the second references a user in the Computer condition using a standard desktop web-browser. The interaction sequence is:

1. User receives an alert via email or text message.

2. User taps or clicks on the URL contained in the alert.

3. User is taken to the first page of the lesson. (Figure 16, Figure 21)

4. User taps the box with the “Play” icon (Figure 16A), or clicks the video directly (Figure 21A) to watch the video (Figure 17, Figure 22)

5. After video concludes, the user is returned to the play screen (Figure 16, Figure 21). The user then taps or clicks the “Continue” button to continue to the multiple choice test. (Figure 16B, Figure 21B)

6. User selects the button corresponding to his/her answer (Figure 18, Figure 23)
7. User receives feedback about the selected answer (Figure 19A, Figure 24A)
8. User taps or clicks the "Continue" button to go to the next video (Figure 19B, Figure 24B)
9. User repeats Steps 4-8 four more times.
10. User receives status message about future lessons. (Figure 20, Figure 25)

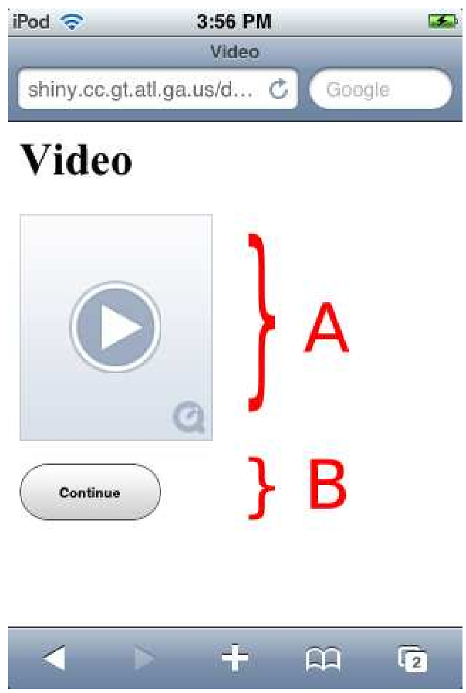


Figure 16: iPhone Interaction Part 1 - A: Video Icon and B: Continue Button

Figure 17: iPhone Interaction Part 2

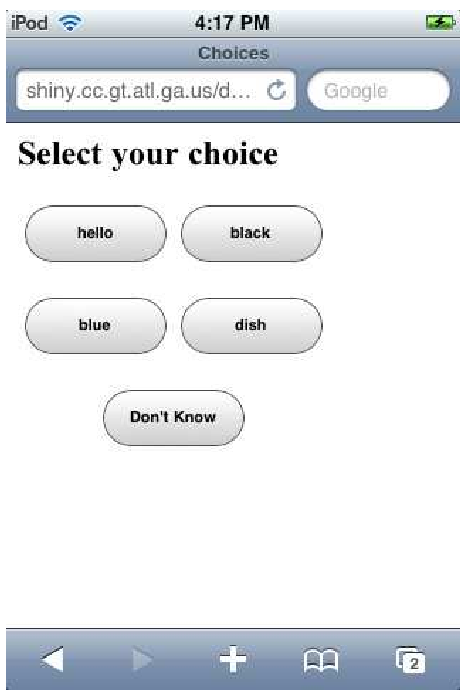


Figure 18: iPhone Interaction Part 3

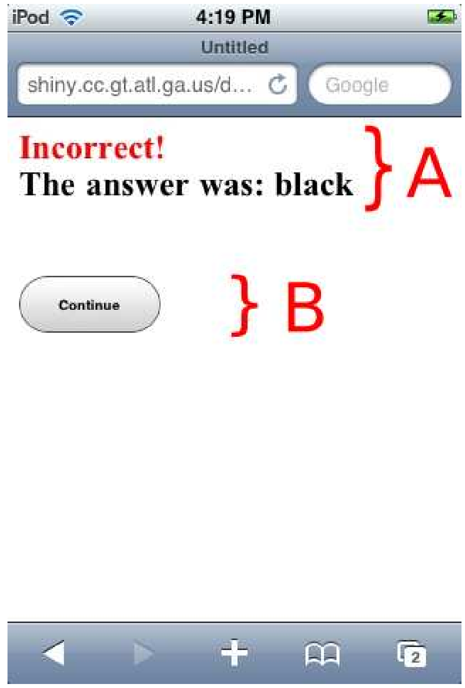


Figure 19: iPhone Interaction Part 4 - A: Feedback to User and B: Continue Button

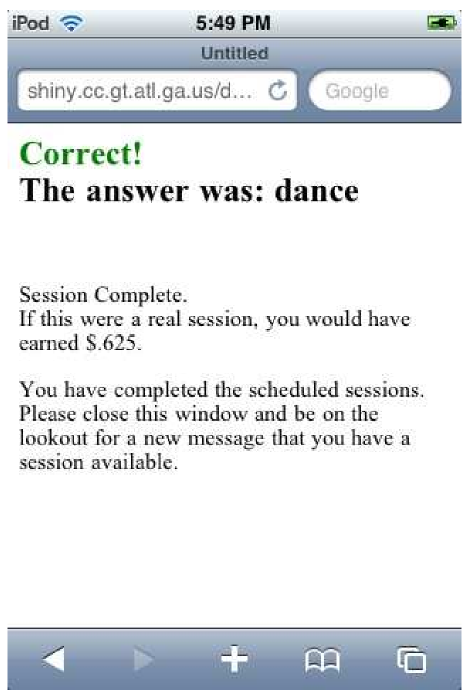


Figure 20: iPhone Interaction Part 5

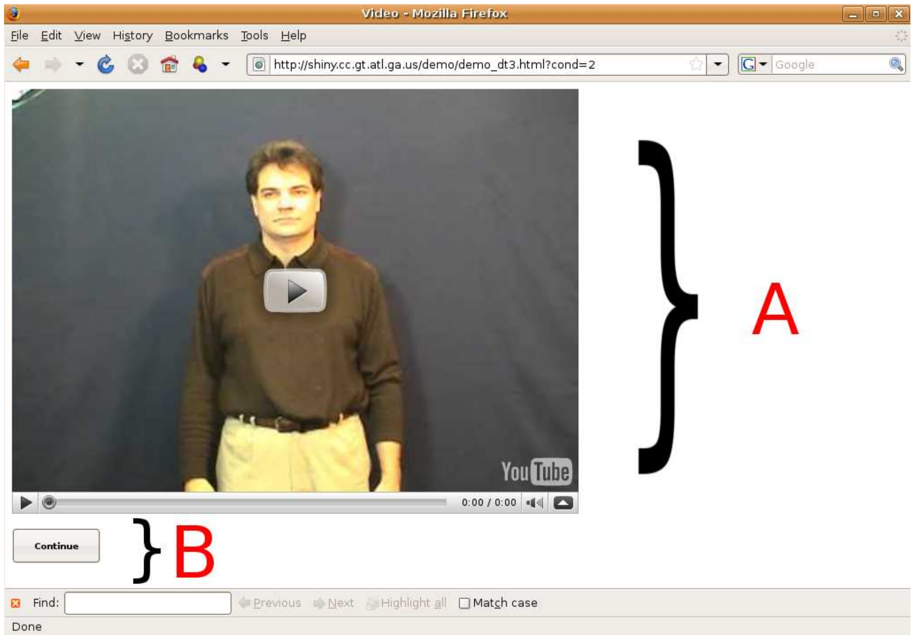


Figure 21: Computer Interaction Part 1 - A: Video Icon and B: Continue Button

Figure 22: Computer Interaction Part 2

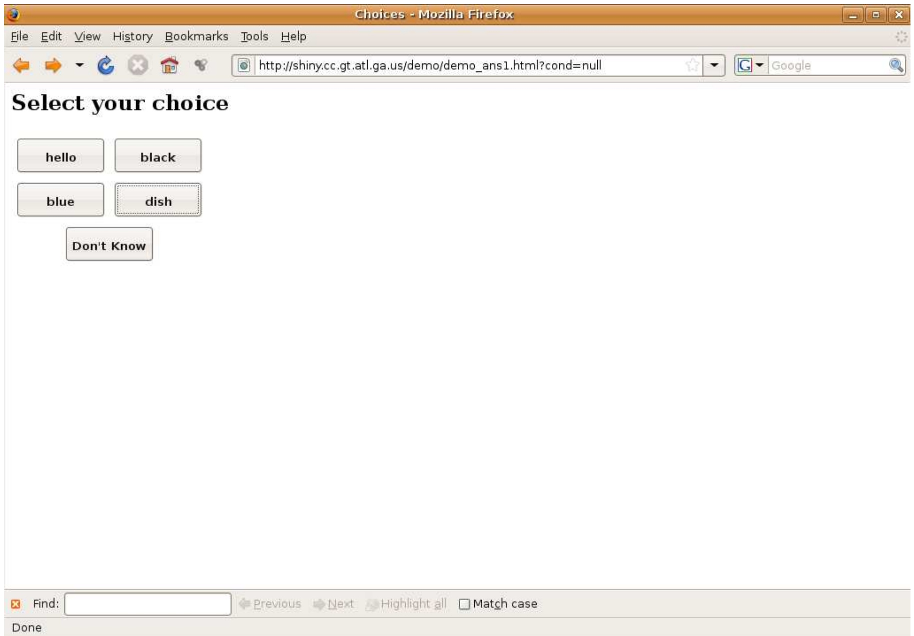


Figure 23: Computer Interaction Part 3

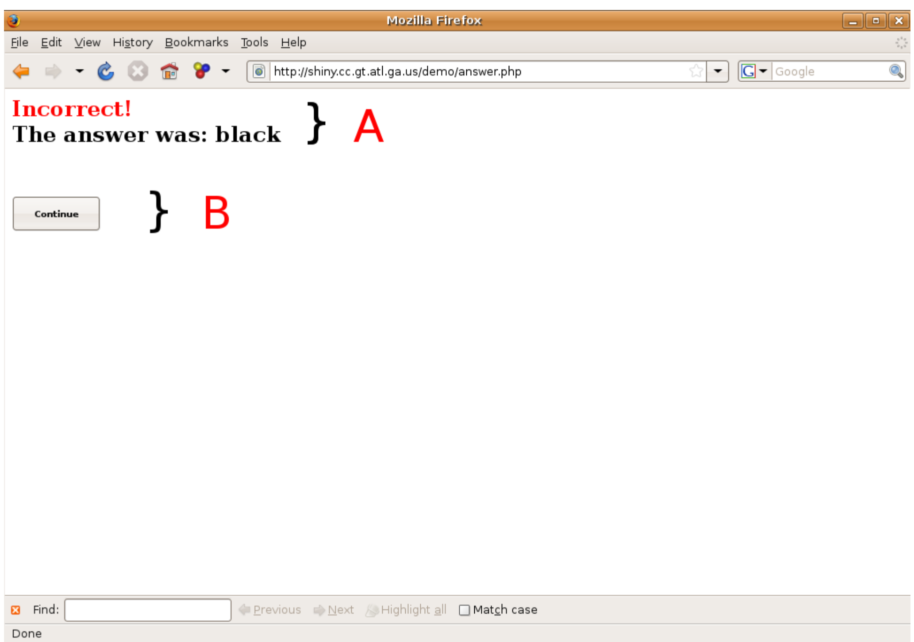


Figure 24: Computer Interaction Part 4 - A: Feedback to User and B: Continue Button

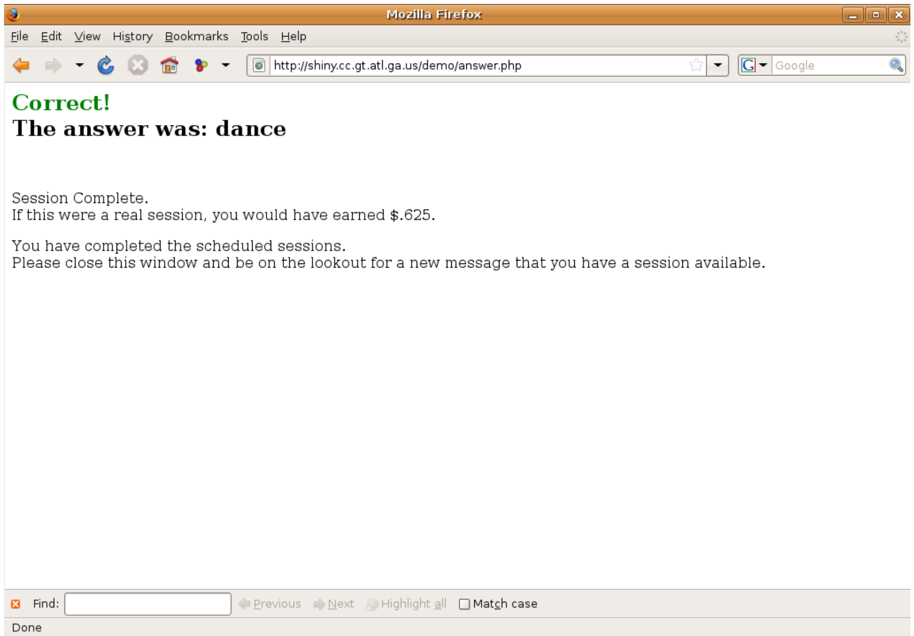


Figure 25: Computer Interaction Part 5

The "Computer" condition webpages used videos stored on YouTube [45]. The YouTube API [44] allowed us to prevent the videos from being watched multiple times and removed the issue of video codecs from the study as YouTube is Flash-based. Since the iPhone does not allow web applications to access the API, I conducted an analysis of logging files from the Apache webserver to assure that participants were not watching the videos multiple times in the "Phone" conditions. Additionally, participants in both conditions were asked to watch the videos only once. Analysis revealed no participant violated this request.

The 80 words to be learned were ranked in order of increasing difficulty as determined by the probability of transitioning from an unknown/forgotten state to a long-term memory state given a correct answer by the participant (see Table 20 in Appendix A). They were then divided among 8 groups to mimic Atkinson's original study [3]. Table 7 shows the resulting groups. In Table 7, the first word listed in each group is the word with the highest probability of transition to long term memory. Thus, the first row of words (i.e., I, drink, banana, etc) are considered the easiest to learn while the last row of words (i.e., want, stop, bad, that, etc) are the hardest to learn. To form a lesson, the word with the highest probability of transitioning to long term memory was chosen from each group. As a lesson consisted of only five words, the groups were chosen in a sequential manner to ensure consistency. Therefore, the words for the first lesson would be chosen from Groups 1-5 in Table 7, and the words for the second lesson would be chosen from groups 6, 7, 8, 1, and 2. If all words in a group were predicted to be in long term memory, a word was chosen at random from the group. If the words were not grouped, then the algorithm would always choose the word with the highest probability of transition to long term memory. It would then repeatedly present the same word until it predicted the word was learned. Grouping the words into eight groups as in the original experiment performed by Atkinson, ensured that the same word was never presented twice in the same lesson. Grouping introduced variability to the lessons.

Table 7: Grouping of Words

Grp 1	**Grp 2**	**Grp 3**	**Grp 4**	**Grp 5**	**Grp 6**	**Grp 7**	**Grp 8**
I	drink	banana	book	eat	tired	look	thirsty
my	soap	love	you	hurry	out	baby	little
down	up	bedroom	this	pants	cold	sleep	in
there	cat	toy	water	food	truck	hungry	go
juice	off	apple	big	shoes	yesterday	dad	medicine
dog	on	now	car	milk	happy	wait	brother
no	finish	more	sick	help	grandmother	sweet	who
home	sister	your	please	grandfather	person	yes	hello
where	hot	mom	school	what	good	jacket	tomorrow
want	stop	bad	that	bathroom	not	careful	thank you

5.1.2 Participants

For this study, I recruited 40 participants over four months in late 2009, 11 females and 29 males with an average age of 24.5 years old. All participants had no prior knowledge of American Sign Language beyond fingerspelling which is sometimes taught in elementary school.

Participants were recruited via email primarily from the Georgia Tech community and were assigned to one of the four experimental conditions. If the participant owned an iPhone, they were randomly assigned to either the Phone Distributed or Phone Massed condition. If the participant did not own an iPhone, they were randomly assigned to either the Computer Distributed or Computer Massed condition. Ten participants were assigned to each of the conditions. There were no statistically significant differences in the mean age of participants assigned to any of the four conditions. Table 8 shows statistics about the ages of each of the four conditions.

Table 8: Age Demographics by Condition

	Condition	**Mean**	**Std. Dev**	**Min**	**Max**
Age (years)	PD	25.0	7.1	18	40
	PM	25.6	5.1	20	33
	CD	23.7	5.2	18	36
	CM	23.7	5.9	18	34
	Total	24.5	5.7		

Recruiting iPhone users and more generally, from the Georgia Tech community, does introduce a population bias. However, having a tech-savvy population who were experts in using their own devices allowed me to eliminate any learning curve that would have occurred in a population using an unfamiliar mobile phone.

As mentioned above, participants were allowed to select which hours they wished to receive alerts and lessons. On average, participants selected a window 11.4 hours long with a maximum of 16 hours and a minimum of five hours. Table 9 shows this information broken down by condition. In Table 9 and following tables, statistically significant differences between pairs of conditions are indicated in the last column with matching symbols. As can be seen in Table 9, participants in the Phone Massed condition chose smaller windows of time for the lessons to be delivered than participants in any of the other conditions. Participants in the Phone Massed condition chose a

Table 9: Delivery Window Statistics – *Statistical Significance shown in last column with corresponding symbols*

		Mean	Std. Dev.	Minimum	Maximum	Stat. Sig.
Delivery Window (seconds)	PD	40499.8	9572.1	21600	50400	
	PM	29249.8	9760.6	18000	49500	$\star, \dagger$
	CD	47963.9	8457.1	28800	57600	$\star$
	CM	46436.9	12211.6	21600	57600	$\dagger$
	Total	41037.6	12228.1			

window of only 8.12 hours compared to the overall average of 11.4 hours. In fact, the difference in the window of time was statistically significant between the Phone Massed participants and the participants in both of the Computer conditions, and this information is shown in the last column of Table 9. Participants in the Phone Massed condition ($M = 29248.8$ seconds) chose smaller windows ($F(3, 36) = 7.079, p = .001$, Tukey post-hoc $p = .001$) than did those in the Computer Distributed ($M = 47963.9$ seconds). Also, participants in the Phone Massed condition ($M = 29248.8$ seconds) chose smaller windows ($F(3, 36) = 7.079, p = .001$, Tukey post-hoc $p = .003$) than did those in the Computer Massed condition ($M = 46436.9$ seconds).

5.2 *Quantitative Results*

Here, I present a detailed analysis of the quantitative results (5.2.1) obtained from logging and from post-tests evaluation. In the following section (5.3), I present the results obtained from participant interviews.

5.2.1 Quantitative Results

There are several ways to analyze the data collected. First, it is useful to examine each of the four conditions: Phone delivery with Distributed lessons (PD); Phone delivery with Massed lessons (PM); Computer delivery with Distributed lessons (CD); and Computer delivery with Massed lessons (CM).

It is also useful to examine differences that may have been related to delivery method (Phone vs. Computer) or presentation (Massed vs. Distributed). Analysis for each of three different configurations (comparison of all four conditions, comparison of Phone vs. Computer conditions, and comparison of Massed vs. Distributed conditions) is presented below. I also present a detailed analyses of the NASA-TLX data obtained in this study. Additionally, I compare the data obtained in this study using the Atkinson algorithm to the data collected in the Data Gathering Study (DGS), described in Chapter 4, which used a naive scheduling algorithm to present the words.

After a discussion on the dependent variables in this study and the types of data analyzed, I present detailed analysis of these categories.

5.2.2 Data Collected

In this study, participants were scheduled to complete 80 lessons over the week (seven consecutive days) of lessons. However, some participants completed fewer lessons. In some cases, this difference in the number of lessons completed meant that the participant did not see all 80 words. In calculating the percentage of questions answered correctly on the final expressive and receptive tests, I adjusted for words that the participants had not seen. For example, if a participant completed only 75 lessons and never saw the word ‘dog’, I did not count an incorrect response to ‘dog’ on the final tests. By adjusting the final scores this way, the percentages more accurately reflect any learning that occurred. A participant could not learn a word they did not see, and on the final tests, there were no instances of a participant correctly guessing or signing a word they did not see during the week of lessons.

The data presented and analyzed included:

1. **Number of Lessons Completed**: Each participant was scheduled to complete 80 lessons over seven consecutive days. Each lesson had five words. However, not all participants kept to the schedule. This variable counts the number of

lessons completed out of the maximum of 80.

2. **Words Seen**: Since not all participants completed all 80 lessons, some participants did not view all 80 words. Thus, if they never saw the words, they had no chance to learn them. This variable counts the number of words the participants saw out of the maximum total of 80 words.

3. **Receptive Percentage Correct**: When scoring the receptive test in which the participant watched a video and wrote down the English translation, the answer was either correct or incorrect. This variable reports the percentage of words the participant answered correctly of the words they had seen. This method of calculating the percentage normalizes the percentage based on the fact that the participant may not have been exposed to all words. Since words on the receptive test were marked as either 'correct' or 'incorrect,' the receptive percentage incorrect can be calculated by merely subtracting the percentage correct from 100.

4. **Expressive Percentage Correct**: The responses to the expressive test (in which the participant tried to sign words upon hearing the English translation) were coded in the same manner as described earlier in Section 4.1.1.3: a response scored as "1" was completely incorrect or unknown, a response scored as a "2" had some elements such as handshape, motion, or body position correct, and a response scored as a "3" was completely correct.

   This variable reports the percentage of words the participant signed correctly (i.e., were rated a "3"). Again, this data is normalized to allow comparison amongst all participants who were exposed to varying numbers of words. Counting only the words which were completely correct (e.g., by not including words that were rated as a "2" or somewhat correct) leads to a conservative estimate of correctness.

5. **Expressive Percentage Incorrect**: Same as above except this variable reports the percentage of words the participant signed completely incorrectly (i.e., were rated a "1") after normalization to account for the fact that the participant may not have been exposed to all words. This metric counts the number of signs which the participants answered completely wrong since it does not include any signs scored at a '2' or a '3'. For example, a participant who got many signs partially correct might have a low incorrect percentage even though they also have a low percentage correct.

6. **Average Lesson Time**: For the lessons that the participant completed, this variable reports the average time that elapsed between when the participant started the lesson and when he finished it.

7. **Average Time to Respond to Alert**: For the lessons that the participant completed, this variable reports the average time that elapsed between the first time an alert was sent and when they began the lesson. Additionally, this variable is only computed for lessons in which an alert was actually sent. For example, if a participant in the Massed condition responded to an alert for the first lesson of the day and then completed all subsequent lessons immediately, the only data used in this variable's calculation would be the time elapsed from the first alert to the beginning of the first lesson.

8. **NASA-TLX Workload Measurement**: This variable is the overall workload score computed using the NASA-TLX measurement.

In the following sections, I use the metrics above to compare conditions in the final learning study (Section 5.2.3). I then compare some of the data from this study which uses the Atkinson algorithm of scheduling to the data obtained in the DGS which uses a naive scheduling algorithm (Section 5.2.4).

5.2.3 Analysis of Learning Study

Using the above metrics, I now compare all four conditions to one another (Section 5.2.3.1). I then present a comparison of the Phone vs. Computer delivery platforms (Section 5.2.3.2) and the Massed vs. Distributed presentation strategies (Section 5.2.3.3). I also discuss the NASA-TLX workload measurements (Section 5.2.3.4) for the entire task as well as the categorical data. I conclude with a brief summary of the findings.

5.2.3.1 Comparison of All Four conditions

Table 10 reports the mean, standard deviation, minimum value, and maximum value for each of the four conditions outlined above. For convenience, the abbreviations listed above for the four conditions are used in reporting the data.

All between-condition measures were analyzed using a one-way ANOVA with an $\alpha = 0.05$ criteria to check for statistical significance. The number of completed lessons $(F(3, 36 = 4.294), p = .011)$, receptive percentage correct $(F(3, 36) = 3.542, p = .024)$, and the average time to respond to an alert $(F(3, 36) = 6.733, p = .001)$ showed statistically significant differences. Tukey post-hoc comparisons were then used to differentiate amongst the four conditions. The Tukey comparison was chosen as it is a conservative measure of significance, the population variances are equal, and the sample sizes were equal. The average length of lesson showed significance $(F(3, 36) = 3.153, p = .037)$ during the ANOVA analysis but the Tukey post-hoc comparison showed no significance.

When looking at the average number of lessons completed, statistically significant differences were found between the Phone Massed ($M = 77.3$) and Computer Distributed ($M = 66.8$) conditions ($p = 0.039$) and between the Computer Distributed ($M = 66.8$) and Computer Massed ($M = 79.3$) conditions ($p = 0.010$). Thus, the participants in the Computer Distributed condition completed significantly fewer lessons

Table 10: Analysis of Conditions 1-4 – *Statistical Significance shown in last column with corresponding symbols*

		Mean	Std. Dev.	Min	Max	Stat. Sig.
Lessons Completed	PD	73.3	8.3	54	80	
	PM	77.3	6.1	61	80	$\star$
	CD	66.8	13.0	47	79	$\star, \dagger$
	CM	79.3	2.2	73	80	$\dagger$
	Total	74.3	9.4			
Words Seen	PD	78.7	2.3	74	80	
	PM	79.0	1.7	75	80	
	CD	76.6	5.1	66	80	
	CM	79.5	1.6	75	80	
	Total	78.45	3.1			
Recept Correct (%)	PD	41.0	18.8	15.0	67.5	$\circ$
	PM	63.0	14.5	35.9	90.0	$\circ$
	CD	45.5	14.0	19.4	67.5	
	CM	47.4	16.7	20.0	72.5	
	Total	49.2	17.6			
Express Correct (%)	PD	21.8	10.5	7.5	35.0	
	PM	32.8	15.8	13.9	72.5	
	CD	25.8	16.0	5.0	62.5	
	CM	28.6	9.7	20.0	47.5	
	Total	27.2	13.4			
Express Incorrect (%)	PD	63.3	10.6	48.6	82.5	
	PM	45.3	17.3	12.5	77.8	
	CD	59.6	17.5	35.0	85.0	
	CM	58.4	13.5	35.0	75.0	
	Total	56.6	16.0			
Avg. Lesson Time (seconds)	PD	352.0	333.2	98.3	1037.2	
	PM	432.1	538.6	98.1	1826.6	
	CD	86.8	42.8	34.9	173.0	
	CM	87.3	64.2	33.6	255.9	
	Total	239.5	344.4			
Avg. Response Time (seconds)	PD	4370.5	2175.7	1319.62	9251.4	
	PM	1676.4	1185.8	60.8	3587.7	$\ddagger$
	CD	6326.7	2597.8	3214.0	11197.5	$\ddagger, \diamond$
	CM	3481.8	3052.6	541.6	8858.0	$\diamond$
	Total	3963.9	2828.4			
NASA-TLX	PD	37.4	14.5	9	55	
	PM	33.5	17.5	0	55	
	CD	34.9	11.2	22	52	
	CM	32.1	8.7	15	43	
	Total	34.5	13.0			

then participants in either the Computer Massed or the Phone Massed conditions. Comparisons between other pairs of conditions were not statistically significant at $p < .05$.

The percentage of correct responses on the receptive tests also showed a statistically significant difference ($p = 0.021$) between conditions Phone Distributed ($M = 41.0\%$) and Phone Massed ($M = 63.0\%$). Thus, the participants in the Phone Distributed condition answered significantly fewer of the receptive questions correctly on average than the participants in the Phone Massed conditions. It is interesting to note that these two groups did not have a statistically significant difference in the number of lessons completed.

The average time to respond to an alert also differed significantly. A significant difference ($p = 0.000$) existed between the Phone Massed ($M = 1676.4$ seconds) and Computer Distributed ($M = 6326.7$ seconds) conditions. There was also a statistically significant difference ($p = .049$) between the Computer Massed ($M = 3481.8$) and Computer Distributed ($M = 6326.7$) conditions. The participants in the Phone Massed condition responded to an alert sent to them, on average, 4650.3 seconds (or approximately 1 hour and 18 minutes) faster than the participants in the Computer Distributed condition did. Likewise, the participants in the Computer Massed condition responded 2844.9 seconds (or approximately 47 minutes) faster than did the participants in the Computer Distributed condition.

5.2.3.2 Comparison of Delivery Method: Phone vs. Computer

I also wanted to investigate the data set for differences in the method of content delivery: Phone vs. Computer. I grouped the data by content delivery and ran a two-tailed t-test on the two groups. Data for mean and standard-deviation is presented in Table 11.

Only two of the variables, average lesson time and the average response time

Table 11: Analysis of Phone vs. Computer – *Statistically Significant conditions shown in bold*

		Mean	Std. Dev
Lessons Completed	Phone	75.5	7.3
	Computer	73.1	11.1
Words Seen	Phone	78.9	2.0
	Computer	78.1	4.0
Recept Correct (%)	Phone	52.0	19.8
	Computer	46.4	15.0
Express Correct (%)	Phone	27.3	14.2
	Computer	27.2	13.0
Express Incorrect (%)	Phone	54.3	16.8
	Computer	59.0	15.2
Avg. Lesson Time (seconds)	Phone	392.1	437.8
	Computer	87.0	53.
Avg. Response Time (seconds)	Phone	3023.5	2195.1
	Computer	4904.2	3121.0
NASA-TLX	Phone	35.5	15.8
	Computer	33.5	9.8

to messages, showed statistically significant differences. Levene's test for equality of variances showed that both variables did not have equal variances, and thus the significance figures for unequal variances are reported here.

The average lesson length showed differences ($t(19.56) = 3.09, p = .006$). The Phone condition had a mean of 392.1 seconds and the Computer having a mean of 87.0 seconds. On average, participants in the Phone conditions required 305 seconds (5.085 minutes) longer to complete a lesson of five words than the participants in the Computer condition. While some of this difference can be explained by differences in operating systems and network latency, the five minute difference is too large to be explained solely by these factors. The interview data (Section 5.3) suggests that participants also became distracted in the Phone conditions which explains the large difference more completely.

The average time to respond to messages showed a difference ($t(34.1) = -2.02, p =$

.034) with the Phone conditions having a mean of 3023.5 seconds (approximately 50 minutes) and the Computer conditions having an average of 4904.2 seconds (approximately 1 hour and 22 minutes).

5.2.3.3 *Comparison of Presentation: Massed vs. Distributed*

In another combination analysis, I grouped the data by the presentation methods. In the Massed conditions, participants were sent an alert and, after finishing the first lesson, were allowed to continue through the day's lessons without stopping. In the Distributed condition, the alerts and lessons were equally distributed throughout the time window the participant had given us. If the participant delayed completing a lesson, the system tried to schedule the remaining lessons into the time available for that day. However, the system enforced a minimum time of four minutes between lessons to differentiate the Distributed from the Massed condition.

Both the Distributed and Massed conditions were programmed to allow the participant to complete up to 12 lessons per day for the first six days and eight on the last day. Lessons that were not completed were carried over and scheduled for the next day. Thus, the system would speed up the sending of the alerts in the Distributed condition if the participant fell behind. Table 12 shows the mean and standard deviation for the measures mentioned in Section 5.2.2.

The number of lessons completed showed statistically significance ($t(25.22) = -2.97, p = .006$) differences. As before, the number of lessons completed showed that variances were unequal using Levene's test. Thus, significance has been reported using the assumption that variances are not equal. Lessons completed differed between Distributed ($M = 70.3$) and Massed ($M = 78.3$) by eight lessons.

With this grouping of data, the percentage of receptive words answered correctly showed a statistically significant difference ($t(38) = -2.25, p = .030$). Participants in the Massed condition answered 55.2% of the questions correctly, while participants

Table 12: Analysis of Distributed vs. Massed – *Statistically Significant conditions shown in bold*

		Mean	**Std. Dev**
Lessons Completed	Distributed	70.3	11.2
	Massed	78.3	4.6
Words Seen	Distributed	77.7	4.0
	Massed	79.3	1.6
Recept Correct (%)	Distributed	43.3	16.3
	Massed	55.2	17.2
Express Correct (%)	Distributed	23.8	13.4
	Massed	30.7	12.9
Express Incorrect (%)	Distributed	61.5	14.2
	Massed	51.8	16.5
Avg. Lesson Time (seconds)	Distributed	219.4	268.3
	Massed	259.7	413.1
Avg. Response Time (seconds)	Distributed	5348.6	2538.9
	Massed	2579.1	2436.7
NASA-TLX	Distributed	36.15	12.7
	Massed	32.8	13.5

in the Distributed condition answered only 43.3% correctly.

Additionally, the response time to alerts differed between these two conditions ($t(38) = 3.52, p = .001$). It is important to remember that this time includes only lessons when an alert was sent. Thus, if a participant was in the Massed condition and completed all lessons in succession after a single alert was sent, only the time between when the alert was sent and the first lesson's beginning was used in this calculation. The participants in the Distributed condition had an average response time of 5348.6 seconds (1 hour and 29 minutes) while participants in the Massed condition had a response time of 2579.1 seconds (43 minutes).

5.2.3.4 NASA-TLX Analysis

While no statistically significant differences were found in the overall NASA-TLX workload (see sections 5.2.3.1, 5.2.3.2, and 5.2.3.3), differences emerged when examining each of the six subcategories of the test. For convenience, these categories are summarized here, with a brief explanation (from Hart and Staveland, [53]).

- **Mental Demand**: How much mental and perceptual activity was required (e.g., thinking, deciding, calculating, remembering, looking, searching, etc)? Was the task easy or demanding, simple or complex, exacting or forgiving?

- **Physical Demand**: How much physical activity was required (e.g., pushing, pulling, turning, controlling, activating, etc)? was the task easy or demanding, slow or brisk, slack or strenuous, restful or laborious?

- **Temporal Demand**: How much time pressure did you feel due to the rate or pace at which the tasks or task elements occurred? Was the pace slow and leisurely or rapid and frantic?

- **Own Performance**: How successful do you think you were in accomplishing the goals of the task set by the experimenter (or yourself)? How satisfied were you with your performance in accomplishing these goals?

- **Effort**: How hard did you have to work (mentally and physically) to accomplish your level of performance?

- **Frustration Level**: How insecure, discouraged, irritated, stressed, and annoyed versus secure, gratified, content, relaxed and complacent did you feel during the task?

Figure 26 shows a breakdown of these six subcategories by experimental condition. As can be seen from the figure, the mental demands of the task dominated the overall score; however, none of the conditions showed a statistically significant difference from any of the other conditions.

Table 13 gives a numerical breakdown of each of the categories of workload measured by the NASA-TLX assessment.

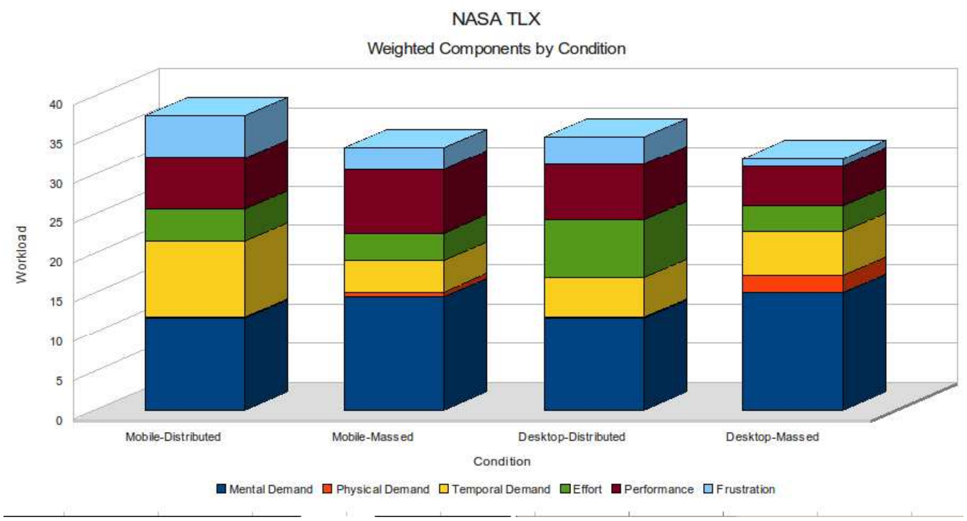


Figure 26: NASA-TLX Category Breakdown

While the differences were not statistically significant, it is interesting to note that the participants in the Phone Distributed condition rated the Temporal demand higher than participants in any other condition ($M = 9.6$ as compared to $M = 4.0$, 5.0, or 5.5). This result reflects that participants in this condition were receiving alerts via text message directly to their iPhone throughout the day. Text messages often incur a burden or urgency to respond [46, 47], and the participants may have felt a time pressure to respond to the alert. Participant 31 (PD), hinted at this when discussing the timing of the alerts and lessons around a fixed time commitment

> "Most of the time they fit in pretty good except for around practice time. Then I came out of practice and I would have like four unread messages so I had to get them in real quick after that. So I could make sure the next one would come in the right amount of time.

Unlike participants who were in the Computer conditions, participants in the Phone conditions received the alerts immediately on their iPhones, perhaps avoiding an "out of sight, out of mind" phenomena for the email alerts sent to Computer participants.

Table 13: NASA-TLX Numerical Breakdown of Categories by Condition

		Mean	**Std. Dev**	**Minimum**	**Maximum**
Mental Demand	PD	11.8	5.7	0.0	18.4
	PM	14.4	9.1	0.0	26.7
	CD	11.7	8.6	1.3	25.7
	CM	15.0	8.1	0.7	24.0
Physical Demand	PD	0.1	.3	0.0	0.9
	PM	0.6	1.8	0.0	5.7
	CD	0.2	0.4	0.0	1.1
	CM	2.2	6.3	0.0	20.0
Temporal Demand	PD	9.6	10.5	0.0	26.7
	PM	4.0	5.4	0.0	17.4
	CD	5.0	3.2	0.0	11.7
	CM	5.5	3.8	0.7	10.4
Effort	PD	4.1	3.5	0.0	9.7
	PM	3.5	3.6	0.0	10.0
	CD	7.3	4.1	0.2	12.8
	CM	3.3	3.7	0.0	13.3
Performance	PD	6.5	4.7	0.6	15.7
	PM	8.1	6.3	0.0	16.3
	CD	7.1	6.4	0.9	18.7
	CM	5.1	2.9	0.7	10.0
Frustration	PD	5.4	6.6	0.0	20.8
	PM	2.8	4.9	0.0	16.0
	CD	3.6	4.5	0.0	12.7
	CM	1.0	1.1	0.0	2.7

Another category which showed an interesting difference was the Effort category. In this category, the participants in the Computer Distributed condition rated Effort higher than those in the other conditions (M = 7.3 vs. M = 4.1, 3.5, or 3.3). I hypothesize that this difference occurred because of the difficulty in receiving alerts. Participants in the Computer Distributed condition not only had to complete the lessons, but they had to check their email multiple times per day to make sure they completed the lessons as the alerts arrived. For participants who were not always at their computers, this requirement could have raised the level of effort they needed to make to participate fully in the study. Participant 13 in the Computer

Distributed condition said that he, "had to get used to checking it [email] more frequently, especially since it doesn't send you a new lesson until a little while after you've completed the previous one. So the first day, I kinda forgot about it during the day..." When asked if the alerts were ever intrusive or annoying, Participant 6 who was also in the Computer Distributed condition said, "The only thing would be if I forgot the alerts were coming or if I was away from the computer for an extended period of time, I'd come back and there would be like six warnings and it was like 'OK!"'. Both of these responses explain why the ratings for effort would be higher for the Computer Distributed condition.

5.2.3.5 Summary

In summary, there were several significant results from the quantitative analysis of the learning study. They are listed below for clarity. All the results below showed statistically significant differences with $p < .05$:

- Comparison of all four conditions:

    1. Participants in the Phone Massed condition completed more lessons than those in the Computer Distributed condition (77.3 lessons vs. 66.8 lessons)
    2. Participants in the Computer Massed condition completed more lessons than those in the Computer Distributed condition (79.3 lessons vs. 66.8 lessons)
    3. Participants in the Phone Massed condition answered correctly a higher percentage of the time on the receptive test then did those in the Phone Distributed condition (63.0% vs. 41.0%).
    4. Participants in the Phone Massed condition answered alerts faster on average than participants in the Computer Distributed condition (1676.4 seconds vs. 6326.7 seconds).

5. Participants in the Computer Massed condition answered alerts faster on average than participants in the Computer Distributed condition (3481.8 seconds vs. 6326.7 seconds).

- Comparison of Phone condition and Computer condition:

  1. Participants in the Phone conditions took longer to complete a lesson than did those in the Computer condition (392.1 seconds vs. 87.0 seconds).
  2. Participants in the Phone conditions answered alerts faster than did those in the Computer conditions (3023.5 seconds vs. 4904.2 seconds).

- Comparison of Massed condition and Distributed condition

  1. Participants in the Massed conditions completed more lessons than did those in the Distributed conditions (78.3 lessons vs. 70.3 lessons).
  2. Participants in the Massed condition answered more questions correctly on the receptive test than did those in the Distributed conditions (55.2% vs. 43.3%).

It is also useful to look at the rank ordering of some of the statistics. For example, when looking at the percentages on both the expressive and receptive tests, the Phone Massed condition always ranks highest (highest percentage of correct responses and lowest percentage of incorrect responses). It is then always followed by the Computer Massed condition, the Computer Distributed condition, and lastly the Phone Distributed condition. If we evaluate these conditions based only on the percentages correct, the Phone Massed condition is the best method of delivery. However, this rank ordering changes when we look at the number of lessons completed and the words seen. For those variables, participants completed the most lessons and saw the most words in the Computer Massed condition followed by the Phone Massed, Phone Distributed, and Computer Distributed.

Next, I compare the data obtained from the four conditions in this study to the naive scheduling algorithm previously discussed in Chapter 4.

5.2.4 Comparison to Data Gathering Study

The data from this study were also compared to the results from the Data Gathering Study (DGS) detailed in Chapter 4. As discussed in Chapter 4, participants in the DGS completed five trials of test–and–train presentations of vocabulary. Each trial contained one showing of each of the 80 words in random order. Therefore, each participant saw each word five times for a total of 400 presentations. Participants in the DGS completed their learning trials over approximately an hour in one day and returned a week later for the receptive and expressive tests described in Section 4.1.1.3. There were 20 participants in the DGS.

Participants in the study outlined in this chapter also had 400 word presentations. However, the number of times a participant saw each word differed due to the use of the Atkinson algorithm as the more difficult words were presented more often while the easier words were presented less fequently. Participants completed their lessons over a week, as detailed in Section 5.1.1. They then returned after a week to be given the receptive and expressive post-tests. There were ten participants per condition (40 total) in this study.

As an example, in the DGS, a participant would have seen the word "I", an easy vocabulary word in which the subject points to him/herself, five times regardless of how many times they correctly or incorrectly identified it. However, in the study in this chapter which uses the Atkinson algorithm, the participant might have seen "I" once and answered the multiple choice correctly. If the Atkinson model predicted the word had been learned, and the system would move on to other, more difficult words. After a week, participants in both studies were tested in the same manner, with generative and expressive tests.

The DGS detailed in Chapter 4 represents what might occur if someone focused on learning ASL via a computer program for an extended time. For this portion of the analysis, I only compared three metrics:

- Receptive Correct Percentage
- Expressive Correct Percentage
- Expressive Incorrect Percentage

Table 14 gives the statistics for the DGS in addition to the four different conditions. While the means and standard deviations were given in Table 10, for ease of comparison, they are duplicated here.

Table 14: DGS vs. Final Study – *Statistical significance shown in last column with corresponding symbols*

		N	**Mean**	**Std. Dev.**	**Min**	**Max**	**Stat. Sig.**
Recept Correct (%)	PD	10	41.0	18.8	15.0	67.5	$\star$
	PM	10	63.0	14.5	35.9	90.0	
	CD	10	45.5	14.0	19.4	67.5	†
	CM	10	47.4	16.7	20.0	72.5	
	DGS	20	65.5	21.8	27.5	95.0	$\star$, †
Express Correct (%)	PD	10	21.8	10.5	7.5	35.0	
	PM	10	32.8	15.8	13.9	72.5	
	CD	10	25.8	16.0	5.0	62.5	
	CM	10	28.6	9.7	20.0	47.5	
	DGS	20	34.8	19.0	2.5	75.0	
Express Incorrect (%)	PD	10	63.3	10.6	48.6	82.5	
	PM	10	45.3	17.3	12.5	77.8	
	CD	10	59.6	17.5	35.0	85.0	
	CM	10	58.4	13.5	35.0	75.0	
	DGS	20	48.9	18.3	12.5	95.0	

Looking at the percentage means, the DGS method of learning produced results very similar to that of the Phone Massed (PM) condition. I then performed an ANOVA ($\alpha = .05$) analysis comparing each of the four conditions to the DGS data.

The ANOVA analysis showed significant differences in the receptive correct percentage ($F(4,55) = 4.710, p = .002$), but no statistically significant differences in either the expressive correct ($F(4,55) = 1.457, p = .228$) or incorrect percentages ($F(4,55), p = .091$).

The Tukey post-hoc comparison after ANOVA analysis showed statistically significant differences between the DGS and the Phone Distributed (PD) condition ($p = .009$) with the DGS generating 24.5% more correct receptive responses than the Phone Distributed condition. The DGS was also statistically significantly different than the Computer Distributed (CD) condition ($p = .049$) with the DGS generating 20.0% more correct receptive responses.

Table 15 shows how much time the participants in each condition spent on instruction. The data for the participants in the DGS does not include the time spent on breaks between trials. This data is reported earlier in the form of 'average time spent on lessons' in Section 5.2.3.1. However, the aggregate reporting of total time spent is useful for comparing to the DGS when participants completed all their lessons in one experiment.

Table 15: Total Time (in minutes) Spent on Instruction – *Statistical significance shown in last column with corresponding symbols*

Experimental Condition	**N**	**Mean (minutes)**	**Std. Dev.**	**Stat. Sig.**
Phone Distributed	10	435.9	422.5	$\star$
Phone Massed	10	516.8	557.9	†
Computer Distributed	10	91.0	31.6	
Computer Massed	10	114.5	85.0	
DGS	20	51.68	6.2	$\star$, †

As can be seen from the standard deviation figures in Table 15, these times are incredibly variable. However, statistically significant differences exist. When running ANOVA analysis of this data, it showed that statistically significant differences in time existed ($F(4, 55) = 6.705, p = .000$). Additionally, the Tukey post-hoc comparison showed that significant differences existed between the DGS and the Phone Distributed condition ($p = .009$) with a difference of 384.2 minutes (or approximately 6.4 hours) spent on lessons. There was also a significant difference ($p = .001$) of 465.1 minutes (or 7 hours and 45 minutes) between the DGS and the Phone Massed condition. (Differences in time spent on lessons between the four conditions themselves are not reported here as they were reported in Section 5.2.3.1 and Table 10).

5.2.4.1 Comparison between DGS and Computer and Phone conditions

I also compared the DGS method of delivery content to the groups of Computer and Phone conditions. Again, ANOVA analysis showed a statistically significant difference in the number of receptive questions answered correctly ($F(2, 57) = 5.269, p = .008$) but no differences in the expressive questions answered correctly or incorrectly. Data for the mean and standard deviation of each of these conditions is listed in Table 16.

The Tukey post-hoc comparison revealed differences between the Computer conditions and the DGS ($p = .007$). The participants in the DGS answered 19.1% more questions correctly on the receptive test than those participants in the Computer conditions. The difference of 13.5% between the DGS and the Phone condition was not statistically significant.

5.2.4.2 Comparison between DGS and Distributed and Massed conditions

Finally, I also used ANOVA to test for differences between the DGS and the groups of Distributed and Massed conditions. The analysis showed a statistically significant difference for the receptive correct percentage ($F(2, 57) = 7.179, p = .002$). Further analysis using the Tukey post-hoc comparison revealed differences between the distributed conditions and the DGS ($p = .001$) with the participants in the

Table 16: DGS vs. Computer and Phone Conditions – *Statistical significance shown in last column with corresponding symbols*

		N	Mean	Std. Dev	Stat. Sig.
Recept Correct (%)	Phone	20	52.0	19.8	
	Computer	20	46.4	15.0	†
	DGS	20	65.5	20.4	†
Express Correct (%)	Phone	20	27.3	14.2	
	Computer	20	27.2	13.0	
	DGS	20	34.8	19.0	
Express Incorrect (%)	Phone	20	54.3	16.8	
	Computer	20	59.0	15.2	
	DGS	20	48.9	21.8	

DGS answering correctly on the receptive test 22.2% of the time more than in the Distributed conditions. While there was an 10.3% difference between the DGS and the Massed conditions, it was not statistically significant. There were no statistically significant differences between the Massed conditions and the DGS or between either the Distributed, Massed, or DGS conditions for the expressive correct and incorrect percentages. Data for these comparisons is listed in Table 17.

Table 17: DGS vs. Distributed and Massed Conditions– *Statistical significance shown in last column with corresponding symbols*

		N	Mean	Std. Dev	Stat. Sig.
Recept Correct (%)	Distributed	20	43.3	16.3	$\star$
	Massed	20	55.2	17.2	
	DGS	20	65.5	21.8	$\star$
Express Correct (%)	Distributed	20	23.8	13.4	
	Massed	20	30.7	12.9	
	DGS	20	34.8	19.0	
Express Incorrect (%)	Distributed	20	61.5	14.2	
	Massed	20	51.8	16.5	
	DGS	20	48.9	21.8	

5.2.4.3 *Summary of DGS vs. Final Study*

When comparing the DGS data to the final study data, the following results showed statistically significant ($p < .05$) differences:

- Comparison of DGS to all four conditions

  1. Participants in the DGS study answered a higher percentage of questions correctly on the receptive test than did participants in the Phone Distributed condition (65.5% vs. 41.0%).
  2. Participant in the DGS answered a higher percentage of questions correctly on the receptive test than did participants in the Computer Distributed condition (65.5% vs. 45.5%).

- Comparison of DGS to Phone and Computer conditions

  1. Participants in the DGS answered more questions correctly on the receptive test than did those in the Computer conditions (65.5% vs. 46.4%).

- Comparison of DGS to Massed and Distributed conditions

  1. Participants in the DGS answered more questions correctly on the receptive test than did those in the Distributed conditions (65.5% vs. 43.3%)

Using some of the metrics, participants using the naive scheduling algorithm in the DGS far outperformed those using the adaptive algorithm in this study. They scored higher on the receptive and expressive tests and took less time to learn, although not all of those metrics showed statistically significant differences. However, these differences also require tradeoffs, particularly in time and environment. The DGS required an hour of uninterrupted time while working on the lessons. The study detailed in this chapter was designed to take advantage of small moments throughout the day. Additionally, the DGS was performed in a laboratory setting with few or no distractions. The study outlined in this chapter took place in-the-wild while participants went about their daily lives. The following interview data helps explain why alternate forms of learning may still be advantageous, particularly for parents of deaf children who have limited time and must fit learning amongst other daily activities.

5.3 Interview Data

During the final testing session, before the participants completed their receptive and expressive tests, I conducted a short interview. These interviews focused primarily on several questions listed here:

1. Describe your days while you were getting alerts and doing lessons. How well did they fit into your day? What other things were you doing while you were doing the lessons?

2. Did you ever find the alerts annoying or intrusive? Did you feel like they came at a bad time or came too frequently?

3. Do you feel like you learned anything? How confident are you in your ASL signs and knowledge you acquired?

4. Did you feel like anything was missing from the lessons? If you were going to continue learning ASL, what would you like to learn next?

5. If you had the ability to request lessons at any time you wished, would you have used that functionality? Did you ever have some spare time when you wished to do lessons but couldn't because the system hadn't sent you an alert yet?

6. (After a description of the Massed and Distributed conditions) What do you think about the <condition you didn't have>? Do you think you would have liked that better or worse than the one you did have?

7. Can you think of anything else you would like us to know about your experience using the software and your time spent learning ASL?

After compiling the interviews, seven themes emerged:

- How to integrate the lessons and alerts into daily life.
- Enjoyment of the system and learning ASL.
- Scheduling issues when attempting to complete the lessons.
- Hypothetical reaction to other methods of scheduling.
- Assessment of learning and performance.
- Additions and changes to the system.
- Structure of content.

Some of these themes offered insights for future versions of learning software, while others helped explain some of the quantitative results presented previously.

5.3.1 Integration with Daily Life

Participants discussed a variety of ways they adapted their lives around the alerts and lessons. Several participants mentioned making an effort to keep pace with the lessons and complete them as soon as possible. Other participants spoke about conflicts between the alerts and daily life. Participant 29, in the Phone Massed condition said, "For the most part I would try to do them as soon as they came unless I was actually doing work or something else in which case I would let them pile up. One day, I didn't finish all the lessons... I started late and things went slow." Participant 4 in the Computer Massed condition expressed similar sentiments, "For most of them, I got them done in the morning, before I left for school. I think on two occasions I did them in the afternoon in between breaks in classes, but yeah, I pretty much did them all in one chunk."

Unsurprisingly, the participants in the Distributed conditions fell behind more often and were more aware of the fact they had not completed all the lessons. Participant 8 in the Phone Distributed condition highlighted this fact, "A lot of times, I got alerts when I wasn't in a position to do anything about it. Then, I wasn't sure how to tell if I was caught up or if that mattered." Participant 31, an athlete who was also in the Phone Distributed condition, gave a description of part of his days when the lessons did not fit well with his schedule:

> "Most of the time they fit in pretty good except for around practice time. Then I came out of practice, and I would have like four unread messages so I had to get them in real quick after that so I could make sure the next one would come in the right amount of time. The last day I felt like they didn't come in quite fast enough because I think I had maybe eight to go and I slept in until like 11 that day so I was expecting them to come in a bit quicker. I got to the next to the last one, and I had like 20 minutes left, and I didn't get a message for the last one."

The participants in the Computer Distributed condition highlighted the scheduling conflicts even more. Unlike the Phone participants who frequently had their iPhone with them, the Computer participants were often away from their computer. Thus, they highlighted how they would adjust their days to make sure the lessons were completed by the appropriate time. For example, Participant 2 singled out the weekends as particularly problematic, "My days vary with what I'm doing. If I was at the computer and working, I'd just go ahead and do it. But there are a few times where I was doing something else, and it was a weekend or I was out or something like that, and I wasn't near a computer and of course they would pile up a little bit." Several participants in the Computer conditions spoke about having to keep track of the alerts and develop a mental "alarm clock" in order to stay on schedule. Participant 6 (CD) demonstrated this:

> Probably have to divvy up the days into two categories. One being when I'm at the lab working in which case they fit in pretty well since I'm always having to check my email in case a vendor sends an email or my boss gives me something new to do. So I would just see the alert in my email and quickly do it. It would only take me about 45 seconds so that worked pretty well. The difficulty being on the weekends when I'm not paying attention to my email, so I'd get like 5 alerts if I haven't been keeping track. Or in the evenings when I'm doing something that I'm not near a computer, it might be problematic, but for the most part, especially during the weekdays, it fit in rather well.

Participant 7 who was in the Computer Massed condition, gave a particularly detailed account of how he got behind in his lessons due to a schedule which did not follow the normal, daily patterns:

> Well, the first several days it caught me during my normal workweek. Typically, I would either get into the office in the morning and maybe go through a couple of the sessions before I started my job or sometimes like in the morning, I'd be doing something I really didn't want to do like work, and I'd do the lessons instead. Others I did during lunch. But then towards the end of the week, my wife went on a girl-cation for a few days so I was Mr. Mommy. Apparently I got behind on the lessons more than I realized. I thought I had missed one day, but actually, apparently I had missed two days because I was changing lots of dirty diapers. I was very occupied with just trying to keep my household together so I would think, 'Oh, the study. Gotta do the study,' but I didn't check my email.

> Typically I live on email. So, I would know as soon as they were available. But in the case of the weekend, I just had to wait until my own mental alarm went off and said, 'Go and do the study.'

Several of the participants mentioned that they made an effort to focus on the lessons 100%. For example, Participant 5 described putting the iPhone down and signing along with every lesson and practicing regularly during the week of lessons. Participant 21 also described signing along with the lessons and also submitted sketches he made of the signs to help himself learn them (Figure 27). This dedication to the lessons was unusual and for many participants, the lessons faded into the background of daily life.

Participants in the Phone conditions highlighted the fact that they often completed the lessons while doing other things and multitasking. Participant 41 "completed each set daily and didn't let it stack up. I normally did it while I was walking from place to place because I usually have some long walks in my day. So I was usually able to finish a bunch during that time." While Participant 37 said she usually completed her lessons right away:

> "There were a couple times I did it while cooking dinner. Other times I was watching TV and would do it during commercial breaks. So I wasn't really multitasking while doing it but at the same time I wasn't 100% focused on sitting down and doing it, especially as the week went on."

Participant 35 who was in the Phone Distributed condition said:

> They [the alerts] came about every 45 minutes. Which was fine except sometimes when I was in the study room studying or watching television. For some reason my phone doesn't get reception there. I'd have to wait for a commercial if it was a really good show and go back to my room.

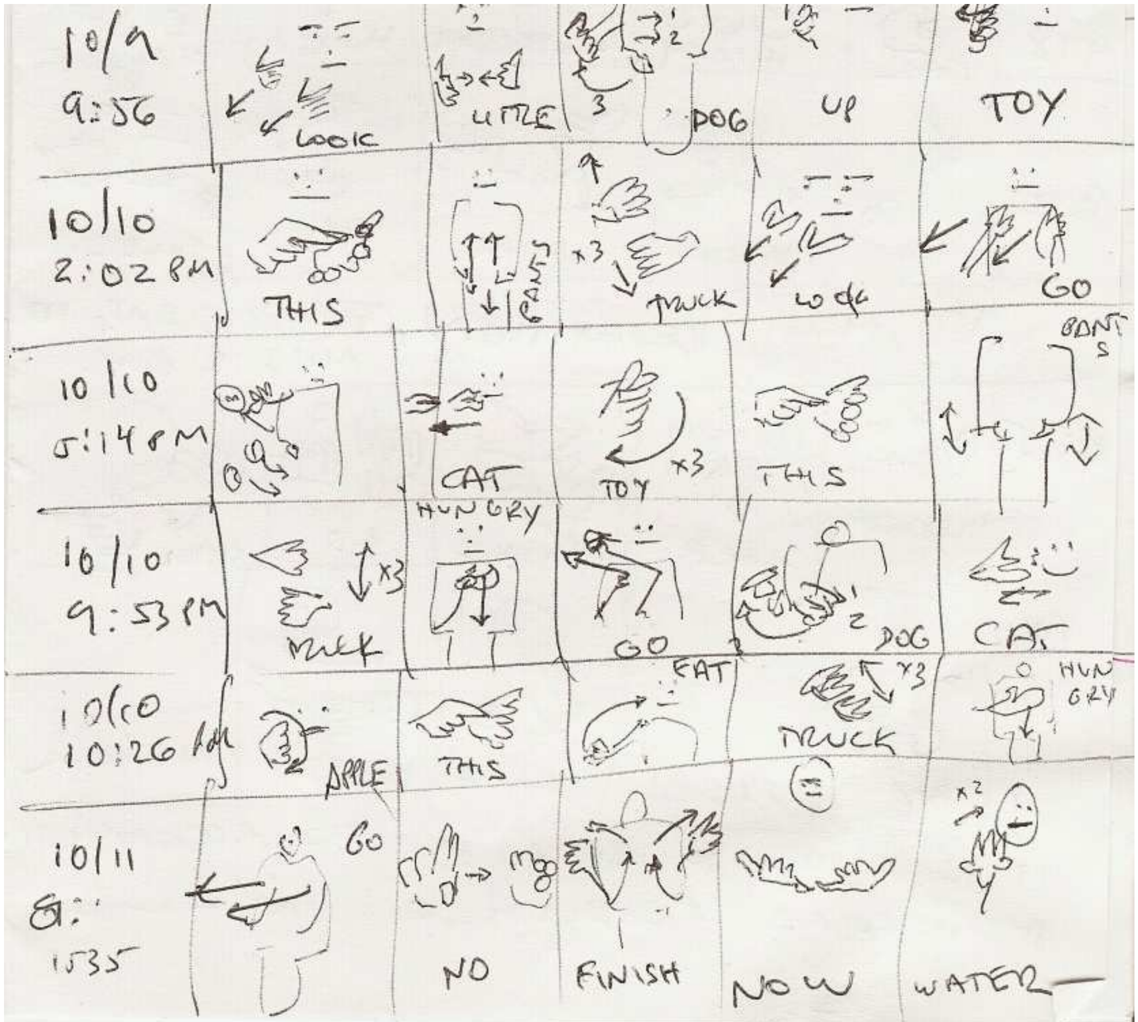


Figure 27: Participant 21's Sign Sketches

> That was no big deal. Most of them usually came when I was outside going to class."

The extra measures taken to ensure learning such as signing along with the videos or drawing signs would add time to the lessons. Likewise, being distracted or multitasking while doing the lessons would add time to the lessons as well. This helps explain why the times for all of the conditions were more than the DGS. In the DGS, the participants focussed only on the lessons and had no outside distractions. I believe that the four conditions described in this chapter are far more indicative of real world behavior than the earlier DGS.

5.3.2 Enjoyment

Many of the participants enjoyed the experience of using the software and learning ASL.

For example, Participant 11, who was in the Computer Distributed condition said,

> It became more like a little game where I was waiting for them. So it started getting fun. It was also a nice distraction to have from boring assignments because it only lasted, I think, about a minute for 5 questions. And I would be doing it in the middle of anything. The last few days I started looking forward to it, like, 'So when can I stop coding and go do a lesson.'

Like the quote from Participant 11 above, Participant 26 also viewed it as a distraction from work: "I was waiting tables one night so any time I had some downtime, I was messing with it. I had fun with it."

Participant 37 who was in the Phone Massed condition drew on the new knowledge she was acquiring and how it affected her enjoyment: "I actually really enjoyed it and was doing little signs when people would say things I knew."

5.3.3 Scheduling Issues

Participants in the Phone Massed condition stated that the alerts were intrusive or annoying more often than those in the other conditions.

For example, in the Phone Massed condition, some of the reactions when asked if the alerts were annoying, intrusive, or distracting were:

- "Yes, in the sense that I would get a text and I would start and as I'm finishing one lesson I would get another text in the middle of the lesson. That happened at least once or twice. Or sometimes I would stop doing them and say I'll do them in an hour because I can't do it now and I would get a text five minutes later. If I get a text once every hour or every two hours that would be a little easier because it's not overwhelming." – *Participant 36*

- "A little bit. Too many of them. There were a couple times when I was out and there weren't going to be any good times to do them. But there was no way to tell the system that. That's the only time I thought it was intrusive." – *Participant 29*

- "Yeah. The spacing of the alerts helped. But I wanted more time between them. If I received two messages per day that would be enough. But I was getting more. It was too many." – *Participant 27*

However, participants in the other conditions rarely shared these sentiments. Often, they commented on the number of alerts but did not find them annoying because they were expected. Some example comments were:

- "Seems like the alerts came about every hour. That was a bit much. But not really intrusive other than having maybe like five emails because I hadn't checked my email in awhile." – *Participant 4, Computer Massed*

- "No. They came right during the hours we had discussed. I picked the times, so it was fine."– *Participant 22, Computer Distributed*

- "Not intrusive. I kinda knew what was coming. So I was kind of prepared for it and knew when it was coming. It's only 30 seconds." – *Participant 35, Phone Distributed*

Additionally, in the Computer conditions where participants received emails rather than text messages, several of the participants used Gmail addresses. The use of Gmail may have affected the perceptions of the intrusiveness of the system. Gmail "threads" email conversations and notifies the receiver of duplicate messages by putting a small number next to the duplicate email indicating how many times it has been received. So if a participant using Gmail missed five alerts, they would only see one message in their inbox with a small number "5" next to it. Additionally, Gmail allows all duplicate messages to be deleted with a single click. For example, when asked about the repetitive alerts, Participant 18 in the Computer Massed condition said, "Not really, because in Google's mail, they all just get put into one thread. So it's not that bad." Likewise, Participant 20 said, "No, it seemed like the reminders came pretty quickly, but I have Gmail, so it stacked them," and Participant 23 said:

> "No, I got it all through Gmail so it stacks the emails from the same sender. If I was using a different client where it didn't stack those emails, then I probably would have had just a ton of the same one. But since I'm using Gmail it just all sits in one spot and once you get started on it, as long as you go through them all, you don't get any new emails.

In addition to the use of different email clients, I speculate that some of this annoyance is due to the fact that individuals frequently had their iPhones with them. It was relatively easy for people in the Computer conditions to ignore the messages. Although they might experience a brief spurt of annoyance when logging into email and finding multiple messages, the emails would not interrupt their day frequently in the manner that text messages would. For this reason, it is important to give the users of mobile devices control over the alerts. I discuss this implication more in Section 5.4.1.2.

5.3.4 Consideration of Other Method of Content Scheduling

When asked about the other method of content scheduling, participants often gave very nuanced and considered answers. Some participants preferred the condition they were in to the hypothetical other condition. For the participants who received Massed lessons, responses generally centered around the ability to do all the scheduled lessons at once and be done for the day. For example, Participant 15 who had the Massed method of content scheduling said, “I think that it [distributed delivery] would have been more intrusive. Because I just sat down and did it once and that was it for the day. I didn’t have to worry about it again.” Similarly, Participant 30 noted, “Because my schedule changes so much day to day that [distributed scheduling] would have been more intrusive to me. I have time right after I get off work to sit down and do them all right then. And I would know I was done so I didn’t have to worry about it for the rest of the day.”

Participants who had the Distributed method of content delivery centered their preference of the Distributed condition around the fact that each lesson did not consume a large amount of time and required only a few minutes to complete. Participant 3 who had the Distributed condition said:

> I think I prefer them spaced out. Because then, you know, it’s a short amount of learning. You can squeeze it in. You know, if I’ve got five minutes between meetings, I can do one right now. Whereas if I had to do like 15 minutes worth at once, I’d have to actually take some downtime, and that would be hard for me and for a lot of people, I think.

Participant 17 expressed the same sentiment, “I think I prefer the method I used. Because that way it didn’t require as much planning to do it. I just did it as it came and it didn’t take too long.”

In the Distributed conditions, many participants expressed a desire to be able to selectively complete more lessons at once time. Six out of the ten participants in the Phone Distributed condition and three of the ten participants in the Computer Distributed condition thought they would prefer being able to complete more lessons in a single sitting. Participant 34 (PD) said, “That would have been nicer. Then I would have been able to get all of them done.”

However, only two participants in the Massed conditions expressed a desire for distributed lessons. Participant 20 (CD) thought that distributed lessons would be OK, but only on a mobile device. "I think if I was doing it with like an iPhone app or something like that, something that was more portable, getting it throughout the day would be, for me, good I think. I was very saturated, sitting down and doing it all at once." The only other participant in the Massed conditions to think that distributed lessons might be better was Participant 40 who said, "I think I might of liked that little bit better just because if I have a lesson I would do it. With the way I had it, it [the massed lessons] was such a commitment." Other than Participants 20 and 40, all the participants in the Massed conditions preferred having massed lessons.

5.3.5 Learning and Performance

Participants uniformly felt that they learned some vocabulary during their lessons. Since they were interviewed before they were given either the expressive or receptive post-tests, their actual performance did not affect their perceptions of their performance. However, almost all participants correctly realized that generating the signs was harder than recognizing them, and most were hesitant about their ability to generate the signs.

Representative responses from all four conditions took the form of:

- "That's a good question. Yeah, it was interesting because I've never really had any exposure to ASL other than like maybe doing the alphabet in like 4th grade. So, yeah, I definitely picked up a little bit. I definitely feel I can recognize signs. Don't know how well I'll do at actually generating signs. That was one thing that I realized. I know that there's some subtlety to some of the signs, and I don't feel like I really captured the subtleties, but I feel like I can discriminate between the signs that were presented." – *Participant 7, Computer Massed*

- "I think I know a good dozen to two dozen signs right now. Or rather, I could recognize them easier than I could produce them at this point." *Participant 2,* *Computer Distributed*

- "I think I learned more than a few. Towards the end I could just look at a sign and before they're even finished, I knew what it was. I could instantly recall some of them." – *Participant 35, Phone Distributed*

- "Yeah I definitely do. I would try to practice them as I was looking at them to try to teach myself to remember them. I probably feel like I forgot a lot of them now though." – *Participant 40, Phone Massed*

Participants also had a range of opinions about whether the Distributed or Massed condition might help them learn more. Participant 14 (CD) felt that the distributed method might be conducive to better learning:

> I think this one [distributed condition] would be more effective because you go back and review it again and again. So rather than learning something for 20 minutes straight and then forgetting about it, going back and reviewing a little bit each time would probably be more effective.

Likewise Participant 2 (CD) said, "For timing purposes, yes I would have preferred [massed delivery]. From a learning standpoint, I imagine that spacing them out more might be better because you get a constant exposure to the lessons throughout the day so it's constant reinforcement."

However, participants in the Massed condition championed the repetition of seeing signs repeated frequently in a small window of time. For example, Participant 19 (CM) said, "The way I did it, with the lessons over and over again... I felt like there was more fluid repetition so it probably stuck in my head better." Similarly, Participant 33 (PM) noted:

> I think having them close together was very helpful because you could remember. Instead of having one and then having another an hour later,

it would be like, 'I know that sign! It was in the last one but I don't remember what it was.' It was really helpful to realize I saw it two lessons ago, and I remember that that's grandfather or whatever.

5.3.6 Additions and Changes to System

The system described in this chapter was a "push" system. That is, the system sent notifications to the participants when lessons were available according to the scheduling algorithm. Participants did not have the capability to "pull" lessons or to request that lessons be sent to them at specific times or intervals. During the post-lessson session, participants were asked if they would have used such a feature.

Regardless of the experimental condition they were in, participants wished for more control and thought they would have used such a feature to fine-tune the delivery of the lessons. Some example responses were:

- "Yeah, there would be sometimes when I couldn't schedule in, when I had a block of free time and I'd be like, 'I wish I could a couple more right now and just get them out of the way' to, in a sense, make up for the other times when I wouldn't be at a computer." – *Participant 6, Computer Distributed*

- "I would've really liked that. I wanted to be able, if I'm in class early and I have ten minutes, I could do a few lessons." – *Participant 35, Phone Distributed*

- "Yes. At the end of the day, that was when I wanted to do the lessons. And I wanted to do a few more but I couldn't. So either a link or go to a web site or something so I could do a couple more. " – *Participant 27, Phone Massed*

- "Not necessarily during the course of the study, but just thinking about how I use technology, that's definitely something I would appreciate. Like, I'm a big fan, I just love podcasts. So, I'm always finding time, like sitting on the MARTA bus or you know, in a waiting room or just killing time when someone's late for a meeting or whatever. I like having something at hand that I can just take these few extra minutes and do something useful or stimulating.” – *Participant 7, Computer Massed*

In spite of wanting the ability to request lessons, several participants were careful to note that they did not want a system which only supported the “pull” mechanism. For example:

- “I like both. I like the reminders as well.” – *Participant 4, Computer Massed*
- “Maybe. The current system worked pretty well for me. Sometimes I would forget that I had one [alert] earlier and then it would send me another to remind me. I’m not sure if I could remember to go into lessons every day.” – *Participant 38, Phone Distributed*
- “I think that would be really good, especially for someone who needed to know it [ASL]. I don’t think I would have remembered to request lessons every day, but then, I really didn’t need to know it. So I can’t really say, but I definitely feel like the reminders help.” – *Participant 39, Phone Distributed*

Another feature that several participants mentioned was a “pause button” which would suspend reminders for a few hours. Participant 40, who was in the Phone Massed condition, best summarized this idea:

> . . . But then when I was doing something else they would come about every hour or so. It was a little bit annoying because I was like, ‘I know, I know. I’ll get to them in a little while. I’ll do it later.’ At the same time if I wasn’t doing anything and thought I would just do it in a while and put it off then it would pop up and remind me. So that was good. So if there was a way to set a reminder for three hours later like a pause button that would be good.”

Given the feedback from the participants, a hybrid system that combines alerts or reminders with the ability to request lessons should be considered for future versions.

5.3.7 Content Structure

Asking the participants about what they wanted to learn next if they were to continue their lessons proved to be a good way to understand how the lessons could be improved in the future. In their work on mathematics education, Jungbauer, Baggett and Ehrefeucht called for identification of "cohesive" elements which are common features of a group of words or concepts which help learners distinguish meaningful differences in material [61, 36]. This idea of cohesion can be seen in participants' requests and comments in this study.

For example, several participants wanted lessons centered around different themes. Participant 13 (CD) noticed that the signs for family members such as 'father,' 'mother,' 'grandmother,' and 'grandfather' were similar and said, "I would introduce some basics from some categories, so I thought it would have been nice to learn like, all family members at the same time. Like when we learned a couple family members they were sort of similar."

The similarity of signs was often a point of confusion and several participants stated they would like clarification or explanation of the differences in similar signs. Participant 7 (CM) said, "Maybe some emphasis on some of the subtle aspects of the signs would have been nice. Also, the relationships between signs. I noticed there were several signs that were clearly related, like father and grandfather for example. An understanding of maybe the thought process behind... if there's something more generalized that would apply to other signs that could be captured in some way."

Other participants wanted examples of phrases or conversations. "Pretty much like seeing how the conversation would go because I have seen signed conversations before and it would be going so fast I couldn't tell how the conversation would work"

Participant 31, PD. Participant 37 (PM) wanted, "Probably some basic sentences. A lot of the basic communication that the guidebook has when you're traveling. Those sentences and questions there to help you. I feel like that would be the next step."

In the next section, I discuss some implications from the above work and synthesize design recommendations. In particular, I separate my recommendations into two key considerations:

1. Design of mobile learning systems (5.4.1)

2. Design of content for mobile learning systems (5.4.2)

5.4 Design Recommendations

As Sharples *et al.* point out, "The use of (mobile) technology is not the target but rather a means to enable activities that were otherwise not possible, or to increase the benefits for the learners" [99]. The goal of the study outlined in Chapter 5 was to evaluate how mobile technologies could be deployed to better support parents of deaf children in their goals of learning ASL. However, the findings of this study lead to some implications for the design of mobile learning systems and the structure of content for mobile learning systems.

5.4.1 Design of Mobile Learning System

In the following section, I distill the findings presented in this chapter to several key components which make for a more effective and user-friendly system. These are:

1. Goals

2. Control

3. Assessment and Feedback

5.4.1.1 Goals

In this software system, having a goal proved motivating to participants. Participants were told they would be trying to complete 80 lessons in seven days, but were not told a specific number of lessons needed to be completed each day. The system was designed to try to schedule 12 lessons per day for the first six days and eight lessons the last day. In reality, participants often fell behind and completed fewer lessons than the 80 for which they were scheduled. Participants in the Massed conditions consistently mentioned how they liked the fact they were "done for the day" once they had reached their goal and completed a day's lessons (see Sections 5.3.4). Several participants also mentioned that they liked "status markers" and used the money total reported at the end of each lesson (see Figures 20 and 25) as a way to gauge how many lessons had been done. Participants knew that if they completed the lesson as they arrived, they would stay on track to complete all lessons in the study (and receive full payment).

Depending on the content to be learned, goals could be implemented in different ways for different systems. If the system described in this chapter were carried forward, goals could be established based on many different metrics. For example, instead of completing 12 lessons (60 signs) each day, participants could be given a total time metric to strive for or a number of correct signs/answers. Regardless of the metric, the pedagogical goal gave learners in this study a sense of accomplishment and allowed them to complete the lessons over the course of a week.

5.4.1.2 Control

The application gave the participants very little control over when and how the alerts and lessons arrived. Participants were allowed to set the starting and ending times for the alerts but were not afforded any other control. Participants could exert control by ignoring alerts, but the periodic reminder alerts rendered this a somewhat useless tactic as text messages and emails were still sent. A per-day granularity of scheduling would be a useful feature to implement in the future, as participants distinguished between weekdays and weekends as well as days when they had different schedules.

In the follow-up interviews, participant comments centered around two features which would have enhanced user control: a "pause" button and a mechanism to request lessons in addition to the ones that were sent. Users requested a pause button to stop alerts from arriving for a certain amount of time. Participants described using this feature for a short period of time (one to three hours) when they were involved with other, temporary activities. Additionally, participants were very enthusiastic about the possibility of being able to request lessons. They envisioned using this feature to give them better control over the scheduling features. Participants described having extra time and wanting to get ahead on lessons by doing them during "down moments." This feature would give participants who are very motivated opportunities to study extra content. Combined with a well defined goal as described above, this feature would give participants a mechanism to be in control and limit the number of alerts they receive.

5.4.1.3 Assessment and Feedback

This study was designed to provide participants immediate feedback after each vocabulary word. The quick feedback cycle provided a way for participants to immediately learn from their mistakes while the sign was still at the forefront of their memory. Although this study used multiple-choice questions, in the future I would implement the input channel to incorporate text input (i.e., typing the English word) as well as more complex multiple choice. Allowing text input would remove the mental prompts that occur when a participant sees the four possible correct answers. An open-ended text response would force the participant to recall the meaning of the word with no help.

Participant 29 suggested a more complex feedback mechanism:

> "It would be cool in the future if there was a way for it to record my gestures. And then it's not just about me watching it. I could flip it around and use the camera or have some other attachment that let me record it. Even if it was just a comparison showing 'here's what it was supposed to be' and 'here's what you did' side by side it would be an interesting thing to look at."

While a video feedback option mechanism is becoming more feasible, particularly with smartphones, participants might consider this a cumbersome and time consuming method of obtaining feedback.

Another possible feedback mechanism would be automated sign language recognition of recorded video. Currently, smartphones do not have the processing capability to support sign language recognition. However, there have been studies which have shown the feasibility of using compressed video (suitable for data networks) for a two-way video conversation in ASL [23, 24], and video transmission of sign language to a system which has sufficient processing power may be feasible.

5.4.2 Design of Content for Learning Systems

Another important consideration when designing a computer aided learning system is how the content will be presented to the learners. During the study outline in this chapter, I identified the following 4 areas as being particularly important:

1. Details
2. Themes or Mnemonics
3. Review
4. Schedules

5.4.2.1 Details and Nuances

Several participants in the Phone conditions commented on the feeling that they were missing details and nuances of the signs. I believe that the lack of details were particularly apparent in the Phone conditions because of the small screen, even though the video and the video platform were high quality. In future iterations of this system, I would ensure that the participants have more details. This detail could take the form of two videos shown sequentially, one with the complete sign as in the current system, followed by a closeup of the hand(s) for the appropriate handshape. In sign language, handshape is the most important detail which might not be seen from a full video. However, depending on the content, "detail" might be spelling, pronunciation, context (such as using a word in a longer expression or sentence), or other details specific to the content being learned.

5.4.2.2 Themes or Mnemonics

During the post-lesson interviews, several participants mentioned their desire for the content to be structured into units or themes. A common example used was a series of signs for family members such as 'mother,' 'father,' 'grandmother,' 'grandfather,' etc. Participants wanted similar signs grouped together to help them recognize the subtle differences between the signs. During the expressive testing, it was common for participants to think out loud and remember signs by remembering how they were different from a reference point. For example, Participant 29 said, "Mother was here [gesturing to the chin area], so grandmother is out from that [moving the hand away from the chin in the correct sign]." The grouping of common material into thematic units would allow learners to make more connections between related material and aide them in achieving their goals.

Similarly, providing mnemonics or explanations of the content would also have helped the learners in this study. Many of the participants in the study recognized that there were similarities in signs but were unable to articulate the precise differences. Again, using the example of family signs, several participants recognized that the signs for 'father' and 'grandfather' started from the forehead while the signs for 'mother' and 'grandmother' started from the chin. However, there is a more generalizable rule in ASL that male signs are signed around the upper part of the head (e.g., forehead, crown of the head, and temple) while female signs are signed around the lower part of the head (e.g., chin or jaw). Like using themes, providing short lessons or explanations of some of these universal rules would also allow the learners to reinforce their own conclusions or categorize the signs for future learning.

5.4.2.3 Review

Given the Atkinson scheduling algorithm used in this study, it became apparent that participants needed ways to review words that had already been "learned." In this study, the only way for a participant to review signs was to learn all words in a given group (see Section 5.1.1.3) at which point the signs would be presented randomly. Participants commented on this lack of review and said they sometimes felt that they learned a word at the beginning of the study but never saw it again. In the DGS, participants had a built-in review as they saw each word during every trial, regardless of whether or not they had learned it.

In the future, I would build in either random review of words already learned, or build in review units as described above. Based on a participant's performance on review words, words could be transitioned back into regular rotation if the participant had forgotten them. These review mechanisms would ensure that words were not forgotten as more words were added to the learner's schedule.

5.4.2.4 Schedules

This study used the Atkinson algorithm as derived in Chapter 4 to schedule when words would be presented to participants. The original experiments using this algorithm showed increased performance over several other methods of presentation. However, the process of deriving the per-word coefficients needed to use this algorithm is extremely laborious. Moreover, the DGS described in Chapter 4 would need to be re-run any time one needed to add more words to the system. I did not see a large enough difference in performance from the naive scheduling algorithm used in the DGS to justify the time necessary to derive and program the Atkinson algorithm.

Going forward, I would replace the Atkinson algorithm with a simpler, easier to implement algorithm. One possible avenue would be to consult with a linguist and use a simple rank ordering scheme with easy words receiving a label of '1', medium difficulty words receiving a label of '2,' and difficult words receiving a label of '3'. The system could then adjust the number of presentations of a word based on its difficulty rating.

Another possible algorithm would be to "retire" words as participants answered them correctly. For example, if a participant answered a word might be placed in a rotation where it was presented once every ten words until a participant correctly identified it several times in a row. It would then be transitioned to being shown once every 20 words until the participants correctly identified it several times in a row. This approach would allow the frequency of a word to be "stepped down" while still ensuring the participant reviewed words that he or she had previously learned. This algorithm could be altered to accommodate different metrics of correctness as needed.

There are many different algorithms which could be implemented in a learning system to structure the content to keep the user engaged and learning. However, the results of this experiment do not support the time and effort needed to calculate and use the Atkinson algorithm.

5.5 Summary

In this chapter, I have presented a study designed to measure the differences in delivery platforms (desktop computers vs. mobile phones) and content scheduling (Massed practice vs. Distributed practice).

The goal of this study was to evaluate two hypotheses:

1. That participants learning ASL utilizing mobile phone platforms as a content delivery mechanism would demonstrate better receptive and generative language abilities than participants who learn using a traditional desktop-based platform.

2. That participants learning ASL using the spacing effect, which presents the material on a distributed schedule instead of a massed schedule, would demonstrate better receptive and generative language abilities.

The results presented in this chapter somewhat supported the hypothesis that individuals could increase their mastery of a second language via mobile devices, as the participants in the Phone Massed condition achieved results similar to those in the DGS on post-test assessment. However, the data collected did not support the hypothesis that participants would learn more in the Distributed conditions. In fact, the participants in the Distributed conditions completed statistically significantly fewer lessons and answered fewer questions correctly on the receptive tests. Moreover, participants in the DGS from Chapter 4 performed better than did participants in any of the four conditions: Phone Distributed, Phone Massed, Computer Distributed, or Computer Massed.

Based on the quantitative results and the interview data presented in this chapter, recommendations for other mobile learning systems can be made. In the next chapter I discuss these recommendations along with proposed future work for evaluating them and incorporating them into an effective mobile learning system.

CHAPTER 6

FUTURE WORK

In this chapter, I build on the results presented in Chapter 5 by presenting limitations of this thesis work and recommendations for future work based off the design recommendations presented in Section 5.4.1.

6.1 Thesis Validity

In the next two sections, I discuss various threats to the internal and external validity of this thesis. While each study has slightly different threats, it is useful to discuss the thesis as a whole.

6.1.1 Internal Validity

The internal validity of an experiment refers to how conclusively we can conclude that changes in dependent variables were brought about through manipulation of independent variables. In their seminal work on experimental methods, Campbell and Stanley [22] identified eight factors which can jeopardize an experiment's internal validity. In the following section, I will discuss each of these eight factors in light of the three studies contained in this dissertation. I will also identify the biggest threats to the overall internal validity of this thesis.

- **History:** Outside factors may influence participants' responses between repeated measures trials or between measurements. The Deaf teenagers study in Chapter 3 had no repeated measures. Being a largely qualitative study, measurements were not effected. However, the two later studies both contained opportunities for history to become a factor. In both studies, events could intervene between the learning phase of the study and the measurement session with interviews and post-tests. However, all participants in all conditions had the same amount of time between the learning phase and the post-test phases of the studies. While historical artifacts may have intervened for some subjects, it is unlikely that they all had intervening issues between the two sessions of the studies.

- **Maturation:** Internal validity can be affected by natural (rather than experimenter imposed) restrictions. For example, participants may get tired or bored during the course of a study, and this ennui may measurably affect their performance. In the study of Deaf teenagers, described in Chapter 3, the teenagers completed the study during regularly scheduled class times. As they were given regular instruction and feedback via an experimenter who was present at all times, there was probably little effect from maturation. Likewise, the Data Gathering Study was completed under laboratory conditions. However, participants certainly could have become tired or bored, as Session 1 of the study lasted over an hour. Participants were required to take regular breaks to combat this effect. The study in this dissertation most affected by maturation was the final learning study. Participants were required to complete lessons over seven consecutive days. This time period afforded a significant amount of self-directed control to the participants. They were not monitored in person by an experimenter, nor were they prompted to complete the task. Moreover, several participants commented on boredom toward the end of the week of lessons, or after many lessons completed in quick succession. Worries about maturational effects were the primary reason that participants in the final study were incentivized by the lesson. A monetary incentive gave participants reason to complete the lessons, but it did not necessarily affect the studiousness or care the participants took while completing the lessons. However, the results of the study are still informative, even if we accept the fact that the participants' learning was somewhat affected by maturation issues. The participants in all four conditions experienced maturation over the same period (one week of lessons) and are thus still comparable. Thus, while there is a threat to internal validity through maturation, it is not large enough to justify skepticism of the results.

- **Testing:** This threat relates to the consequences of pre-testing subjects to ensure that the subject has no prior knowledge of the subject before the experiment. Pre-testing can contaminate the participants' knowledge and lead to a loss of internal validity. In the studies described in Chapters 4 and 5, the subjects were not pre-tested on their ASL knowledge. However, participants were required to have no knowledge of ASL beyond basic fingerspelling.

- **Instrumentation:** Internal validity can be affected by changing methods of measurement or human observers during the time of the experiment. In the study of Deaf teenagers, detailed in Chapter 3, all interviews were conducted with the same interviewer. However, a different ASL interpreter was used for the interview sessions for the two classes described in Section 3.1.2.3. In the Data Gathering Study (Chapter 4) and the learning study (Chapter 5), the methods of measurement (in this case, the software used in the experiments) was not changed during the course of the study. Moreover, in the learning study, the same individual conducted all exit interviews. Thus, instrumentation is not a large threat to the internal validity of this dissertation.

- **Statistical Regression to the Mean:** This threat concerns subjects which were recruited based on particularly high (or low) scores on a selection metric. Re-testing the participants will almost always lead to a score that is less extreme (i.e., closer to the mean). In the studies described in this dissertation, the participants were not recruited based on previous performance, so this is not a threat to the internal validity of the work in this dissertation.

- **Selection:** This threat concerns differences amongst the experimental conditions in a study. In all studies in this dissertation, participants in all conditions were recruited from the same initial pool. For example, in the Deaf teenagers study (Chapter 3), all participants were recruited from the same class at the Atlanta Area School for the Deaf. In the DGS and ASL learning studies, participants were recruited via email from the same subject pools. They were assigned semi-randomly to conditions, but all drew from the same population base. Therefore, there is a low threat to internal validity based on selection bias.

- **Experimental Mortality:** Internal validity can be affected when participants drop out of a study. In particular, if participants remove themselves from the experiment in one experimental condition more than the others, results can be significantly skewed. In the study described in Chapter 5, only one participant withdrew from the study. No participants withdrew from the Data Gathering Study. Therefore, there is little threat to this thesis's internal validity through experimental mortality.

- **Selection Interactions:** This threat to internal validity occurs when one or more of the above factors combine to bias the results of a study. No such interactions have been identified for any of the studies in this dissertation.

6.1.2 External Validity

External validity is the extent to which a study's findings can be generalized. Two main threats to external validity will be examined below. Overall, threats to external validity are much more prevalent in this thesis than threats to internal validity.

- **Population Validity:** This threat to external validity occurs when an interaction between the the population used in the experiment and the independent variable can occur. In the DGS and the mobile learning study (Chapters 4 and 5), the intervention is eventually intended for parents of deaf children. However, the initial testing and evaluation was conducted using mostly students. This discrepancy presents several issues including technological familiarity and non-scheduled time activities. Parents of deaf children are much less likely to have access to high technology or the financial resources to acquire technology such as smart phones and the associated high-cost data plans. Additionally, the individuals who participated in these studies were extremely technologically savvy and adapted quickly to new and unfamiliar software. This same familiarity cannot be assumed with the targeted user demographic. Students and people associated with academic institutions tend to have very flexible schedules compared with those employed in blue-collar or white-collar jobs. This difference could lead to different findings, particularly in the Distributed and Massed conditions. Parents may not have time to devote significant blocks of time to lessons. Similarly, if they are away from their desks for long periods of time, the software may have to be adapted to accommodate this fact.

- **Ecological validity:** Ecological validity concerns the set of environmental conditions under which an experiment occurred and comprises many different factors [16]. For example, the "experimenter effect" may be of concern if only one researcher elicits a particular response, and it cannot be duplicated by other researchers. In this dissertation, the largest concerns of ecological validity are the novelty effect and treatment diffusion. The novelty effect occurs if a participant responds (or does not respond) to the novelty or newness of an intervention. The mobile learning system evaluated in the learning study in Chapter 5 is sufficiently novel that parents of deaf children may react to it very differently than the tech-savvy population with which it was originally tested. Also of concern are the manners in which participants in the massed conditions occasionally adopted features of the distributed conditions. Massed condition participants could have paused in the midst of their lessons, thus acting effectively as distributed participants. However, post-hoc analysis of the spacing between lessons did not support this hypothesis.

Many of the recommendations above should be incorporated into future systems, but with careful study to measure their effectiveness. In the next section, I describe the next steps to continue progress on a mobile learning system for parents of deaf children.

6.2 Future Work

There are many ways this work could be extended in the future, largely based on the analysis presented in Sections 5.4.1 and 5.4.2.

A first step would be to enhance the content of the system to include a more complete and nuanced understanding of ASL. At a minimum, in the future, the system should include ASL tips, grammar basics, and example sentences or phrases to assist the user in learning the nuances of ASL as described in Section 5.4.2.1. Additionally, the content should be structured to include conceptual units to assist in learning as described in Section 5.4.2.2.

The m-learning system should be enhanced for more control by the user, including a pause button and a hybrid system including both pushing and pulling of content as described in Section 5.4.1.2.

After these system and content changes are made, there are several studies which would give a clear picture of the efficacy of a m-learning system. Even though the Atkinson algorithm of scheduling did not produce statistically significant differences compared to the naive scheduling algorithm used in the Data Gathering Study (DGS), it does not mean that the approach is without merit. I believe an adaptive presentation algorithm could enhance content delivery, but it must be less time and labor intensive than the Atkinson algorithm's implementation. A study to evaluate a simple presentation algorithm as described in Section 5.4.2.4 would help clarify the necessity and benefit of adaptive presentation while taking into account the inherent difficulty of the content to be learned.

Both the DGS data obtained in Chapter 4 and the results obtained from the study in Chapter 5 serve as important benchmarks in evaluating future changes and improvements to a mobile ASL learning system for parents of deaf children. While the Phone Massed paradigm of instruction returned the best results, a paradigm similar to the Phone Distributed one should not be ruled out, particularly if it suits the parents' needs or schedules the best. A good starting point may be something similar to the Phone Massed condition, but allow the parents to "fine tune" it with a shorter session length or through use of other control features as described in Section 5.4.1.2.

CHAPTER 7

CONCLUSION

The original thesis of this dissertation was

I hypothesize that participants who learn American Sign Language vocabulary:

1. *utilizing mobile phone platforms as a content delivery mechanism will demonstrate better receptive and generative language abilities than participants who learn using a traditional desktop-based platform as measured by post-intervention tests.*

2. *using the spacing effect, which presents the material on a distributed schedule instead of a massed schedule, will demonstrate better receptive and generative language abilities as measured by post-intervention tests.*

The work in this dissertation has supported the hypothesis that we could utilize mobile phones to support individuals in learning American Sign Language. Participants in all conditions detailed in Chapters 4 and 5 learned some ASL vocabulary words and were able to both express and recognize words. However, the algorithm used did not enhance learning for the participants when compared to a naive scheduling algorithm.

In this work, I have made the following contributions:

1. Research into the role that mobile communication technologies currently play in the social lives of Deaf teenagers and with whom they communicate (Chapter 3).

2. An investigation of algorithms for ASL vocabulary learning (Chapter 4).

3. Comparison of mobile devices and traditional desktop/laptop computers as platforms for ASL vocabulary learning (Chapter 5).

4. Assessment of different scheduling methods for ASL vocabulary learning (Chapter 5)

In Chapter 3, I discussed a study of Deaf teenagers at the Atlanta Area School for the Deaf and their use and preferences of mobile communication technologies. The findings of this study led to the work discussed in Chapters 4 and 5. This study showed that communication, either remote or face-to-face, is not an issue within the Deaf community. When co-located, Deaf individuals sign to one another; when remotely located, they have a common set of tools, such as instant messaging on a computer or Sidekick, which work quite well for their communication needs. These findings are supported by Hogg, Lomicky, and Weiner's work on computer-mediated communication in the Gallaudet community [57]. They also found the most popular device was a Sidekick, although they note that as more phones become available with video capabilities, the Sidekick may be supplanted.

Communication problems arise when Deaf and hearing individuals need to communicate with one another. The findings of my study indicate that remote communication is not necessarily the dominant issue. As discussed earlier in Section 3.2.3, while not at school, the teenagers communicate primarily with hearing friends and family. As shown in Figure 5, most members of their families with whom they communicate with regularly are hearing. Additionally, most teens reported using ASL the majority of time with hearing family in face-to-face communication (Figure 6). Figure 9 summarizes the categories and methods of communication utilized by the teenagers for both Deaf and hearing communication partners.

Given the difficulties faced when communicating with hearing individuals, there appear to be two ways to address the communication divide. One way is to focus on helping Deaf individuals use written English through augmented communication devices, phrase books, automated translators, etc. Another way is to focus on assisting hearing people in learning ASL. The research presented in this document has focused on the latter strategy. This method has the added benefit of being appropriate for any age. While teaching English to deaf individuals is difficult until they reach school age and often requires native language (i.e., ASL) fluency, teaching ASL to hearing people can be undertaken at any age. The benefits of teaching children to communicate early using BabySign have been touted in popular culture for children who are not deaf (e.g., [18, 103, 91]). As mentioned in Chapter 2, fluency in a first language impacts fluency in other languages, and it is critically important for a child's development to acquire language as early as possible. Given this background, I decided to concentrate on applications for teaching parents or other hearing adults survival level ASL vocabulary.

After this decision, the question became how best to deliver the material. Since many parents of deaf children are busy with other children, full time jobs, etc., I felt that it was important to maximize instruction time and spend the time available to the parents on an efficient method of instruction. This need for efficiency led to the use of the Atkinson algorithm of second language learning described in Section 2.3. This model predicts a word's transition in memory from an unknown or forgotten state to short term memory to long term memory based on the intrinsic difficulty of a word and the learner's response when presented with a word to be learned.

The study presented in Chapter 5 used the measures of difficulty obtained in the Data Gathering Study for the same 80 ASL vocabulary words. However, this study was designed to investigate the spacing of lessons either throughout the day or massed together at one time and the delivery mechanism of a mobile platform (such as a smartphone) or a more traditional desktop or laptop computer. All content was prioritized based on the probability of a word transitioning from its current state (unknown/forgotten or short term memory) into long term memory where it was considered to be permanently learned. This study found that the Massed method of content delivery on the iPhone platform was better than a Distributed method of content delivery on an iPhone platform or either method of content delivery using a traditional desktop or laptop platform. Additionally, participants provided many suggestions to improve future systems and make them more usable.

Chapter 6 presented some final recommendations for the design of future mobile learning systems as well as how to proceed in the design of a specific system for parents of deaf children who wish to learn ASL.

APPENDIX A

SUPPLEMENTARY DATA TABLES

Table 18: Vocabulary Words Used in Study

Word	Exp % Correct	Rec. % Correct	T_1-T_5 % Correct	Learnability
bad	0.27	0.00	0.66	0.31
good	0.00	0.25	0.72	0.32
not	0.00	0.30	0.67	0.32
that	0.11	0.18	0.69	0.32
tomorrow	0.00	0.38	0.71	0.36
sister	0.00	0.38	0.74	0.37
bathroom	0.15	0.43	0.59	0.39
now	0.09	0.33	0.78	0.40
what	0.00	0.44	0.76	0.40
home	0.08	0.43	0.70	0.40
happy	0.09	0.38	0.76	0.40
sweet	0.15	0.29	0.80	0.41
brother	0.00	0.50	0.77	0.42
stop	0.00	0.43	0.86	0.43
where	0.33	0.27	0.71	0.43
please	0.25	0.31	0.81	0.45
help	0.10	0.50	0.77	0.45
careful	0.38	0.17	0.83	0.45
hot	0.23	0.43	0.74	0.46
school	0.11	0.64	0.66	0.46
wait	0.13	0.58	0.77	0.49
who	0.00	0.75	0.73	0.49
go	0.13	0.58	0.83	0.51
grandmother	0.62	0.14	0.79	0.51
sleep	0.17	0.50	0.89	0.51
your	0.11	0.70	0.75	0.52
no	0.11	0.64	0.82	0.52
toy	0.11	0.64	0.85	0.53
yes	0.25	0.63	0.75	0.54
truck	0.00	0.87	0.78	0.54
my	0.5	0.25	0.92	0.55
yesterday	0.56	0.36	0.81	0.57
thank you	0.40	0.40	0.93	0.57

(Continued on next page)

(Table 18 continued)

Word	Exp % Correct	Rec. % Correct	T_1-T_5 % Correct	Learnability
finish	0.08	0.86	0.81	0.58
dog	0.18	0.78	0.79	0.58
mom	0.33	0.64	0.79	0.58
hello	0.21	0.67	0.88	0.58
hungry	0.31	0.63	0.84	0.58
more	0.56	0.45	0.77	0.59
water	0.25	0.75	0.81	0.60
milk	0.45	0.56	0.80	0.60
person	0.25	0.75	0.82	0.60
grandfather	0.29	0.83	0.75	0.62
medicine	0.33	0.73	0.83	0.62
apple	0.25	0.88	0.78	0.63
sick	0.00	1.00	0.91	0.63
out	0.40	0.60	0.93	0.64
off	0.22	0.82	0.91	0.64
shoes	0.78	0.36	0.81	0.64
cat	0.36	0.80	0.81	0.65
there	0.58	0.50	0.90	0.65
hurry	0.29	0.80	0.91	0.66
love	0.29	0.85	0.93	0.68
this	0.60	0.60	0.88	0.69
want	0.33	0.82	0.97	0.70
tired	0.25	0.92	0.96	0.70
thirsty	0.50	0.75	0.91	0.71
look	0.22	1.00	0.95	0.72
I	0.40	0.80	0.98	0.72
car	0.67	0.77	0.75	0.72
juice	0.27	1.00	0.92	0.72
on	0.25	1.00	0.95	0.73
jacket	0.50	0.79	0.94	0.73
cold	0.62	0.71	0.93	0.75
dad	0.45	1.00	0.84	0.76
soap	0.64	0.78	0.92	0.77
bedroom	0.43	1.00	0.93	0.78
eat	0.58	0.86	0.97	0.80
big	0.88	0.67	0.87	0.80
food	0.89	0.64	0.92	0.81
in	0.67	0.91	0.89	0.81
up	0.90	0.70	0.90	0.83
pants	0.71	1.00	0.91	0.87
banana	0.71	1.00	0.98	0.89

(Continued on next page)

(Table 18 continued)

Word	Exp % Correct	Rec. % Correct	T_1-T_5 % Correct	Learnability
little	0.89	0.91	0.92	0.90
drink	0.75	1.00	0.98	0.90
down	0.89	0.91	0.94	0.90
you	0.82	1.00	0.95	0.91
book	1.00	0.89	0.96	0.94
baby	0.89	1.00	0.97	0.94

Table 19: Parameters and Chi-squared Values for Words

Word	chi-square	a	c	f
apple	22.3039	0.4568	0.4567	0.9999
baby	1.0123	0.4541	0.9999	0.2535
bad	14.3369	0.2958	0.2956	0.9999
banana	0.3645	0.8952	0.8951	0.9999
bathroom	26.6566	0.2666	0.2662	0.9999
bedroom	6.0879	0.6288	0.6286	0.9999
big	7.9874	0.3755	0.555	0.7269
book	1.9835	0.8642	0.8571	0.0681
brother	16.597	0.4157	0.4156	0.9999
careful	16.6818	0.1015	0.4752	0.283
car	19.1468	0.4218	0.4217	0.9999
cat	25.1833	0.5157	0.5157	0.9999
cold	8.778	0.3087	0.9999	0.6659
dad	20.879	0.4481	0.4481	0.9999
dog	19.7602	0.4298	0.4295	0.9999
down	8.8079	0.4362	0.9999	0.2202
drink	0.3645	0.8952	0.8951	0.9999
eat	0.8669	0.847	0.8469	0.9999
finish	6.588	0.3899	0.4330	0.9210
food	2.2145	0.3956	0.632	0.3495
good	11.607	0.3347	0.3345	0.9999
go	9.8141	0.4716	0.4715	0.9999
grandfather	21.007	0.379	0.3789	0.9999
grandmother	16.4904	0.3987	0.3977	0.9999
happy	29.5389	0.4200	0.4200	0.9999
hello	7.7791	0.2598	0.5166	0.3530
help	19.0139	0.3991	0.3992	0.9999
home	19.5019	0.3919	0.3919	0.9999
hot	16.0044	0.3644	0.3644	0.9999
hungry	18.4799	0.4747	0.4747	0.9999
hurry	1.7075	0.6844	0.6844	0.9999
in	12.2426	0.9248	0.3074	0.9999
i	0.0863	0.9464	0.9463	0.9999
jacket	8.2953	0.1565	0.6773	0.2468
juice	8.1421	0.2218	0.9999	0.4877
little	5.1567	0.6646	0.6645	0.9999
look	2.6992	0.7611	0.7611	0.9999
love	1.1156	0.7083	0.7083	0.9999
medicine	8.3631	0.4407	0.4408	0.9999
milk	11.9057	0.4209	0.4205	0.9999
mom	19.0406	0.3624	0.3624	0.9999

(Continued on next page)

(Table 19 continued)

Word	chi-square	a	c	f
more	20.7516	0.4039	0.4039	0.9999
my	3.9364	0.7948	0.7001	0.1752
no	13.9415	0.413	0.4127	0.9999
not	20.1072	0.2396	0.2498	0.9669
now	11.4346	0.4255	0.4255	0.9999
off	8.7135	0.3359	0.6510	0.5786
on	3.3804	0.2526	0.7172	0.1650
out	4.1987	0.6837	0.6836	0.9999
pants	8.1770	0.9999	0.3239	0.0001
person	25.1424	0.3092	0.4526	0.7763
please	11.2188	0.3458	0.4204	0.7798
school	32.6986	0.3534	0.3533	0.9999
shoes	24.4934	0.4529	0.4530	0.9999
sick	3.5767	0.2919	0.5487	0.0914
sister	27.363	0.3875	0.3874	0.9999
sleep	4.9527	0.4572	0.6659	0.2368
soap	2.4544	0.7155	0.7154	0.9999
stop	5.8658	0.1594	0.5741	0.5063
sweet	10.5995	0.3958	0.3956	0.9999
thank you	3.0184	0.0245	0.6242	0.0644
that	20.274	0.2794	0.2792	0.9999
there	6.2714	0.5234	0.5297	0.0815
thirsty	8.6684	0.8947	0.6273	0.1596
this	9.5842	0.6012	0.6011	0.9999
tired	1.6311	0.8021	0.8021	0.9999
tomorrow	21.6315	0.323	0.3229	0.9999
toy	5.5483	0.5084	0.5083	0.9999
truck	21.2465	0.495	0.4951	0.9999
up	4.5171	0.6379	0.6379	0.9999
wait	18.3497	0.4162	0.4160	0.9999
want	2.0047	0.1063	0.8615	0.0690
water	14.4451	0.5004	0.5003	0.9999
what	19.7526	0.3440	0.3439	0.9999
where	24.7338	0.3652	0.3652	0.9999
who	21.5589	0.3925	0.3924	0.9999
yes	16.5989	0.3678	0.3676	0.9999
yesterday	17.4856	0.4521	0.4521	0.9999
you	7.2723	0.6201	0.7808	0.1840
your	22.5643	0.3831	0.3829	0.9999

Table 20: U/F to LTM Transition Probability

Word	U/F to LTM transition probability
i	89.56
drink	80.13
banana	80.13
book	74.07
eat	71.73
tired	64.34
look	57.93
thirsty	56.12
my	55.64
soap	51.19
love	50.17
you	48.42
hurry	46.84
out	46.74
baby	45.41
little	44.16
down	43.62
up	40.69
bedroom	39.53
this	36.14
pants	32.39
cold	30.87
sleep	30.44
in	28.43
there	27.72
cat	26.59
toy	25.84
water	25.04
food	25.00
truck	24.51
hungry	22.53
go	22.24
juice	22.18
off	21.87
apple	20.86
big	20.84
shoes	20.52
yesterday	20.44
dad	20.08
medicine	19.43
dog	18.46

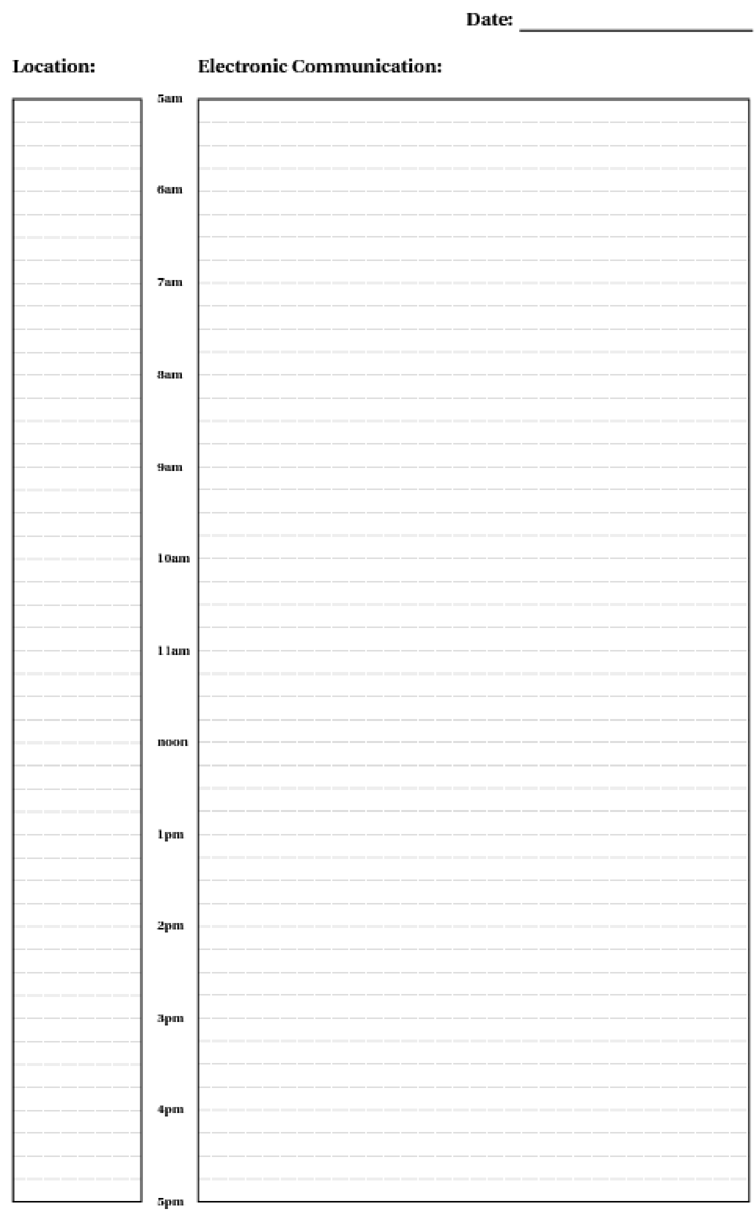
(Continued on next page)

(Table 20 continued)

Word	U/F to LTM transition probability
on	18.12
now	18.11
car	17.79
milk	17.70
happy	17.64
wait	17.31
brother	17.28
no	17.04
finish	16.88
more	16.31
sick	16.02
help	15.93
grandmother	15.86
sweet	15.66
who	15.40
home	15.36
sister	15.01
your	14.67
please	14.54
grandfather	14.36
person	13.99
yes	13.52
hello	13.42
where	13.34
hot	13.28
mom	13.13
school	12.49
what	11.83
good	11.20
jacket	10.60
tomorrow	10.43
want	9.16
stop	9.15
bad	8.74
that	7.80
bathroom	7.10
not	5.99
careful	4.82
thank you	1.53

APPENDIX B

SUPPLEMENTARY STUDY DEVICES


Date: ____________________

Location:

Electronic Communication:

5am
6am
7am
8am
9am
10am
11am
noon
1pm
2pm
3pm
4pm
5pm

Figure 28: Communication Logging Aid (front)

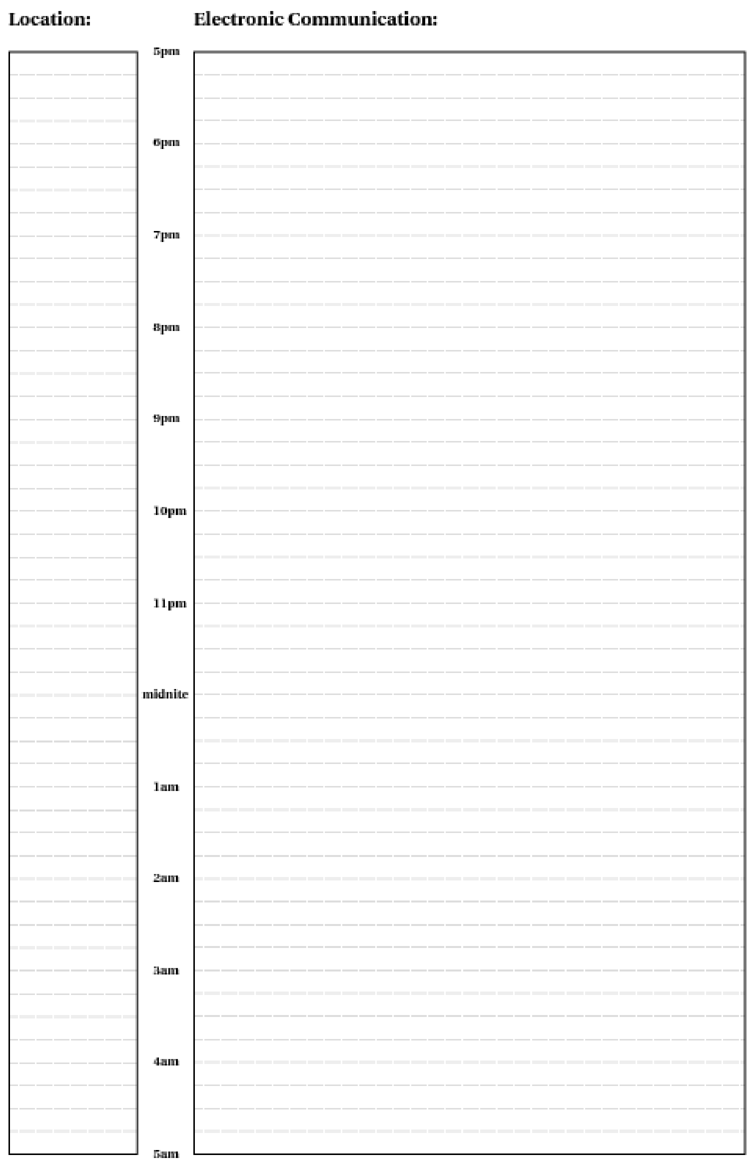


Figure 29: Communication Logging Aid (back)

REFERENCES

[1] AdMob, "Admob mobile metrics report." `http://metrics.admob.com/wp-content/uploads/2010/01/AdMob-Mobile-Metrics-Dec-09.pdf`, December 2009. Last retreived Feb 20, 2010.

[2] Anderson, J., "Multiple-choice," *Medical Teacher*, vol. 26, no. 2, pp. 110–113, 2004.

[3] Atkinson, R. C., "Optimizing the learning of a second language vocabulary," *Journal of Experimental Psychology*, vol. 96, pp. 124–129, 1972.

[4] Atkinson, R. and Crothers, E., "A comparison of paired associate learning models having different acquisition and retention axioms," *Journal of Mathematical Psychology*, vol. 1, pp. 285–315, 1964.

[5] Baddeley, A. D., *Essentials of Human Memory*. Psychology Press Ltd, 1999.

[6] Bahrick, H. P., "Semantic memory content in permastore: Fifty years of memory for Spanish learned in school," *Journal of Experimental Psychology: General*, vol. 113, pp. 1–29, 1984.

[7] Bailes, C. N., Erting, C. J., Erting, L. C., and Thumann-Prezioso, C., "Language and literacy acquisition through parental mediation in American Sign Language," *Sign Language Studies*, vol. 9, pp. 417–456, Summer 2009.

[8] Bakken, F., *The Inside Text: Social, Cultural and Design Perspectives on SMS*, ch. SMS Use Among Deaf Teens and Young Adults in Norway, pp. 161–174. Kluwer Academic Publishers, 2004.

[9] Barkhuus, L. and Vallgårda, A., "Saying it all in 160 characters: Four classes of SMS conversations," Tech. Rep. TR-45, IT University of Copenhagen, Cophenhagen, Denmark, April 2004.

[10] Baron, N. S., "Instant messaging and the future of language," *Communications of the ACM*, vol. 48, pp. 29–31, July 2005.

[11] Baron, N. S., Squires, L., Tench, S., and Thompson, M., "Tethered or mobile? use of away messages in instant messaging by American college students," in *Front Stage – Back Stage: Mobile Communication and the Renegotiation of the Social Sphere*, (Grimstead, Norway), Springer–Verlag, June 2003.

[12] Baynton, D. C., *Forbidden Signs: American Culture and the Campaign Against Sign Language*. University of Chicago Press, 1996.

[13] Bellugi, U. and Fischer, S., "A comparison of sign language and spoken language," *Cognition*, vol. 1, pp. 173–198, 1972.

[14] Blinkoff, R. and Barranca, M., "Mobile lifestyles," Receiver Magazine 6, Vodafone Group, 2002.

[15] Bowe, F., "Deaf and hard of hearing Americans' instant messaging and email use: A national survey," *American Annals of the Deaf*, vol. 147, pp. 6–10, October 2002.

[16] Bracht, G. and Glass, G., "The external validity of experiments," *American Education Research Journal*, vol. 5, pp. 437–474, 1968.

[17] Brashear, H., Henderson, V., Park, K.-H., Hamilton, H., Lee, S., and Starner, T., "American Sign Language recognition in game development for deaf children," in *Assets '06: Proceedings of the 8th international ACM SIGACCESS conference on Computers and Accessibility*, pp. 79–86, 2006.

[18] Brody, J. E., "From birth, engage your child with talk," *New York Times*, p. D7, September 29, 2009.

[19] Bruce, D. and Bahrick, H., "Perceptions of past research," *American Psychologist*, vol. 47, pp. 319–328, 1992.

[20] Bryant, D. P. and Bryant, B., *Assistive Technology for People with Disabilities*. Allyn and Bacon, 2003.

[21] Caccamise, F. and Newell, W., *An overview of the sign communication proficiency interview (SCPI): History, development, methodology, and use.* Rochester, NY: Rochester Institute of Technology, National Technical Institute for the Deaf, 1997.

[22] Campbell, D. and Stanley, J., *Experimental and Quasi-Experimental Designs for Research*. Boston: Houghton-Mifflin, 1963.

[23] Cherniavsky, N., Cavender, A., Riskin, E., and Ladner, R., "Variable frame rate for low power mobile sign language communication," in *Proceedings of ASSETS 2007: The Ninth International ACM SIGACCESS Conference on Computers and Accessibility*, pp. 163–170, 2007.

[24] Chon, J., Cherniavsky, N., Riskin, E., and Ladner, R., "Enabling access through real-time sign language communication over cell phones," in *43rd Annual Asilomar Conference on Signals, Systems, and Computers*, 2009.

[25] Clark, H. and Brennan, S., *Perspectives in Socially Shared Cognition*, ch. Grounding in Communication, pp. 127–149. American Psychological Association, 1991.

[26] Cohen, P. N. and Petrescu-Prahova, M., "Gendered living arrangements among children with disabilities," *Journal of Marriage and Family*, vol. 68, pp. 630–638, August 2006.

[27] Commission on Education of the Deaf, "Toward equality: Education of the deaf." Report to Congress and the President, February 1988.

[28] Costabile, M. F., De Angeli, A., Lanzilotti, R., Ardito, C., Buono, P., and Pederson, T., "Explore! possibilities and challenges of mobile learning," in *CHI '08: Proceeding of the twenty-sixth annual SIGCHI conference on Human factors in computing systems*, pp. 145–154, 2008.

[29] Crain, S. and Lillo-Martin, D., *An Introduction to Linguistic Theory and Language Acquisition*. Blackwell Publishing, 1988.

[30] Crystal, D., *Language and the Internet*. Cambridge University Press, 2001.

[31] Dempster, F. and Farris, R., "The spacing effect: Research and practice," *Journal of Research and Development in Education*, vol. 23, no. 2, pp. 97–101, 1990.

[32] Ebbinghaus, H., *Memory: A Contribution to Experimental Psychology*. New York: Teachers College, Columbia University, 1913. Originally published 1885.

[33] Ellis, K. and Blashki, K., "Children, Australian Sign Language and the web: The possibilities," in *Proceedings of AUSWEB04*, 2004.

[34] Ellis, K. and Kathy, "The digital playground: Kindergarten children learning sign language through multimedia," *AACE Journal*, vol. 15, no. 3, pp. 225–253, 2007.

[35] Emmorey, K., *Language, Cognition, and the Brain: Insights from Sign Language Research*. Lawrence Erlbaum Associates, Inc., 2002.

[36] Espinosa, R. and Baggett, P., "A design for accessing multimedia information using cohesion," *COMPUTER AND SYSTEMS SCIENCES*, vol. 129, p. 275, 1994.

[37] Fenson, L., Marchman, V. A., Thal, D. J., Dale, P. S., Reznick, J. S., and Bates, E., *MacArthur-Bates Communicative Development Inventories*. Brookes Publishing, 2004.

[38] Fischer, S. D., Delhorne, L. A., and Reed, C. M., "Effects of rate of presentation on the reception of American Sign Language," *Journal of Speech, Hearing, and Language Research*, vol. 42, no. 3, pp. 568–582, 1999.

[39] Freeman, B., Dieterich, C. A., and Rack, C., "The struggle for language: Perspective and practices of urban parents with children who are deaf and or hard of hearing," *American Annals of the Deaf*, vol. 147, pp. 37–44, December 2002.

[40] Friess, S., "Silence=deaf," *POZ Magazine*, vol. April, pp. 60–63, 1998.

[41] Frishberg, N., "Home sign," in *Gallaudet encyclopedia of deaf people and deafness* (Cleve, J. V., ed.), vol. 3, pp. 128–131, McGraw Hill, 1987.

[42] Gallaudet Research Institute, "Regions regional and national summary report of data from the 1999-2000 annual survey of deaf and hard of hearing children and youth," tech. rep., Gallaudet University, Washington, D. C., January 2001.

[43] Goldin-Meadow, S., *The Resilience of Language: What Gesture Creation in Deaf Children can Tell Us about How All Children Learn Language*. Psychology Press, 2003.

[44] Google, "Youtube apis and tools." `http://code.google.com/apis/youtube/overview.html`. Last retrieved January 14, 2010.

[45] Google, "Youtube: Broadcast yourself." `http://www.youtube.com`. Last retreived January 14, 2010.

[46] Grinter, R. and Eldridge, M., "y do tngrs luv 2 txt msg?," in *Proceedings of the Seventh European Conference on Computer-Supported Cooperative Work*, pp. 219–238, 2001.

[47] Grinter, R. and Eldridge, M., "Wan2tlk?: everyday text messaging," in *Proceedings of the conference on Human factors in computing systems*, pp. 441–448, ACM Press, April 2003.

[48] Grinter, R. E. and Palen, L., "Instant messaging and teen life," in *Proceedings of CSCW '02*, (New Orleans, LA), pp. 21–30, November 2002.

[49] Groot, P., "Computer assisted second language vocabulary acquisition," *Language Learning & Technology*, vol. 4, pp. 60–81, May 2000.

[50] Hamilton, H. and Holzman, T., "Linguistic encoding in short–term memory as a function of stimulus type," *Memory and Cognition*, vol. 17, no. 5, pp. 541–550, 1989.

[51] Hamilton, H. and Lillo-Martin, D., "Imitative production of verbs of movement and location: A comparative study," *Sign Language Studies*, vol. 50, pp. 29–57, 1986.

[52] Harkins, J. and Bakke, M., *Oxford Handbook of Deaf Studies, Language, and Education*, ch. Technologies for Communication: Status and Trends, pp. 407–419. Oxford University Press, 2003.

[53] Hart, S. G. and Staveland, L. E., "Development of NASA–TLX (task load index): Results of empirical and theoretical research.," in *Human Mental Workload* (Hancock, P. A. and Meshkati, N., eds.), pp. 139–183, Elsevier Science Publishers, 1988.

[54] Henderson, V., Lee, S., Brashear, H., Hamilton, H., Starner, T., and Hamilton, S., "Development of an American Sign Language game for deaf children," in *Interaction Design and Children*, pp. 70–79, June 2005.

[55] Henderson-Summet, V., Grinter, R. E., Carroll, J., and Starner, T., "Electronic communication: Themes from a case study of the deaf community," in *Proceedings of INTERACT'07*, vol. 4662 of *LNCS*, pp. 347–360, Springer, 2007.

[56] Henderson-Summet, V., Weaver, K. A., Westeyn, T. L., and Starner, T., "American Sign Language vocabulary: computer aided instruction for non-signers," in *Assets '08: Proceedings of the 10th international ACM SIGACCESS Conference on Computers and Accessibility*, pp. 281–282, 2008.

[57] Hogg, N. M., Lomicky, C. S., and Weiner, S. F., "Computer-mediated communication and the Gallaudet University community: A preliminary report," *American Annals of the Deaf*, vol. 153, p. 96, Spring 2008.

[58] Holt, J., Traxler, C., and Allen, T., "Interpreting the scores: A user's guide to the 9th edition Stanford achievement test for educators of deaf and hard-of-hearing students," Tech. Rep. Gallaudet Research Institute Technical Report 97-1, Gallaudet University, Washington, DC, 1997.

[59] Hott, L. and Garey, D., "Through deaf eyes." PBS Documentary, 2007.

[60] Isaacs, E., Walendowski, A., Whittaker, S., Schiano, D. J., and Kamm, C., "The character, functions, and styles of instant messaging in the workplace," in *CSCW '02: Proceedings of the 2002 ACM conference on Computer supported cooperative work*, pp. 11–20, ACM Press, 2002.

[61] Jungbauer, J., Baggett, P., and Ehrenfeucht, A., "A browsing system based on multimedia cohesion," Tech. Rep. CU-CS-803-96, University of Colorado, 1996.

[62] Kam, M., Agarwal, A., Kumar, A., Lal, S., Mathur, A., Tewari, A., and Canny, J., "Designing e-learning games for rural children in India: a format for balancing learning with fun," in *DIS '08: Proceedings of the 7th ACM conference on Designing interactive systems*, pp. 58–67, 2008.

[63] Klima, E. and Bellugi, U., *The Signs of Language*. Cambridge, MA: Harvard University Press, 1979.

[64] Lang, H., *A Phone of Our Own: The Deaf Insurrection Against Ma Bell*. Gallaudet University Press, 2000.

[65] Lang, H., "Review of 'Cochlear Implants in Children: Ethics and Choices'," *Sign Language Studies*, vol. 3, no. 1, pp. 90–93, 2002.

[66] Laurent Clerc National Deaf Education Center, "Laurent Clerc: Apostle to the deaf people of the new world," in *http://clerccenter.gallaudet.edu/Clerc_Center/About_the_Clerc_Center/Laurent_Clerc_The_Man.html*, Gallaudet University, Retrieved Dec 31, 2009 2007.

[67] Lee, S., Henderson, V., Hamilton, H., Brashear, H., Starner, T., and Hamilton, S., "A gesture–based american sign language game for deaf children," in *CHI '05 Extended Abstracts*, pp. 1589–1592, ACM Press, April 2005.

[68] Lenhart, A., Madden, M., and Hitlin, P., "Teens and technology," tech. rep., Pew Internet and American Life Project, Washington, DC, July 2005.

[69] Lindquist, D., Denning, T., Kelly, M., Malani, R., Griswold, W. G., and Simon, B., "Exploring the potential of mobile phones for active learning in the classroom," in *SIGCSE '07: Proceedings of the 38th SIGCSE technical symposium on Computer science education*, pp. 384–388, 2007.

[70] Ling, R., "We release them little by little: Maturation and gender identity as seen in the use of mobile telephony," *Personal Ubiquitous Comput.*, vol. 5, no. 2, pp. 123–136, 2001.

[71] Liu, Y., Hu, F., and Li, H., "Understanding learners' perspectives on m-learning: results from a survey," in *EATIS '09: Proceedings of the 2009 Euro American Conference on Telematics and Information Systems*, pp. 1–3, 2009.

[72] Lofland, J. and Lofland, L. H., *Analyzing Social Settings: A Guide to Qualitative Observation and Analysis*. Wadsworth Publishing, 1995.

[73] MacDougall, J., "Legal education and information issues: A national needs assessment," tech. rep., Justice Canada, Public Legal Education and Information Program, 2000.

[74] Marchman, V. and Bates, E., "Continuity in lexical and morphological development: A test of the critical mass hypothesis.," *Journal of Child Language*, vol. 21, pp. 339–366, 1994.

[75] Mayberry, R., "The critical period for language acquisition and the deaf child's language comprehension: A psycholinguistic approach," *Bulletin d'Audiophonologie: Annales Scientifiques de L'Universite de Franche-Comte*, vol. 15, pp. 349–358, 1998.

[76] Meadow-Orlans, K., Mertens, D., and Sass-Lehrer, M., *Parents and Their Deaf Children.* Washington, DC: Gallaudet University Press, 2003.

[77] Miller, M., "Sign iconicity and receptive vocabulary testing," *American Annals of the Deaf*, vol. 152, pp. 441–449, Winter 2008.

[78] Mitchell, R. E. and Karchmer, M. A., "Chasing the mythical ten percent: Parental hearing status of deaf and hard of hearing students in the United States," *Sign Language Studies*, vol. 4, pp. 138–163, Winter 2004.

[79] Nardi, B., Whittaker, S., and Bradner, E., "Interaction and outeraction: Instant messaging in action," in *Proceedings of CSCW'00*, (Philadelphia, PA), pp. 79–88, 2000.

[80] National Institute on Deafness and Other Communication Disorders, "Cochlear implants." `http://www.nidcd.nih.gov/health/hearing/coch.asp`. Last retrieved January 1, 2010.

[81] NetMarketShare, "Mobile browsing explodes in december." `http://www.netmarketshare.com/operating-system-market-share.aspx?qprid=61&sample=13`. Last retreived February 20, 2010.

[82] Newell, W. and Caccamise, F., *Skills important for effective sign language communication and sign communication proficiency interview (SCPI) rating levels.* Rochester, NY: Rochester Institute of Technology, National Technical Institute for the Deaf, 1997.

[83] Newport, E., "Maturational constraints on language learning," *Cognitive Science*, vol. 14, no. 1, pp. 11–28, 1990.

[84] Northern, J. and Downs, M., *Hearing in Children.* Lippincott Williams and Wilkins, 2002.

[85] Padden, C. and Humphries, T., *Deaf in America: Voices from a Culture.* Harvard University Press, 1988.

[86] Palen, L., Salzman, M., and Youngs, E., "Going wireless: behavior & practice of new mobile phone users," in *Proceedings of the 2000 ACM conference on Computer supported cooperative work*, pp. 201–210, ACM Press, 2000.

[87] Papousek, H. and Papousek, M., *Handbook of Infant Development*, ch. Intuitive Parenting: A Didactic Counterpart to the Infant's Integrative Competence, pp. 669–720. Wiley, 1987.

[88] Paredes, R., Ogata, H., Saito, N., Yin, C., Yno, Y., Oishi, Y., and Ueda, T., "Loch: supporting informal language learning outside the classroom with handhelds," in *IEEE International Workshop on Wireless and Mobile Technologies in Education, 2005. WMTE 2005.*, Nov. 2005.

[89] Parsons, D., Ryu, H., and Cranshaw, M., "A study of design requirements for mobile learning environments," in *Proceedings of Sixth Intl. Conf. on Advanced Learning Technologies (ICALT '06)*, pp. 96–100, IEEE Computer Society, 2006.

[90] Patel, S., Kientz, J., Hayes, G., Bhat, S., and Abowd, G., "Farther than you may think: An empirical investigation of the proximity of users to their mobile phones," in *Proceedings of UbiComp '06*, vol. LNCS 4206, pp. 123–140, Springer Verlag, 2006.

[91] Paul, P., "Hands across America: Is baby sign language an essential or a rip off?." babble, April 7, 2008.

[92] Power, D., Power, M. R., and Rehling, B., "German deaf people using text communication: Short message service, TTY, relay services, fax, and email," *American Annals of the Deaf*, vol. 252, pp. 291–301, Summer 2007.

[93] Power, M. R. and Power, D., "Everyone here speaks TXT: Deaf people using SMS in Australia and the rest of the world," *Journal of Deaf Studies and Deaf Education*, vol. 9, no. 3, pp. 333–343, 2004.

[94] RNID, "Opportunity blocked: The employment experiences of deaf and hard of hearing people," tech. rep., Royal National Institute for the Deaf, 2006.

[95] Sánchez, J. and Salinas, A., "Science problem solving learning through mobile gaming," in *MindTrek '08: Proceedings of the 12th international conference on Entertainment and media in the ubiquitous era*, pp. 49–53, 2008.

[96] Schiano, D. J., Chen, C. P., Isaacs, E., Ginsberg, J., Gretarsdottir, U., and Huddleston, M., "Teen use of messaging media," in *CHI '02: CHI '02 extended abstracts on Human factors in computing systems*, pp. 594–595, ACM Press, 2002.

[97] Schmaling, C. and Monaghan, L., eds., *HIV/AIDS and Deaf Communities.* Forest Books, 2006.

[98] Sebrechts, M. M., "From testing to training: Evaluating automated diagnosis in statistics and algebra," in *Lecture Notes in Computer Science: Intelligent Tutoring Systems*, vol. 608/1992, pp. 559–566, 2006.

[99] Sharples, M., Arnedillo-Sanchez, I., Milrad, M., and Vavoula, G., "Mobile learning: Small devices, big issues," in *Technology-Enhanced Learning*, Springer Netherlands, 2009.

[100] Sharples, M., Corlett, D., and Westmancott, O., "The design and implementation of a mobile learning resource," *Personal Ubiquitous Computing*, vol. 6, no. 3, pp. 220–234, 2002.

[101] Singleton, J. and Newport, E., "When learners surpass their models: The acquisition of American Sign Language from inconsistent input," *Cognitive Psychology*, vol. 49, pp. 370–407, 2004.

[102] Smith, H., Rodgers, Y., and Brady, M., “Managing one’s social network: Does age make a difference?,” in *Proceedings of INTERACT ’03*, pp. 551–558, 2003.

[103] Solomon, W., “Baby sign language a growing movement.” azcentral.com, April 25, 2006.

[104] Spencer, P. and Lederberg, A., *Communication and Language: Discoveries from Atypical Development*, ch. Different Modes, Different Models: Communication and Language of Young Deaf Children and their Mothers, pp. 203–230. Harvard University Press, 1997.

[105] Stokoe, W., “Sign language structure: An outline of the visual communication system of the American deaf,” *Studies in linguistics: Occasional papers*, vol. 8, pp. 3–79, 1960.

[106] Strong, M. and Prinz, P., “A study of the relationship between American Sign Language and English literacy,” *Journal of Deaf Studies and Deaf Education*, vol. 2, no. 1, pp. 37–46, 1997.

[107] Sutel, S., “Technology spawns new devices to help deaf and hearing to communicate.” Associated Press, http://www.cbc.ca/cp/health/071106/x110619A.html, November 6 2007.

[108] Valli, C. and Lucas, C., *Linguistics of American Sign Language*. Gallaudet University Press, 1992.

[109] Wagreich, B., “Electronic mail for the hearing impaired and its potential for other disabilities,” *IEEE Transactions on Communications*, vol. COM–30, no. 1, pp. 58–65, 1982.

[110] Wells, G., *Language Development in Preschool Years*. Harvard University Press, 1985.

[111] Wireless Intelligence, “Subscriber statistics end Q1 2007.” `http://www.gsmworld.com/news/statistics/pdf/gsma_stats_q1_07.pdf`, April 2007.

[112] Yule, G., *The Study of Language*. Cambridge University Press, 2nd ed., 1996.