# THE STRUCTURE OF SIGN LANGUAGE LEXICONS:

# INVENTORY AND DISTRIBUTION

# OF HANDSHAPE AND LOCATION

Lorna Rozelle

A dissertation submitted in partial fulfillment of the

requirements for the degree of

Doctor of Philosophy

University of Washington

2003

Program Authorized to Offer Degree: Department of Linguistics

University of Washington

Graduate School

This is to certify that I have examined this copy of a doctoral dissertation by

Lorna Rozelle

and have found that it is complete and satisfactory in all respects,
and that any and all revisions required by the final
examining committee have been made.

Chair of Supervisory Committee:

______________________________
Sharon Hargus

Reading Committee:

______________________________
David Corina

______________________________
Sharon Hargus

______________________________
Ellen Kaisse

______________________________
Richard Wright

Date: ______________________

University of Washington

Abstract

THE STRUCTURE OF SIGN LANGUAGE LEXICONS:
INVENTORY AND DISTRIBUTION
OF HANDSHAPE AND LOCATION

Lorna Rozelle

Chairperson of the Supervisory Committee:

Professor Sharon Hargus
Department of Linguistics

This dissertation is a quantitative cross-linguistic study of phonological properties of the lexicons of four unrelated and geographically diverse sign languages: American Sign Language, Korean Sign Language, New Zealand Sign Language and Finnish Sign Language. The ultimate goal of this research project is to describe the structure of the inventories and lexicons of naturally occurring sign languages; the more immediate goal of this dissertation is to describe systemic properties and patterns in the lexicons of these four languages. Although much is known about the cross-linguistic properties of inventories and lexicons of spoken languages, similar characterizations of sign language inventory structures have yet to be proposed. Sign language inventory and lexicon structure is investigated with respect to the phonological constituents of *handshape*, the configuration of the hand and fingers, and *location*, the place on the body or in space where the sign is articulated. *Duets*, that is, pairs of handshape and location that occur simultaneously within a sign, are another phonological resource investigated. The inventories of handshapes, locations and duets in these four languages are described. The distributions of these resources throughout the lexicon is determined and are found to be remarkably similar cross-linguistically. Finally, patterns of dependence between properties of signs are examined, in particular, correlations between handshape and location.

# Table of Contents

# List of Examples

## Acknowledgments

I wish to thank many people for their help in the preparation of this dissertation. Sharon Hargus provided perpetual and unstinting encouragement, open-mindedness, generous consultations, and intellectual honesty. Without her confidence and enthusiasm, this project would never have been begun, let alone completed. My committee members, Richard Wright, Ellen Kaisse and David Corina contributed many valuable and constructive comments.

A Fulbright Fellowship to Finland in 1992-93 gave me the opportunity to begin cross-linguistic study of sign languages. In Finland, many people helped me learn Suomalaisen Viittomakieli, including Leena Savolainen at the Kuurojen Liitto, Terhi Rissanen at Turun Kristillinen Opisto, Fred Karlsson at the University of Helsinki, and many others. Members of the Finnish Deaf community welcomed me, including Vesa Pikkuhookana, Helena Tanskanen, Jarmo Harju, and Soini Sunni. I also received financial assistance from a University of Washington Graduate School Arts and Humanities Dissertation Fellowship.

As for academic assistance, Haewon Cho translated the Korean Sign Language dictionary; Soohee Kim and Emily Curtis answered questions on spoken Korean, while Hansook Lee answered questions on Korean Sign Language. Onno Crasborn, Harry van der Hulst and Els van der Kooij, the creators of SignPhon, the database of Sign Language of the Netherlands, allowed me access to their work. Jacob Wegelin provided statistical consulting. Ian Maddieson helped me find information on segment frequency in spoken language. I had useful discussions with John Goldsmith, in which he pointed me toward information theory. Kevin Karplus taught me how to calculate dependence by means of mutual information and explained how to calculate its significance. Clifford Wong wrote the mutual information significance program given in Appendix 2 and the sequential broken stick program. Without the mutual information technique and its implementation I would not have been able to analyze dependence.

As for personal assistance, Laurel Balf-Johnson, Maria Galvão, Lynne Kershner and Kevin Rochlin, and Bob Southron took care of my loved ones and me. Maria Galvão, Susan McBurney and Alice Taff provided friendship as well as professional advice. Elmer and Carol Rozelle and Ruth and Walter Turner taught me to value education. And Clifford Wong, Eva Rozelle, and Rachel Rozelle let me put my dissertation first for a while, even though we know what is really important. So much support made it possible for me to complete this work.

Finally, thanks to the members of the Deaf communities in Seattle and Finland for teaching me about your beautiful languages.

## Dedication

To Cliff, Rachel and Eva

# Chapter 1: Introduction

This dissertation is a quantitative cross-linguistic study of phonological properties of the lexicons of four unrelated and geographically diverse sign languages: American Sign Language (ASL), Korean Sign Language (KSL), New Zealand Sign Language (NZSL) and Finnish Sign Language (SVK).[1] The ultimate goal of this research project is to describe the structure of the inventories and lexicons of naturally occurring sign languages; the more immediate goal of this dissertation is to describe systemic properties and patterns in the lexicons of these four languages. Although much is known about the cross-linguistic properties of inventories and lexicons of spoken languages, similar characterizations of sign language inventory structures have yet to be proposed. Sign language inventory and lexicon structure is investigated with respect to the phonological constituents of *handshape*, the configuration of the hand and fingers, and *location*, the place on the body or in space where the sign is articulated. *Duets*, that is, pairs of handshape and location that occur simultaneously within a sign, are another phonological resource investigated. The inventories of handshapes, locations and duets in these four languages are described. The distributions of these resources throughout the lexicon is determined and are found to be remarkably similar cross-linguistically. Finally, patterns of dependence between properties of signs are examined, in particular, correlations between handshape and location.

## 1.1 Organization of the dissertation

Chapter 1 introduces the reader to the fundamentals of sign language phonology, in particular, to the parameters of handshape and location, as well as analogues in sign language of phonological constituents of spoken language. A summary of cross-linguistic findings in the analysis of spoken language inventory structure is presented. Various markedness criteria, which will be important in Chapter 5, are gathered from different sources, and an attempt is made to reconcile them. Finally, the relevance of quantitative analysis of sign language to linguistics is discussed.

---

[1] Sign languages are usually referred to by an acronym that includes the name of the country or area in which the language is used, often written in the language of that area. SVK is an acronym for *Suomalaisen viittomakieli*, which means "Finnish Sign Language," and is occasionally abbreviated FinSL in some other places.

Chapter 2 discusses the methodology used in this study. Databases were constructed for the languages ASL, KSL, NZSL, and SVK based on published dictionaries. The way in which these databases were constructed is explained, including information on the dictionaries and software used, decisions made during the database creation, and the statistical and mutual information methods used in data analysis.

Chapter 3 examines the handshape parameter. Handshape is inventoried in ASL, KSL, NZSL and SVK; handshape inventories of Old Finnish Sign Language (VSVK), Sign Language of the Netherlands (NGT: Nederlandse Gebarentaal) and Italian Sign Language (Lingua Italiana dei Segni: LIS) are examined too. The lexical distributions for these four inventories are determined and compared cross-linguistically. It is found that the rank-frequency distribution of handshape is remarkably uniform cross-linguistically, closely approximating an exponential decay curve. Dependence between the variables of handshape and sign type is investigated.

Chapter 4 examines the location parameter. Location is inventoried in the four languages. The lexical distributions are determined and compared. In particular, non-body locations, such as neutral space, that is the location in front of the torso, and locations on the nondominant hand in Type 2 and Type 3 signs are investigated. Again, the rank-frequency distributions are found to be remarkably similar cross-linguistically, although in the case of location, they approximate a hyperbolic curve. Dependence between the variables of location and number of hands is investigated.

Chapter 5 examines duets, that is, pairs of handshape and location that occur simultaneously within a sign. The two-dimensional space defined by handshape – location duets is presented and the rank-frequency distributions of duets graphed. Next, the distribution through the handshape – location space is investigated to determine whether handshapes are uniformly distributed across locations or whether some distributional pattern exists.

Chapter 6 summarizes and discusses the results and provides suggestions for further research.

## 1.2 Background on sign language

The term *sign language* refers to a set of languages in the visual/manual modality; they are articulated with the hands and body and perceived with the eyes, as opposed to spoken languages, which are articulated with the vocal tract and perceived with the ears. These fully-formed, natural languages are mutually unintelligible and have grammatical structures parallel to but independent of spoken languages. In the past, when the languages of the Deaf were mentioned,[2] they were believed to be parasitic on spoken language, no more than the spoken language encoded, like writing, semaphores or Morse code.

Sapir (1921:21) said that some might speculate that in gesture languages "the ideas are directly conveyed by an utterly unrelated symbolic process or by a quasi-instinctive imitativeness." However, he goes on to assert that, "Such an interpretation would be erroneous. The intelligibility of these vaguer symbolisms can hardly be due to anything but their automatic and silent translation into the terms of a fuller flow of speech." Thus, while recognizing that information is conveyed in gesture languages and that direct imitation is inadequate for communication, Sapir concludes that full communication requires translation of these gestures by the perceiver into spoken language.

Bloomfield (1933:39) stated,

> "Some communities have a *gesture language* which upon occasion they use instead of speech. Such gesture languages have been observed among the lower-class Neapolitans, among Trappist monks (who have made a vow of silence), among the Indians of our western plains (where tribes of different language met in commerce and war), and among groups of deaf-mutes. It seems certain that these gesture languages are merely developments of ordinary gestures and that any and all complicated or not immediately intelligible gestures are based on the conventions of ordinary speech. … Linguistic forms, however, result, for the most part, in far more accurate, specific, and delicate co-ordination than could be reached by non-linguistic means. Apparent exceptions, such as elaborate systems of gesture, deaf-and-dumb language, signaling-codes, the use of writing, telegraphy, and so on, turn out, upon inspection, to be merely derivatives of language."

Both Sapir and Bloomfield are partially right: all of these communication systems necessarily originate from human linguistic capacity. However, in the case of Deaf sign languages, they are not "derivatives," as Bloomfield calls them, or "transfers," as Sapir calls them (Sapir, 1921:19), because they are not based upon any spoken language, but rather they are direct developments of this capacity, that is, autonomous languages.

---

[2] Following the conventional notation, Deaf with an uppercase "D" refers to the community of people who share a (signed) language and culture; deaf with a lowercase "d" refers to the audiological condition of diminished hearing.

Even if the expressive capacity and independent status of signed languages is recognized, it is not necessarily the case that these languages possess their own linguistic structure. Even if sign languages are not parasitic on spoken languages, it is possible that they are the product of other cognitive facilities, and as such, do not have a grammar. Thus, the form of an expression would vary without principle depending upon the signer, his pantomimic ability, and his familiarity with his partner. Communication would be accomplished through mutual cooperation and convention between the conversationalists. Another possibility is that sign languages do have grammars, but that the modality difference between spoken and signed languages permits vastly different grammars. Over the past four decades, these assumptions have been refuted.[3] Signed languages have been shown to have the same structural levels as spoken language: phonetics, phonology, morphology, syntax, and semantics. Moreover, the same theoretical tools can be used for their structural analysis.

Ironically, the difference in modality that initially obscured the relevance of signed languages to linguistics is ultimately the property that makes them such a fertile resource for investigation of some of the most intriguing questions in linguistics. What are true linguistic universals, as opposed to properties specific to the modality of speech? Are there modality-specific grammatical differences? What parts of language result from human linguistic capacity and what are attributable to more general cognitive abilities? What grammatical model is general enough to account for both signed and spoken languages without over-generation? How does the brain organize sign language, which is both a visual-spatial and a linguistic task? Does modality affect language acquisition? Does the creole-like nature of sign language transmission help us understand the structure of early languages? Can observing the birth of new sign languages give us insight into the origin of language?

---

[3] For the reader interested in reading about the development of the field of sign language research, *The Signs of Language* (Klima and Bellugi, 1979) is a good starting point. This book was the first detailed survey of sign language research to reach a wide audience, and it still remains one of the most interesting introductions. *The Signs of Language Revisited* (Emmorey and Lane, 2000), which provides a recent look at many of the same topics discussed in the first book, as well as indicating some new directions taken by the field of sign language research, would be an interesting second compilation to read.

### 1.3 The phonological elements of a sign

Before 1960, when Stokoe's work (1960) appeared, it was generally believed that not only were signs iconic representations of their predominantly concrete referents, but also that they were holistic, atomic gestures. Therefore, although signs and words are analogous functionally, they were considered to be formally different. Since phonology is closely related to the phonetic resources used in the formation of words, it is reasonable to expect that the greatest modality differences would be found in this level. Stokoe had the insight that signs can be analyzed into sublexical units. He wrote (1965:vii),

> "Each sign of this language has three things which distinguish it from all other signs in the language. Let us call these things *aspects* since they are ways of looking at something that can happen all at once. The three aspects of a sign are (1) the place where it is made, (2) the distinctive configuration of the hands or hands making it, and (3) the action of the hand or hands."

These aspects are now commonly called *location*, *handshape*, and *movement*. By showing that linguistic analysis was possible even at the level of the sign, Stokoe commenced the linguistic study of sign language.[4] After Stokoe, other aspects, sometimes called minor parameters, have been proposed. Among them are *orientation*, the direction the fingers and palms are facing, *contact*, the region of the hand or hands that contacts the body or the other hand, *hand arrangement*, the role played by each hand in a sign, and *non-manual components*, face and mouth movements that accompany or are part of a sign.

#### 1.3.1 Handshape and location

Handshape and location were two of the phonological components present in the earliest analysis of sign structure (Stokoe, 1960), and they remain present in the most recent analyses (Brentari, 1998). There is abundant evidence that handshape and location are major elements in sign production. While ASL was the first language whose phonology was studied, subsequent investigation into other sign languages has confirmed that these sublexical elements are relevant to all sign languages.

---

4 Tervoort's (1953) dissertation, Structural analysis of a visual language used by a group of deaf children, apparently was the first linguistic work to describe a sign language (Sign Language of the Netherlands) and to argue that it is actually a language.

### 1.3.1.1 Phonological evidence

Every sign is produced with a particular handshape at a particular location with a particular movement, and substituting one for another can change the meaning of a sign.[5] Thus, there are minimal pairs differing only in handshape, only in location, and only in movement as shown in (1-1), (1-2), and (1-3).[6] The signs CANDY and APPLE, in (1-1), are both made with one hand performing the same movement at the cheek. The handshapes are similar in that both have the index finger extended, the thumb opposed and the remaining fingers closed. However, in CANDY the index finger is straight, while in APPLE it is curved.[7] The signs APPLE and ONION, in (1-2), differ only in location; all other parameters are the same. APPLE is articulated at the cheek and ONION at the side of the eye. The SVK signs KAHDEKSAN_VUOTIAS *eight year old person* and KAHDEKSAN_MARKKAA *eight marks* (a monetary unit), in (1-3), differ only in movement. In KAHDEKSAN_VUOTIAS, the movement is inward and rotating, while in KAHDEKSAN_MARKKAA, the movement is outward and rotating. (A list of the handshapes and locations, together with their HamNoSys notations, is provided in Appendix A.)

---

[5] In this dissertation, a particular location is identified as such if there is contact between the hand and the location or it there is just proximity. There are signs in ASL and SVK which vary apparently freely between proximity and actual contact. I do not know if this alternation occurs in KSL and NZSL.

[6] The system of notation used is the Hamburg Notation System, abbreviated HamNoSys (Prillwitz, 1987). It is discussed in section 2.1.2.1 and illustrated in Appendix A.

[7] It is conventional to indicate a sign by its gloss in the dominant spoken language of the area in which it is used. This gloss is written in small capital letters. I will gloss ASL and NZSL signs with English words, SVK signs with Finnish words, and KSL signs with Korean words written in the Roman alphabet. English translations are also provided for the Finnish and Korean signs. When one sign needs two words for its gloss it is notated WORD_WORD to distinguish it from a sign language compound, which is notated WORD+WORD.

(1-1) Handshape: minimal pairs in ASL

| CANDY | APPLE |
|---|---|
| handshape: 49<br>location: ξ<br>movement: rotation<br> | handshape: 4B9<br>location: ξ<br>movement: rotation<br> |

(1-2) Location: minimal pairs in ASL

| APPLE | ONION |
|---|---|
| handshape: 4B9<br>location: ξ<br>movement: rotation<br> | handshape: 4B9<br>location: υ□<br>movement: rotation<br> |

(1-3) Movement: minimal pairs in SVK

| KAHDEKSAN_VUOTIAS *eight year old* | KAHDEKSAN_MARKKAA *eight marks* |
| --- | --- |
| handshape: < | handshape: < |
| location: π | location: π |
| movement: rotating inward | movement: rotating outward[8] |
|  |  |

Other phonological evidence for handshape and location as phonologically relevant parameters is found in the process of compound formation in ASL and SVK, which typically involve assimilation between the members. Both handshape and location participate in the regressive and progressive assimilation processes shown in (1-4). Sandler (1989) notes that regressive handshape assimilation occurs in the ASL sign MIND + DROP = *faint*, where MIND has the handshape 49, and DROP has the changing handshape ; η 78A. The handshape of DROP spreads to MIND, so that both members of the compound share the same handshape. Likewise, the SVK compound KUURO *deaf* + KERHU *club* = *deaf club* exhibits regressive handshape assimilation. The three SVK compounds TIETÄÄ *know* + ALUE *area* = *theory*, PORO *reindeer* + ALUE *area* = *Lapland*, and YLIOPPILAS *student* + TALO *house* = *university* exhibit progressive location assimilation. When the second members of these compounds, ALUE *area* and TALO *house*, are signed independently, they are articulated in neutral space, in particular, in the area in front of the chest. When they are members of these compounds they are articulated at forehead height. *Lapland* and *theory* are articulated near the side of the forehead without contacting it. *University* is articulated in front of the forehead without contacting it.

---

8 In Finland I was taught the mnemonic that in the sign about age, the rotation is toward the signer because years keep coming to us, while in the sign about money, the rotation is away from the signer because money keeps going away from us.

(1-4) Compound assimilation

| first member | second member | compound |
|---|---|---|
| MIND<br>handshape: 49 | DROP<br>handshape: ; η 78A | *faint*<br>handshape: ; η 78A |
| KUURO *deaf*<br>handshape: 59 | KERHU *club*<br>handshape: 4≅B | *deaf club*<br>handshape: 4≅B |
| TIETÄÄ *know*<br>location: σ□ | ALUE *area*<br>location: π | *theory*<br>location: σ□ |
| PORO *reindeer*<br>location: σ□ | ALUE *area*<br>location: π | *Lapland*<br>location: σ□ |
| YLIOPPILAS *student*<br>location: σ | TALO *house*<br>location: π | *university*<br>location: σ |

#### 1.3.1.2 Psycholinguistic evidence

The psychological reality of handshape and location is supported by evidence from many studies. Two early studies are reported in Klima and Bellugi (1979), henceforth, K&B. The first is a memory experiment. Hearing subjects, when asked to recall a spoken list of unrelated words, not only omit forgotten words, but also substitute words. These are called intrusion errors, and the study of these errors shows that hearing people encode words in short-term memory phonologically, rather than, for example, semantically. A sample of intrusion errors made by hearing speakers of English is shown in (1-5). If the stimulus words were being encoded semantically, for example, the intrusive errors might have been *elect* for *vote*, *coffee* for *tea*, and *Pepsi* for *Coke*.

(1-5) Intrusion errors made by English speakers

| word presented (spoken) | erroneous response (written) |
|---|---|
| vote | boat |
| tea | tree |
| Coke | coat |

Compare the intrusive errors in (1-6) that ASL signers made when they had to recall in written English a signed list of ASL signs. As discussed above, CANDY and APPLE are minimal pairs differing only in handshape, and ONION and APPLE differ only in location.[9] These intrusion errors suggest not only that signers encode signs phonologically in their short-term memory, but also that the parameters that were altered, including handshape, location, orientation and movement, are phonologically relevant.[10]

(1-6) Intrusion errors made by ASL signers

| presented sign (signed) | erroneous response (written) | mistaken parameter |
|---|---|---|
| ONION | apple | location |
| CANDY | apple | handshape |

The second study is parallel to a slip of the tongue study in spoken language. The fact that sublexical features can be metathesized, anticipated or perseverated is evidence of their relevance to phonological organization. In this study, a corpus of 131 signing errors, called slips of the hands, was compiled. If ASL signs were indeed holistic gestures, the only type of systematic error would involve entire signs. Out of 131 errors, only 9 are of this type. Sixty five involve substitution of a handshape parameter, thirteen involve substitution of a location parameter, and eleven involve substitution of a movement parameter. An example involving an exchange of handshape is shown in (1-7), and an exchange of location is shown in (1-8). In each slip of the hand, other parameters, such as movement and orientation, are correct.

---

9 This error in location is presented in K&B as having occurred in another experiment.

10 The argument is not circular; it is possible to construe other types of errors for signers to make, such as semantic ones: *sweet* could have been written for *candy*, for example.

(1-7) Slip of the hand involving handshape

| | SICK | BORED |
|---|---|---|
| intended expression |  |  |
| erroneous articulation with handshape metathesis of 78□A 49 |  |  |

(1-8) Slip of the hand involving location

| | RECENTLY | EAT |
|---|---|---|
| intended expression |  |  |
| erroneous articulation with location metathesis of ξ ψ |  |  |

Neurolinguistic studies also provide evidence for the psychological reality of phonological parameters. Impaired signing, as produced by signers with aphasia (Poizner et al., 1987; Whittemore, 1987; Corina, 1998) or Parkinson's Disease (Brentari and Poizner, 1994; Brentari et al., 1995) or by signers undergoing direct cortical stimulation (Corina et al., 1999), dissolves according to the partitions delineated by the phonological parameters of handshape, location and movement. Likewise, other psycholinguistic studies that have investigated the on-line activation of phonological parameters through priming experiments show that handshape and location are involved in facilitation or inhibition of responses (Corina and Emmorey, 1990; Corina, 2000b; Corina and Hildebrandt, 2002).

### 1.3.2 Other parameters

In addition to handshape and location, there are other aspects of sign formation that are regarded as phonologically relevant. In some cases these other aspects have been notated in the databases that have been created for the studied languages, but they have not been inventoried or quantitatively analyzed as handshape and location have.

Stokoe (1960) included movement as the third relevant parameter. While the existence of movement in sign articulation is undisputed, the phonological status of movement is controversial. Movement as a phonological constituent is crucial to the theories of Liddell and Johnson (1989), Sandler (1989), and Perlmutter (1991). These theories view movement as segmental; Perlmutter (1991) places movement at the nucleus of the syllable. In contrast, other researchers, such as Stack (1988), Wilbur (1990; 1993), Hayes (1993), van der Hulst (1993) and Uyechi (1994), deny any phonological status to movement. In these theories, movement is an epiphenomenon, the inevitable result of a change in location, handshape or orientation. Movement was not transcribed in the databases created for this dissertation.

Battison (1978) added orientation to Stokoe's three sign parameters. Orientation is viewed as a relatively minor parameter, and in theoretical work is represented as a dependant of handshape (Sandler, 1989 and others; Hulst, 1996a). Orientation was transcribed in the databases created for this dissertation, but it was not inventoried or quantitatively analyzed. Contact also is considered a minor parameter, with controversial status, either as an important, multivalent feature (Friedman, 1976; Mandel, 1982) or as a redundant property that is the result of phonetic implementation (Kooij, 1997). Contact was transcribed but not inventoried or analyzed.[11]

In addition to the activity of the hands, non-manual components, especially mouth movements, occur during sign production. Some non-manual components are bound morphemes that convey syntactic information, such as the raised brow and backward tilted head that marks a topic in ASL (Liddell, 1980). Other non-manual components mark category, such as the mouth movements that distinguish some nouns from verbs in SVK (Rissanen, p.c.). Often the mouth movement is a necessary accompaniment to a lexical item. This mouthing can be derived from

---

11 Contact data were used to answer question such as "Do all instances of 39 have ulnar contact?"

the spoken language gloss of the sign, such as the mouthing [v] that accompanies the ASL sign HAVE, and the mouthing [vuo] that accompanies the SVK sign VUOSI *year*. Or the mouthing can be unrelated to spoken language, such as the mouthing [pa] that accompanies the ASL sign FINALLY, and the mouthing [pi] that accompanies the SVK sign AITO *true*. Non-manual components were transcribed when either visible in the drawing or indicated in the text, but they were not investigated.

An essential aspect of sign formation is sometimes called *hand arrangement* (Klima and Bellugi, 1979). This broad term, referred to by Battison (1978) as sign “type”, describes the number of hands used in the articulation of a sign and the relationship between the two hands when both are employed. In brief, a sign can be articulated with only one hand (Type 0/X) or with both hands. If a sign is articulated with both hands, either both hands move, in which case they must have the same handshape, location orientation and movement (Type 1), or one hand moves and acts upon the other, which remains still. In this case, either the hands have the same handshape (Type 2) or different handshapes (Type 3). In addition, the handshape of the non-active hand in Type 3 signs is restricted to a small set. This typology, which was created for ASL, characterizes almost all monomorphemic signs in ASL, and it appears to be a good typology for other sign languages, as well, and in particular, for the four languages investigated here. Sign type was transcribed for every sign and is well-investigated in this dissertation. It is discussed further in section 2.1.2.2.3.

### 1.3.3 Organization of constituents

The previous sections have presented a phonological decomposition of signs into constituent units: handshape, location, movement, and so on. These units have been further decomposed in numerous different analyses into features, which are directly analogous to phonological features in spoken languages. However, the analogy between signed and spoken language phonology is not clear for constituents on intermediate levels of analysis. For example, is there a sign language analog to the mora, syllable, and segment of spoken language? The theory that basic aspects of linguistic structure are innately present in all humans predicts that phonological concepts proven relevant to spoken language phonology should have reflexes in any human language, including signed languages. Researchers following this theory search for and speak of syllables, segments and other phonological structures. Other researchers, such as Edmondson (1987), argue that concepts developed for the phonological analysis of spoken language, in particular, the segment, are irrelevant at best and misleading at worst for the analysis of signs.

The earliest phonological analysis (Stokoe, 1960) did not use the term “phoneme” for any formational constituent “to avoid false analogy” (Stokoe, 1960:30), instead using the term “chereme” (from the Homeric Greek *cher* ‘handy’) for the parameters of handshape, location and movement. (In contrast, Battison (1978) chooses to use the word “phonology” to highlight similarities between speech and sign.) These three meaningless elements combined to form a meaningful element, the sign. The question then is, are these parameters analogous to phonemes, or are they analogous to dimensions of spoken language sounds, such as place of articulation? In various chapters of K&B, for example, it is stated that handshape, location and movement are phonemes, and that a major difference between spoken and signed languages is that spoken languages arrange phonemes sequentially while signed languages arrange them simultaneously.

Some later phonological analyses do not dispute the notion that these parameters are phonemic, but rather assert that there is sequential internal structure to signs (Liddell, 1984a; Sandler, 1986; Sandler, 1987; Perlmutter, 1988; Liddell and Johnson, 1989; Sandler, 1989). The development of autosegmental phonology for spoken language (Goldsmith, 1976) presented the theoretical possibility of simultaneously occurring tiers, so that phonological information, such as nasality or tone, could extend over several segments. Sandler, for example, has as sequential phonemes L and M (for location and movement); handshape is autosegmentally represented on its own tier, and association lines connect it to the appropriate sequential L and M slots. Another phonological constituent from spoken language whose existence has been argued for in signed language is the syllable. Some syllable models are compatible with the notion of segments (Liddell, 1984a; Sandler, 1986; Sandler, 1987; Perlmutter, 1988; Sandler, 1989; Perlmutter, 1992; Perlmutter, 1993)[12], while others are not, in that the only sequential elements are the syllables themselves, which are based on features (Wilbur, 1990; Brentari, 1992; Wilbur, 1993; Corina, 1996).

In some models, sign constituents, such as handshape, location and movement, are subphonemic, akin to major class feature groupings such as place or manner of articulation of

---

[12] In these theories, the segmental elements of L and M are organized into different syllable types, such as LM, ML, LML, MLM, and L. It is suggested that syllables are determined by sonority, which is defined in various ways, usually with reference to visual salience. See the references in this paragraph for more discussion.

spoken language segments. Liddell (1984b) suggests this analysis and comments on the combinatorial possibilities that arise from it:

> "Stokoe's proposal that handshape, movement and location are phonemic in ASL is a very appealing and long-held idea. However, the entire segment, rather than these aspects of a segment, is the ASL unit which carries out the contrastive functions as a phoneme. A preliminary look at the number of possible contrastive segments in ASL suggests that the number will be considerably larger than that found in spoken languages. If this result is born out after a thorough analysis, it would represent a very interesting modality difference."

Van der Hulst (1985; 1995) also expresses this view. He concludes that most sign language morphemes are monosegmental. Likewise, Channon (2002) concludes that simple signs are best represented as monosegmental.

Among the models that accord phonemic status to at least some of the parameters of handshape, location and movement, there is considerable discussion of the analogy between these parameters and the segment types of consonant and vowel. According to many researchers, movement is analogous to the vowel (Chinchor, 1978; Liddell, 1984a; Perlmutter, 1992; Sandler, 1993b; Brentari, 1998). Handshape is represented autosegmentally on a separate tier, analogous to tone (Sandler, 1989; Brentari, 1990b; Corina, 1990). Based on sonority arguments, location in these models is analogous to the consonant. In contrast, Corina (2000a) draws an analogy between handshape and consonant on the basis of the susceptibility of handshape to error in paraphasic signing. Siedlecki and Bonvillian (1993) suggests that location, as a phoneme class, might be similar to vowels, in that location, like vowels, is acquired early and accurately by children.

### 1.4 Spoken language inventory structure

Much of what is known about the size and structure of the inventories of spoken language is derived from a database created the UCLA Phonetics Laboratory called the UCLA Phonological Segment Inventory Database (UPSID) (Maddieson, 1984).[13] This ongoing project is a database containing phonological segment inventories of a large (317 languages), genetically

---

[13] This section refers to the 1984 version of UPSID. Maddieson and Precoda (1989) discusses a revised version of the database.

representative sample of the world's languages, on which cross-linguistic generalizations about inventory structure can be based. This section summarizes some of the findings on spoken language inventory structure presented in Maddieson (1984) that are relevant to this dissertation.

According to UPSID, inventory size varies widely. The distribution is not normal; it is positively skewed and platykurtic, that is, the high end tail is longer than the low end tail, both tails are heavy, and the peak is low. The number of segments varies between 11, for Rotokas and Mura, and 141, for !X⊗. The inventory size of seventy percent of the languages lies between 20 and 37 segments. The mean inventory size is 31, while the median is between 28 and 29.

According to Maddieson (1984:7), "Whether the tendency to have from 20 to 37 segments means that this is an *optimum* range is an open question. It seems likely that there is an upper limit on the number of segments which can be efficiently distinguished in speech, and a lower limit set by the minimum number of segments required to build an adequate vocabulary of distinct morphemes. But these limits would appear to lie above and below the numbers 37 and 20, respectively."

Indeed, it is reasonable to assume that these limits lie above 141, the number of segments in !X⊗, and below 11, the number of segments in Rotokas and Mura. Presumably, speakers of !X⊗ are able to distinguish all of their phonemes, and speakers of Rotokas and Mura have an adequate vocabulary. Maddieson (1984) points out that comparative studies of Khoisan languages (Baucom, 1974; Traill, 1978) show that large inventories are a stable feature of these languages that has persisted over time. Likewise, languages with small inventories exhibit no difficulties in forming sufficient contrastive morphemes, such as an unmanageably large number of homonyms or morphemes with too many segments. Indeed, Hawaiian, with only 13 segments, has a mean morpheme length of only 3.5 segments. Likewise, comparative studies of language families with small inventories show small inventories to be a stable, persistent feature of these languages (Grace, 1959). If these very small or very large inventories were problematic for speakers, there would be pressure for them to add or lose segments to move closer to the range of 20 to 37 segments. Since this does not occur, extreme inventories must be capable carriers of language.[14]

Maddieson (1984) also discusses possible principles governing the relation between inventory size and structure with respect to several possible constructions. One possibility is that

[14] Note that Mura, Rotokas and Hawaiian are not tonal languages. Indeed, contrastive tone or stress is more likely to occur in languages with large inventories (Maddieson, 1984: 21).

*compensation* governs the composition of inventories. According to this hypothesis, simplicity of one aspect of an inventory necessitates complexity of another. One aspect of inventory structure considered is the ratio between vowels and consonants, which varies between 0.065 and 1.308 with a mean of 0.36. The trend is that while larger inventories are dominated by consonants, large inventories also have a larger number of vowels than smaller inventories, which is opposite what compensation predicts. Stop and affricate inventories are examined separately to see if there is an inverse relationship between number of places of articulation and number of manners of articulation.[15] Also, stop and affricate inventories are compared to see if there is an inverse relationship between number of stops and number of affricates in an inventory. In all three cases there is either no correlation (stop places versus manners), or a positive correlation (affricate places versus manners and stops versus affricates), contrary to the prediction of the hypothesis. Segment inventory size is compared with suprasegmental involvement, such as contrastive stress and tone, with the result that larger inventories are more likely to use suprasegmental features than smaller inventories. Finally, segment and syllable inventory sizes are compared, to determine if elaboration of the segment inventory is met with compensatory phonotactic restrictions on syllable shapes. As with all of the other possible instances of compensation, it is found that complexity of one sort coincides with complexity of another sort.

Note that Maddieson's results about the lack of compensatory complexity apply only to the phonological component. There is no claim about whether or not compensatory complexity occurs in other levels of linguistic analysis. For example, Nettle (1995) examines segment inventory size versus word length. Based on data from ten languages, the relationship $L = 29.35\, S^{-0.43}$ between inventory size $S$ and word length $L$ is derived, supporting the hypothesis that as inventory size increases, word length decreases.[16] Maddieson's point is that phonological systems as a whole can be either simple or complex; balance does not need to be achieved.

The second relation between inventory size and composition discussed was the hypothesis that small inventories contain common segments while large inventories contain rare segments. In particular, the hypothesis that a smaller inventory is more likely to contain a common segment, while a larger inventory is more likely to contain an uncommon segment, was

---

[15] The hypotheses about compensatory complexity in this paragraph are all explicitly made and tested in Maddieson (1984) on pages 17 through 21.

[16] Thanks to Hargus (p.c.) for pointing out that Nettle (1995) is in accord with Maddieson's findings.

tested. However, when the frequency of thirteen of the most common segments in a sample of 57 small and 54 large inventories was examined, this generalization was not found to be true.

On the other hand, implicational statements about particular segments are generally sustained by the UPSID data, though not necessarily unanimously. These statements are of the type shown in (1-9). These observations, however, cannot be consolidated into a single hierarchy.

(1-9) Implicational generalizations about the co-occurrence of segments

a. / p / implies / k / occurs.

b. Nasal consonants imply stops and affricates occur at the same place.

c. Voiceless nasals and approximates imply their voiced counterparts occur.

d. Mid vowels imply high and low vowels occur.

Moreover, prohibitions on segment co-occurrence, of the type shown in (1-10), also make it impossible to form a single hierarchy. These implicational generalizations and co-occurrence restrictions hold for the great majority of the languages in UPSID, but there are a few exceptions.

(1-10) Prohibitions on the co-occurrence of segments

a. Laryngealized plosives and (voiced) implosives do not co-occur.

b. Voiceless lateral fricatives and approximants do not co-occur.

c. / ÷ / and / f / do not co-occur. / B / and / v / do not co-occur.

While some of these co-occurrence restrictions involve salient phonetic contrasts, some of them involve minimally contrastive segments. This observation leads to a third hypothesis about inventory structure: inventories are structured to maximize phonetic salience. Segments do not co-occur contrastively unless they are sufficiently distinct. However, maximization of phonetic distinctness alone, while operational in inventory composition, is not the principle governing inventory structure. Instead, Maddieson (1984:16) suggests that certain dimensions of contrast are preferentially selected independent of phonetic distinctness, and other dimensions of contrast are added only after the primary dimensions are utilized to a sufficient extent.

In summary, while there are generalizations about inventory size and structure in spoken language, they cannot be consolidated into a single theory that can adequately predict inventory composition. Most notably, the idea of compensation is not validated; the opposite appears to hold: complexity in one aspect accompanies complexity in another aspect. Similar issues will be addressed in regard to sign language inventories. In particular, the notion of compensation will be investigated to determine if simplicity in one aspect of sign formation, such as the handshape inventory, implies complexity in another aspect, such as the location inventory. As in spoken language, it will be seen that the sizes of the handshape and location inventories are directly related; in contrast, the sizes of the duet inventories will be shown to be remarkably similar for all four languages investigated.

### 1.5 Markedness

The concept of linguistic markedness was developed in conjunction with distinctive feature theory by Jakobson and Trubetzkoy in the early twentieth century (Jakobson et al., 1952; Trubetzkoy, 1967 [1939]). With binary features, phonemes often occur in pairs; one member has a feature that the other lacks. For example, the pair of bilabial stops /b/ and /p/ differ only in that /b/ has the property of being voiced; thus, /b/ is marked, while /p/, which does not have this additional property, is unmarked. Frequently, the notion that a particular *set* of phonemes is marked, while another, disjoint set is unmarked is put forth.

Although actual correlates of markedness are difficult to pinpoint, markedness is intuitively appealing and useful in phonological theory. For example, a central tenet of Optimality Theory (OT) is the claim that markedness is a part of Universal Grammar (McCarthy and Prince, 1993; Prince and Smolensky, 1993). According to OT, phonology is the competition between faithfulness constraints, which strive to maintain existing differences in underlying representations, and markedness constraints, which strive to simplify output forms according to language-universal propensities. These markedness constraints are often quite specific to the oral-aural modality, such as NO-LAR, a constraint banning the laryngeal place of articulation. It seems unlikely that humans would have a biological endowment of two sets of markedness constraints, one referring to the oral-aural modality and the other to the manual-visual modality. Yet, since signed languages parallel spoken languages in their function, structure and acquisition, it is hard to see how innate markedness constraints are crucial for spoken languages but unnecessary for signed languages. Alternatively, constraints can be formulated in a sufficiently abstract manner so as to be applicable to either modality. For example, Kirchner (1998) adduces a principle of effort minimization that he calls LAZY that is theoretically applicable to sign languages. See Rozelle (1998) for further discussion of the impact of sign language on Optimality Theory.

Sign language linguists have sought to classify phonological parameters, especially handshape, according to their markedness. Either the set of handshapes is divided into two disjoint subsets of marked versus unmarked handshapes, or the set is ordered on a continuum from more marked to less marked. Then a correspondence between the set of unmarked handshapes and some linguistic phenomenon under investigation is drawn. If it is the case that the same set of handshapes is implicated in each phenomenon, a significant result has been obtained. Additionally, it is posited that markedness criteria are universal; the same set of unmarked handshapes should function in the same way cross-linguistically.

Lists of markedness criteria for ASL have been compiled from numerous sources; see, for example, Brentari (1990a) and Sandler (1996). There are four arenas in which markedness is said to play a role, shown in (1-11). The next four sections examine each of these criteria with respect to handshape and, as appropriate, location. Unfortunately, as will be seen, these four criteria do not provide an unambiguous unmarked set. In fact, even within a criterion, such as handshape acquisition by signing children, researchers do not agree on an unmarked set.

(1-11) Criteria for markedness in ASL handshapes

| | Unmarked handshapes: |
|---|---|
| acquisition and impairment | Children and aphasics make fewer errors on unmarked handshapes and substitute them for other shapes. |
| productive and perceptual ease | Unmarked handshapes are motorically easier to produce and are easier to discriminate than other shapes. |
| phonology | Unmarked handshapes have special phonological uses. |
| frequency | Unmarked handshapes occur more frequently both language internally and cross-linguistically. |

### 1.5.1 Acquisition and impairment

Unmarked handshapes and locations are hypothesized to be acquired by children earlier than marked handshapes, and when an incorrect element is substituted, this substitution is predicted to be drawn from the unmarked set. There have been numerous studies of signing children's acquisition of phonological parameters, particularly handshape and location (McIntire, 1977; McIntire, 1980; Hamilton, 1986; Boyes Braem, 1990; Siedlecki and Bonvillian, 1993;

Bonvillian and Siedlecki, 1996; Siedlecki and Bonvillian, 1997; Conlin et al., 2000; Marentette and Mayberry, 2000). Most of these studies have been conducted on American Sign Language, but Takkinen (1990) investigates the acquisition of Finnish Sign Language. These studies indicate that children's errors align with phonological parameters. Results show that handshape is the most difficult parameter to master, that handshapes are acquired in a particular order, and that handshape substitutions occur in a predictable manner. This evidence indicates that children are acquiring a structured, phonological system. It provides further support for the utility and independence of the parameters of location and handshape, as well as potential evidence for the classification of handshapes as marked or unmarked. Examples of handshape and location errors are shown in (1-12). In (1-12 a) the handshape 28 is substituted for the correct handshape 78. In (1-12 b) the location 3 on the nondominant hand (H2[17]) is substituted for the correct location □.[18]

(1-12) Handshape and location errors in child language acquisition

| | |
|---|---|
| a. MOTHER<br>substituted handshape: 28<br>location: ζ<br>movement: contacting<br> | b. CRACKER<br>handshape: 2<br>substituted location: 3 on H2<br>movement: contacting<br> |

[17] Whether a sign is made with the right or left hand is not phonologically distinctive in any known sign language, although dominance can switch for morphological and syntactic reasons. Right-handed people use their right hand when signing one-handed signs, and left-handed people use their left. The left hand in left-handed people and the right hand in right-handed people is called the *dominant hand*, the *strong hand*, or *H1*; the other hand is called the *nondominant hand*, the *weak hand*, or *H2*.

[18] See Appendix A for a directory of HamNoSys notation.

Differences in accuracy in the production of handshape and location across a number of studies are shown in (1-13). Handshapes are articulated with a low level of accuracy, while locations are articulated fairly well, an observation in line with the suggestion made by Kantor (1980) that the order of acquisition of phonological parameters is first location, then movement, and lastly handshape.

(1-13) Accuracy in children's production or perception of phonological parameters

| | SL | type | ages yr;mo | hs | mov | loc |
|---|---|---|---|---|---|---|
| Conlin, et al. (2000) | ASL | production | 0;7-1;5 | 25% | 54% | 81% |
| Marentette, et al. (2000)[19] | ASL | production | 1;0-2;1 | 27% | 57% | 83% |
| Siedlecki, et al. (1993; 1997) Bonvillian, et al. (1996) | ASL | production | 0;6-1;6 | 50% | 61% | 84% |
| Takkinen (1990) | SVK | production | 2;4-3;3 | 81% | 76% | 94% |
| Hamilton (1986) [20] | ASL | perception | 6-9 | 95% | 94% | 97% |

Several explanations have been advanced by Siedlecki and Bonvillian (1993) and expanded on by others for the early acquisition of location. The gross motor skills required for location articulation are acquired earlier than the fine motor skills required for the articulation of

19 The percentage for location accuracy is the average of the horizontal location accuracy (89%), which encodes how far away from the body a sign is produced, and the vertical location accuracy (74%), which encodes near which body part a sign is produced. No total location accuracy rate was calculated (Marentette, p.c.).

20 This article presented the results of a perception test performed by children age six to nine. Thirty children viewed 30 stimuli each, for a total of 900 responses. There were 125 discrimination errors in all. Of these errors, 29 were made on pairs differing only in location, 50 on pairs differing only in movement, and 46 on pairs differing only in handshape. While the accuracy of handshape perception by these older children is much higher than the accuracy of handshape production by the younger children studied in the other articles, the pattern is still the same: handshape is not as accurate as location.

handshapes. In particular, by the age of one year, children are able to point to and reach for an object. Memory for spatial information develops early, and by this age, children have also acquired a basic body schema for how their body is organized and how it functions (Neisser, 1991; Butterworth, 1992). These factors, together with possible visual, tactile and kinesthetic feedback, allow a child to reproduce on her own body locations that she sees articulated on the body of another signer.[21]

McIntire (1977), Boyes Braem (1990), Siedlecki and Bonvillian (1997), Conlin et al. (2000), and Marentette and Mayberry (2000) propose stages of handshape acquisition in ASL, but they do not agree on an order of acquisition. However, this set of first handshapes is approximately {78 49 2 / 29 }; this set either includes or is followed in a second stage by {3 3≅B : ; 48 }, where 2 and 29 are considered allophonic in Boyes Braem (1990) and Siedlecki and Bonvillian (1997), and 3 is included in the earliest acquired set by Siedlecki and Bonvillian (1997) and Marentette and Mayberry (2000).[22] In addition, when a child does not articulate the correct handshape for a sign, the substituted handshape is chosen from this set of first handshapes. According to these researchers, these handshapes are the least marked handshapes in ASL.[23]

The pattern of accurate location articulation and error-ridden handshape articulation is repeated in situations where language is impaired in signing adults experiencing phonological paraphasia. Corina (2000a) reports on the formational errors produced by three aphasic signers. There were abundant handshape errors but only one location error; movement and orientation errors were also rare. Corina suggests an explanation for this pattern. In many phonological models (Corina and Sagey, 1989; Sandler, 1989; Brentari, 1990b), location and movement are autosegmentally represented together on a tier, while handshape is represented on its own

---

[21] I would like to ask, however, whether the categories for locations were sufficiently narrow to capture errors. The transcription of location in the databases created for this dissertation are narrower. In these articles, if the location parameter of adult target signs were transcribed more narrowly, perhaps children would have evidenced more location errors.

[22] It is interesting that 3 is not universally considered to be one of the earliest acquired handshapes since 3 is the most frequent handshape is every sign language examined in this dissertation.

[23] Due to the paucity of location errors, no model of location acquisition emerges.

separate tier, thus accounting for its independent behavior. It has been suggested that movement is analogous to vowels (Perlmutter, 1991), so perhaps handshape is analogous to consonants. In spoken language, consonants are particularly susceptible to disruption in aphasia, while the syllable is preserved (Shankweiler and Harris, 1966).

However, in contrast to the language acquisition evidence, on a number of occasions in Corina (2000a), a late-acquired handshape, such as 59, 48□, or 69□□, was substituted for an early-acquired handshape, such as 49 or 78, in impaired signing. In spoken language, impaired speech substitutes unmarked forms for marked forms. Corina (2000a) proposes that this pattern of substitution is evidence for a feature-based definition of markedness, such as that developed by Brentari (1990b), in which handshapes with fully open configuration, such as 59, 48□, or 69□□, have the unmarked feature value of [+peripheral].

### 1.5.2 Productive and perceptual ease

Unmarked handshapes are hypothesized to be easier to produce than marked handshapes. The most extensive production study is Ann (1993), which addresses the question of whether "easy" handshapes are common and "hard" handshapes are rare in the lexicons of two sign languages, ASL and Taiwan Sign Language. She presents a detailed account of the physiology of the hand and, based on this physiology, assigns difficulty scores to handshapes. "Easy" handshapes receive scores of zero, "hard" handshapes receive scores of one, two or three, and "impossible" handshapes receive scores of four or more.[24] Generally, the easy handshapes occur more often than expected and the hard handshapes occur less often than expected.[25] This study is particularly useful since it uses information outside linguistics for establishing a class of

---

[24] "Impossible" handshapes are included just because they are a combinatorial possibility, not a physiological possibility. For example, it is possible for the index finger and the pinky to be fully extended while the remaining fingers are fully closed; however, it is physiologically impossible for the middle and ring fingers to be fully extended while the remaining fingers are closed. Handshapes that approximate this description actually have the middle or ring finger bent or do not have the other fingers fully closed.

[25] "Expectation" in this context is a statistical concept. The expected value is calculated by multiplying the row count (for example, the number of signs with a particular handshape) by the column count (for example, the number of signs with easy handshapes) and dividing by the grand total of the whole table.

handshapes, and then applies this independently defined classification to a linguistic question, that of handshape frequency in the lexicon, thus verifying a suspected relation between two markedness criteria. However, the set of unmarked, that is, easy, handshapes so defined is quite large. Ann identifies 44 handshapes in the dictionary she used (Stokoe et al., 1965); out of these, 21 are easy and 23 are hard.

Perceptual ease refers to both discrimination and disambiguation of stimuli. Battison (1978:36) says that the unmarked handshapes ( 3 2 29 4≅B ; 49 78 by his claim) are "maximally distinct, basic geometric shapes … maximally distinct in both articulatory and perceptual terms," though no evidence is presented for this. The goal of Lane, et al. (1980) and Stungis (1981) is to determine a featural analysis of handshape based on visual salience. Subjects were asked to identify and discriminate handshape stimuli amidst visual noise, and confusion groupings were determined. The four least confused handshapes were 78, 3, 3≅B and ; (Lane et al., 1980). However, 49, which by criteria of productive ease, early acquisition, and phonological processes is one of the least marked handshapes, ranked 19 out of 20, in Lane et al. (1980), that is, only one other handshape was misidentified more often than 49. In Stungis (1981) it ranked 17 out of 20, only 59, 68, and 69 were misidentified more often. In fact, the five handshapes misidentified least often were 59□□□, 78, 39X, 3≅B, and ;. By other criteria, 59□□□ is considered to be one of the most highly marked handshapes.

The perception of location has received less attention than handshape, but one interesting study is Poizner and Lane (1978). The stimuli in this study were possible but non-occurring ASL signs presented amidst visual noise. Deaf and hearing subjects were asked to identify the location. The two groups of subjects had similar confusion clusters, that is, similar groups of locations that were incorrectly identified for each other. Unsurprisingly, these confusion clusters produced an almost perfect representation of the topography of the body. Also, averaging over all locations, both groups correctly identified location about 62% of the time. What was surprising, however, was that the two groups differed on which locations were identified correctly most often. Deaf subjects were most accurate in identifying locations listed most frequently in the ASL dictionary (Stokoe et al., 1965), while hearing subjects were most accurate in identifying locations used less frequently. Presumably, Deaf subjects are skilled at identifying the most frequent locations in the ASL lexicon because they have trained on these locations, while hearing subjects have not. But the question remains: why is the ASL lexicon structured in a way that does not take advantage of what must be purely perceptually based abilities? Is the lexicon structured instead to facilitate production?

### 1.5.3 Phonological status of unmarked handshapes

There are four phonological phenomena in which researchers have suggested that a set of unmarked handshapes plays a special role; these are listed in (1-14), and each is discussed in this section.

(1-14) Special phonological status of unmarked handshapes

| | |
|---|---|
| H2 in Type 3 signs | Unmarked handshapes are allowed on H2 in Type 3 signs; marked are not. |
| Handshape change | Unmarked handshapes allow handshape change; marked do not. |
| Classifiers | Unmarked handshapes are used as classifiers; marked are not. |
| Contact | Unmarked handshapes are less restricted in how they contact a location. |

One proposed property of unmarked segments is that they are distributed more widely than marked segments. As discussed in section 1.3.2, Battison (1978) noted that in Type 3 signs, that is, signs in which the hands bear different handshapes, the nondominant hand is more restricted in its choice of handshapes. According to Battison, this set has the seven elements {3 2 29 3≅B ; 49 78 }, often called BASCO15 for the letter of the ASL fingerspelling alphabet or the ASL numeral that the handshape represents. The set of handshapes used on the nondominant hand in Type 3 signs in the ASL database were found to be {3 39 2 29 3≅B ; 49 48 78}, which is quite similar to BASCO15, since Battison identifies B as 39, but includes 3 as alternate form. The single difference is the inclusion of 48, which is used in one Type 3 sign in the ASL database, THEN.

There are two potential problems with this method of identifying unmarked handshapes. The first problem is that the set of handshapes allowed on the nondominant hand in Type 3 signs varies widely cross-linguistically, as seen in (1-15) for the languages in this study. Some of these handshapes, such as 48□X, 49□or <, are considered marked by other criteria. This variability is problematic because the set of unmarked handshapes is supposed to be universal.

(1-15) Handshapes used on the nondominant hand in Type 3 signs

| ASL | KSL | NZSL | SVK |
|---|---|---|---|
| 3 | 3 | 3 | 3 |
| 2 | 2 | 2 | 2 |
| 29 | 29 | 29 | 29 |
| 49 | 49 | 49 | 49 |
| 78 | 78 | 78 | |
| 3≅B | 3≅B | 3≅B | |
| | 38 | 38 | 38[26] |
| ; | | ; | |
| | : | : | |
| 28 | 28 | 28 | |
| | 3B | 3B | |
| 39 | | | |
| 48 | | | |
| | < | | |
| | 48□X | | |
| | 49□ | | |
| | 59 | | |
| | | 3B8 | |
| | | 69 | |
| | | 3≅A | |
| | | | 3≅ |

The second problem is that in Type 3 signs, the nondominant hand plays a very different phonological role than the dominant hand. The dominant hand is an active articulator; it moves, it contacts a location, it can produce a handshape change, and it can assume any handshape in the inventory. In contrast, the nondominant hand is considered to be the location at which the sign is

---

[26] The handshape 38 is not used as a handshape on the dominant hand in any sign in the SVK database. It is used on the nondominant hand in a Type 3 sign only once, in the sign LAPASET *mittens*, in which the dominant 49 hand outlines a mitten on the nondominant 38 hand. In contrast, 3≅, which is also used only once on the nondominant hand in a Type 3 sign, is used on the dominant hand in VANHA *old*, RUOTSI *Sweden* and ORANSSI *orange*, which feature handshape change, and LUOKKA *class*, which does not.

articulated. It cannot move, contact a location, or produce a handshape change; nor can it assume any handshape in the inventory. The restricted phonological role of the nondominant hand is discussed more fully in section 2.1.2.2.3. In light of the very different role of the nondominant hand in Type 3 signs, it is not obvious that restrictions on this location are relevant to restrictions on the active articulator. Nevertheless, this markedness characteristic is one of the easier criteria to apply because it is actually possible to establish the set of handshapes used on the nondominant hand in Type 3 signs in a well-defined manner.

In some signs, the handshape is dynamic; it changes during the articulation of the sign. This change in handshape is suggested to be a diagnostic of markedness (Battison, 1978). In the databases created for this study, the three types of signs involving handshape change listed in (1-16) were encountered.[27]

(1-16) Types of handshape change

| **type** | **description** | **example** |
|---|---|---|
| handshape-internal change | Slight bending and wiggling of fingers, possibly repeated; intermediate handshapes are not distinctive. | ASL: SHRIMP 49□<br>SVK: RANTA *beach* 78 |
| contrastive | The handshape changes from one contrastive handshape to another contrastive handshape. | ASL: JOB 49□η 39<br>SVK: TAKSI *taxi*<br>4Aη59□□□ |
| contour | A basic underlying handshape changes from one position to another, such as open to closed or straight to curved; usually the exact handshape of one terminus is not important. | KSL: GOONGUMHADA *curious*<br><□ η 7≅<br>NZSL: SENSITIVE<br>7≅□B□B η <□ |

[27] Changes in handshape have been extensively studied (Friedman, 1977; Klima and Bellugi, 1979; Wilbur, 1987; Liddell and Johnson, 1989; Sandler, 1989; Perlmutter, 1992; Corina, 1993; Brentari, 1998). Because there are a great many different terms for different types of handshape change, it is necessary to read carefully to determine exactly what is meant by each author.

Both marked and unmarked handshapes participate in handshape internal change, as shown by the ASL examples of SHRIMP 49□, WEIRD 79□X, WHY 78□A□A□A, and NOODLE 59.

In handshape contours, the same basic handshape is used for both beginning and ending handshapes, but there is a simple feature change for the selected finger or fingers, from open to closed or from straight to bent, or vice versa (Sandler, 1989). According to Brentari (1998), unmarked handshapes are more likely to participate in handshape contours.[28] There are examples in all four databases of handshape contours involving handshapes that are typically considered unmarked, such as the KSL sign JIDA *fall* 78 η 2. However, there are plenty that involve putatively marked handshapes. KSL has quite a few, such as GAT DA *identical* 4≅A η :A, GOONGUMHADA *curious* <□ η 7≅, and SSAL *rice* <□B□B η 78. Thus, it is not obvious in the present study that unmarked handshapes, which are supposed to be generally more frequent in the lexicon, are more frequent in handshape contours.

Distinguishing handshape contour from handshape contrast signs can be difficult. For example, is it necessary for the ASL sign HATE, which begins <□, to end 78, or can it end 78□A? Signs that have definite beginning and ending handshapes are often borrowings from the fingerspelled representation of the corresponding word in the spoken language, such as the SVK sign TAKSI *taxi* 4A η 59□□□, where 4A is fingerspelled “T” and 59□□□ is “X” in the old Finnish fingerspelling alphabet. These signs, of course, use highly marked handshapes, some of which only appear in fingerspelled loan signs such as these.

It was also suggested (Brentari, 1990b) that classifier forms are more likely to be unmarked handshapes. However, in ASL several common classifiers and Size and Shape Specifiers (SASS; see K&B) have handshapes that are commonly considered marked, such as the form for vehicles 68, for round flat objects <, for long thin objects 49□, and for human legs 69.

[28] Brentari (1998: 159) limits handshape contours to “open or closed variants of the same underlying handshape.”

Contact is also a proposed diagnostic of markedness. Battison (1978) claims that the handshapes BASCO15 “have greater variety in how they may contact the body or the other hands in order to form signs; the more marked handshapes have greatly restricted points of contact.” While this suggestion is interesting, a quantitative study is required to verify it. Compare, for example, the putatively unmarked and marked handshapes, 49 and 59□□□. In ASL, 49 can contact a location at the fingertip of the index finger, as in THINK, the backs of the middle, ring and pinky fingers, as in WEEK, the ulnar side of the index, as in MOUSE, the ulnar side of the hand, as in MINUTE, the back of the index, as in DRY, the back of the hand, as in PUZZLED, the non-ulnar side of the index, as in MONTH, the distal phalange of the index, as in BEGIN, and the whole index, as in APPEAR. In contrast, 59□□□ can contact a location on the ulnar side of the hand, as in ROCKET, the ulnar side of the middle finger, as in RAT, the back of the hand at the metacarpophalangeal joint of the index finger, as in CIGAR, and the tips of the index and middle finger, as in RELIGION. In this informal survey, 49 has eight different contact places, while 59□□□ has only five. However, a survey like this is not an accurate way to determine whether this difference is significant. As will be seen in Chapter 3, 49 is a very common handshape, while 59□□□ is a very rare handshape. A statistical analysis is necessary to determine whether the difference in number of places of contact is not just due to the fact that 49 occurs more often. On the other hand, perhaps a statistical analysis would show that a candidate set of unmarked handshapes really do have more places of contact than another set; however, such an analysis will not be undertaken in this dissertation.

### 1.5.4 Frequency of unmarked handshapes

The frequency with which a handshape is used both within a language as well as cross-linguistically is thought to be another diagnostic of markedness. Woodward (1982; 1985; 1987) produced a series of articles examining type frequency of certain handshapes in nine sign languages. He hypothesizes that the stages of handshape acquisition proposed by Boyes Braem (1973; 1990) correspond to markedness; thus, these stages should also correspond to cross-linguistic frequency of occurrence. In separate analyses, he investigated the frequency of certain classes of handshapes: single finger extension handshapes, single finger contact handshapes, two finger extension handshapes, and stage three handshapes, which are handshapes from Boyes Braem's third acquisitional stage, which he labels D, H, V, I, Y, K, 3, W, following Stokoe's (1960) notation.[29] He found that single finger extension handshapes are more frequent than single finger contact handshapes; they are also more frequent than two finger extension handshapes or stage three handshapes. Within one of these handshape groupings, such as single finger extension, handshapes using the index finger are more common than handshapes using the pinky, which are more common than handshapes using the middle finger, which are more common than those using the ring finger. Furthermore, he found that among stage three handshapes, the following frequency ordering holds: V > H > Y > I > 3 > W > K > D. Ordering of shared handshapes across the four languages considered in this dissertation is discussed in Chapter 3.

### 1.6 Quantitative linguistic aspects of the lexicon

Beyond determination of an inventory (whether this inventory is of features, segments, syllables, morphemes, words or meanings) is inquiry into how often these elements appear. Quantitative linguistics looks at not just what elements exist, but also the frequency with which they occur. Têšitelová (1992:11) states, "For deeper understanding of an object or phenomenon and thus also language it is necessary to know not only its qualitative but also its quantitative side." Frisch et al. (1997) states, "The statistical patterns of language are systematic linguistic data which must be accounted for in linguistic theory." Furthermore, Goldsmith (1998) observes, "A study of frequency can often be tantamount to a search for lurking generalizations."

While it might be argued that quantitative lexical patterns might not be part of the grammar of native speakers, a study by Frisch and Zawaydeh (2001) confirmed the psychological reality of an abstract, gradient consonant phonotactic constraint in Jordanian Arabic that had been uncovered originally through statistical analysis of dictionary corpora (McCarthy, 1994; Frisch et al., 1997). The phonological constraint against adjacent consonants with homorganic places of articulation in Arabic roots is well known. However, this constraint is not categorical; it is a gradient constraint whose relative lexical frequency correlates with its acceptability, as

---

[29] It is not clear in these articles exactly which handshapes are included in the frequency counts. For example, the handshape in which the index finger is extended could refer to simply 49, or it could also refer to 4, 48, 4≅, 4A, 48A, 49A, 4A≅, 4B, 48B, 49B, 4≅B. This illustrates the problem with using Stokoe notation to refer to handshapes. Even if it were indisputable which handshapes are contrastive in ASL, it is likely that other languages have other contrasts.

demonstrated by the native speakers' gradient judgments of novel verbs containing gradient violations.

Statistical properties of the lexicon are sometimes used to support phonological theories. For example, in a gated sign recognition experiment (Emmorey and Corina, 1990), signers could identify signs after only 35% of the sign had appeared, about 240 milliseconds, as compared to 85% of a spoken English word, about 330 milliseconds. Emmorey (2002) suggests that this quick recognition occurs because ASL has few signs that share a given initial handshape and location, the parameters that require no temporal span for their expression. In another example, Brentari (1998) concludes, based on a variety of evidence, that all two-handed signs are more complex prosodically than one-handed signs. One supporting argument is the claim that signs articulated with only one hand, Type 0/X signs, bear more marked handshapes than signs articulated with two hands, Type 1 signs. Another example is Siple (1973; 1978; 1980), which examines the effect of visual constraints on sign language, in particular, the fact that visual acuity is highest at the focal point and lower on the periphery. It is claimed that since Deaf sign perceivers focus on the face of the signer, more marked handshapes ought to be used at locations on the face and more unmarked handshapes on the torso. These are thought-provoking theories, but the absence of actual statistical confirmation weakens them.[30]

In the quantitative linguistics of spoken language, various units of population have been analyzed. Phonological statistics for spoken language uses distinctive features (Krámský, 1976a), phonemes (Hayden, 1950; Denes, 1963; Roberts, 1965; Kucera and Monroe, 1968), syllables (Kucera and Monroe, 1968), or higher level units of phonological analysis, such as stress and intonation (Adams, 1969).[31] Trubetzkoy (1967 [1939]) discusses phoneme frequency, as well as the distinction between type, the occurrence of an element in an inventory, such as a word in a lexicon or a segment in an inventory, and tokens, the occurrence of an element in a corpus of speech or text.

Among the earliest and most prominent studies is the work by Zipf (1935; 1949) on word frequency distribution, in which Zipf discovered an extraordinarily robust empirical law that

---

[30] Siple (1973; 1978; 1980) simply states, "A look at the Dictionary of ASL (Stokoe et al., 1965) will show that these predictions are confirmed." See section 5.2.5 for further discussion.

[31] Most of these are references to quantitative linguistic research on English. For a comprehensive overview of publications in quantitative linguistic research in many languages, see Tε(šitelová (1992).

remains controversial to this day.[32] The frequency of occurrence of words in a text is calculated, and these frequencies are ranked from largest to smallest. Zipf discovered that there exists a relationship between the rank of a word, *r*, and the frequency of this word, *f*, where $f = C/r^{a}$, where *C* and *a* are constants. In the ideal case, $a = 1$, and the frequency is inversely proportional to the rank. Equations of this form are called *power laws*. They have the property that when the logarithm of *r* is plotted against the logarithm of *f*, the graph is a straight line. This is to say that $\log(f)$ is a linear equation in $\log(r)$, with slope $-a$ and intercept $\log(C)$, since $f = C / r^{a}$ implies that $\log(f) = \log(C / r^{a})$, which implies that $\log(f) = \log(C) - a\log(r)$. This rank-frequency distribution, called *Zipf's law*, is robust cross-linguistically, and it appears in other domains, such as physics, biology, and demography.

Miller, in his 1965 introduction to Zipf (1935) observed that, "Faced with this massive statistical regularity, you have two alternatives. Either you can assume that it reflects some universal property of human mind, or you can assume that it represents some necessary consequence of the laws of probabilities." Zipf believed the former was true, and he explained it by his Principle of Least Effort. He claimed that people act so as to minimize their average rate of work. In the case of language, Zipf argues that the conflicting demands of the speaker and the hearer, both of whom wish to minimize their efforts, produce a compromise solution whose form is a version of Zipf's law. (The resolution of competing demands lies at the heart of many current linguistic enterprises, such as functional linguistics and Optimality Theory.) Others, such as Mandelbrot (1953; 1977), Herdan (1966) and Li (1992) have criticized Zipf's principle as being unformalizable and have explained Zipf's law as the inevitable statistical result of a non-interacting system of independent symbols, analogous to the ideal gas model. Other researchers, such as Günther et al. (1996), acknowledge this possibility, but show that interdependent and interacting components also lead to phenomena that exhibit Zipf's law; the Principle of Least Effort entails interacting components. Still others, such as Bak (1988; 1996), view the uniform power law behavior of these disparate events as manifestation of a process of self-organization that evolves over time, and use such examples in the construction of a theory of complex, self-organizing systems, a theory that will be revisited in section 6.2.4.[33]

---

32 This law was actually first found by Pareto (1897) in economics.

33 The controversy does not end here. See Horgan (1995) for a scathing review of complexity.

Zipf's law is relevant to this dissertation because in Chapters 3, 4 and 5, handshape, location, and handshape by location duet frequencies are computed, ranked and graphed. All of the languages surveyed show remarkably similar graphs; however, not all are power laws; some are exponential. I do not resolve the controversy surrounding Zipf's law in this dissertation, but I add to the debate. [34]

### 1.7 Previous quantitative studies of sign language

This section presents an overview of quantitative studies on signed languages in order to situate the investigative range of this dissertation relative to previous work. There have been two types of quantitative studies in sign language. One type is a cross-linguistic vocabulary comparison, usually for the purpose of studying historical relations or more generally, lexical similarity between sign languages. Each of these studies uses a list of 100 to 300 signs, similar to the Swadesh (1955) word list developed for spoken language vocabulary comparison. Using methods from glottochronology, Woodward (1978) compares French Sign Language (FSL) and ASL. Since Old FSL was introduced into America in the early 19$^{th}$ century, and modern ASL developed from Old FSL, it would be expected that FSL and ASL would have a high percentage of cognate signs. Yet Woodward found that only 60% of the signs are cognates, a very low percentage, as 90% would be the expected value for a such fairly recent divergence in spoken languages. Woodward attributes this greater-than-expected difference to the presence of one or more sign languages already existing in America, which blended with FSL.[35] Other historical-comparative linguistic studies by Woodward compare language varieties in Costa Rica, in India, Pakistan and Nepal, in Thailand, and in Thailand and Viet Nam (Woodward, 1991; 1993; 1996; 2000) in order to determine the structure of sign language families.

Kyle and Woll (1985) reports on a vocabulary comparison of fifteen different sign languages. Languages are paired, and signs for a set of concepts are compared. For each language pair, 35% to 40% of the tested concepts have signs which are articulated similarly.

---

[34] Interestingly, Zipf (1935) also theorized about phoneme frequency. He proposed that less complicated phonemes occur more frequently, the theory Ann (1993) tested for unmarked handshapes, as discussed in section 1.5.2. Unfortunately, Zipf did not have a metric for assessing phoneme complexity (Trubetzkoy, 1967 [1939]).

[35] See section 2.1.1 for histories of ASL, KSL and NZSL.

Guerra Currie et al. (2002) compares Mexican Sign Language (LSM) with French Sign Language (FSL; which is related historically to LSM), Spanish Sign Language (LSE; Mexico shares a spoken language and a similar culture with Spain, and LSE is related to FSL) and Japanese Sign Language (Nihon Syuwa: NS; which is unrelated culturally and linguistically). For LSM and FSL, 38% of the tested concepts have similarly articulated signs; for LSM and LSE, 33% have similar signs, and for LSM and NS, 23% have similar signs. In a similar study, Sasaki (2001) compares the vocabulary of NS with KSL and Taiwan Sign Language, three historically related languages, to determine the degree of influence of NS on KSL and Taiwan sign language, was well as to study how signs have changed since the Japanese colonization of Taiwan and Korea.

McKee and Kennedy (2000) attempts to establish how closely related NZSL is to ASL, Australian Sign Language (Auslan), and British Sign Language (BSL). For the three historically related languages, NZSL, Auslan, and BSL, between 79% and 87% of the tested concepts have signs with identical or similar articulation. When ASL, a historically unrelated language, is included, this figure drops to 26% to 32%. All of these similarity percentages are far greater than would be expected in spoken language comparisons. Greenberg (1957:37) states, “where the percentage of [lexical] resemblance between languages is very high, say 20 percent or more, some historic factor, whether borrowing or genetic relationship, must be assumed.” Yet, even sign languages that are historically unrelated and used in geographically distant, culturally diverse countries have similarity figures greater than 20%. Explanations for the cross-linguistic similarity of sign languages include iconic potential, that is, direct expression of a visual aspect of a referent by the physical form of a sign, the creole-like nature of sign language transmission, and modality constraints on the form of signs and phonetic resources.

Another type of quantitative lexical analysis, more similar to the work undertaken in this dissertation, examines phonological resources in the lexicon. These phonological resources are compared cross-linguistically; in contrast, in the analyses discussed above, the forms of signs with equivalent meanings are compared. The productive ease study of Ann (1991) discussed in section 1.5.2 is such a study. Recall that she devised a method for assigning ease ratings to handshapes based on the physiology of the hand, and then counted the number of handshapes of each type in ASL and Taiwan Sign Language to determine whether easy handshapes occur more frequently than hard handshapes in these two languages.

The frequency studies of Woodward (1982; 1985; 1987) discussed in section 1.5.4 are another example of this second type of cross-linguistic quantitative analysis. Recall that he found that single finger extension handshapes are more frequent than single finger contact handshapes, two finger extension handshapes or stage three handshapes. Also, handshapes using the index finger are more common than handshapes using the pinky, which are more common than handshapes using the middle finger, which are more common than those using the ring finger. Among stage three handshapes, the following frequency ordering holds: V > H > Y > I > 3 > W > K > D. Since the focus of these articles is relative markedness as expressed by frequency orderings, and since this question is divided by handshape type, complete inventories and orderings are not put forth, so the kind of whole-inventory comparison done in this dissertation is not possible.

Woodward (1982; 1985; 1987) also examines how handshapes combine with locations, a topic considered in Chapter 5. He divides locations into four categories: hand or neutral space, face, trunk and arm. He claims that if a handshape, such as an extended index finger, can appear on the trunk, it can also appear on the arm; if a handshape can appear on the face, it can also appear on the trunk and arm; if a handshape can appear on the hand or neutral space, it can also appear on the face, trunk and arm. In two of these papers, he merely shows what an ideal implicational hierarchy looks like and comments that the data fit this pattern. Thus, this research, though intriguing, is limited by the lack of quantitative analysis necessary to support the claims.

As will be seen in Chapter 4, vastly more signs occur on the hand and in neutral space than on the arm, so the fact that a handshape does not occur at a particular location might be due to the rarity of that location, not due to a proscription on co-occurrence. Indeed, this hypothesis as stated is not verifiable statistically. For all four languages, the number of signs occurring at each location in Woodward's proposed hierarchy decreases according to this hierarchy, as shown in (1-17). The alternative hypothesis to Woodward's proposed distribution is independence: the handshapes are distributed across these four location groups without regard for the particular group. Because of the decreasing numbers of signs at each location, the independent distribution itself satisfies Woodward's hierarchy. The question of the distribution and dependence of handshape and location will be revisited in Chapter 5.

(1-17) Number of signs at each location group in Woodward's hierarchy

| | ASL | KSL | NZSL | SVK |
|---|---|---|---|---|
| hand and $\pi$ | 286 | 271 | 346 | 316 |
| face | 77 | 84 | 64 | 104 |
| torso | 32 | 26 | 44 | 55 |
| arm | 10 | 3 | 7 | 7 |

To my knowledge, the only other quantitative linguistic work on a signed language comparable to this dissertation is Pietrandrea (1998; 2000) on Italian Sign Language (Lingua Italiana dei Segni: LIS) in which LIS handshape and location inventories and their frequencies are determined. Pietrandrea (2000) presents a rank-frequency distribution graph for LIS handshape. Her results for handshape and location inventory and distribution are discussed in sections 3.1.6, 3.2.2.7, 4.2.2.6 and 6.2.2. Furthermore, she does statistical analysis on the distribution of handshape in one-handed versus two-handed signs as well on the frequency of one-handed versus two-handed signs articulated in neutral space versus with body contact, although it is not possible to directly compare her results to the results in this dissertation due to differences in intent and methodology.

# Chapter 2: Methodology

This chapter discusses the methodology used to investigate the central questions of this dissertation: What are the handshape and location inventories? Do they vary cross-linguistically? How are these parameters distributed in the lexicon? Do these distributions vary cross-linguistically? How do the parameters of handshape and location combine into duets? First, the construction of the database is detailed, including language selection, dictionary descriptions, software used, and database structure. Second, the analytical methods that were used are explained, including information theory and the mutual information significance program.

## 2.1 Databases

In order to address questions about the inventory and distribution of phonological parameters I constructed databases of four sign languages. Each database contains approximately 600 randomly chosen signs transcribed from drawings or photographs from a published dictionary. As sign language dictionaries are typically smaller than spoken language dictionaries, 600 signs represents 17% to 50% of the dictionary entries.[36]

### 2.1.1 Languages and dictionaries

The languages were chosen to be historically unrelated and geographically diverse so that a degree of typological spread is attained. My familiarity with ASL and SVK led me to select these two languages. KSL and NZSL were selected because in addition to forming an independent set, they have large, inexpensive, easily obtainable dictionaries. The following languages were also inspected, and are included for comparison sake: Old Finnish Sign Language (Vanha Suomalaisen Viittomakieli: VSVK), Sign Language of the Netherlands (Nederlandse Gebarentaal: NGT), and Italian Sign Language (Lingua Italiana dei Segni: LIS). These languages are not historically independent or geographically diverse. VSVK is of course related to SVK; NGT and LIS are members of the French Sign Language family, of which ASL is also a member.

Types (occurrences in the lexicon) rather than tokens (occurrences in use) are counted. Sign language dictionaries are readily available, while sign language corpora are not. In addition,

[36] This sample size was chosen as a compromise between statistical validity and work load feasibility.

it would be necessary for the corpora to be transcribed and translated because I do not know all four sign languages. Even if transcribed corpora were available, comparison would be difficult because of the great variety of transcription systems. It is possible that frequency in conversation might differ from frequency in the lexicon; compare, for example, the low type frequency of the voiced interdental fricative in English to its high token frequency due to its occurrence in the word *the*. In spite of such exceptions, token and type frequency probably do not diverge greatly. Herdan (1966:58) states, “If the Saussurian axiom of the independence of sound and meaning is true, there should be no significant difference between the frequency distribution of phonemes (letters) in samples from the dictionary and from running text. As I have shown [(Herdan, 1960)], this is actually so.” However, the question of type versus token frequency for sign languages is open.

#### 2.1.1.1 American Sign Language

ASL was chosen as one of the languages because of my knowledge of this language and because of the large amount of research that has already been conducted on it. ASL is the primary means of communication of approximately one-half million Deaf people in the United States and parts of Canada. ASL is related to a family of sign languages that have descended from Old French Sign Language. Little is known about the language or communication system of American deaf people in the eighteenth century.[37] In 1817, Thomas Hopkins Gallaudet and Laurent Clerc founded the first American school for the deaf in Hartford, Connecticut. Clerc, a Deaf Frenchman, was a student of Sicard, the director of the Royal Institute of the Deaf in Paris, the school founded by Abbé de l’Epée that used French Sign Language (FSL) as its means of instruction. Clerc introduced his sign language to the students attracted to the new American school. Thus, the sign language of early nineteenth century France combined with whatever sign languages and communication systems existed in America at that time to create what became American Sign Language. Students of this school, as well as hearing people trained as teachers, spread throughout North America, carrying ASL with them. Other European countries also established schools for the deaf following the French model, and often employing Deaf French teachers. In this way, FSL was introduced into many countries, combining with the extant indigenous sign languages or communication systems to form new sign languages. Over time,

[37] See Groce (1985) for a interesting account of the Deaf people and their language on Martha’s Vineyard from the eighteenth to twentieth centuries.

these new sign languages diverged, becoming mutually unintelligible in many cases, but they are historically related languages (Lane, 1984).

Since a great deal of ASL sign language research has consulted Stokoe's (1965) dictionary (DASL), it would have been practical to base the ASL database on it. However, DASL does not contain drawings or photographs of signs. Instead each sign is transcribed using a notational system Stokoe created for ASL after analyzing the language and proposing phonemic classifications. This classification might not be appropriate for distinguishing handshapes and locations relevant to other languages. Also, because I was transcribing languages other than ASL, I wanted to be sure that the method of database construction was as uniform as possible for all four languages. I chose to use *The American Sign Language Handshape Dictionary* (Tennant and Brown, 1998) because it is a recent dictionary containing more than 1,600 signs illustrated with large, clear drawings. The dictionary is organized first by hand arrangement, then by handshape. In addition to a drawing and English gloss for each entry, information on handshape, orientation, location, movement and non-manual signal is included. Although I based my transcription on the drawing, it was useful to have this extra information to consult as necessary. There are approximately five entries per page and 330 pages, so I transcribed the first and last sign on each page for a total of 656 signs in the ASL database, representing about 41% of the dictionary.

There is an old dictionary of ASL (Long, 1918). It would have been interesting to be able to compare modern ASL with old ASL, especially as I created a database based on an Old Finnish Sign Language dictionary (Hirn, 1910), as discussed in section 2.1.1.4. Unfortunately, the Long dictionary does not contain any drawings or photographs; instead, for each entry, a verbal description of the sign is given. These descriptions are not sufficient for me to produce a transcription of each sign; in addition, this dictionary would introduce a large difference in the method of database construction.

#### 2.1.1.2 Korean Sign Language

Korean Sign Language was chosen as an Asian sign language representative. KSL has been used in Korea at least since 1889 (Grimes, 1996). It is related to Japanese Sign Language (Nihon Syuwa: NS) and Taiwan Sign Language, but it is distinct from them. The relation between KSL and NS is similar to the relation between ASL and French Sign Language. The Japanese provided education for the deaf relatively early, establishing a school for the deaf in Kyoto in 1878 (Grimes, 1996). They established a deaf school in Korea in 1913 (Hansook Lee, p.c.) and in Taiwan in 1915 (Ann, 1998), introducing NS to these countries during the colonial period preceding World War II. Since this time KSL has developed independently of NS (Miyamoto, 2000). In 1947, Beak-Won Yoon invented signs for the Korean alphabet, Hangul, and in 1963, the first sign language textbook was published by the Seoul School for the Deaf (Hansook Lee, p.c.). In 1982, the first dictionary of KSL, the *Standard Sign Language Dictionary*, by Seung-Gook Kim, with approximately 2,200 entries, was published.[38]

I used the 1993 edition of this dictionary. Each sign is illustrated with a clear drawing. The signs are arranged alphabetically according to the Korean gloss of the sign. A native speaker of Korean who was a graduate student at the University of Washington translated the Korean gloss for each sign into English, and transliterated it into the roman alphabet. There are three or four entries per page, so I transcribed the first sign on each page, for a total of 614 signs, representing about 21% of the dictionary.

#### 2.1.1.3 New Zealand Sign Language

New Zealand Sign Language was chosen to be one of the languages included because it is from a different geographical area than the other three and because the spoken language is English, so that the text of the dictionary was accessible to me. Despite the fact that the United States and New Zealand have the same spoken language, their sign languages are historically unrelated. NZSL is related to British Sign Language (BSL), which is unrelated to French Sign Language. At the time that the United States and many European countries were following the example of the French by using sign language, in particular, FSL, for Deaf education, the English did not advocate the use of sign language. They attempted to educate by oral methods, that is, teaching Deaf people to speak and lip read. Thus, FSL did not influence BSL the way it did many other sign languages. Later, in the years between 1825 and 1880, sign language was accepted as a means of educating the Deaf, and it was during this time that many British people emigrated to New Zealand, where they established schools that used BSL as the language of instruction (Collins-Ahlgren, 1990).

---

[38] Other sources claim the dictionary has 5,700 entries (Lee, p.c.) or 6,000 entries (Kim et al., 1996; Lee et al., 1997). The count of 2,200 does not include numerals or Hangul. *Korean Sign Language For the Guide* (Son, 1988) is another book with an extensive glossary that is already translated into English, but it appears to be intended as a language learning text rather than a general-purpose dictionary.

I used *A Dictionary of New Zealand Sign Language* (Kennedy, 1999) as the basis of the NZSL database. This is a very large dictionary of approximately 4,000 signs. The signs are arranged according to handshape. There were six entries per page and 687 pages of signs. I transcribed the last entry on each page, for a total of 688 signs, representing 17% of the dictionary. There is a great deal of formational information included for each entry: a drawing, a verbal description of the sign's form, and a HamNoSys transcription of the sign.

#### 2.1.1.4 Finnish Sign Language

Finnish Sign Language is the primary language of about 5,000 people in Finland. The first school for the deaf was formed in the 1850's by C. O. Malm, who was deaf. He had studied at a deaf school in Sweden, and he used Swedish Sign Language at the school he founded. The Swedish Sign Language Malm introduced changed quickly as it merged with the existing sign language dialects (Rissanen, 1987). Today, SVK is related to but unintelligible with Swedish Sign Language, as well as Portuguese Sign Language, which was similarly influenced by Swedish Sign Language. Swedish Sign Language is unrelated to the FSL family, and does not share its origins with any other sign language; hence, SVK is unrelated to ASL (Grimes, 1996).

The dictionary that forms the basis of the SVK database is *Suomalaisen viittomakielen perussanakirja* (Kuurojen Liitto, 1998). Each entry in this dictionary is illustrated with one to three clear photographs. The signs are arranged according to handshape. I translated the Finnish into English. Instead of giving just a single gloss for a sign, all of the possible translations of the sign into Finnish are given, along with Finnish sentences illustrating the various meanings of the sign. Where appropriate, information on sign variants is included. There are 1219 entries in the dictionary, numbered 1 through 1219. I transcribed the even-numbered entries for a total of 609 signs, representing 50% of the dictionary.

There is an excellent dictionary of Old Finnish Sign Language (Hirn, 1910). It could not be included as one of the primary languages because it is related to Finnish Sign Language, and because the dictionary is much smaller than the others, containing only 355 signs. The entries are illustrated with photographs of Hirn and glossed in Finnish and Swedish. No other linguistic information is provided. In addition, the photographs are not always clear enough to discern the handshape. However, I created a database of this dictionary, transcribing all 355 signs, and included analysis of this database for the sake of comparison whenever possible.

#### 2.1.1.5 Other databases

In addition to the five databases I created for this dissertation, I was fortunate to have the opportunity to examine the SignPhon database of Sign Language of the Netherlands (Nederlandse Gebarentaal: NGT) created by researchers at Leiden University (Crasborn, 1998; Crasborn et al., 1998). This is a very large database with more than 3,000 entries. For each entry there is a narrow phonetic description of the manual articulation of the sign, details about the signer, iconicity, morphology, and so on.[39] NGT was not eligible to be one of the investigated languages because it is related to ASL. Also, the transcription system differed greatly from mine. As will be seen in Chapter 3, the four sign languages I examined have from 34 to 49 handshapes according my transcription; SignPhon recognizes 112 different handshapes in a sample of 3305 signs. This transcription is probably overly narrow. Nevertheless, NGT was analyzed and compared to the other languages whenever possible.

In addition to the NGT data, I am also able to incorporate data from Italian Sign Language (Lingua Italiana dei Segni: LIS) gleaned from Pietrandrea (1998). This paper includes a quantitative study of handshape and location that is similar though more limited than the study done in this dissertation. The NGT and LIS data are significant because they exhibit the same behavior as the ASL, KSL, NZSL and SVK data, demonstrating that this behavior is not an artifact of the transcription method.

### 2.1.2 Structure of the databases

#### 2.1.2.1 Software

The database program used for this dissertation was FileMaker Pro 5. It was chosen because it was the only database program that allowed specification of different fonts for different fields. This feature is essential in order to specify an ordinary font, such as Times New Roman, for the gloss field, and a sign language font, HamNoSys, for the fields transcribing phonetic properties of the sign, such as handshape and location.

The fields included in the databases with their descriptions are shown in (2-1). There are additional fields not listed here that exist to accommodate the few compounds with more than two members, but these fields were rarely used. In (2-2) is a sample page from the SVK database.

---

[39] I have been told that each sign took about twenty minutes to enter.

(2-1) Database fields

| | |
|---:|---|
| # | A number that uniquely identifies each entry; often the page number |
| gloss | Gloss of the sign in the spoken language of the country |
| English | English translation of the gloss |
| type | Sign type: 0, 1, 2, 3, or combination for compounds |
| bor | Is the sign a borrowing from spoken language? |
| hs | Handshape (initial handshape if sign includes a handshape change) |
| hsch | Final handshape if sign includes a handshape change |
| contact | Part of the active hand that contacts the location or the other hand |
| loc | Location (initial location if sign includes a location change) |
| loch | Final location if sign includes a location change |
| or | Orientation (initial if sign includes an orientation change) |
| orch | Final orientation if sign includes an orientation change |
| hs2 | Second handshape in compound signs |
| hsch2 | Final handshape in compound signs |
| cont2 | Second contact in compound signs |
| loc2 | Second location in compound signs |
| locch2 | Final location in compound signs |
| or2 | Second orientation in compound signs |
| memo | Notes on the signs |

(2-2) Sample page of SVK database

| # | gloss | English | type | bor | hs | hsch | contact | loc | locch | or | orch | hs2 | hsch2 | cont2 | loc2 | lch2 | or2 | memo |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 2 | musta | black | 0 | | 3 | | □□□□□ □ | υ | | H – | | | | | | | | cover one eye |
| 4 | tietokone | computer | 0/1 | | 3 | | □□□□□ □ | σ | | H – | | 78 | | | | | H ¯ | min pair: hattu and tietää? |
| 6 | Ranska | France | 0 | | 3 | | | σ□ | | H ϕ | | | | | | | | also with 3 handshape |
| 8 | elokuva | film | 0 | | 3 | | | υ | | O ϕ | | | | | | | | |
| 10 | ei tiedä | don't know | 0 | | 3 | | □□□□□ | σ□ | | H – | | | | | | | | |
| 12 | unohtaa | forget | 0 | | 3 | 3B | □□□□□ | σ□ | | H – | | | | | | | | |
| 14 | kasvojen iho | skin of the face | 0 | | 3 | | □□□□□ □ | ξ | | H – | | | | | | | | |
| 16 | kiltti | gentle | 0 | | 3 | | □□□□□ □ | ξ | | H – | | | | | | | | min pair |
| 18 | tyyny | pillow | 0/1 | | 3 | | □ | ξ | | H – | | 49 | | | | | Θδ | pt 2: SASS |
| 20 | pyytää | request | 0 | | 3 | | □□□□□ | ξ | | H | | | | | | | | |

| | | | | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | | | □ | | | — | | | | | | | | |
| 22 | myöhemmin | later | 0 | | 3 | | □ | ξ | | H — | | | | | | | | |
| 24 | viime | previous | 0 | | 3 | | □ | | | H — | | | | | | | | |
| 26 | kotiin | go home | 0/0 | | 3 | | □□□□□ □ | ξ | | H — | | 3 | 3A | | | | Hδ | assimilated compound? |
| 28 | äsken | just | 0 | | 3 | 3A | □ | ξ | | H — | | | | | | | | |
| 30 | hyvä | good | 0 | | 3 | | □□□□□ | ψ | | H — | | | | | | | | |
| 32 | hyvä huomenta | good morning | 0/1 | | 3 | | □□□□□ | ψ | | H — | | 3 | | | | | Θ | — |
| 34 | homo | homosexual | 0 | | 3 | | | ψ | | H — | | | | | | | | |

Signs were labeled as borrowings if they incorporated elements of the written form of the sign as glossed into the spoken language of the country. For example, ASL has many initialized signs, in which the handshape corresponds to the fingerspelling of the first letter of the spoken English word that serves as a gloss for that sign (Battison, 1978). For example, the sign ROCKET, uses the handshape 59□□□, which is the handshape in the fingerspelling alphabet for R. SVK and NZSL also have borrowing of this type, even though the fingerspelled alphabet for NZSL is quite different from the one used in the United States and Finland.[40] In contrast, no borrowings from the fingerspelled Hangul alphabet occurred in the KSL database. However, KSL has borrowings from the written form of Chinese characters, discussed in section 3.1.2.

All of the handshape, location, orientation and contact fields in the databases were transcribed using the Hamburg Notation System (HamNoSys) (Prillwitz, 1987). This phonetic notational system is the one used most widely by linguists, particularly lexicographers. I chose it because it is used and known by others, because it is not language-particular, because it permits transcription of reasonably narrow phonetic detail, and because it has its own fairly transparent font.[41] I modified some of the symbols for my own use. For example, I notated the so-called "baby C" as 4≅B instead of = in order to clarify the selected fingers (see Appendix A).

#### 2.1.2.2 Construction and decisions

To create an entry, I examined the drawing or photograph and determined the handshape, location, orientation and contact. Because I wanted my transcription method to be uniform across all four languages, I did not rely upon the NZSL HamNoSys notation or on any other description included in an entry. On a few occasions, when I was unsure of a handshape or location because of an unclear picture, I referred to the written description if one existed. If I was still unsure, I put "?" in the field in question.

---

[40] There are minor differences between the fingerspelled alphabets of ASL and SVK, both being similar to the International Fingerspelling Alphabet. However, the old Swedish fingerspelling alphabet formerly was used in Finland, a version of which is still used in Sweden today. This alphabet is quite different from the current one. A few signs use handshapes borrowed from this old alphabet.

[41] Stokoe Notation System (Stokoe, 1960) is better known, but as it was developed for ASL it could miss handshapes and locations contrastive in other languages. Signwriting (Sutton, 2003) is another way to

In the case of the NZSL dictionary, I sometimes had a surfeit of information, in that there is a drawing, a written description, and HamNoSys notation, which do not always agree. When this occurred, I chose the parameter description agreed upon by two out of the three sources of information. For example, the NZSL entry FEW is shown in (2-3). The handshape in the drawing and the written description agree, but the HamNoSys notation that was given for this sign labeled the handshape as :□□ instead of = or =9 or 4≅B, which would have been more plausible.

(2-3) NZSL entry FEW

| **handshape shown** | **description** | **HamNoSys given** |
|---|---|---|
|  | The right fist is held up, palm left, blade forward, with the forefinger and thumb extended and the tips a little way apart | :□□ |

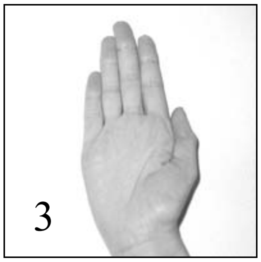
#### 2.1.2.2.1 Handshape

Certain handshapes required finer judgments. For example, HamNoSys does not have a way to distinguish the handshape intermediate between 3 and 38. Illustrated on top pages of one of the sections of the NZSL dictionary are the handshapes shown in (2-4 a), with the metacarpophalangeal and interphalangeal joints flexed so that the distal phalange of the thumb contacts the side of the hand, and (2-4 c), with the metacarpophalangeal and interphalangeal joints extended.

write signed languages. It is not language-particular, but as it was developed for general rather than linguistic use, it cannot transcribe sufficient phonetic detail for this purpose.

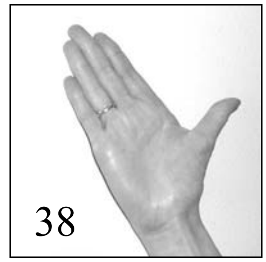
(2-4) Problematic handshapes

a.




b.

c.




The handshape in (2-4 b), however, has the metacarpophalangeal joint flexed and the interphalangeal joint extended, so that the distal phalange of the thumb does not contact the side of the hand. Furthermore, this intermediate handshape is quite variable, as the interphalangeal joint is often not fully extended. Since I distinguish only two handshapes, 3 and 38, I had to assign this intermediate handshape to one or the other symbol. I chose to assign it to the symbol 3; 38 is used only if the metacarpophalangeal joint is extended. In addition, in signs in which the thumb tip contacts a location, the handshape in (2-4c) is used exclusively; (2-4b) is never used. Also, when the ulnar edge of the hand contacts a location, the metacarpophalangeal joint is hyperflexed so the thumb is not in the way. In these signs also, the position of the distal phalange varies as the interphalangeal joint varies from straight to flexed. The HamNoSys notation in the NZSL dictionary apparently is inconsistent in its assignation of this intermediate handshape to one symbol or the other.

#### 2.1.2.2.2 Location

The identification and representation of body location was more straightforward. I decided to use not just a general body location but also a diacritic to indicate if the articulation was at the side of the location.[42] I am aware that such fine distinction in location might not be contrastive, and that location articulations can vary, but I preferred to begin with a narrower transcription and consolidate categories later if necessary. So the ASL sign RED has the location ψ, while the ASL sign JEALOUS has the location ψ□.

I did not distinguish different ways in which a location is contacted or moved to. Thus, signs with movement from a location (Sandler’s syllable type LM), movement to a location (ML),

---

[42] Stokoe (1965) only indicates the general location, such as the mouth.

glancing movement (MLM), and so on, were all said to have location L, as are signs with proximity to a location. I did not, for example, notate a sign such as KNOW in ASL as having neutral space as its first location and the forehead as its second. Even a sign with one location and a movement to a second location (LML) could be represented as having only one location, if both locations were within the same major body area, as in the ASL sign IMPROVE, which has the location □. Stokoe (1965) just specifies a location and has a symbol denoting "movement toward" or "movement away". Battison (1978:48) states, "No sign may be specified for more than two locations which themselves must be located in the same major body area. The only exceptions to this are compound signs or signs derived from compound signs."

#### 2.1.2.2.3 Type 2 and Type 3 signs

Type 2 and Type 3 signs were problematic for the creation of the database, and this problem is connected to location. Because the location parameter is inventoried and the frequency of use in the lexicon of each location is calculated, it is necessary to know how to categorize signs articulated on the nondominant hand. There are three issues. First, how should sign type be assigned? Second, should the location of Type 2 and Type 3 signs be considered neutral space or the nondominant hand? Third, should the nondominant hand location be considered one location or many different locations?

The nondominant hand does not function independently of the dominant hand in general, as noted for ASL by many researchers. In fact, there are stringent restrictions on the role of the nondominant hand in monomorphemic signs (Stokoe, 1960; Stokoe et al., 1965; Battison, 1974; Friedman, 1977). The interaction between the dominant and nondominant hands is circumscribed by the Symmetry and Dominance Conditions (Battison, 1974; 1978), given in (2-5).

(2-5) Symmetry and Dominance Conditions

Symmetry Condition: If the two hands move independently during the articulation of a sign, then they must have the same handshape, place of articulation, movement, and either the same or opposite orientation.

Dominance Condition: If the two hands have different handshapes, then only the dominant hand can move, and the handshape of the nondominant hand is restricted to a small set of handshapes.

To the Symmetry and Dominance Conditions, I have added another, called the Contact Condition (Rozelle, 1998). Battison alludes to this restriction when he notes, "Most of the signs in ASL that have non-identical hands require that the hands contact each other in a particular way during the sign, which requires close coordination of movement and timing." However, he does not formulate a third condition, nor does he extend it to include signs in which the two hands have identical handshapes but in which one hand does not move.[43] A simple, physically possible sign disallowed by the Contact Condition is shown in (2-7).

(2-6) Contact Condition

If one hand is moves and the other remains still (Type 2 or Type 3), there must be contact (or proximity) between the two hands at some time during the articulation of the sign.[44]

(2-7) A sign disallowed by the Contact Condition


It is important to note that these three conditions are phonotactic constraints, hence are phonological. They do not represent physiological limitations since they prohibit from the lexicon a large variety of gestures that are not only physically possible and articulatorily simple, but also actually occur under certain special morphological or syntactic conditions. The

43 Crasborn (1995) mentions contact, but he interprets Battison (1974) incorrectly (p. 25). In the Symmetry Condition, Battison says, "if both hands move independently … (as opposed to one or both being static) …." Crasborn takes this to mean "without contacting each other," which is wrong in light of the parenthetical remark. However, contact is crucial for the Dominance Condition.

44 Likewise, if both hands are moving and have different handshapes, there must be contact (or proximity) between the two hands continuously throughout the articulation of the sign. These signs are polymorphemic or are derived historically from polymorphemic forms.

Symmetry Condition disallows as a lexical item the gesture with 3 on the nondominant hand and 2 on the dominant hand with both hands moving downward simultaneously. The Dominance Condition disallows the gesture in which 49 on the dominant hand contacts < on the nondominant hand. Although these handshapes occur freely on the nondominant hand in both languages, such gestures combining configurations of the dominant and nondominant hands do not occur as monomorphemic signs in any sign language studied so far. The Contact Condition disallows even simpler gestures from the lexicon. For example, if both hands have the handshape <, allowable possibilities are for both hands to move downward without contacting each other (ASL DECIDE) or for the nondominant hand to remain still, while the dominant hand contacts the nondominant hand and then moves outward (ASL POSTPONE). However, the simple gesture in which both hands have the handshape <, where the nondominant hand is still, the dominant hand moves, moving downward, for example, as shown in (2-7), and there is no contact between them, does not occur as a monomorphemic lexeme.

The Symmetry and Dominance Conditions are incorporated in Battison's typology (1978), which was first discussed in section 1.3.2. Signs of Type 0 and Type X use only one hand. Type 0 signs are articulated in neutral space, while Type X signs contact the body. Type 1 signs use two hands; they can be articulated in neutral space or they can contact the body. In Type 1 signs, the hands have identical handshapes, symmetrical or opposite locations and orientations, and identical synchronous or alternating movements. Type 2 and Type 3 signs both involve two hands, but in these signs only the dominant hand moves, while the nondominant hand remains still. In Type 2 signs the two hands bear the same handshape. In Type 3 signs, the two hands can bear different handshapes, but the handshape of the nondominant hand is restricted to a small set of "unmarked" handshapes. In addition, Battison includes Type C for signs that are compounds of two or more of the other types. This typology is summarized in (2-8). (The HamNoSys notation for neutral space is π. The question of location of Type 2 and Type 3 signs will be addressed in section 2.1.2.2.3.)

(2-8) Typology of signs

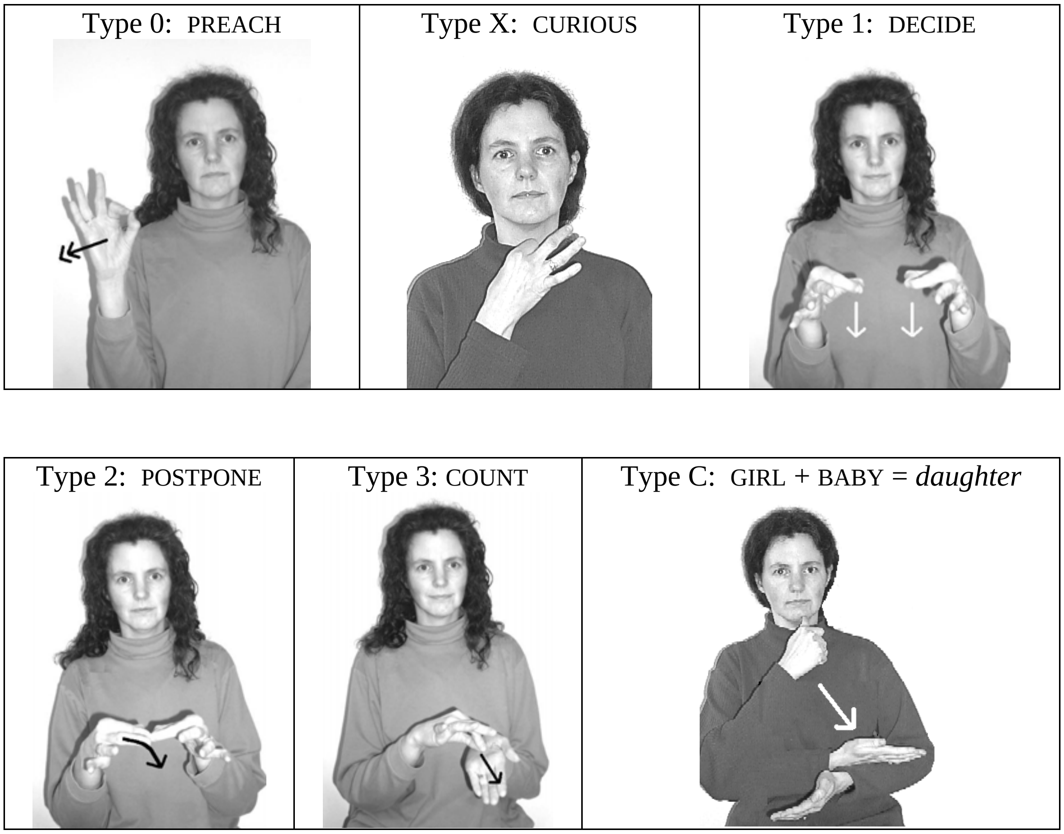
| Type | hands used | location | handshape | movement |
|---|---|---|---|---|
| **0** | H1 only | neutral space | unrestricted | unrestricted |
| **X** | H1 only | body | unrestricted | unrestricted |
| **1** | H1 and H2 | $\pi$ or body | H1 = H2 | H1 = H2 |
| **2** | H1 and H2 | ( $\pi$? H2? ) | H1 = H2 | H2 is still |
| **3** | H1 and H2 | ( $\pi$? H2? ) | H1 ≠ H2;<br>H2 restricted | H2 is still |

Examples of signs of each type from four languages are shown in (2-9); photographs are shown in (2-10). Type C is an additional category for compounds combining two or more of the other five types.

(2-9) Battison's typology: examples from each language

| Type | ASL | KSL | SVK | NZSL |
|---|---|---|---|---|
| **0** | PREACH | AMA *maybe* | HEVONEN *horse* | MUST |
| **X** | CURIOUS | GIDARIDA *wait for* | OLLA *be* | CRAVE |
| **1** | DECIDE | EUNHENG *bank* | HIENO *fine* | PROPER |
| **2** | POSTPONE | SUNMOOL *gift* | ARVATA *guess* | SEW |
| **3** | COUNT | UKJIRO *by force* | NIMI *name* | ORDER |
| **C** | GIRL + BABY<br>*daughter* | WHITE + PULL<br>*cotton* | TIETÄÄ + MACHINE<br>*computer* | FREE + VOTE<br>*democracy* |

(2-10) ASL: Examples of each sign type




Although this classification was created for ASL, it also holds for KSL, NZSL, SVK and VSVK. 98% of ASL signs fall into one of these five categories, as do 96% of KSL signs, 97% of NZSL signs, and 99% of SVK signs. Signs that do not fall into any category have numerous forms. Surveying all four databases, fifteen signs have two moving hands with two different handshapes, as in the ASL sign HELP and the SVK sign HISSI *elevator*.[45] Three signs have two simultaneous locations, as in the variant of the ASL sign SICK, which has the locations $\sigma$ and O. Two signs have a change in dominance, whereby the dominant hand acts upon the nondominant base hand, as in a Type 3 sign, and then the nondominant hand acts upon the dominant hand, as in the ASL sign JESUS. In eight signs, the nondominant hand supports the dominant hand, forearm or elbow, as in the ASL sign EVENING or the SVK sign LIPPU *flag*. In two signs, both hands move and have identical handshapes, but only one hand has a hand-internal movement, as in the NZSL sign HELP.

---

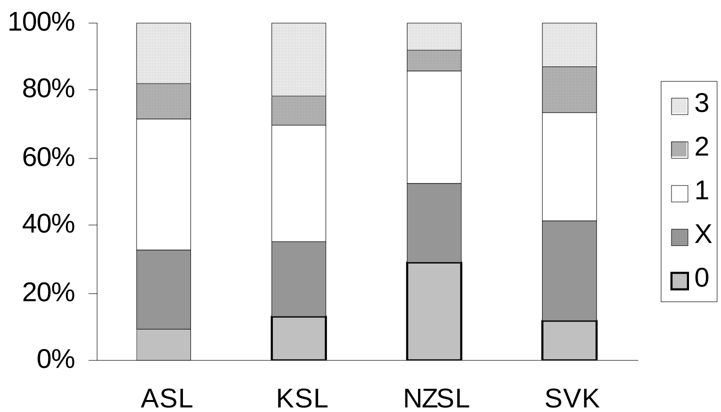
45 These signs are arguably not monomorphemic.

The frequency of each sign type is given in (2-11) and graphed in (2-12); compound signs are omitted.

(2-11) Frequency of sign types table

| | ASL | KSL | NZSL | SVK |
|---|---|---|---|---|
| 0 | 0.0918 | 0.1306 | 0.2893 | 0.1179 |
| X | 0.2363 | 0.2185 | 0.2339 | 0.2964 |
| 1 | 0.3906 | 0.3492 | 0.3375 | 0.3232 |
| 2 | 0.1035 | 0.0879 | 0.0589 | 0.1357 |
| 3 | 0.1777 | 0.2138 | 0.0804 | 0.1268 |

(2-12) Frequency of sign types graph




The Korean Sign Language database contains a large number of Type C signs, 26% of the total KSL database versus about 6% of the other three languages. Since a description of what phonological changes accompany the compounding process in KSL was not available to me, I could judge which signs are linguistic compounds and which are collocations.[46] Therefore, all signs with two or more members were transcribed as compounds. An example of what appears to be a collocation rather than a compound is the entry in the KSL dictionary glossed as HAMDAE

46 See K&B, Liddell (1983; 1984a) and Liddell and Johnson (1985; 1986) for a description of some phonological processes that occur in ASL compounds. See Rozelle (1996a; 1996b; 1998) for a description of some phonological processes that occur in SVK compounds.

*fleet*. This entry has five members: *water* + *wave* + *army* + *ship* + *go*. From entries like this one, it appears that the KSL dictionary was intended as a resource for people wanting to translate from spoken Korean to Korean Sign Language, so that it was prepared by taking a list of spoken Korean words and requesting the KSL sign for each. Likewise, compounds represent 26% of the VSVK database, in which possible collocations (in VSVK, not in spoken Finnish) such as HEDELMÄTARHA *orchard* and MAAILMAN RAKENNE *world structure*, both with three members, are found.

Type 2 and Type 3 signs are not well-defined, in that a sign like SOME in ASL could be considered a Type 2 sign, because both hands bear the same handshape, or a Type 3 sign, because the nondominant hand assumes one of the seven unmarked handshapes allowed by the typology. I refer to these signs as Type 2.5. Should a sign be considered Type 3 only if the handshapes differ? Should a sign be considered Type 2 only if the hands bear marked handshapes? Before the creation of the databases I did not know what handshapes other sign languages allow on the nondominant hand when the two hands bear different handshapes; therefore, while inputting signs I called any sign Type 2 if the nondominant hand did not move while bearing the same handshape as the dominant hand. After the databases were compiled, I examined all Type 3 signs in each language to determine what handshapes are allowed on H2. Then I reviewed all Type 2 signs bearing these handshapes. I retained the Type 2 label if in addition to the two hands having the same handshape, they have the same or opposite orientations, and they have the same or opposite points of contact. If the orientations or points of contact differed, then the sign was reclassified as a Type 3 sign in which the hands coincidentally bear the same handshape. Thus, the ASL sign SCHOOL in (2-13 a) would retain its Type 2 classification because the hands have the same handshape, same contact, and opposite orientations, while the sign SOME in (2-13 b) would be reclassified as a Type 3 sign because the two hands have different orientations and points of contact. Then I reviewed all remaining Type 2 signs to confirm that even those Type 2 signs articulated with marked handshapes satisfied the criteria that orientations and contacts are the same or opposite.

(2-13) Type 2.5 signs in ASL

a. Type 2: SCHOOL

b. Type 3: SOME

The refined typology that permits classification of Type 2.5 signs is shown in (2-14).

(2-14) Battison's typology refined

| **Type** | **hands used** | **location** | **handshape** | **movement** | **orientation and contact** |
|---|---|---|---|---|---|
| **0** | H1 only | $\pi$ | unrestricted | unrestricted | unrestricted |
| **X** | H1 only | body | unrestricted | unrestricted | unrestricted |
| **1** | H1 and H2 | $\pi$ or body | H1 = H2 | H1 = H2 | H1 = H2 |
| **2** | H1 and H2 | ( $\pi$? H2? ) | H1 = H2 | H2 is still | H1 = H2 |
| **3** | H1 and H2 | ( $\pi$? H2? ) | H1 ≠ H2; H2 restricted | H2 is still | unrestricted |

The location of signs in which the hand or hands contact the body at a place other than the nondominant hand is clear. This location is certainly phonologically relevant; the question of neutral space as a location is somewhat more problematic. Neutral space is traditionally considered the unmarked place of articulation. Although this position is probably correct, explicit arguments for it have not been adduced. Since differences in points in neutral space are used to make morphological and syntactic distinctions (K&B), the status of neutral space is an important issue. However, in none of the languages under consideration are there lexemes distinguished only by where in neutral space they are articulated. Following Liddell (1990; 1995), 'points in space' are assumed to be extra-linguistic, or, at least, out of the realm of phonology.[47] Any sign in which the moving hand or hands does not contact any part of the body, including the passive nondominant hand, has neutral space as its location.

What is the location of Type 2 and Type 3 signs? There are two possibilities. They can have neutral space, $\pi$, as their location; thus, the interaction between the dominant and nondominant hands is unrelated to the location parameter. Alternatively, these signs can be analyzed as having the nondominant hand as their location. Sandler (1993a) presents thirteen morphological and phonological arguments to show that the nondominant hand behaves differently in Type 1 signs than in Type 2 and Type 3 signs, thus, providing evidence in favor of representing the nondominant hand as a location. These arguments are not repeated here. Instead, two more arguments in favor of representing the nondominant hand as a location are presented.

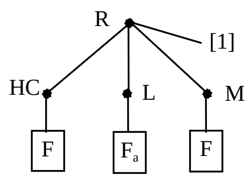
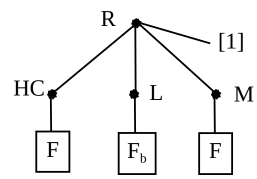
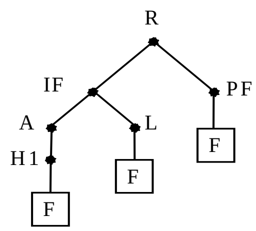
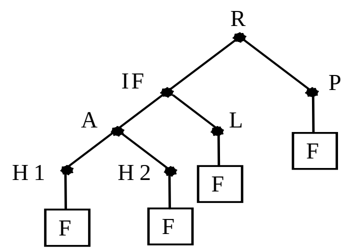
The first of the two new arguments in favor of representing the nondominant hand as a location is that if the nondominant hand is so represented, then the Contact Condition follows immediately. Type 2 and Type 3 signs require contact or proximity between the dominant and nondominant hands because "location" implies contact or proximity between the articulator and the location. An analysis in which the nondominant hand is not a location cannot account for this phonotactic constraint without additional stipulation.

The second argument is more theoretical. In models that represent the nondominant hand in Type 2 and Type 3 signs as a separate articulator, the representational differences between Type X signs and Type 3 signs are significant (Brentari, 1990b; Brentari, 1993; Hulst, 1993; Hulst, 1996b; Brentari, 1998). Other models view the nondominant hand as a location in Type 2 and 3 signs and seek to eliminate its representation in all or some types of signs (Stokoe et al., 1965; Sandler, 1989; Perlmutter, 1991; Rozelle, 1992; Rozelle, 1996a; Rozelle, 1998). Simplifications of Brentari (1998) and Rozelle (1998) are diagrammed in (2-15).

---

[47] Note that it is possible for a sign with two active hands to have contact between them, such as in the ASL sign BECOME, in which the 3 handshapes on both hands rotate while the hands maintain palm contact throughout the articulation of the sign. The nondominant hand in this case is an active articulator, not a location; the location is neutral space. The sign BECOME contrasts with the ASL sign COOK, in which the nondominant 3 hand does not move, serving as the base upon which the dominant 3 hand acts.

(2-15) Representational differences between Type X and Type 3 signs

| | Type X<br>old HELP and HOSPITAL | Type 3<br>modern HELP and RAAMATTU |
|---|---|---|
| H2 as location |  |  |
| H2 as articulator |  |  |

If the nondominant hand is represented as another articulator, Type X signs have a dominant hand node, H1, and a location node, while Type 3 have a dominant hand node, a location node and a nondominant hand node, H2. Yet, there does not appear to be any great difference between, for example, the ASL Type 0 sign HOSPITAL and the SVK Type 3 sign RAAMATTU *bible*. In both signs, the dominant 59 handshape traces an X or cross. In HOSPITAL, this tracing is on the upper arm, □, while in RAAMATTU, this tracing is on the back of the nondominant 3 hand, orientated palm in, fingers up. Indeed, diachronic change shows that the representational dichotomy between these two signs is false. In the past, the ASL sign HELP was articulated at the elbow, but it is now articulated at the nondominant hand. In models that represent the nondominant hand as a separate articulator, this change is major, from a simpler sign with no nondominant hand node and the elbow as location to a more complex sign with a nondominant hand node and neutral space as location. If the nondominant hand is seen as the location, then both old and new HELP are one-handed signs; the location has simply moved from the arm to the hand.

Under the assumption that Type 2 and Type 3 signs have the nondominant hand as their location, there are two options for the formal representation of this location. The representation can be fully specified for its handshape features, or the representation can be simply “hand” ( | ), which can be conceived of as the same type of node for both Type 2 and Type 3 signs or as different.[48] The elaboration of the “hand” location is realized later or in some other manner.[49] These possibilities are listed in (2-16), with potential formal representations shown in (2-17).

(2-16) Possible analyses of the nondominant hand location of Type 2 and 3 signs

a. an unelaborated node “hand”, |

b. individually, fully-specified handshapes, $h_1$, $h_2$, … $h_n$, e.g., BASCO15.

(2-17) Formal representation of Type 2 and Type 3 location possibilities

| hand: $\vert$ | individually specified handshapes: $h_1$, $h_2$, … |
|---|---|
| loc | loc $\begin{pmatrix} \alpha\ F_1 \\ \beta\ F_2 \\ \dots \end{pmatrix}$ |

Since the handshape of the nondominant hand in Type 2 signs is a copy of the handshape of the dominant hand, while the handshape of the nondominant hand in a Type 3 sign varies independently of the handshape of the dominant hand, it is reasonable to require that the

---

[48] If both Type 2 and Type 3 signs have the same node, | , then when calculating the frequency of use of different locations, all Type 2 and Type 3 signs would be counted as having the same location, | . If the hand nodes of Type 2 and Type 3 signs are considered different sorts of location, then they would be counted as two separate locations, $\vert_{\text{type 2}}$ and $\vert_{\text{type3}}$ in the frequency calculation.

[49] In Type 2 signs, the handshape features of the nondominant hand can simply be copied from the dominant hand. In Type 3, it is difficult to propose a source from which these features could be derived, as there are a very few signs that differ only in H2 handshape (Brentari (1998) gives the examples of TRY-ON and PUT-ON-SHOE), although Perlmutter (1991) argues that it is perhaps predictable by some means.

representation of a Type 2 sign be simpler than that of a Type 3 sign. With this restriction, there are four possibilities for the locations of Type 2 and Type 3 signs, shown in (2-18).[50]

(2-18) Possibilities for locations of Type 2 and Type 3 signs

| | Type 2 | Type 3 |
|---|---|---|
| analysis 1 | $\vert$ | $\vert$ |
| analysis 2 | $\vert_{\text{Type 2}}$ | $\vert_{\text{Type 3}}$ |
| analysis 3 | $\vert$ | $h_1, h_2, h_3 \ldots$ |
| analysis 4 | $h_1, h_2, h_3 \ldots$ | $h_1, h_2, h_3 \ldots$ |

The third possibility in which Type 2 signs all have $\vert$ as their location, while Type 3 signs have different locations corresponding to the configuration of the nondominant hand, conforms to the phonological arguments given above (Sandler, 1993a; Brentari, 1998; Rozelle, 1998). An additional argument, which will be more presented in Chapter 6, is that by counting the nondominant hand in Type 2 signs as one location, but differentiating the nondominant hand location in Type 3 signs by handshape, later computations of location rank-frequency yield consistent results language-internally as well as cross-linguistically. All the other possibilities, including allowing neutral space as a possible location for either just Type 2 signs or for both types, do not produce a clear picture.

---

[50] One possibility not explored is that of a complex location, with the nondominant hand as the location for the dominant hand, and neutral space as the location of the nondominant hand or of the two-handed composite. This possibility is not so interesting for a combination with neutral space, as neutral space is usually represented with an empty location node. However, for signs such as NOSE-TO-THE-GRINDSTONE or SURGERY-ON-BODY-PART, in which the dominant hand contacts the nondominant hand while the nondominant hand contracts a part of the body, this analysis is of location is reasonable (Brentari, 1998). Many such forms violate the Symmetry and Dominance conditions and are, in fact, morphologically complex or derived historically as polymorphemic forms. (The ASL sign INTERNALIZE is a monomorphemic example.) They do not occur as monomorphemic lexical items; hence, the complexity of their location ought to be represented on another linguistic level. See Rozelle (1998) for further discussion.

To sum up, a sign is classified as Type 2 or Type 3 if it employs two hands, but only one hand moves during its articulation. If the handshapes are the same, and both hands also have the same orientation and point of contact, the sign is classified as Type 2. If the handshapes differ, or if the orientations or points of contact differ, the sign is classified as Type 3. The location of Type 2 and Type 3 signs is the nondominant hand, H2. The final typology is shown in (2-19).

(2-19) Final typology

| Type | hands used | location | handshape | movement | orientation and contact |
|---|---|---|---|---|---|
| **0** | H1 only | $\pi$ | unrestricted | unrestricted | unrestricted |
| **X** | H1 only | body | unrestricted | unrestricted | unrestricted |
| **1** | H1 and H2 | $\pi$ or body | H1 = H2 | H1 = H2 | H1 = H2 |
| **2** | H1 and H2 | H2 | H1 = H2 | H2 is still | H1 = H2 |
| **3** | H1 and H2 | H2 | H1 ≠ H2; H2 restricted | H2 is still | unrestricted |

## 2.2 Analysis

After the databases were created, they were imported into Microsoft Excel 2000 from FileMaker Pro 5. The spreadsheet format of Excel permits freer manipulation of the data. The functions of sorting, charting, pivot table, and statistical analysis were essential to the data analysis.

For the analysis, not all of the signs in the database could be used. Borrowings were excluded from all analyses. While borrowings can provide insights into the structure of the phonology of native elements, they can also form their own separate phonological subsystems.[51]

When examining the handshape parameter, signs involving handshape change were excluded for three reasons. First, it is possible that in handshape change signs the handshape parameter behaves differently than in non-handshape change signs. (Perhaps there is a special set of handshapes used in these signs or a special distribution). Second, notating these handshapes is problematic. It is difficult to notate the beginning and ending handshapes because it is hard to

---

51 See, for example, Battison (1978) and Brentari (2001).

determine actual handshape in many drawings unless both initial and terminal forms are included. Also, the beginning or ending handshape is often indeterminate in that one handshape is the target, and the other is just the onset or offset. Nor was it feasible to notate only one handshape and indicate how it changes, such as opening or closing movement. Note that signs with handshape internal movements (handshape contour rather than handshape contrast signs; see section 1.5.3) were included, such as the slight bending at the metacarpophalangeal joints of 78 in the SVK sign RANTA *beach*, or of 59 in the ASL sign NOODLE. These minor handshape changes are repeated, and the resulting intermediate handshape is not used in any sign without a handshape change.

Likewise, signs involving a location change were excluded from the examination of location, though there were very few monomorphemic signs with a change in location.

Compounds were excluded from all analyses. The presence of morphological complexity can result in a phonological form illicit morpheme-internally (e.g., *[vd] but lived [λIϖ + d] *lived*). While sequential polymorphemic signs were excluded, simultaneous polymorphemic forms were not excluded because I cannot reliably segment a sign into morphemes in a language I do not know. An example is the SVK sign HISSI *elevator*, which violates the Symmetry and Dominance Conditions because both hands move although they bear different handshapes. The violation exists because this sign was historically a bimorphemic form involving two classifiers. Today it is usually articulated as a one-handed sign (Type 0) or as a symmetrical two-handed sign (Type 1). However, I do not necessarily have this sort of information for all the languages investigated in this dissertation, and it is possible that another sign language does allow such forms monomorphemically.

### 2.2.1 Statistical analysis

For most of the quantitative analysis of these data, I used common statistical tests. The Excel statistical analysis package included many of these tests, such as linear regression, paired-sample *t*-test and the chi-square test for independence in a contingency table. To test whether the orderings of two ranked lists are independent or whether they vary in the same or opposite directions, I used the Spearman rank correlation test (Snedecor and Cochran, 1982).

### 2.2.2 Mutual information analysis

In Chapter 5 duet tables for each language are created. These are large *n* x *m* arrays in which the rows are labeled with the *n* handshapes used in the language and the columns are labeled with the *m* locations used in the language. In each cell in this array is the number of signs, $n_{ij}$ in the language that use that particular duet of handshape and location. One of the questions investigated in Chapter 5 is whether the distribution of handshape is independent of location, or whether certain handshapes co-occur with certain locations. It is possible to test for independence of handshape and location by means of a chi-square test on the *n* x *m* contingency table if the data are grouped in such a way that $n_{ij} \geq 5$; if $n_{ij} < 5$, the chi-square test is invalid. For example, if handshapes are grouped into marked versus unmarked sets, and if locations are grouped into contact versus non-contact, then independence can be tested via the chi-square test. However, in order to test independence without grouping, some other means is necessary. This is because the number of cells in the duet tables range from approximately 1200 to 2300 (number of handshapes times number of locations) and each database contains only about 600 signs per language; thus, most of the cells are empty.

To solve this problem, I use information theory (Shannon and Weaver, 1949), which in linguistics is used in speech recognition. Even within more traditional phonology, Goldsmith (1998) argues that the central notions of information theory, entropy and mutual information, are "the natural quantitative measures of many of the concepts used by phonologists." Entropy is a quantitative measure of variation; less variation means lower entropy. It is also a measure of randomness; less randomness means lower entropy (Manning and Schütze, 1999).

For example, consider a language with just four handshapes, $h_1$, $h_2$, $h_3$, $h_4$, that appear with equal frequency, 0.25. What is the least number of questions necessary to guess which handshape is used in a given sign? Whichever handshape is used, the answer is two: (1) Is it $h_1$ or $h_2$? If the answer is "yes," then the second question is: (2) Is it $h_1$? Otherwise, the second question is: (2) Is it $h_3$? This general procedure is the best regardless of the number of handshapes if the frequencies are uniform, and the number of questions is $\log_2 n$, where *n* is the number of handshapes.[52] Now consider a language with four handshapes that do not have equal frequencies. Say $f(h_1) = 0.5$, $f(h_2) = 0.25$, $f(h_3) = 0.125$, $f(h_4) = 0.125$. Now the best question to ask first is, "Is it $h_1$?" because the handshape will be ascertained with only one question half of

---

52 It is possible in the case of four handshapes to ask the three questions, "Is it $h_1$?, Is it $h_2$?, Is it $h_3$?" with the same results, but for eight, it is more efficient to halve the set.

the time. If the answer is no, the second question should be, “Is it $h_2$?” because then the handshape will be ascertained with two questions a quarter of the time. If the answer is no, one needs to ask, “Is it $h_3$?”, so that three questions are needed the remaining quarter of the time. On average, (1 * 0.50) + (2 * 0.25) + (3 * 0.25) = 1.75 questions are needed. Notice that the entropy is smaller for the second language with the less random, more structured set of handshape distributions.

Entropy is defined in (2-20). P(x) is the probability that X = x, where X is a random variable over a set of symbols ***X***, where x $\in$ ***X***. In this dissertation, ***X*** could be the set of handshapes (or locations or duets) in a language; p(x) is interpreted as the relative frequency with which handshape x occurs in the database of that language. Entropy can be thought of as the weighted average of the logarithm of the probability. The negative sign before the summation is there because positive numbers are easier to comprehend.

(2-20) Entropy H

$$H(X) = -\sum_{x \in X} p(x) \log_2 p(x)$$

If the value of one variable is known, how much is known about the other variable? Mutual entropy is the amount of information that one variable conveys about the other variable. For example, if it is known that handshapes $h_1$ and $h_2$ only occur at locations $l_1$ and $l_2$, while handshapes $h_3$ and $h_4$ only occur at locations $l_3$ and $l_4$, then knowing what the handshape is gives some information about what the location is. On the other hand, if $h_1$, $h_2$, $h_3$, $h_4$ occur randomly with $l_1$, $l_2$, $l_3$, $l_4$, then knowing the handshape does not give any information about the location. Mutual information is defined in (2-21). In this example, as in Chapter 5, ***X*** is the set of handshapes, and ***Y*** is the set of locations. P(x) is the lexical frequency of handshape x, p(y) is the lexical frequency of location y, and p(x, y) is the lexical frequency of the duet (x, y).

(2-21) Mutual information I

$$I(X;Y) = \sum_{x,y} p(x,y) \log_2 \left( \frac{p(x,y)}{p(x)p(y)} \right)$$

If handshape and location are independent, then the probability at which a particular duet occurs is just the product of the handshape probability and the location probability. Thus,

$p(x, y) = p(x)p(y)$, so that $\log_2\left(\frac{p(x, y)}{p(x)p(y)}\right) = \log_2(1) = 0$, and I(X:Y) = 0. The greater I(X:Y) is , the greater the dependency between the variables X and Y is. However, the data are very sparse in the duet array, because so many handshape-location combinations are not attested, so it is possible for a non-zero mutual information to occur spuriously. For example, if a handshape is so rare that it occurs only once in the data, it will appear to have a special affinity for the location at which it is articulated, producing a non-zero term in the summation.[53]

In order to compensate for the scarcity of the data, I used a technique used in bioinformatics for sequencing amino acids in proteins (Karplus, 1995).[54] This technique calculates the mutual information of the actual data as well as the mutual information of 1,000 scrambled data sets that are based on the actual data. The actual mutual information is compared to that of the 1,000 scrambled sets. If it is much greater than would be expected based on this random sampling, it can be assumed that the two data sets possess nonzero mutual information, and are thus not independent.

For each language, I prepared a long list of ordered pairs ($x_i$, $y_i$), one pair for each eligible sign $S_i$ in the database,[55] where $x_i$ is the handshape of the sign $S_i$ and $y_i$ is the location of the sign $S_i$. I input these data into a program designed to calculate the mutual information for these actual data as well as for 1,000 other data sets. The source code is in Appendix B. The program took the handshape data, given by the first element of the ordered pair, and scrambled the order of the handshapes. Thus, the actual frequencies of the handshapes and the locations are preserved (the marginal probabilities, p(x) and p(y)), but new pairs of (handshape, location) are formed. The mutual information of this new, scrambled data set is computed and recorded. Note that whatever dependency that might have existed between the handshape and the location is destroyed by this scrambling process. A nonzero mutual information value would occur only through chance

---

53 0*log (0) is defined to be 0. An alternative to this exceptional definition is to incorporate pseudocounts, that is, to add a very small nonzero amount to each cell in the duet table based on, for example, the expectation set by the duet table resulting from the pooling of all the database (Cline, 2000). Since all cell counts are nonzero, p(x, y) is always nonzero, and log (p(x, y)) is defined. Either method is acceptable. I chose the simpler one.

54 I wish to express my gratitude to Kevin Karplus for teaching me this technique and to Clifford Wong for writing the mutual information significance computer program.

55 Recall that compounds, borrowings and signs with handshape or location change are excluded from this analysis.

association.  This procedure is repeated 1,000 times.  The actual mutual information value is compared to the values of the scrambled ones.  If the actual mutual information value is greater than, for example, all but 50 of the scrambled trials, we can say that the variables are independent, with $p<0.05$.  This technique can, of course, be applied to the sort of non-sparse tables with cells counts greater than five to which the chi-square test is applicable.  When both the chi-square test and the mutual information test are applied to the same data sets, the p values are extremely close, confirming the validity of this method.[56]

[56] For example, in ASL if handshapes are grouped into marked and unmarked sets (on the basis of the universal markedness criteria, section 3.3), and body locations are grouped into hand, arm, face, and torso, a 2 x 4 table is created.  Chi-square analysis of this table gives $p=0.549$, while the mutual information significance program gives $p=0.558$.

# Chapter 3: Handshape

In this chapter, the handshape inventories of ASL, KSL, NZSL, SVK and VSVK are presented and compared. The handshape inventories are ranked according to lexical frequency, and the rank-frequency distributions of all four are shown to be remarkably well approximated by an exponential decay curve. The set of handshapes common to the four main languages is determined, and the ordering of these shared handshapes is compared cross-linguistically. The data are pooled to provide an idea of the "International Handshape Alphabet," and it is shown that the lexical rank-frequency distribution of the pooled data is also an exponential decay curve. Lastly, dependence between handshape and sign type (Type 0, 1, 2 or 3) is investigated with respect to markedness. Is it the case that handshapes occur freely within different types of signs, or do certain sign types attract certain handshapes?

## 3.1 Inventories

The databases in this dissertation have been built on a sample of the entries from a dictionary; furthermore, the dictionary itself represents only a sample of the lexical items available to users of a language. Thus, it is possible that a handshape not included in the inventory of a language presented in this chapter might actually occur. However, since these inventories were created on the basis of a large sample, it is unlikely that there are many omissions. The handshapes present in the databases of ASL, KSL, NZSL, SVK and VSVK, which were created for this study, are investigated in depth. Where possible, the databases for NGT and LIS are consulted and the results compared.

### 3.1.1 American Sign Language (ASL)

The American Sign Language database uses 35 handshapes for native, monomorphemic signs that do not incorporate a handshape change. They are shown in (3-1), organized by selected finger. The first column has zero selected fingers. The next four columns have one selected finger: the index, the middle and the pinky for the second, third and fourth columns, and the index finger in a closed formation for the fifth column. The sixth column has two selected fingers. The last three columns have all fingers selected; the seventh column has the fingers compact, the eighth has them spread, and the last has them closed.

(3-1) ASL: handshape inventory

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| 2 | 49 | 78□A | 49□ | < | 59 | 3 | 78 | ; |
| 29 | 48 | 7≅□A | 48□ | :□ | 69 | 39 | 79 | ;A |
| 28 | 4B9 | | | | 68 | 38 | 7B8 | |
| | 48□B | | | | 6B9 | 3B | 78B | |
| | 4≅A | | | | 69□□ | 3A | 7≅B | |
| | | | | | 6□□□ | 3B≅□ | | |
| | | | | | | 3≅B | | |
| | | | | | | 3≅A | | |

There are 87 borrowings in the ASL database, representing 13% of the signs. One of these signs is PSYCHOLOGY, which is an iconic representation of the Greek letter Ψ, so it perhaps is not truly a borrowing. The other 86 borrowings are initialized signs. An initialized sign is a sign in which the handshape corresponds to the fingerspelling of the first letter of the spoken English word that serves as a gloss for that sign (Battison, 1978). The handshapes in (3-2) are used only in borrowings.[57]

---

[57] Whether an initialized sign is indeed a 'borrowing' in the linguistic sense is unclear. Even a complete fingerspelling of an English word is just a manually coded representation of the spelling of the word, not the English word itself (Davis, 1989; 1994). Sign language borrowing parallel to spoken language borrowing would involve the use and adaptation in ASL of a sign from a foreign sign language. This situation does occur, for example, when ASL signers use a country's own sign for itself, rather than the native ASL sign. However, this type of sign language-to-sign language borrowing has not been well investigated. For example, the recent book *Foreign Vocabulary in Sign Language* (Brentari, 2001) focuses exclusively on spoken language-to-sign language borrowing. Nevertheless, these signed borrowings from spoken language have properties that distinguish them as a group from other signs (Battison, 1978; Brentari and Padden, 2001).

(3-2) ASL: novel handshapes used only in borrowings

| Fingerspelled letter | Handshape | Sign |
|---|---|---|
| R | 59□□□ | RELIGION |
| K or P | 6□□□ | PERSONALITY |
| D | <□□□ | DICTIONARY |
| W | 79□X | WORLD |
| T | 2□□□ | TRANSLATE |
| N | 2□□□ | NORTH |
| M | 2□□□ | MONDAY |
| M | 3A9□X | MEDICINE |
| E | 39X | EAST |

The handshapes in (3-3) are used only in signs having a dynamic handshape. In some signs the handshape changes from one well-defined handshape into another, such as in the sign BETTER, which begins with the handshape 3 then changes to 2. In other signs either the beginning or the ending handshape is not well-defined, such as in the sign LIKE, which clearly ends with <□, but begins with 78 or 7≅□A. Handshape change occurs in 106 signs, 16.2% of the ASL sample.

(3-3) ASL: novel handshapes used only in handshape change signs

| | | | | |
|---|---|---|---|---|
| 48□B | <□ | A68B | 3A8 | 78□A□A□A |
| | <□A | | | 7≅A |
| | 7≅□B□B | | | |
| | : | | | B |

The ASL database contains 31 compound signs, 4.7% of the database. These signs use seventeen different handshapes. However, these seventeen handshapes also appear in monomorphemic signs, so they contribute no novel handshapes to the inventory in (3-1).

One of the other handshape inventories that has been adduced for ASL is presented in (3-4). This inventory is from Klima and Bellugi (1979), based on Stokoe, et al. (Stokoe et al., 1965). *The Dictionary of American Sign Language on linguistic principles* ((Stokoe et al., 1965); hence, DASL) lists nineteen handshapes. Sixteen of these handshapes are notated by the letter of the roman alphabet that most closely resembles the fingerspelled handshape for that letter. Otherwise, 78 is represented by 5, 68 by 3, and 78□A or <□ by a special symbol. DASL also includes diacritics to indicate whether the thumb is extended and whether the selected fingers are curved. K&B terms the nineteen handshapes represented by a single grapheme ‘primes’ and the others ‘subprimes’. The primes are identified in (3-4) by being boxed. K&B states that the subprimes vary sometimes allophonically, but sometimes freely. It is not clear whether the primes are supposed to be phonemes, or whether they have attained the status of primes simply because they have names, since these handshapes correspond to letters in the ASL fingerspelled alphabet. Thus, the list in (3-4) appears to be a phonetic inventory only.

(3-4) ASL: handshape inventory from K&B, based on DASL

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | | 78□ | 4□ | | | 799□ | | | |
| 2 | 4 | A | | < | 59 | X | 3 | 78 | ; |
| 29 | 49 | | 48□ | : | 69 | <□□□ | 39 | 79 | ;A |
| 28 | 48 | | | <□ | 68 | | 38 | 78B | |
| 2□□□ | 49B | | | | 6B9 | | 3A | 7≅A | |
| | 4≅A | | | | 59B | | 3B≅ | | |
| | 48B | | | | 68 | | 39X | | |
| | | | | | 6□□□ | | | | |
| | | | | | 59□□ | | | | |
| | | | | | □ | | | | |
| | | | | | 68□□ | | | | |
| | | | | | 69□□ | | | | |

### 3.1.2 Korean Sign Language (KSL)

The Korean Sign Language database uses 44 handshapes for native, monomorphemic signs that do not incorporate a handshape change. They are shown in (3-5).

(3-5) KSL: handshape inventory

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| 2 | 4 | 49□ | 49□ | : | 59 | 3 | 78 | ; |
| 29 | 49 | | 48□ | :A | 58 | 39 | 79 | ;A |
| 28 | 48 | | 49A□ | < | 69 | 38 | 7≅B | |
| | 4≅ | | | <□ | 68 | 3B | 7≅A | |
| | 4X | | | | 5A9 | 3A | | |
| | 49B | | | | 59B | 38B | | |
| | 49A | | | | 69B | 3≅B | | |
| | 48X | | | | 6≅A | 3≅A | | |
| | 4≅B | | | | | | | |
| | 4≅A | | | | | | | |
| | 4≅□ | | | | | | | |
| | B | | | | | | | |

There are eight borrowings in the KSL database, representing only 1.3% of the signs. The handshapes in (3-6) are used only in borrowings.

(3-6) KSL: novel handshapes used only in borrowings

<□A□A 5□□□ 79□X

These borrowings resemble the ASL initialized signs in that they are a representation of a written form of a spoken word. This representation is more direct in that some aspect of the sign mimics a visual element, most commonly the written Chinese character used for the spoken Korean gloss. This type of borrowing from the spoken language occurs also in Taiwan Sign Language (Ann, 1998). For example, the Chinese character for *mountain* and the corresponding KSL sign are shown in (3-7 a). The character for *field* and the KSL sign are shown in (3-7 b). In these signs, the handshape itself resembles the form of the Chinese character.

(3-7) KSL: borrowings from written Chinese

| gloss | character | handshape | sign |
|---|---|---|---|
| a. SAN *mountain* | 山 | 49□ |  |
| b. BAT *field* | 田 | 79□X |  |

In SACHON *cousin*, the handshape does not resemble the Chinese character, but instead the shape of the character is traced in the air with the index finger. Similarly in MITUR *meter*, the letter *M* from the Roman alphabet is traced in the air with the index finger. In the KSL database there is one example of an initialized sign, like the ASL borrowings, that uses a handshape from the ASL fingerspelling alphabet. The sign is YOO EN *United Nations* and has the handshape of the fingerspelled *R*. Although there is a fingerspelled alphabet for the Korean alphabet, there are no signs in the KSL database that incorporate handshapes from the Korean fingerspelled alphabet.

The handshapes in (3-8) are used only in signs having a dynamic handshape. In some signs the handshape changes from one well-defined handshape into another, such as in the sign CHONG *gun*, which begins with the handshape 48 then changes to 48□. In other signs either the beginning or the ending handshape is not well-defined, such as in the sign JAESAN *property*, which clearly begins with <, but ends with 3≅ or 3≅B. Fifty-six signs, 9.1% of the KSL sample, contain a handshape change.

(3-8) KSL: novel handshapes used only in handshape change signs

5≅A <□B□B?□

:A□□ 48□X :□

The KSL database contains 137 compound signs, 22.3% of the database. These signs use 35 different handshapes. All but the three handshapes in (3-9) also appear in monomorphemic signs.

(3-9) KSL: novel handshapes only used in compound signs

;B <X 39□X

### 3.1.3 New Zealand Sign Language (NZSL)

The New Zealand Sign Language database uses 49 handshapes for native, monomorphemic signs that do not incorporate a handshape change. The handshape inventory is shown in (3-10).

(3-10) NZSL: handshape inventory

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| 2 | 4 | 4□ | 49□ | : | 69 | 79□X | 3 | 78 | ; |
| 29 | 49 | 49□ | 78□A | :A | 59 | | 38 | 79 | ;B |
| 28 | 48 | 48□ | 7≅□A | < | 69B | | 39 | 7B | |
| | 4A | | | <B | 5≅A | | 3≅ | 7B8 | |
| | 4X | | | | 69□□ | | 3B | 7≅B | |
| | 49B | | | | 6□□□ | | 3A | | |
| | 49A | | | | | | 3X | | |
| | 4≅A | | | | | | 38B | | |
| | 4≅B | | | | | | 3X8 | | |
| | 48□B | | | | | | 3A8 | | |
| | | | | | | | 3≅B | | |
| | | | | | | | 3≅A | | |

There are 42 borrowings in the NZSL database, representing 6.1% of the signs. These signs use twenty different handshapes. The fingerspelling alphabet used in New Zealand is different from the aphabets used in the United States and in Finland. The New Zealand fingerspelling uses two hands in the formation of each letter, while the US and Finnish alphabets use one hand. The handshapes in (3-11) are used exclusively in borrowings. The handshape 59□□□ is probably a borrowing from New Zealand gesture; the dictionary uses the phrase, “Cross your fingers for luck.”

(3-11) NZSL: novel handshapes used only in borrowings

| Fingerspelled letter | Handshape | Sign |
|---|---|---|
| S | 49□B | SATURDAY, SOCIAL WORK |
| none | 59□□□ | HOPE, WISH |
| M | 79B□X | MILO |
| M | 39□X | MORMON, MUSHROOM, MOKO |

The handshapes in (3-12) are used only in signs having a dynamic handshape. One hundred twenty four signs, 18.0% of the NZSL sample, contain a handshape change.

(3-12) NZSL: novel handshapes used only in handshape change signs

| | | | | | |
|---|---|---|---|---|---|
| 2□ | 4≅ | <□ | ;A | 5≅ | 7≅ |
| | 48□X | <□B | :B | 68 | 7≅A |
| | 48□X | ?□ | :□□□ | 6B8 | |
| | | | | :B□□ | |
| | | | | :A□□ | |

The NZSL database contains 55 compound signs, 8.0% of the database. These signs use 25 different handshapes. All of these handshapes also appear in monomorphemic signs, except for the handshape in (3-13), which is used for the second member of the sign WEDDING RECEPTION = WEDDING + FEAST. The sign FEAST is not in the dictionary.

(3-13) NZSL: novel handshape used only in a compound

58 □A□A

### 3.1.4 Finnish Sign Language (Suomalaisen Viittomakieli: SVK)

The Finnish Sign Language database uses 34 handshapes for native, monomorphemic signs that do not incorporate a handshape change. The handshape inventory is shown in (3-14).

(3-14) SVK: handshape inventory

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| 2 | 49 | 49□ | < | 59 | 3 | 78 | ; |
| 29 | 48 | 48□ | : | 69 | 39 | 79 | ;A |
| 28 | 49B | | | 68 | 3≅ | 7≅ | |
| | 4X | | | 69B | 3B | 7B | |
| | 4A | | | | 3A | 7≅B | |
| | 49A | | | | 3≅B | | |
| | 4≅B | | | | 3≅A | | |
| | 48□B | | | | | | |
| | 4≅□B | | | | | | |

The eight handshapes in (3-15) are used only in borrowings. The handshape 78□ is used in only one sign in the SVK dictionary, JEESUS *Jesus*, which has the same form as the ASL sign, and most likely represents a borrowing. (In contrast, 78□ is used in 22 signs in the ASL dictionary.) The other seven handshapes are from the Finnish fingerspelling alphabet and are used in initialized signs.[58] There are only eight signs that incorporate fingerspelling, 1.3% of the SVK sample. An example of an SVK initialized sign is EHKÄISYTABLETTI *contraceptive pill*, which uses the handshape 39X.

---

[58] The standard fingerspelled alphabet used in Finland today is similar to that used in the United States. F, H, M, N and P are slightly different, and T is quite different, 4A instead of 2□□□. A different fingerspelled alphabet, similar to the one currently used in Sweden, used to be common. It is still used by older people, and the handshapes of some signs, such as TAKSI *taxi*, are drawn from it.

(3-15) SVK: novel handshapes used only in borrowings

| Fingerspelled letter | Handshape | Sign |
|---|---|---|
| X | 59□□□ | TAKSI *taxi* |
| K | 6□□□ | KESKIVIIKO *Wednesday* |
| P | <□□□ | PERJANTAI *Friday* |
| M | 39□X | MARKKA *mark* |
| W | 79□X | W.C. *toilet* |
| I | 4□ | TIISTAI *Tuesday* |
| E | 39X | EHKÄISYTABLETTI *contraceptive pill* |
| none | 78□A | JEESUS *Jesus* |

The handshapes in (3-16) are used only in signs having a dynamic handshape. In these signs the handshape changes from one well-defined handshape into another, such as in the sign JOSKUS *sometimes*, which begins with the handshape :B, then changes first to 49, then to 69, then to 79. In other signs either the beginning or the ending handshape is not well-defined, such as in the sign HELPPO *easy*, which clearly ends with :A, but begins with 4≅, 48, or 4≅A. Seventy-nine signs, 13% of the SVK sample, contain a handshape change.

(3-16) SVK: novel handshapes used only in handshape change signs

59A :A :B 4≅ 4≅A

The SVK database contains 41 compound signs, or 6.7% of the database. These signs use seventeen different handshapes; however, all of these seventeen also appear in monomorphemic signs, so they contribute no novel handshapes to the inventory in (3-14).

### 3.1.5 Old Finnish Sign Language (Vanha SVK: VSVK)

Because the dictionary of Old Finnish Sign Language (Hirn, 1910) contains only 355 signs, the database for this language is much smaller than those of the other four databases, even though all 355 signs were transcribed. For this reason, and because it is of course related to SVK, VSVK is not included in the lexical analyses performed on the other four languages. Based on this sample, the VSVK database uses 28 handshapes for native, monomorphemic signs that do not incorporate a handshape change. The handshape inventory is shown in (3-17)[59].

(3-17) VSVK: handshape inventory

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| 2 | 4B | : | 49 | 59 | 3 | 7 | ; |
| 29 | 49 | :A | | 59B | 38 | 7B | ;A |
| 28 | 49B | | | 69 | 39 | 78 | |
| 2 | 49A | | | | 3B | 7≅B | |
| | 4≅A | | | | 3A | 7≅A | |
| | | | | | 3≅B | | |

The only sign in the VSVK dictionary that is clearly a borrowing is KYSYMYS *question*. It is a fingerspelled loan sign based upon the spoken Swedish word *fråga*. This borrowing accounts for only 0.3% of the VSVK database. The handshape, shown in (3-18), represents the letter *F* in the Swedish fingerspelling alphabet, which was formerly used in Finland.

(3-18) VSVK: novel handshape used only in a borrowing

3B

The handshapes in (3-19) are used only in signs having a dynamic handshape. As with the other languages, it is often the case that the beginning or the ending handshape is not well-defined, as in the sign OTTAA *take*, which begins with the fingers of 78 very slightly curled and ends with 29. The modern SVK sign has this same handshape change, although with movement

---

[59] Note that the handshape 2 is used in only one VSVK sign, VESI *water*. The latest dictionary of Finnish Sign Language (Kuurojen Liitto, 1998), which was used in creating the SVK database, shows the

toward the body rather than away from it. The initial handshape can be 78, 7B8, 78B, 7≅B, and so on. Twenty-five signs, 7% of the VSVK sample, contain a handshape change.

(3-19) VSVK: novel handshapes used only in handshape change signs

3≅ 3≅A

The VSVK database contains 50 compound signs, or 14.1% of the database. These signs use 25 different handshapes, 22 of which also appear in monomorphemic signs. The three new handshapes are shown in (3-20).

(3-20) VSVK: novel handshapes used only in compounds

79 79□X 48

The handshapes 48 and 79 are shared by the other four languages.[60] Their absence from the VSVK handshape inventory based in non-compound signs is perhaps an artifact of the small size of the sample. However, 79, as well as 79□X, are used to iconically represent different widths of flowing water in the signs in (3-21); therefore, these handshapes should perhaps be considered Size and Shape Specifiers (SASS) rather than phonological primitives of monomorphemic lexical signs.

(3-21) VSVK: signs with SASSs

| | | |
|---|---|---|
| NORO | *little stream* | 49 |
| PURO | *brook* | 69 |
| JOKI | *river* | 79□X |
| JOKI | *river* | 79 |
| VIRTA | *current* | 78 |

---

handshapes 48□ or 2 for VESI. However, the previous dictionary (Kuurojen Liitto, 1973) shows VESI with the older 2□ handshape. I also observed this older form of VESI in Finland in 1993.

[60] Note, however, that 48 is used in only one monomorphemic, non-handshape-changing sign in SVK, VIDEOIDA *video*, which is a more recently added lexical item.

### 3.1.6 Italian Sign Language (Lingua Italiana dei Segni: LIS)

The corpus studied by Pietrandrea (1998) uses 41 handshapes for signs that are not compounds and do not incorporate a handshape change. I do not know whether this corpus includes non-native signs or other types of polymorphemic signs, such as classifier formations. The handshape inventory shown in (3-22) includes only signs without a handshape change.

(3-22) LIS: Handshape inventory

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| 2 | 49 | 78□A | 49□ | :A | 69 | 3 | 78 | ; |
| 29 | 4 | 7≅□A | 48□ | < | 68 | 3A | 7≅B | ;A |
| 28 | 48 | 7≅□B□B | | <A | 59 | 3≅B | 79 | |
| | 4B | 78□A | | :A□□ | 69B | 3A8 | 7≅A | |
| | 4≅B | | | | 6≅A | 3≅A | | |
| | 49B | | | | 68B | | | |
| | 48□A | | | | 59□□□ | | | |
| | 4≅A | | | | 59□□□ | | | |
| | | | | | 49□□ | | | |

The inventory reported in Pietrandrea's work includes 13 "handshapes" that are dynamic; these were excluded from the above inventory.

### 3.1.7 Comparison

Some basic facts about the handshape inventories of these five languages are summarized in (3-23). In ASL a common method for creating new signs is to substitute the fingerspelled initial of an English word for the handshape of an already existing sign in a related semantic field. For example, the signs LANGUAGE and GRAMMAR use the handshapes for L and G, respectively, but are almost identical in all other formational parameters. Thus, ASL has a large number of borrowings. In contrast, SVK, which uses fingerspelling rarely, has a small number of borrowed signs that involve handshape.[61]

---

[61] It is not possible to conclude that SVK is less influenced by spoken language than ASL because SVK signers often mouth a part of the spoken Finnish word that is the usual gloss for a particular sign. Sometimes this mouthing is an integral part of the sign. For example, signers mouth [vuo] while signing VUOSI *year*. At other times the mouthing component is part of a productive morphological process. For

(3-23) Comparison of inventory composition

| | ASL | KSL | NZSL | SVK | VSVK | LIS |
|---|---|---|---|---|---|---|
| Inventory size | 35 | 44 | 49 | 34 | 28 | 41 |
| Borrowings | 13.0 % | 1.3 % | 6.1 % | 1.3 % | 0.3% | ? |
| Handshape change | 16.2 % | 9.1 % | 18.0 % | 13.0 % | 7.0% | 6.1% |
| Compounds | 4.7 % | (22.3 %) | 8.0 % | 6.7 % | (14.1%) | 5.4% |

The handshape inventories of ASL, KSL, NZSL and SVK vary in size from 34 to 49, with an average of 40 and a standard deviation of 7. As was reported in section 1.4, the segment inventory size for spoken languages in UPSID ranges from 11 to 141, with a average of 31 (Maddieson, 1984). If phonetic handshape inventories are compared to phonemic spoken language inventories, and if it is reasonable to attempt to compare results from a sample of only four languages to a sample of 317, it appears that handshape inventories are larger than spoken language inventories. Note that handshape is only one of the parameters of sign formation; location is another. As will be shown in the next chapter, these four languages use on average 20 distinct locations. What phonological material of sign language is decided to be analogous to the phoneme will determine how total inventory size of sign languages is computed. Perhaps these 20 locations should be added to the average of 40 handshapes for an inventory of 60 elements. Alternatively, since handshape and location co-occur, perhaps the inventory consists of 20 x 40 = 800 elements. In either case, it appears that sign languages have more phonetic materials available for morpheme formation.

The 22 handshapes shown in (3-24) are used in all four languages. The relationship between this set and what might be considered to be the set of unmarked handshapes is discussed in section 3.3.

---

example, LENTOKONE *airplane* and LENTÄÄ *fly* differ only in their mouthing. In LENTOKONE, the word is mouthed, and in LENTÄÄ, the lips are pursed while air is blown out (Rissanen, p.c.). This sort of borrowing is not detectable in a lexical study of handshape.

(3-24) Shared handshapes: handshapes used in all four languages

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| 2 | 49 | 49□ | < | 59 | 3 | 78 | ; |
| 29 | 48 | 48□ | | 69 | 39 | 79 | |
| 28 | 49B | | | 69B | 3B | 7≅B | |
| | | | | | 3A | | |
| | | | | | 3≅B | | |
| | | | | | 3≅A | | |

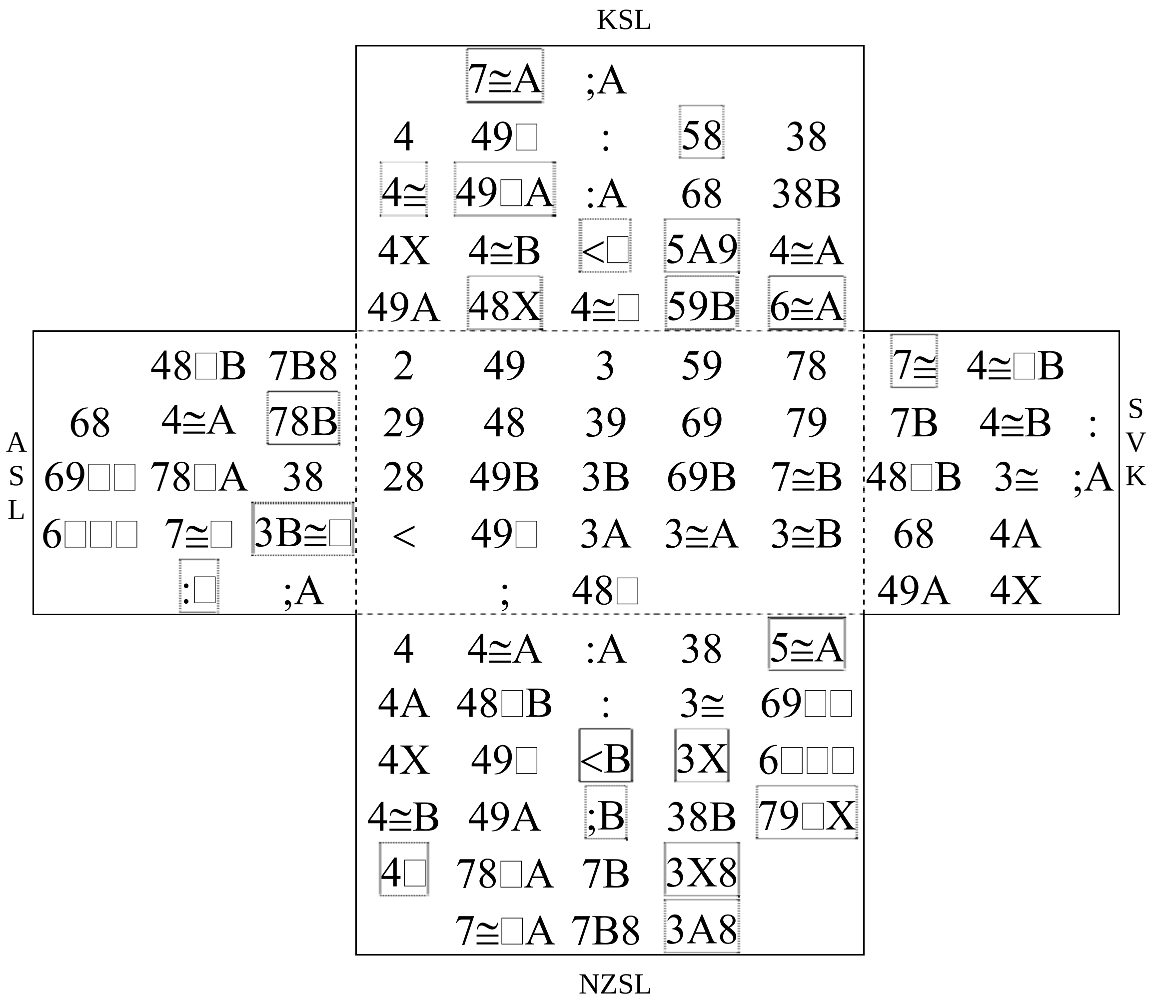
The 21 handshapes in (3-25) occur in only one language each.

(3-25) Handshapes occurring in only one language

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| 4≅ | 49□A | <B | 58 | 79□X | 3X | 7≅ | ;B |
| B48X | 4□ | <□ | 5A9 | | 3A8 | 78B | |
| | | :□ | 59B | | 3X8 | 7≅A | |
| B | | | 5≅A | | 3B≅□ | | |
| | | | 6≅A | | | | |

The diagram in (3-26) shows the handshape inventories of all four languages. The rectangle in the center of the diagram contains the shared handshapes. The handshapes unique to each language are outlined with dotted lines. Other handshapes in the peripheral rectangles for each language are shared with at least one other language.

(3-26) Handshape inventories of all four languages




### 3.2 Distributions of handshapes in a lexicon

After determining the size and content of the handshape inventories, the next step is to investigate the way in which a language uses its phonological resources. In this case the use of a phonological resource, the handshape, is considered within the lexicon. In theory two languages could have the same inventories, but they could utilize their inventories in very different ways. In the next two sections, the issues of rank and distribution are examined.

#### 3.2.1 Handshape rank

To determine the ranking of the handshapes it is necessary to count how many signs are formed with each handshape. As in the determination of inventory, only monomorphemic, native signs without a handshape change were counted. It is possible for two language with the same handshape inventory to rank these inventories in different ways.[62] For example, consider two hypothetical signs languages, SLA and SLB, both having the same inventory {3 49 2 78}. Suppose that SLA ranks these four handshapes {3 49 2 78}, while SLB ranks them {78 2 49 3}. This is to say that 3 is the most common handshape in SLA, 49 is the second most common, 2 the third most common, and 78 the least common. Thus, the lexicons of these languages would differ in that 3 would be most prevalent in lexicon A while 78 would be most prevalent in lexicon B.

Since the four languages investigated do not share all the same handshapes, it is difficult to compare handshape rankings cross-linguistically. For this reason, only the rankings of the shared handshapes are considered. Many of the shared handshapes are among the most common handshapes. The 22 shared handshapes are used in 86% of the ASL database, 85% of the KSL database, 79% of the NZSL database, and 87% of the SVK database (while these 22 handshapes account for 65%, 51%, 46% and 67% of the inventories, respectively). However, it is not the case that the shared handshapes are necessarily the most common ones. For example, 3≃A occurs in only one sign in ASL, NZSL and SVK, and 79 occurs in only two signs in NZSL and KSL. The rankings of the shared handshapes are shown in (3-27).

---

62 Krámský (1976b) presents a quantitative comparison of vowel frequency in 29 languages with the same vowel inventory, [i u e o a].

(3-27) Ranking of shared handshapes

| rank | ASL | KSL | NZSL | SVK |
|---|---|---|---|---|
| 1 | 3 | 3 | 3 | 3 |
| 2 | 49 | 49 | 49 | 49 |
| 3 | 78 | 29 | 78 | 78 |
| 4 | 28 | 28 | 2 | 28 |
| 5 | 3A | 59 | 29 | 29 |
| 6 | 29 | < | 28 | 3B |
| 7 | 69 | 3B | 7≅B | 69 |
| 8 | 4B9 | 49B | < | 7≅B |
| 9 | < | 3≅B | 3A | 2 |
| 10 | 59 | 3A | 39 | 59 |
| 11 | 3≅B | 2 | 69 | 69B |
| 12 | 6B9 | 48□ | 49B | 49B |
| 13 | 2 | 7≅B | 3B | 39 |
| 14 | 39 | 78 | 3≅B | 3A |
| 15 | 48□ | ; | ; | < |
| 16 | 7≅B | 49□ | 49□ | 48□ |
| 17 | 79 | 48 | 69B | 3≅B |
| 18 | 49□ | 69 | 59 | 79 |
| 19 | 48 | 39 | 48□ | 49□ |
| 20 | ; | 59B | 48 | ; |
| 21 | 3B | 3≅A | 79 | 3≅A |
| 22 | 3≅A | 79 | 3≅A | 48 |

The shared handshapes are not ranked the same for any two languages, as shown in (3-27). While 3 and 49 are the first and second most frequent handshapes in each language, there is already variation in what is the third most frequent handshape. Yet, it seems to be possible to say that some handshapes are frequent in all four languages and others are infrequent. To formalize this observation, the Spearman rank correlation test is used to determine whether rankings are independent or whether they vary in the same or opposite directions (Snedecor and Cochran, 1982). The rank correlation coefficient, $r_s$, ranges from –1 for complete discordance to +1 for complete concordance of two rankings. Applying the Spearman rank correlation test to the data in (3-27) shows that all pairs of rankings are highly concordant, as detailed in (3-28). Thus, though the inventories are of different size and the shared handshapes are distributed throughout the entire ranked inventory, the shared handshapes are ranked very similarly across these four languages.

(3-28) Spearman rank correlation test for handshape:

$$r_s = \sum_{i=1}^{n}(x_i - \overline{X})(y_i - \overline{Y})^2 \Big/ \sqrt{\sum_{i=1}^{n}(x_i - \overline{X})^2 \sum_{i=1}^{n}(y_i - \overline{Y})^2} \qquad \overline{X} = \overline{Y} = 11.5, \quad n = 22$$

$X = \{x_1, x_2, \ldots, x_{22}\}$ and $Y = \{y_1, y_2, \ldots, y_{22}\}$ are sets of rankings of the 22 shared handshapes for the two languages being compared.

It must be that $r_s > 0.537$ for correlation to be significant at the 0.01 level.

| | ASL | KSL | NZSL | SVK |
|---|---|---|---|---|
| ASL | 1 | 0.616 | 0.743 | 0.723 |
| KSL | | 1 | 0.625 | 0.613 |
| NZSL | | | 1 | 0.806 |
| SVK | | | | 1 |

### 3.2.2 Frequency distribution

Consider three hypothetical sign languages with the same inventories and non-differing handshape rankings {3 4 ; <}. In sign language A, all handshapes are used with equal frequency in the lexicon. In theory, this is a reasonable strategy; if a language has committed articulatory effort into learning to produce a handshape and perceptual effort into learning to distinguish and identify it, maximal use should be made of that resource. This scenario is shown in (3-29) for a hypothetical language with a lexicon of one hundred signs using four handshapes.

(3-29) SLA: hypothetical language A using four handshapes

| rank | handshape | count | frequency |
|---|---|---|---|
| 1 | 3 | 25 | 0.25 |
| 2 | 49 | 25 | 0.25 |
| 3 | ; | 25 | 0.25 |
| 4 | < | 25 | 0.25 |
| sum | | 100 | 1.00 |

Handshapes are ranked in the first column from most to least common; since all handshapes are used equally often, the order is moot and the ranking is arbitrary.[63] The second column shows the handshape, the third column shows the number of signs using the handshape, and the fourth column shows the frequency with which the handshape is used. The frequency $f_i$ of handshape $h_i$ is defined as $f_i = n_i / N$, where $n_i$ is the number of signs using handshape $h_i$ and N is the total number of signs. Note that $\sum_{i=1}^{N} f_i = 1$.[64]

In hypothetical sign languages B and C, the handshapes are used with unequal frequencies, as shown in (3-30) and (3-31). In SLB, the frequencies decrease linearly. In SLC,

---

[63] There are other methods for assigning serial rankings to a list of frequencies. The method used here is sometimes called "order" instead of "rank". A second method assigns the same rank to all elements occurring with identical frequency, without regard for the number of elements occurring with that frequency. In this case, each of the handshapes in (3-29) would be have the same rank, for example, "1". A third method also assigns the same rank to all elements occurring with identical frequency, but this method takes into account the number of elements occurring with this frequency. Sequential ranks are assigned to the set of arbitrarily arranged elements with identical frequencies; then these ranks are averaged. For example, all four handshapes in (3-29) would be assigned the rank (1 + 2 + 3 + 4) / 4 = 2.5. See Têšitelová (1992) for further discussion.

[64] It is possible to use the absolute frequency of occurrence of an element, that is, the number of elements (signs) in the sample (lexicon) bearing the feature in question (a particular handshape). To facilitate cross-linguistic comparison as well as later computations of entropy, relative frequency, that is, absolute frequency divided by sample size, is used.

the first two handshapes, which are unmarked, are used frequently, while the last two handshapes, which are marked, are used infrequently.

(3-30) SLB: hypothetical sign language B

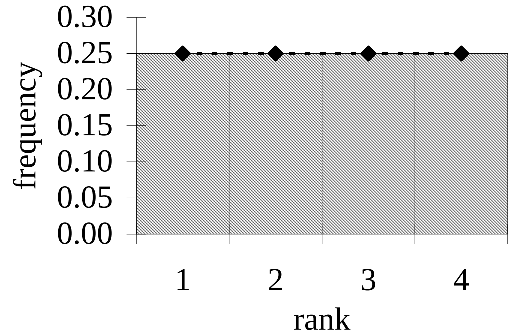
| rank | handshape | count | frequency |
|---|---|---|---|
| 1 | 3 | 40 | 0.40 |
| 2 | 49 | 30 | 0.30 |
| 3 | ; | 20 | 0.20 |
| 4 | < | 10 | 0.10 |
| sum | | 100 | 1.00 |

(3-31) SLC: hypothetical sign language C

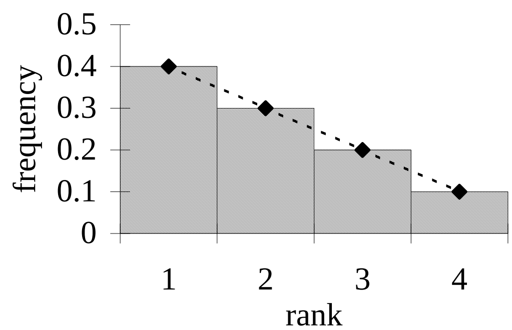
| rank | handshape | count | frequency |
|---|---|---|---|
| 1 | 3 | 40 | 0.40 |
| 2 | 49 | 40 | 0.40 |
| 3 | ; | 10 | 0.10 |
| 4 | < | 10 | 0.10 |
| sum | | 100 | 1.00 |

These data can be expressed graphically by means of a rank-frequency graph. A rank-frequency graph is a two-dimensional graph that plots the rank on the horizontal x-axis and the frequency on the vertical y-axis. Rank-frequency charts for SLA, SLB and SLC are shown in (3-32), (3-33), and (3-34).

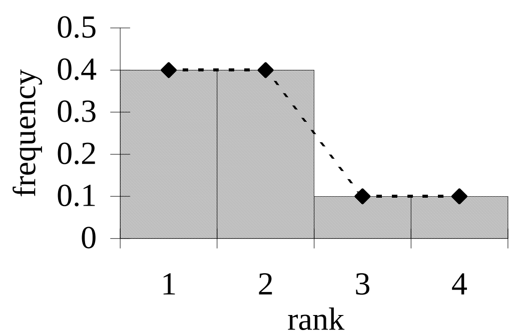
(3-32) Rank-frequency graph for SLA




(3-33) Rank-frequency graph for SLB




(3-34) Rank-frequency graph for SLC




With the notion of rank-frequency as a tool, we can consider the following questions. How are handshapes distributed in the lexicons of ASL, KSL, NZSL and SVK? Do these distributions vary? If they do vary, how can the distributions be compared?

#### 3.2.2.1 ASL: handshape rank-frequency

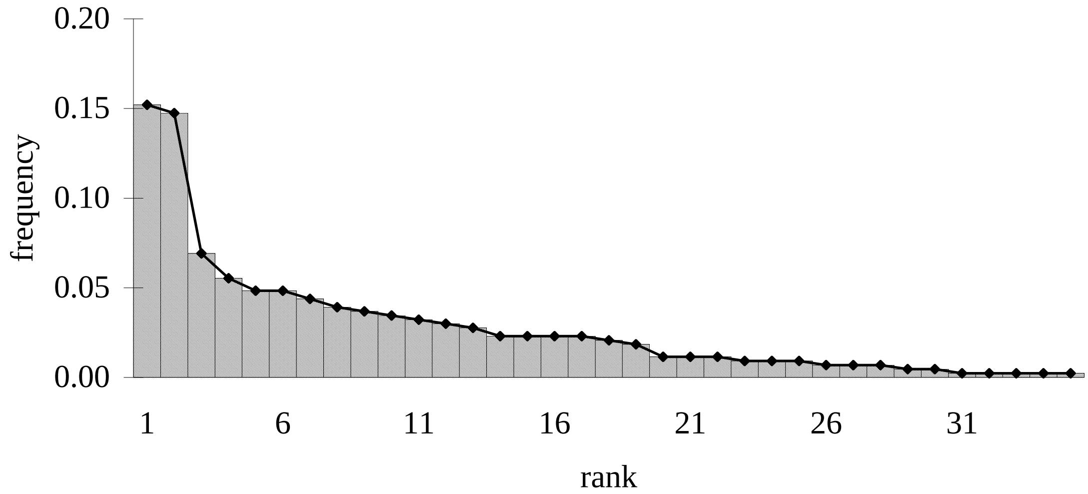
The distribution of handshape in the ASL lexicon is shown in (3-35). Note that the five most common handshapes, {3 49 78 28 3A}, comprise 47% of non-handshape changing signs in the ASL database, while the five least common handshapes, {4≅A 3B≅□ 3B 3≅A 7≅□B}, comprise only 1.2% of the sample. In the sample, each of these five least common handshapes appears in only one monomorphemic, non-handshape-changing sign, THANKSGIVING, GENIUS, HOW, MAGAZINE, EARTH, respectively. Note that it is not the case in ASL, nor in the other three languages, that the signs using rare handshapes are obviously recent additions to the lexicon as they do not refer to more recently encountered referents, such as, for example, *computer* or *cell phone*.

(3-35) ASL: handshape frequency distribution

| rank | hs | count | freq | rank | hs | count | freq |
|---|---|---|---|---|---|---|---|
| 1 | 3 | 66 | 0.1521 | 21 | 69□□ | 5 | 0.0115 |
| 2 | 49 | 64 | 0.1475 | 22 | 4B8□ | 5 | 0.0115 |
| 3 | 78 | 30 | 0.0691 | 23 | 68 | 4 | 0.0092 |
| 4 | 28 | 24 | 0.0553 | 24 | 49□ | 4 | 0.0092 |
| 5 | 3A | 21 | 0.0484 | 25 | 38 | 4 | 0.0092 |
| 6 | 29 | 21 | 0.0484 | 26 | 6□□□ | 3 | 0.0069 |
| 7 | 69 | 19 | 0.0438 | 27 | 48 | 3 | 0.0069 |
| 8 | 4B9 | 17 | 0.0392 | 28 | :□ | 3 | 0.0069 |
| 9 | < | 16 | 0.0369 | 29 | 78B | 2 | 0.0046 |
| 10 | 59 | 15 | 0.0346 | 30 | ; | 2 | 0.0046 |
| 11 | 3≅B | 14 | 0.0323 | 31 | 4≅A | 1 | 0.0023 |
| 12 | 6B9 | 13 | 0.0300 | 32 | 3B≅□ | 1 | 0.0023 |
| 13 | 2 | 12 | 0.0276 | 33 | 3B | 1 | 0.0023 |
| 14 | 7B8 | 10 | 0.0230 | 34 | 3≅A | 1 | 0.0023 |
| 15 | 78□A | 10 | 0.0230 | 35 | 7≅□B | 1 | 0.0023 |
| 16 | 39 | 10 | 0.0230 | sum | | 434 | 1.0000 |
| 17 | ;A | 10 | 0.0230 | | | | |
| 18 | 48□ | 9 | 0.0207 | | | | |
| 19 | 7≅B | 8 | 0.0184 | | | | |
| 20 | 79 | 5 | 0.0115 | | | | |

The handshape rank-frequency graph for ASL is shown in (3-36). Recall that the x-axis refers to the rank of the handshape, while the y-axis refers to the frequency at which that handshape occurs, as given in (3-35).

(3-36) ASL: handshape rank-frequency graph

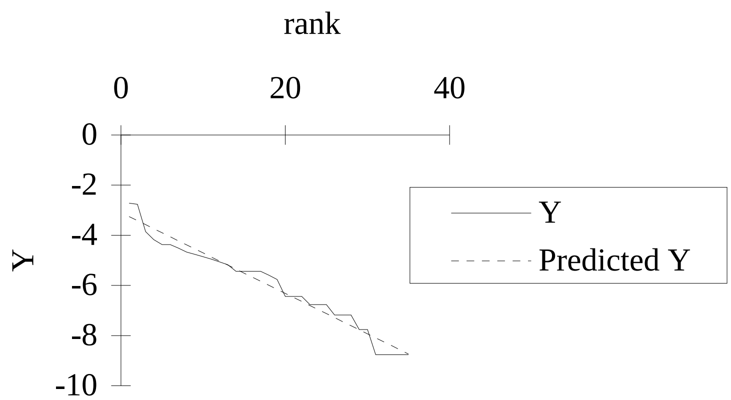



The rank-frequency curve is an exponential decay curve, that is, it is the graph of an exponential function of the form $y = a\ 2^{-c\,x}$. When the handshape rank is plotted against the logarithm base two of the handshape frequency, the resulting graph is close to linear, as shown in (3-37).[65] Linear regression applied to x = rank and Y = $\log_2$(frequency) results in an adjusted correlation coefficient R square value of 0.9627, indicating that handshape frequency declines with close fidelity to the exponential decay law, where $y = 0.12(2^{-0.16x})$ for the ASL graph. An R square value above 0.70 is considered a strong correlation.

---

65 $y = a\ 2^{-cx}$ implies $\log(y) = \log(a2^{-cx})$. Since $\log(a2^{-cx}) = \log a + \log(2^{-cx}) = \log a + (-cx)\log 2 = \log a - cx$, it is the case that $\log(y) = \log a - cx$, which is a linear equation of the form $Y = mx + b$, where $Y = \log y$, the intercept $b = \log a$, and the slope $m = -c$

(3-37) ASL: $Y = \log_2 y$ where x = rank and y = frequency




### 3.2.2.2 KSL: handshape rank-frequency

The distribution of handshape in the lexicon of Korean Sign Language is shown in (3-38). The five most common handshapes are { 3 49 29 28 59}; they comprise 56% of non-handshape changing signs in the KSL database. As in ASL, the least common handshapes each occur in only one monomorphemic, non-handshape-changing sign. The five least common handshapes comprise only 1.2% of the sample. The thirteen handshapes that appear once, the handshapes ranked 32 through 44, occur in the signs MOONGCHIDA *be united*, BOTONG *average*, KKIDO *prayer*, DDEUGAEJIL *knitting*, GOGEUB *seniority*, JAMBA *jacket*, GYOTONGSOONGYUNG *traffic cop*, GUHKKURO *backward*, GYOMI *copulation*, GONG *ball*, INSAM *ginseng*, and SASI *all year*, respectively. In some of these signs, the handshape represents some iconic aspect of the referent, such as the zipper pull in JAMBA *jacket*, the knitting needles in DDEUGAEJIL *knitting*, and the plant's roots in INSAM *ginseng*. Therefore, these signs might be examples of lexicalized Size and Shape Specifiers or classifier constructions whose handshapes are not phonological components of monomorphemic signs, as was suggested for the handshapes 79 and 79□X in VSVK.

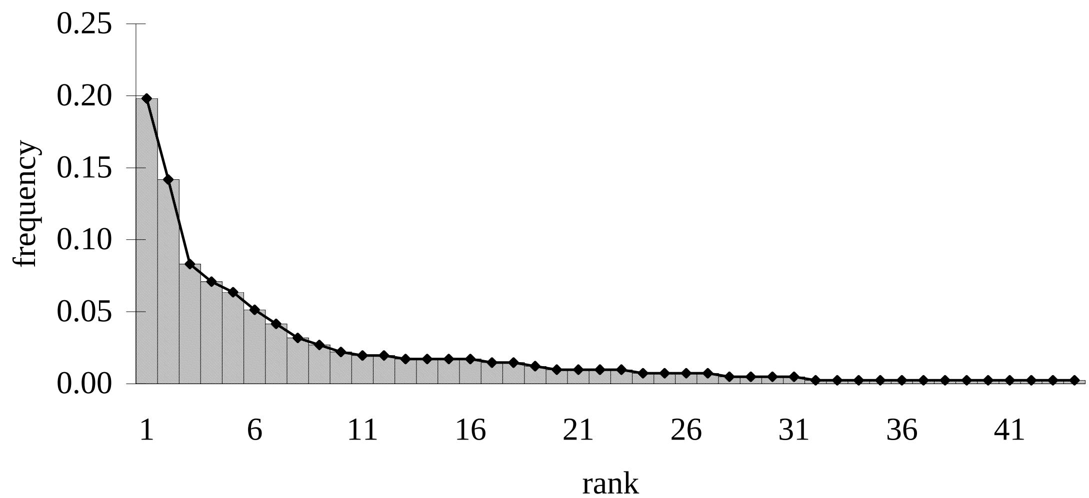
(3-38) KSL: handshape frequency distribution

| rank | hs | count | freq |
|---|---|---|---|
| 1 | 3 | 81 | 0.1980 |
| 2 | 49 | 58 | 0.1418 |
| 3 | 29 | 34 | 0.0831 |
| 4 | 28 | 29 | 0.0709 |
| 5 | 59 | 26 | 0.0636 |
| 6 | < | 21 | 0.0513 |
| 7 | 3B | 17 | 0.0416 |
| 8 | : | 13 | 0.0318 |
| 9 | 49B | 11 | 0.0269 |
| 10 | :A | 9 | 0.0220 |
| 11 | 3≅B | 8 | 0.0196 |
| 12 | 3A | 8 | 0.0196 |
| 13 | 2 | 7 | 0.0171 |
| 14 | 48□ | 7 | 0.0171 |
| 15 | 7≅B | 7 | 0.0171 |
| 16 | 78 | 7 | 0.0171 |
| 17 | ; | 6 | 0.0147 |
| 18 | ;A | 6 | 0.0147 |
| 19 | 49□ | 5 | 0.0122 |
| 20 | 38 | 4 | 0.0098 |
| 21 | 4 | 4 | 0.0098 |
| 22 | 48 | 4 | 0.0098 |
| 23 | 69 | 4 | 0.0098 |
| 24 | 39 | 3 | 0.0073 |
| 25 | 4≅A | 3 | 0.0073 |

| rank | hs | count | freq |
|---|---|---|---|
| 26 | 4≅B | 3 | 0.0073 |
| 27 | 59B | 3 | 0.0073 |
| 28 | 3≅A | 2 | 0.0049 |
| 29 | 49A | 2 | 0.0049 |
| 30 | 4B≅ | 2 | 0.0049 |
| 31 | 69B | 2 | 0.0049 |
| 32 | <□ | 1 | 0.0024 |
| 33 | 38B | 1 | 0.0024 |
| 34 | 4≅ | 1 | 0.0024 |
| 35 | 48□X□X | 1 | 0.0024 |
| 36 | 49A□ | 1 | 0.0024 |
| 37 | 49□ | 1 | 0.0024 |
| 38 | 4X | 1 | 0.0024 |
| 39 | 58 | 1 | 0.0024 |
| 40 | 5A9 | 1 | 0.0024 |
| 41 | 68 | 1 | 0.0024 |
| 42 | 6≅A | 1 | 0.0024 |
| 43 | 7≅A | 1 | 0.0024 |
| 44 | 79 | 1 | 0.0024 |
| sum | | 409 | 1.0000 |

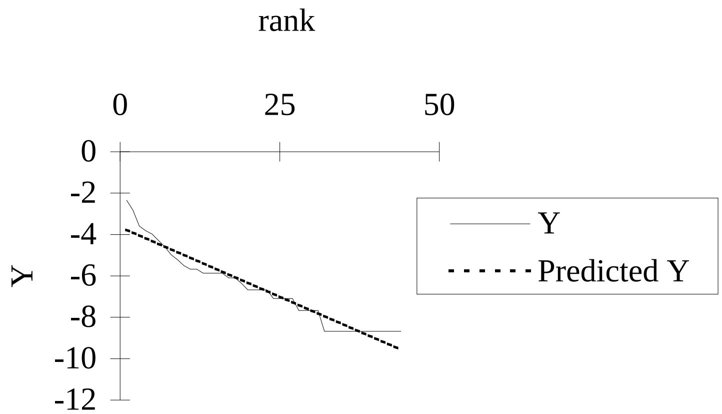
The handshape rank-frequency graph for KSL is shown in (3-39). Recall that the x-axis refers to the rank of the handshape, while the y-axis refers to the frequency at which that handshape occurs, as given in (3-38). Note that the handshape whose frequency is represented by, for example, the sixth bar for the graph of KSL in (3-39) is not necessarily the same handshape represented by the sixth bar in (3-36), the ASL handshape rank-frequency graph.

(3-39) KSL: handshape rank-frequency graph




The rank-frequency graph of KSL is also an exponential decay curve, just as the ASL rank-frequency graph was. When the handshape rank is plotted against the logarithm base two of the handshape frequency, the resulting graph is close to linear, as shown in (3-40). Linear regression applied to x = rank and Y = $\log_2$(frequency) results in an adjusted correlation coefficient R square value of 0.9337, indicating that handshape frequency declines with close fidelity to the exponential decay law, where $y = 0.08(2^{-0.13x})$ for the KSL graph.

(3-40) KSL: $Y = \log_2 y$ where x = rank and y = frequency




### 3.2.2.3 NZSL: handshape rank-frequency

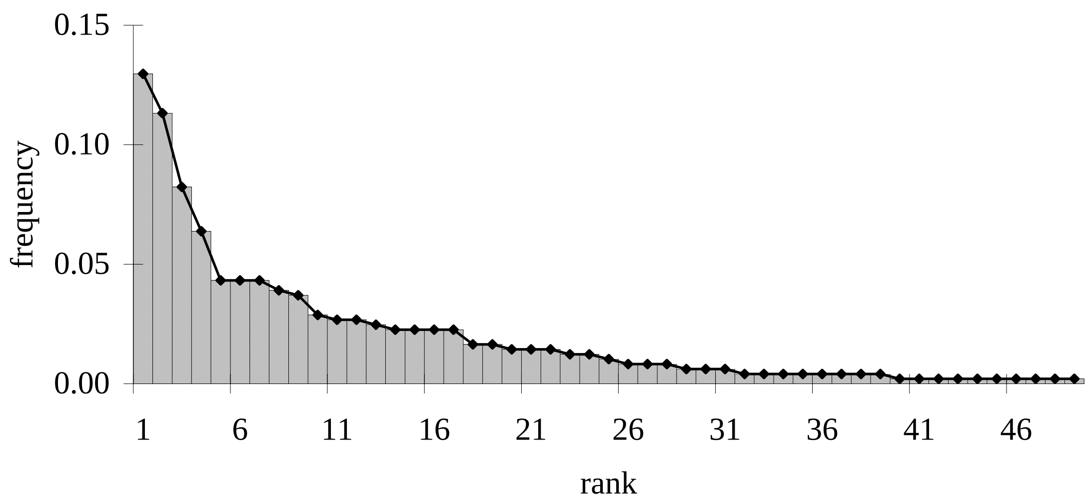
The distribution of handshape in the lexicon of New Zealand Sign Language is shown in (3-41). The five most common handshapes are {3 49 78 2 38}; they comprise 43 % of non-handshape changing signs in the NZSL database. The least common handshapes each occur in only one monomorphemic, non-handshape-changing sign. The five least common handshapes comprise only 1.0% of the sample. The ten handshapes that appear once, the handshapes ranked 40 through 49, occur in the signs TRIPLE, DOCTOR, LAZY, SOUR, BE DEFEATED, DRAWER, LUNG, BET, FINGER, and PLUM, respectively. Again, in some of these signs, the handshape represents some iconic aspect of the referent, as in TRIPLE, DRAWER, LUNG and FINGER.

(3-41) NZSL: handshape frequency distribution

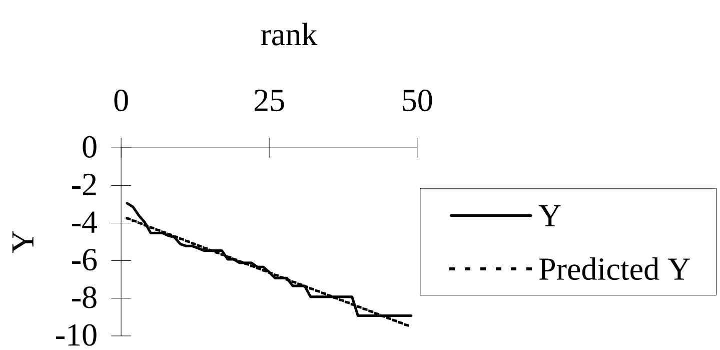
| rank | hs | count | freq | rank | hs | count | freq |
|---|---|---|---|---|---|---|---|
| 1 | 3 | 63 | 0.1296 | 26 | 69B | 4 | 0.0082 |
| 2 | 49 | 55 | 0.1132 | 27 | 59 | 4 | 0.0082 |
| 3 | 78 | 40 | 0.0823 | 28 | :A | 4 | 0.0082 |
| 4 | 2 | 31 | 0.0638 | 29 | 78□A | 3 | 0.0062 |
| 5 | 38 | 21 | 0.0432 | 30 | 48□ | 3 | 0.0062 |
| 6 | 29 | 21 | 0.0432 | 31 | 48 | 3 | 0.0062 |
| 7 | 28 | 21 | 0.0432 | 32 | 7B8 | 2 | 0.0041 |
| 8 | 4X | 19 | 0.0391 | 33 | 79 | 2 | 0.0041 |
| 9 | 7≅B | 18 | 0.0370 | 34 | 69□□ | 2 | 0.0041 |
| 10 | < | 14 | 0.0288 | 35 | 6□□□ | 2 | 0.0041 |
| 11 | 4 | 13 | 0.0267 | 36 | 4A | 2 | 0.0041 |
| 12 | 3A | 13 | 0.0267 | 37 | 4≅A | 2 | 0.0041 |
| 13 | 39 | 12 | 0.0247 | 38 | 3A8 | 2 | 0.0041 |
| 14 | 69 | 11 | 0.0226 | 39 | 3≅ | 2 | 0.0041 |
| 15 | 49B | 11 | 0.0226 | 40 | 79□X | 1 | 0.0021 |
| 16 | 3B | 11 | 0.0226 | 41 | 5≅A | 1 | 0.0021 |
| 17 | 3≅B | 11 | 0.0226 | 42 | 49□ | 1 | 0.0021 |
| 18 | 4≅B | 8 | 0.0165 | 43 | 4□ | 1 | 0.0021 |
| 19 | ; | 8 | 0.0165 | 44 | 3X8 | 1 | 0.0021 |
| 20 | 7B | 7 | 0.0144 | 45 | 3X | 1 | 0.0021 |
| 21 | <B | 7 | 0.0144 | 46 | 38B | 1 | 0.0021 |
| 22 | : | 7 | 0.0144 | 47 | 3≅A | 1 | 0.0021 |
| 23 | 49A | 6 | 0.0123 | 48 | ? | 1 | 0.0021 |
| 24 | 49□ | 6 | 0.0123 | 49 | ;B | 1 | 0.0021 |
| 25 | 4B8 | 5 | 0.0103 | sum | | 486 | 1.0000 |

The handshape rank-frequency graph for NZSL is shown in (3-42). This graph is also an exponential decay curve, as shown by the linearity of the graph in (3-43). Linear regression applied to x = rank and Y = log $_2$ (frequency) results in an adjusted correlation coefficient R square value of 0.9736, indicating that handshape frequency declines with close fidelity to the exponential decay law, where $y = 0.08(2^{-0.12x})$ for the NZSL graph.

(3-42) NZSL: handshape rank-frequency graph




(3-43) NZSL: Y = log $_2$ y where x = rank and y = frequency



#### 3.2.2.4 SVK: handshape rank-frequency

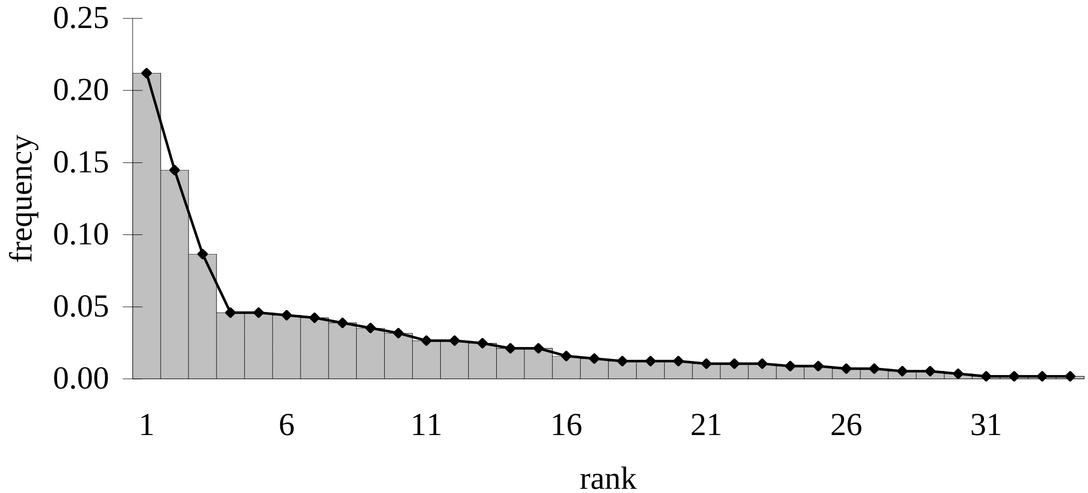
The distribution of handshape in the lexicon of Suomalaisen Viittomakieli is shown in (3-44). The five most common handshapes are {3 49 78 28 29}; they comprise 52% of non-handshape changing signs in the SVK database. The five least common handshapes, {3≅ 3≅A 48 7B 4A}, each occur in only one monomorphemic, non-handshape-changing sign, and they comprise only 1.0% of the sample. These five handshapes occur in the signs LUOKKA *class*, SAIRASTUA *get sick*, VIDEOIDA *video*, AHKERA *diligent*, and PELÄTÄ *fear*, respectively. In none of these signs does the handshape clearly represent an iconic aspect of the referent. However, 48, which was used in VSVK in only one compound sign, KOETTAA *trial*, is used in the sign VIDEOIDA *video*, which must have been a more recent addition to the SVK lexicon.

(3-44) SVK: handshape frequency distribution

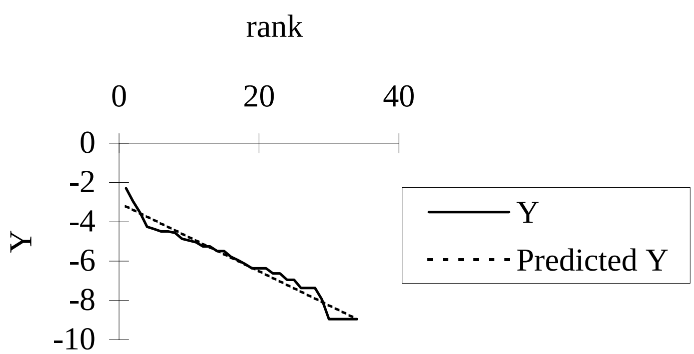
| rank | hs | count | freq |
|---|---|---|---|
| 1 | 3 | 101 | 0.2040 |
| 2 | 49 | 64 | 0.1293 |
| 3 | 78 | 43 | 0.0869 |
| 4 | 78B | 26 | 0.0525 |
| 5 | 69 | 24 | 0.0485 |
| 6 | 29 | 22 | 0.0444 |
| 7 | 3B | 22 | 0.0444 |
| 8 | 28 | 21 | 0.0424 |
| 9 | 4X | 17 | 0.0343 |
| 10 | 2 | 16 | 0.0323 |
| 11 | 69B | 15 | 0.0303 |
| 12 | 49B | 13 | 0.0263 |
| 13 | 59 | 13 | 0.0263 |
| 14 | 39 | 11 | 0.0222 |
| 15 | 3A | 11 | 0.0222 |
| 16 | ;A | 9 | 0.0182 |
| 17 | 4B8□ | 8 | 0.0162 |
| 18 | < | 7 | 0.0141 |
| 19 | : | 6 | 0.0121 |
| 20 | 48□ | 6 | 0.0121 |
| 21 | 68 | 6 | 0.0121 |
| 22 | 3≅B | 5 | 0.0101 |
| 23 | 49A | 5 | 0.0101 |
| 24 | 4B≅□ | 4 | 0.0081 |
| 25 | 79 | 4 | 0.0081 |
| 26 | 4≅B | 3 | 0.0061 |
| 27 | 49□ | 3 | 0.0061 |
| 28 | 7≅ | 3 | 0.0061 |
| 29 | ; | 2 | 0.0040 |
| 30 | 3≅ | 1 | 0.0020 |
| 31 | 3≅A | 1 | 0.0020 |
| 32 | 48 | 1 | 0.0020 |
| 33 | 7B | 1 | 0.0020 |
| 34 | 4A | 1 | 0.0020 |
| | | 495 | 1.0000 |

The handshape rank-frequency graph for SVK is shown in (3-45). This graph is also an exponential decay curve, as shown by the linearity of the graph in (3-46). Linear regression applied to x = rank and Y = $\log_2$(frequency) results in an adjusted correlation coefficient R square value of 0.9608, indicating that handshape frequency declines with close fidelity to the exponential decay law, where $y = 0.17(2^{-0.12x})$ for the SVK graph.

(3-45) SVK: handshape rank-frequency graph




(3-46) SVK: $Y = \log_2 y$ where x = rank and y = frequency




### 3.2.2.5 VSVK: handshape rank-frequency

The distribution of handshape in the lexicon of Vanha Suomalaisen Viittomakieli is shown in (3-47). The five most common handshapes are {3 49 29 3A 49A}; they comprise 61% of non-handshape changing signs in the VSVL database. The five least common handshapes comprise 2.5% of the sample. The four handshapes that appear only once, {2□ 59B 7 7B}, occur in the signs VESI *water*, ISTUA *sit*, VALEHDELLA *lie*, and PARTA *beard*, respectively.

(3-47) VSVK: handshape frequency distribution

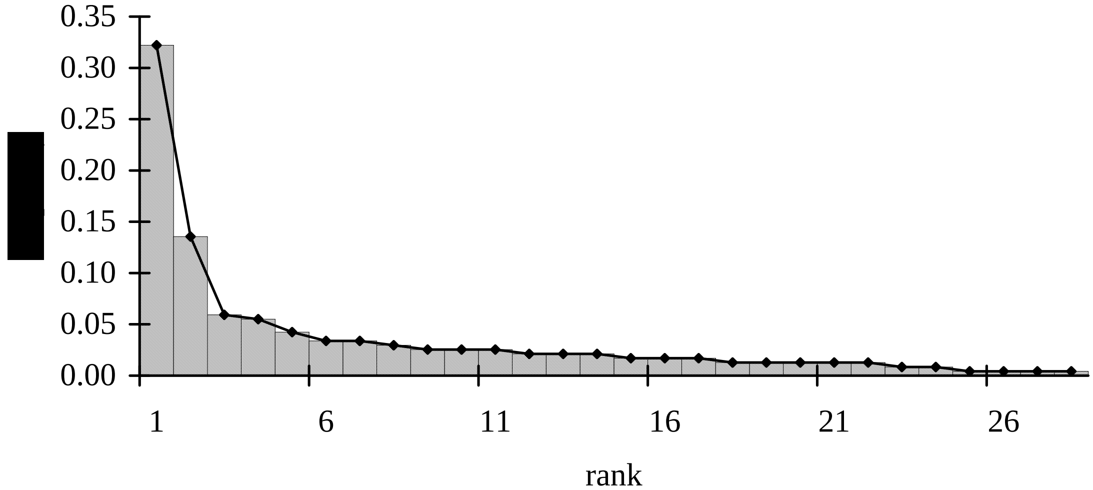
| **rank** | **hs** | **count** | **freq** | **rank** | **hs** | **count** | **freq** |
|---|---|---|---|---|---|---|---|
| 1 | 3 | 76 | 0.3220 | 16 | ; | 4 | 0.0169 |
| 2 | 49 | 32 | 0.1356 | 17 | 2 | 4 | 0.0169 |
| 3 | 29 | 14 | 0.0593 | 18 | 3≅B | 3 | 0.0127 |
| 4 | 3A | 13 | 0.0551 | 19 | 39 | 3 | 0.0127 |
| 5 | 49A | 10 | 0.0424 | 20 | 3B | 3 | 0.0127 |
| 6 | 4B | 8 | 0.0339 | 21 | 49□ | 3 | 0.0127 |
| 7 | 78 | 8 | 0.0339 | 22 | 49B | 3 | 0.0127 |
| 8 | 59 | 7 | 0.0297 | 23 | ;A | 2 | 0.0085 |
| 9 | 69 | 6 | 0.0254 | 24 | 4≅A | 2 | 0.0085 |
| 10 | 7≅B | 6 | 0.0254 | 25 | 2□ | 1 | 0.0042 |
| 11 | 78A | 6 | 0.0254 | 26 | 59B | 1 | 0.0042 |
| 12 | : | 5 | 0.0212 | 27 | 7 | 1 | 0.0042 |
| 13 | 28 | 5 | 0.0212 | 28 | 7B | 1 | 0.0042 |
| 14 | 38 | 5 | 0.0212 | sum | | 236 | 1.0000 |
| 15 | :A | 4 | 0.0169 | | | | |

The handshape rank-frequency graph for VSVK is shown in (3-48). This graph is also an exponential decay curve, as shown by the linearity of the graph in (3-49). Linear regression applied to x = rank and Y = $\log_2$(frequency) results in an adjusted correlation coefficient R square value of 0.8760, indicating that handshape frequency declines with reasonable fidelity to the exponential decay law, where $y = 0.17(2^{-0.10x})$ for the VSVK graph.[66]

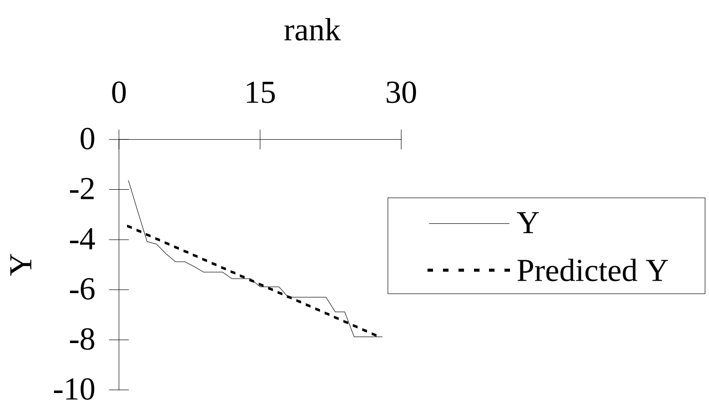
---

[66] The fact that the handshape rank-frequency graph of VSVK is not as close a match to an exponential decay curve has several possible explanations. The sample size for the other languages is almost twice as large as that for VSVK. The photographs are older, thus harder to distinguish, especially for thumb

(3-48) VSVK: handshape rank-frequency graph




(3-49) VSVK: Y = log $_2$ y where x = rank and y = frequency



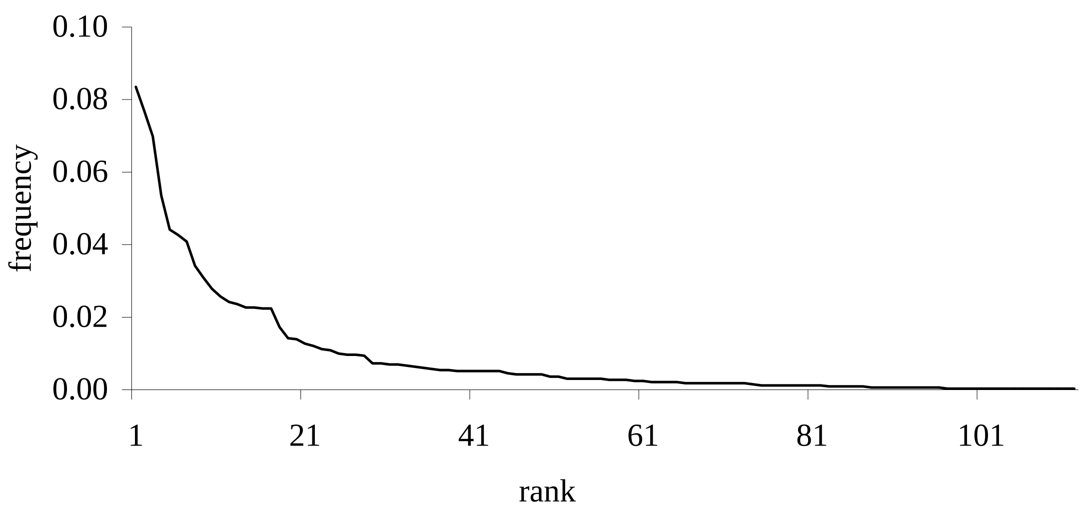

---

extension in 3 versus 38. (Only the SVK and VSVK dictionaries use photographs for the entries; ASL, KSL and NZSL use drawings.) As the language was a younger language one hundred years ago, it might be that younger languages have different handshape distribution patterns. Or it might be that modern languages, whether they are old or young, have this handshape distribution pattern, but languages from the past do not.

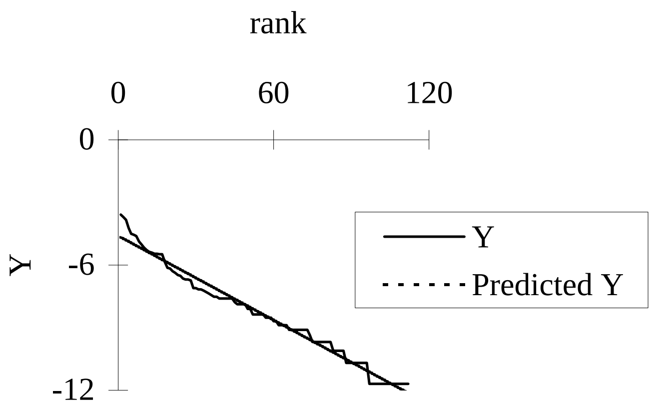
### 3.2.2.6 NGT: handshape rank-frequency

As discussed in Chapter 2, the database for Sign Language of the Netherlands (Nederlandse Gebarentaal: NGT) was created by such different means as to make direct comparison of inventory size and composition impossible. For example, the NGT database, SignPhon, recognizes 112 different handshapes in a sample of 3305 signs. However, it is possible to investigate the rank-frequency graph of NGT for this sample, as shown in (3-50). Despite the great differences in the methods of production for the databases in this dissertation and the SignPhon database, as well as the difference in scale in the graphs (because there are so many handshapes distinguished, each handshape accounts for a smaller fraction of the total), the graphs are remarkably similar. The NGT graph is also an exponential decay curve, as shown by the linearity of the graph in (3-51). Linear regression applied to x = rank and Y = $\log_2$ (frequency) results in an adjusted correlation coefficient R square value of 0.9793, indicating that handshape frequency declines with close fidelity to the exponential decay law, where $y = 0.04(2^{-0.07x})$ for the NGT graph.

(3-50) NGT: handshape rank-frequency graph



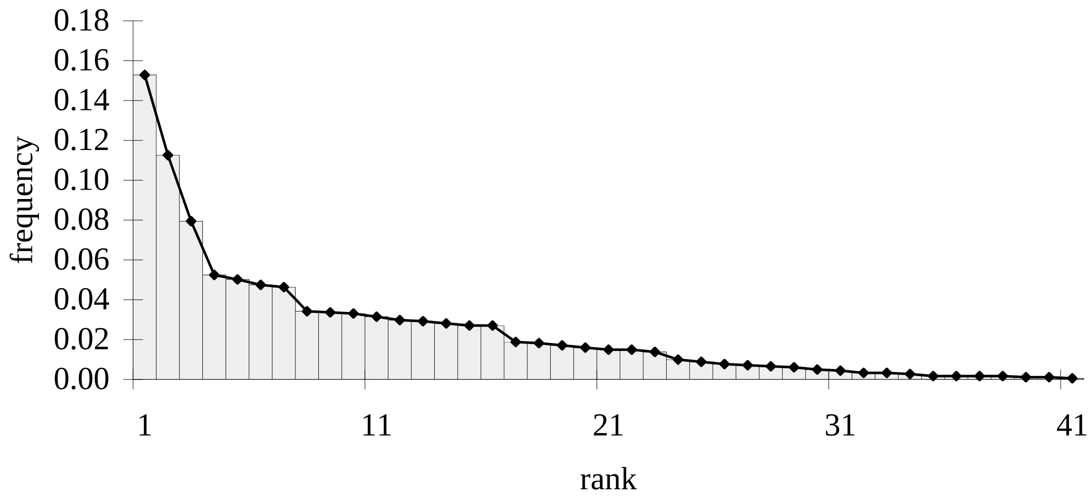
(3-51) NGT: $Y = \log_2 y$ where x = rank and y = frequency



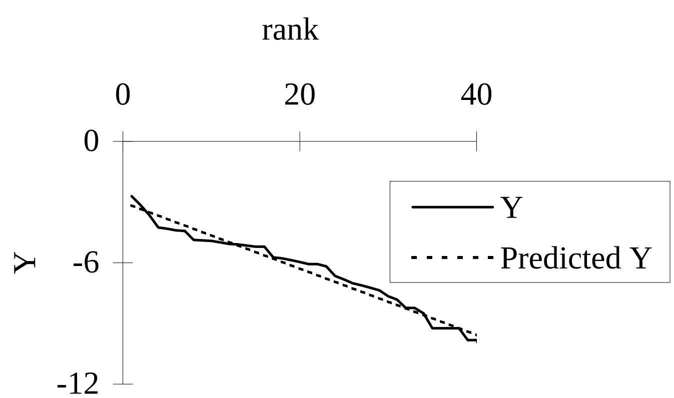

### 3.2.2.7 LIS: handshape rank-frequency

As discussed in Chapter 2, I did not create the Italian Sign Language (Lingua Italiana dei Segni: LIS) database myself. Since differences in methodology might yield different results, no direct comparison of inventory size and composition is made. Yet, in spite of the potential differences that could arise from the different methodologies used, LIS patterns remarkably like ASL, KSL, NZSL, and SVK. The five most common handshapes, {3 49 78 69 29}, comprise 45% of non-handshape changing signs in the LIS database, while the five least common handshapes, {4≅A 59□□□ 7≅□B□B :A 59□□□}, comprise only 0.6% of the sample. The LIS rank-frequency graph is shown in (3-52). Like the other rank-frequency graphs, it is also an exponential decay curve, as shown by the linearity of the graph in (3-53). Linear regression applied to x = rank and $Y = \log_2$ (frequency) results in an adjusted correlation coefficient R square value of 0.9642, indicating that handshape frequency declines with close fidelity to the exponential decay law, where $y = 0.12(2^{-0.16x})$ for the LIS graph.

(3-52) LIS: handshape rank-frequency graph



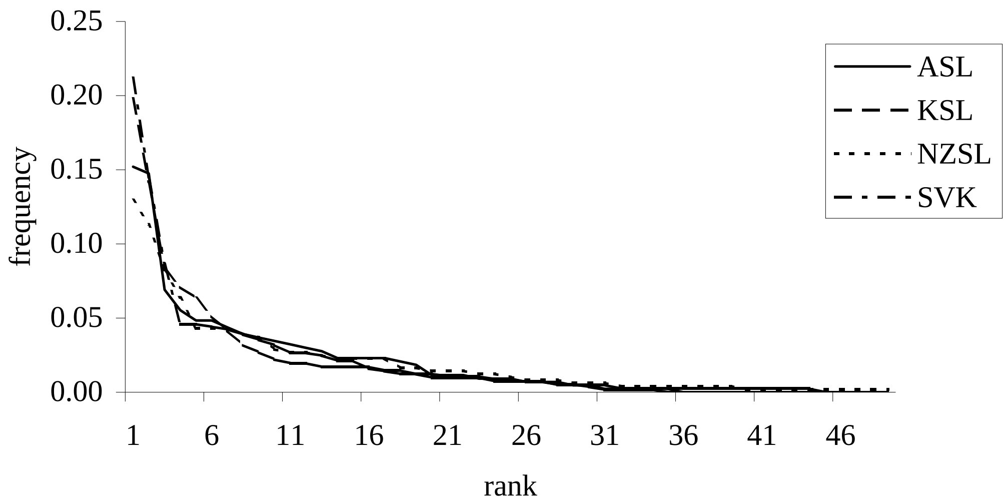

(3-53) LIS: Y= log $_2$ y where x = rank and y = frequency




### 3.2.3 Comparison

The rank-frequency graphs of ASL, KSL, NZSL, and SVK have been shown to be exponential decay curves. How similar are these curves to each other? Is it the case that although all are the same type of curve, they are nevertheless significantly different instantiations of $y = a2^{-cx}$, or is the variation attributable to sampling error? The rank-frequency graph of all four languages is shown in (3-54).[67]

(3-54) Rank-frequency graph for all four languages




Examining two languages at a time, the frequencies associated with each rank were paired, and the signed differences between these frequencies were tested to ascertain whether they can be considered a random sample from a population with mean $\mu = 0$ by using the paired-sample $t$ test. The results of this pair-wise comparison show at the 0.001 level of significance that the differences between each pair of distributions is not significant. Thus, all curves are very similar to each other.

### 3.2.4 Universal handshape pool

Linguists have a reasonably accurate view of the range of sounds available for use in spoken language. This idea is summarized in the chart of the International Phonetic Alphabet

---

[67] Since the languages do not have the same number of handshapes in their inventories, it is necessary to normalize them in some way. I have augmented the handshape frequency distribution tables of ASL, KSL, and SVK with zeroes. For example, ASL contains 35 handshapes while KSL contains 44. In the data used for comparing ASL with KSL, ASL handshapes ranked 36 through 44 are assigned the frequency zero, that is, in ASL the (non-occurring) thirty-sixth handshape occurs with frequency zero. There are other ways to normalize the distribution data, but this was chosen as it is the simplest.

(IPA). The IPA chart was based upon decades, if not centuries, of research into the sounds of hundreds of the world's languages. Pooling the handshapes used in these four sign languages provides a very rough idea of what an analogous chart might look like for signed language. The "International Handshape Alphabet" in (3-55) is an approximation to the set of handshapes available for use in the world's sign languages.

(3-55) "International Handshape Alphabet"

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| 2 | 4 | 4□ | 49□ | : | 69 | 79□X | 3 | 78 | ; |
| 29 | 49 | 49□ | 78□A | :A | 68 | | 38 | 79 | ;B |
| 28 | 48 | 48□ | 7≅□A | :□ | 58 | | 39 | 7≅ | ;A |
| | 4≅ | 49A□ | | < | 59 | | 3≅ | 7B | |
| | 4A | | | <B | 69B | | 3B | 7B8 | |
| | 4X | | | <□ | 6≅A | | 3A | 78B | |
| | 49B | | | | 59B | | 3X | 7≅B | |
| | 49A | | | | 5A9 | | 38B | 7≅A | |
| | 48X | | | | 5≅A | | 3A8 | | |
| | 48□B | | | | 6□□□ | | 3X8 | | |
| | 4≅A | | | | 69□□ | | 3≅B | | |
| | 4≅B | | | | | | 3B≅□ | | |
| | 4≅□B | | | | | | 3≅A | | |

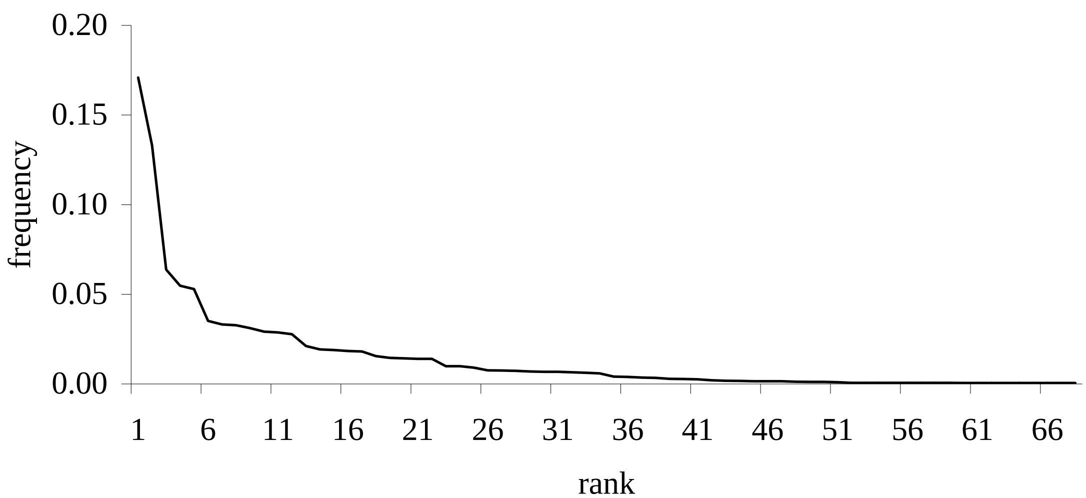
Moreover, pooling the lexicons of the four languages provides an approximation of the universal distribution of handshape, as shown in (3-56).

(3-56) Pooled data: handshape frequency distribution

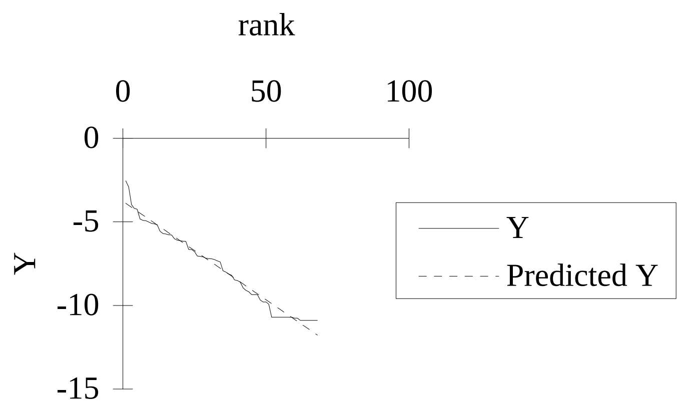
| rank | hs | freq | rank | hs | freq | rank | hs | freq |
|---|---|---|---|---|---|---|---|---|
| 1 | 3 | 0.1709 | 26 | :A | 0.0076 | 51 | 3A8 | 0.0010 |
| 2 | 49 | 0.1330 | 27 | 4≅B | 0.0075 | 52 | 58 | 0.0006 |
| 3 | 78 | 0.0639 | 28 | 78□A | 0.0073 | 53 | <□ | 0.0006 |
| 4 | 29 | 0.0548 | 29 | 4B8□ | 0.0069 | 54 | 4≅ | 0.0006 |
| 5 | 28 | 0.0530 | 30 | 49A | 0.0068 | 55 | 48□X□X | 0.0006 |
| 6 | 2 | 0.0352 | 31 | 7B8 | 0.0068 | 56 | 49A□ | 0.0006 |
| 7 | 59 | 0.0332 | 32 | 79 | 0.0065 | 57 | 5A9 | 0.0006 |
| 8 | < | 0.0328 | 33 | 48 | 0.0062 | 58 | 6≅A | 0.0006 |
| 9 | 69 | 0.0312 | 34 | 68 | 0.0059 | 59 | 7≅A | 0.0006 |
| 10 | 3A | 0.0292 | 35 | 7B | 0.0041 | 60 | ?□ | 0.0006 |
| 11 | 49B | 0.0288 | 36 | 69□□ | 0.0039 | 61 | 3B≅□ | 0.0006 |
| 12 | 3B | 0.0277 | 37 | <B | 0.0036 | 62 | ;B | 0.0005 |
| 13 | 3≅B | 0.0212 | 38 | 4≅A | 0.0034 | 63 | ? | 0.0005 |
| 14 | 39 | 0.0193 | 39 | 3≅A | 0.0028 | 64 | 3X | 0.0005 |
| 15 | 4X | 0.0190 | 40 | 6□□□ | 0.0028 | 65 | 3X8 | 0.0005 |
| 16 | 69B | 0.0184 | 41 | 4B8 | 0.0026 | 66 | 4□ | 0.0005 |
| 17 | 7≅B | 0.0181 | 42 | 4B≅□ | 0.0020 | 67 | 5≅A | 0.0005 |
| 18 | 38 | 0.0156 | 43 | 59B | 0.0018 | 68 | 79□X | 0.0005 |
| 19 | : | 0.0146 | 44 | :□ | 0.0017 | sum | | 1.000 |
| 20 | 78B | 0.0143 | 45 | 3≅ | 0.0015 | | | |
| 21 | 48□ | 0.0140 | 46 | 4A | 0.0015 | | | |
| 22 | ;A | 0.0140 | 47 | 7≅ | 0.0015 | | | |
| 23 | ; | 0.0100 | 48 | 4B≅ | 0.0012 | | | |
| 24 | 49□ | 0.0100 | 49 | 38B | 0.0011 | | | |
| 25 | 4 | 0.0091 | 50 | 49□ | 0.0011 | | | |

The handshape rank-frequency graph for the pooled data, a sample of 1824 signs from all four languages, is shown in (3-57). This graph is also an exponential decay curve, as shown by the linearity of the graph in (3-58). Linear regression applied to x = rank and Y = $\log_2$(frequency) results in an adjusted correlation coefficient R square value of 0.9759, indicating that handshape frequency declines with close fidelity to the exponential decay law, where $y = 0.07(2^{-0.12x})$ for the pooled data.

(3-57) Pooled data: handshape rank-frequency graph




(3-58) Pooled data: Y= $\log_2$ y where x = rank and y = frequency



## 3.3 Handshape by type

In this section the distribution of handshape across the four types of signs will be investigated. Recall that in a Type 0 sign only one hand is used; in a Type 1 sign two hands are used, and both move and have the same handshape; in a Type 2 sign the two hands have the same handshape, but only one moves; in a Type 3 sign the two hands can have different handshapes, but only one moves. The following questions are investigated. Are handshapes distributed evenly in Type 0 and Type 1 signs, that is, are the handshapes used with the same frequency in one-handed and two-handed signs? What does the distribution of handshapes in Type 1 versus Type 3 signs reveal about the difference between nondominant hand locations and other body locations, or about the similarity between these two location categories? Type 1 and Type 2 signs both use two hands with the same handshapes; is the distribution of handshapes used on the dominant hand the same in these two types? Type 2 and Type 3 signs both have a non-moving nondominant hand; is the distribution of handshapes used on the dominant hand the same for these two types? These are fundamental questions about the phonological structure of signs. Cross-linguistic unity or variation in the answers to these questions can inform theory about what must be present universally in sign structure.

Another issue to be investigated is whether markedness plays a role in the distribution of handshape across sign types, or, as will be examined in Chapter 5, across locations. As discussed in section 1.5, there are a number of markedness diagnostics, which unfortunately do not identify a unique set of unmarked handshapes, either cross-linguistically or language-internally. In this dissertation I use the two markedness criteria for which the databases constructed here can offer new information. The first criterion is frequency; unmarked handshapes are supposed to be the most frequent handshapes language-internally as well as cross-linguistically. The most frequent handshapes in the pooled data are shown in (3-59). The top six handshapes were chosen because they account for half of all signs. This set will be referred to as the universal unmarked handshape set. The second criteria is phonological; the handshapes allowed on the nondominant hand in Type 3 signs are supposed to be unmarked. As seen in 2.1.2.2.3, the handshapes allowed on the nondominant hand in Type 3 signs vary cross-linguistically; they are shown in (3-60), in no particular order. These sets will be referred to as language-specific unmarked handshape sets. Note that with the exception of 3≅ and 3A8, the handshapes in each language's set of unmarked handshapes occur in all languages; that is, they are among the 22 shared handshapes, as discussed in section 3.1.7.

(3-59) Universal unmarked set: most frequent handshapes

| handshape | frequency |
|---|---|
| 3 | 0.17 |
| 49 | 0.13 |
| 78 | 0.06 |
| 29 | 0.05 |
| 28 | 0.05 |
| 2 | 0.04 |
| sum | 0.50 |

(3-60) Language-specific unmarked sets: handshapes allowed on H2

| language | handshapes allowed on H2 in Type 3 signs |
|---|---|
| ASL | 3 2 49 78 3≅B ; 39 48 28 |
| KSL | 3 2 49 29 78 3≅B 38 : 28 3B < 59 49□ |
| NZSL | 3 2 49 29 78 3≅B 38 ; : 28 3B 3B8 69 3≅A |
| SVK | 3 2 49 29 38 3≅ |

The first issue considered is whether there is a dependency between the type of sign and the handshape used on the dominant hand or whether handshapes are distributed without regard for sign type. The mutual information significance program described in section 2.2.2 was used to ascertain how much information knowledge of one variable, such as sign type (Type 0, 1, 2, or 3), conveys about the other variable, such as handshape of the dominant hand. A data set consisting of ordered pairs of the form (handshape, type) was created for each language. For example, the ordered pair corresponding to the ASL sign MOTHER is (78, 0), since MOTHER is a one-handed sign (Type 0) with handshape 78. According to this program, ASL and NZSL show a strong dependency between type and handshape (ASL: I(hs; type)=0.269; p=0.001, NZSL: I(hs; type)=0.283, p=0.000). KSL shows a weaker dependence with I(hs; type)=0.281, p=0.050. In contrast, SVK shows independence with I(hs; type)=0.180, p=0.151.[68] To find the source of the dependence it is necessary to examine each language individually.

### 3.3.1 ASL: handshape by type

For ASL, four new data sets were created. The first data set excludes Type 0 signs, the second excludes Type 1 signs, the third excludes Type 2 signs, and the fourth excludes Type 3 signs. The results are shown in (3-61); the first row of this table shows the p value for all types. The only data set that shows independence between type and handshape is the one excluding Type 2 signs, with p=0.066. Also, the three data sets that include only two sign types, one of which is Type 2, are dependent (Types 0,2: p=0.000; Types 1,2: p=0.022; Types 2,3: p=0.003). The three data sets that include only two sign types and exclude Type 2 signs are independent (Types 0,1: p=0.076; Types 0,3: p=0.283; Types 1,3: p=0.124). Thus, it appears that handshapes are distributed similarly in Type 0, 1, and 3 signs, but differently in Type 2 signs. Handshapes are used with the same frequencies throughout signs of Type 0, 1, and 3, but with different frequencies in signs of Type 2. In other words, knowing the type of a particular sign imparts no extra information about the handshape that sign has for the independent set of Types 0, 1 and 3.

---

[68] Recall that the p value was calculated by creating 1,000 scrambled data sets that have the same handshape and type distributions as the original data. The mutual information of the original data set is compared to the mutual information values of the scrambled data sets. For KSL, only 50 out of the 1,000 scrambled data sets have a greater mutual information value. Thus, the probability of a data set having a mutual information value as high as the original set by chance alone is just 0.050. For ASL and NZSL, none of the thousand scrambled data sets had a mutual information value as high as the original data; hence, I report p=0.000. For SVK, 151 out of the 1,000 scrambled data sets have a higher mutual information value; hence, p=0.151, and the two variables are independent of one another.

(3-61) ASL: type by handshape mutual information probabilities

| Types included | p value | Are hs and type dependent variables? |
|---|---|---|
| 0 1 2 3 | 0.000 | dependent |
| 1 2 3 | 0.006 | dependent |
| 0 2 3 | 0.010 | dependent |
| **0 1 3** | **0.066** | **independent** |
| 0 1 2 | 0.001 | dependent |

What is special about the handshapes used in Type 2 signs in ASL? The handshapes used in ASL Type 2 signs in decreasing order of frequency are {3 49 28 29 59 38 :□ < 2 49B 69B }. The first four handshapes are used in 69% of Type 2 signs. A comparison of the observed and expected counts in (3-62) suggests that in ASL, "unmarked" handshapes predominate in Type 2 signs, whether the universal or the language-specific unmarked handshape set is used.[69] To test this observation, each sign in the database was labeled as having a marked or unmarked handshape according to the universal as well as the language-specific unmarked sets in (3-59). Type 2 signs were labeled as such and Type 0, 1 and 3 signs were labeled together as Type "013". Chi-square analyses indicate that Type 2 signs have a different distribution of marked and unmarked handshapes than do the other types of signs when the universal unmarked set is used, with p=0.005. When the language-specific unmarked set is used, p=0.237, showing independence when types and handshapes are so grouped. In this case, the universal unmarked set is better able to capture the dependence between type and marked versus unmarked handshape than the language-specific set.

---

69 $\chi^2 = \sum \frac{(o-e)^2}{e}$, where *o* is the observed count and *e* is the expected count, which is calculated by multiplying the row count by the column count and dividing by the grand total of the whole table.

(3-62) ASL: marked / unmarked handshape by Type 013 / Type 2

**universal unmarked set**

| observed | Type 013 | Type 2 | sum |
|---|---|---|---|
| Marked | 197 | **10** | 207 |
| Unmarked | 179 | **26** | 205 |
| sum | 376 | 36 | 412 |

| expected | Type 013 | Type 2 |
|---|---|---|
| Marked | 189 | 18 |
| Unmarked | 187 | 18 |

p=0.005 dependent

**language-specific unmarked set**

| observed | Type 013 | Type 2 | sum |
|---|---|---|---|
| marked | 185 | 14 | 179 |
| unmarked | 191 | 22 | 233 |
| sum | 376 | 36 | 412 |

| expected | Type 013 | Type 2 |
|---|---|---|
| marked | 182 | 17 |
| unmarked | 194 | 19 |

p=0.237 independent

Note that regardless of markedness criteria, Type 2 signs use fewer marked handshapes than expected. If maximizing a sign language perceiver's opportunity for discerning a marked handshape were a priority, one would expect that Type 2 signs would have more marked handshapes rather than fewer, since the copy of the marked handshape on the nondominant hand arguably doubles the perceiver's opportunity to grasp the more complex signal.[70]

### 3.3.2 KSL: handshape by type

For KSL, four new data sets were created. The first data set excludes Type 0 signs, the second excludes Type 1 signs, the third excludes Type 2 signs, and the fourth excludes Type 3 signs. The results are shown in (3-63); the first row of this table shows the p value for all types. The only data set that shows independence between type and handshape is the one excluding Type 1 signs, with p=0.981. Thus, it appears that handshapes are distributed as expected in Type 0, 2, and 3 signs, but not in Type 1 signs. Also, the three data sets that include only two sign types and exclude Type 1 signs are also independent (Types 0,2: p=0.635; Types 0,3: p=0.991; Types 2,3: p=0.762). Two of the three data sets that include two sign types, one of which is Type 1, are dependent (Types 0,1: p=0.000; Types 1,3: p=0.000). However, the set containing Types 0 and 2 is independent, with p=0.635.

---

[70] As suggested by Kaisse (p.c.), it is perhaps the case that the perceiver's attention is divided by having to look at two hands instead of one.

(3-63) KSL: Type by handshape mutual information probabilities

| Types included | p value | dependency |
|---|---|---|
| 0 1 2 3 | 0.050 | dependent |
| 1 2 3 | 0.048 | dependent |
| **0 2 3** | **0.981** | **independent** |
| 0 1 3 | 0.007 | dependent |
| 0 1 2 | 0.006 | dependent |

What is special about the handshapes used in Type 1 signs? Contrary to expectations, it appears that in KSL the distinction between marked and unmarked handshapes is not relevant, at least not according to either the universal, frequency-based definition or the language-specific, phonological definition. Each sign in the database was labeled as having a marked or unmarked handshape according to the universal as well as the language-specific unmarked sets in (3-59). Then chi-square analyses were preformed on the two 2 x 4 tables of marked/unmarked by Type0/1/2/3 counts. Neither definition of markedness produced evidence of dependence. For the universal set of unmarked handshapes, p=0.9862, while for the language-specific set of unmarked handshapes, p=0.367.

In another analysis, Type 1 signs were labeled as such and Type 0, 2 and 3 signs were labeled together as Type “023”. Chi-square analyses indicate that the Type 1 signs have the same distribution of marked and unmarked handshapes as the other types of signs, regardless of markedness definition. For the universal unmarked set, p=0.984, and for the language-specific unmarked set, p=0.172. The actual versus expected sign counts are given in (3-64).

(3-64) KSL: marked / unmarked handshape by Type 023 / Type 1

**universal unmarked set**

| observed | Type 023 | Type 1 | sum |
|---|---|---|---|
| Marked | 116 | 63 | 179 |
| Unmarked | 132 | 72 | 204 |
| sum | 248 | 135 | 383 |

| expected | Type 023 | Type 1 |
|---|---|---|
| Marked | 116 | 63 |
| Unmarked | 132 | 72 |

p=0.984 independent

**language-specific unmarked set**

| observed | Type 023 | Type 1 | sum |
|---|---|---|---|
| marked | 79 | **34** | 113 |
| unmarked | 169 | 101 | 270 |
| sum | 248 | 135 | 383 |

| expected | Type 023 | Type 1 |
|---|---|---|
| marked | 73 | 40 |
| unmarked | 175 | 95 |

p=0.367 independent

An examination of the mutual information calculation for KSL type versus handshape shows that one handshape, 49, contributes far more to the mutual information value than any other handshape. Recall that mutual information is defined as the sum of the sum of the log-odds ratio, $\log_2\left(\frac{p(x,y)}{p(x)p(y)}\right)$, weighed by the probability of occurrence, $p(x, y)$, so that $I(X;Y) = \sum_{x,y} p(x,y)\log_2\left(\frac{p(x,y)}{p(x)p(y)}\right)$. For KSL, I(hs; type)=0.29, of which 0.06 is contributed by 49. The expected frequency of 49 in Type 1 signs is 20, but actually only 4 signs of Type 1 have this handshape. When the count of Type 1 signs with handshape 49 is increased from 4 to 20, handshape is independent of type, with p=0.5310. Another handshape that appears with unexpected frequencies in different sign type is 3. It is expected to occur in 27 Type 1 signs, but actually occurs in 39. However, this deviation does not affect the mutual information value to such a great extent, so that 3 only contributes 0.02 to the mutual information value. It seems that the handshape 49 has such an unusual distribution, particularly in Type 1 signs, that it skews the overall pattern.

### 3.3.3 NZSL: handshape by type

For NZSL four new data sets were created. The first data set excludes Type 0 signs, the second excludes Type 1 signs, the third excludes Type 2 signs, and the fourth excludes Type 3 signs. The mutual information significance program was run on each set. The results are shown in (3-65); the first row of this table shows the p value for all types. The only data set that shows independence between type and handshape is the one that excludes Type 1 signs, with p=0.148. Also, the three data sets that include only two sign types and exclude Type 1 signs are also independent (Types 0,2: p=0.123; Types 0,3: p=0.162, and Types 2,3: p=0.233). The three data sets that include only two sign types, one of which is Type 1, are all dependent (Types 0,1: p=0.018; Types 1,2: p=0.028; Types 1,3: p=0.037). Thus, it appears that handshapes are distributed as expected in Type 0, 2, and 3 signs, but not in Type 1 signs.

(3-65) NZSL: Type by handshape mutual information probabilities

| Types included | p value | dependency |
|---|---|---|
| 0 1 2 3 | 0.000 | dependent |
| 1 2 3 | 0.008 | dependent |
| **0 2 3** | **0.086** | **independent** |
| 0 1 3 | 0.006 | dependent |
| 0 1 2 | 0.001 | dependent |

What is special about the handshapes used in Type 1 signs in NZSL? It appears that in NZSL, “unmarked” handshapes predominate in Type 1 signs, regardless of whether the universal or language-specific set of unmarked handshapes is used. Each sign in the database was labeled as having a marked or unmarked handshape according to the universal as well as the language-specific unmarked sets in (3-59). Type 1 signs were labeled as such and Type 0, 2 and 3 signs were labeled together as Type “023”. Chi-square analyses indicate that Type 1 signs have a different distribution of marked and unmarked handshapes than do the other types of signs. For the universal unmarked set, p=0.006, and for the language-specific unmarked set, p=0.014. The actual versus expected sign counts are given in (3-66), from which it can be seen that Type 1 signs use fewer than expected marked handshapes. Although the results for KSL in (3-64) were not significant, the tendency in this case was also for Type 1 signs to use fewer than expected marked handshapes. This result is in contradiction to the suggestion made for ASL that two-handed signs ought to use more marked handshapes in order to provide the sign perceiver with a double dose of the more complex signal.

(3-66) NZSL: marked / unmarked handshape by Type 023 / Type 1

**universal unmarked set**

| observed | Type 023 | Type 1 | sum |
|---|---|---|---|
| Marked | 168 | **66** | 234 |
| Unmarked | 129 | 88 | 217 |
| sum | 297 | 154 | 451 |

| expected | Type 023 | Type 1 |
|---|---|---|
| Marked | 154 | 80 |
| Unmarked | 143 | 74 |

p=0.006 dependent

**language-specific unmarked set**

| observed | Type 023 | Type 1 | sum |
|---|---|---|---|
| marked | 124 | **46** | 170 |
| unmarked | 173 | 108 | 281 |
| sum | 297 | 154 | 451 |

| expected | Type 023 | Type 1 |
|---|---|---|
| marked | 112 | 58 |
| unmarked | 185 | 96 |

p=0.014 dependent

Note that the uniform behavior of the handshape distribution in Type 0, Type 2, and Type 3 signs in KSL and NZSL supports a representation that treats these types of signs in a uniform way, as discussed in section 2.1.2.2.3.

### 3.3.4 SVK: handshape by type

The mutual information between handshape and type for SVK was I(hs;type)=0.180. Of the thousand scrambled data sets created by the mutual information significance program, 151 of them have a greater mutual information, thus, p=0.151, indicating that type and handshape are independent variables. There is no dependency to find and explain. Indeed, when the mutual information calculation is examined, there is only one large value of $p(x,y)\log_2\left(\frac{p(x,y)}{p(x)p(y)}\right)$, for the occurrence of 3 in Type 2 signs. 3 is expected to occur in 13 Type 2 signs, but it actually occurs in 23. Recall that since 3 occurs on the nondominant hand in Type 3 signs, 12 signs with 3 on H2 which were originally classified as Type 2 were reclassified as Type 3 on account of differing orientations and contacts H1 and H2. In spite of this reclassification, 3 still occurs at greater than expected frequency in Type 2 signs, while 3 occurs at slightly less than expected frequency in Type 3 signs (the observed count of 3 in Type 3 signs is 12 while the expected count is 14).

### 3.3.5 One versus two active hands

Signs of Type 0, 2 and 3 all have only one active hand, and signs of Type 1 have two active hands. It has been argued that Type 0, 2 and 3 signs have a uniform representation, where the nondominant hand in Types 2 and 3 is simply a location contacted by the dominant hand (Stokoe et al., 1965; Sandler, 1989; Perlmutter, 1991; Rozelle, 1992; Sandler, 1993a; Rozelle, 1996a; Rozelle, 1998). This section investigates how handshape is distributed across signs with one active hand, Types 0, 2, and 3, versus signs with two active hands, Type 1.

In ASL, handshape is independent of the number of active hands. The mutual information, I(hs; active hands)=0.094, is insignificant, with p=0.075. As noted in section 3.3.1, Type 2 signs differ from the other signs in that they use more unmarked handshapes than expected, but this effect is muted by the larger numbers of Type 0 and Type 3 signs. In SVK, also, handshape is independent of the number of active hands. The mutual information, I(hs; active hands)=0.074, is insignificant, with p=0.082. This result is as expected, since SVK shows no dependence whatsoever between handshape and sign type.

In contrast, both NZSL and KSL show dependence between handshape and the number of active hands. In NZSL, I(hs; active hands)=0.125, with p=0.008. As in section 3.3.3, signs with one active hand have more marked handshapes than expected, while signs with two active hands have fewer marked handshapes than expected, regardless of which markedness criterion is used, with p=0.006 for the universal unmarked set and p=0.020 for the language-specific unmarked set. In KSL, I(hs; active hands)=0.172, with p=0.000. As in section 3.3.2, this dependence is not accounted for by markedness. Signs with one versus two active hands occur independently of whether they bear a marked or an unmarked handshape, with p=0.984 for the universal unmarked set, and p=0.172 for the language-specific unmarked set. Instead, the dependence is accounted for by the extreme distribution of the handshape 49. It alone contributes 0.05 to the mutual information value, appearing in 52 signs with one active hand and only 4 signs with two active hands, compared to the expected values of 36 and 20. When this handshape is excluded, the variables of handshape and number of active hands are independent; I(hs; active hands)=0.1400, p=0.052.

### 3.4 Summary

The first part of this chapter overviewed the handshape inventories of four unrelated languages. In spite of differences in the content of these inventories, the manner in which these handshape resources are utilized is essentially the same across all four languages. When only the shared handshapes are considered, it is seen that they are ordered very similarly across all four languages. Even when the entire inventories are considered, the use of these resources is similar cross-linguistically. Roughly, each language uses a few elements very frequently, about five handshapes in about 50% of the signs, and it uses a lot of elements very rarely. More precisely, the rank-frequency graphs of sign language follow an exponential decay curve, $y = a(2^{-bx})$, where $0.08 \le a \le 0.17$ and $0.12 \le b \le 0.16$. The rank-frequency graphs of LIS and NGT exhibit the same behavior, despite the different methods used for obtaining the data. I propose that the following is a property of all natural sign languages: the handshape rank-frequency distribution is an exponential decay curve.

In the second part of this chapter, the distribution of handshape across sign types was analyzed. This analysis reveals three insights. The first is that Type 2 and Type 3 signs do not distinguish themselves as different from the other sign types. Signs having one active hand pattern together, suggesting that the nondominant hand is similar to other location categories. The second insight is that contrary to predictions based on ease of perception, it is not the case that signs that furnish two opportunities to perceive a handshape, Type 1 and Type 2 signs, are more likely to bear marked handshapes. In fact, the opposite tendency appears to be true. The third insight is that neither of the two unmarked sets was especially successful in identifying the source of dependence in all cases.

# Chapter 4: Location

In this chapter, the location inventories of ASL, KSL, NZSL, SVK and VSVK are presented and compared. First, the body location inventories are ranked according to lexical frequency, and the rank-frequency distributions of all four are shown to be a good approximation of an exponential decay curve, just as the handshape rank-frequency distributions were. Then, all locations, including neutral space and nondominant hand locations are included, and the rank-frequency distributions of all four languages are shown to be better modeled by a hyperbolic curve. The set of locations common to the four main languages is determined, and the ordering of these shared locations is compared cross-linguistically. The data are pooled to provide an idea of the “International Location Alphabet,” and it is shown that the lexical rank-frequency distributions of the pooled data are also modeled by an exponential decay curve for body locations and a hyperbolic curve for all locations. Finally, dependence between the location of a sign and number of hands used in for the sign is investigated. Is it the case that a location can freely host either one-handed or two-handed signs, or do certain locations prefer one-handed signs while other locations prefer two-handed signs?

## 4.1 Inventories

### 4.1.1 American Sign Language

The American Sign Language database uses 27 non-hand body contact locations for monomorphemic signs that do not incorporate a location change. They are shown in (4-1), organized roughly vertically, beginning with the location ρ and continuing downward to the location □. A directory of HamNoSys notation for location is given in Appendix A.

(4-1) ASL: location inventory

σ σ□ ρ
υ θ
ϖ ϖ□ ω
ψ ψ□ ξ
ζ
{
| |□ □
} }□ □
~ ~□ □
ο □
□ □□□ □

Including all signs, the ASL database uses the eleven additional locations shown in (4-2). These locations include neutral space, π, used in Type 0 and Type 1 signs (signs articulated with one and two hands, respectively), as well as the hand location, |, used in Type 2 signs and the hand locations {2 29 49 48 3 39 3≅B ; 78 }, used as the base hand in Type 3 signs. The symbol | represents the location of the non-dominant hand having the default handshape, a duplicate of the handshape of the dominant hand. (See section 2.1.2.2.3 for a discussion of location in Type 2 and Type 3 signs).

(4-2) ASL: neutral space and hand locations

π 2 49 3 ; 78
| 29 48 39
3≅B

The fourteen handshapes in (4-3) are used in Type 2 signs as duplicate handshapes.

(4-3) ASL: duplicate handshapes used as the hand location | in Type 2 signs

| | | | | | |
|---|---|---|---|---|---|
| 2 | 49 | :□ | 59 | 3 | 78 |
| 29 | 4B9 | < | 69 | 39 | |
| 28 | | | 6B9 | 38 | |

As discussed in section 1.5.3, Battison (1978) claims that there are seven handshapes that can be used on the non-dominant hand as locations for Type 3 signs, BASCO15 in Stokoe notation, or {3 2 29 3≅B ; 49 78 } in HamNoSys. In this database, the phonetic contrast between 3 and 39 was notated, although in all three signs using the location 39, PROFESSION, PURE and RIDE, the place of contact on 39 is the ulnar side of the hand; the thumb is folded across the palm, perhaps to keep it out of the way. 48 is used in the sign THEN as the location on the non-dominant hand, with 49 as the handshape on the dominant hand.

The locations in (4-4) are used only in signs having a change in location; thus, they are not included in the inventory in (4-1). In some signs the location changes from one well-defined location to another, such as in the sign MAN, which begins at the location σ then moves to location ∼. Many signs either begin or end without body contact, such as the sign KNOW, which begins in the area in front of σ and ends by contacting σ. These are not considered location change signs.[71] Location change occurs in 25 signs, 3.8% of the ASL sample. Neither borrowings nor compound signs employ any locations not used in monomorphemic signs.

---

71 A few signs are difficult to classify, such as DEAF, which can be considered to have two contacts on location ξ or to have a location change from ω to ζ. I have depended upon the picture of the sign as well as whatever written description was given in the dictionary to guide my decision.

(4-4) ASL: novel locations used only in location change signs

| **location** | **sign** |
|---|---|
| ζ□ | BACHELOR |
| o□ | LANDLORD, QUEEN, CHRIST |
| □ | POWER, IMPROVE |

As location has been less studied than handshape, fewer discussions of location inventories exist. DASL, however, lists the location inventory in (4-5) for ASL. DASL makes fewer location distinctions; for example, the eye and nose regions are treated as two distinct locations in the databases created for this dissertation, but they are collapsed into one region in the DASL inventory. In addition to the eleven locations in (4-5), DASL also distinguishes neutral space and the weak hand as locations. All of these DASL locations are included in the ASL location inventory in (4-1).

(4-5) Location inventory from DASL

| **location** | **description** |
|---|---|
| στ | forehead or brow, upper face |
| υϖ | mid-face, the eye and nose region |
| θ | face or whole head |
| ξωσ□ | cheek, temple, ear, side-face |
| ζ | chin, lower face |
| \| | neck |
| }~ο | trunk, body from shoulders to hips |
| □ | upper arm |
| □□□ | elbow, forearm |
| □ | wrist in pronated position |
| □ | wrist in supinated position |

### 4.1.2 Korean Sign Language

The Korean Sign Language database uses 27 locations for non-hand body contact monomorphemic signs that do not incorporate a location change. They are shown in (4-6).[72]

[72] Some of the locations are iconically motivated. For example, BAEGAE *pillow* has location θφ, the only occurrence of this location. Nevertheless, locations such as these were considered as part of the phonological inventory since their meanings were not transparent, as might be the case if the sign meant 'back of the head.'

(4-6) KSL: location inventory

<table>
<tr><td>σ</td><td>σ□</td><td>ρ</td><td>θφ</td></tr>
<tr><td>υ</td><td>υ□</td><td>θ</td><td>θ□</td></tr>
<tr><td>ϖ</td><td>ϖ□</td><td>ω</td><td></td></tr>
<tr><td>ψ</td><td>ψ□</td><td>ξ</td><td></td></tr>
<tr><td>ζ</td><td></td><td></td><td></td></tr>
<tr><td>{</td><td></td><td></td><td></td></tr>
<tr><td>|</td><td>|□</td><td>|φ</td><td>□</td></tr>
<tr><td></td><td>}□</td><td></td><td>□</td></tr>
<tr><td>~</td><td>~□</td><td></td><td></td></tr>
<tr><td>ο</td><td>ο□</td><td></td><td></td></tr>
<tr><td></td><td>□□</td><td></td><td></td></tr>
</table>

Including all signs, the KSL database uses the fifteen additional locations shown in (4-7).

(4-7) KSL: neutral space and hand locations

<table>
<tr><td>π</td><td>2</td><td>49</td><td><</td><td>3</td><td>78</td></tr>
<tr><td>|</td><td>29</td><td>48□X</td><td>:</td><td>38</td><td></td></tr>
<tr><td></td><td>28</td><td>49□</td><td></td><td>3B</td><td></td></tr>
<tr><td></td><td></td><td></td><td></td><td>3≅B</td><td></td></tr>
</table>

The eleven handshapes in (4-8) are used in Type 2 signs as duplicate handshapes.

(4-8) KSL: duplicate handshapes used as the hand location | in Type 2 signs

<table>
<tr><td>29</td><td>49</td><td><</td><td>59</td><td>3</td><td>78</td><td>;</td></tr>
<tr><td>28</td><td></td><td>:</td><td></td><td>3B</td><td></td><td></td></tr>
<tr><td></td><td></td><td>:A</td><td></td><td></td><td></td><td></td></tr>
</table>

There are only three signs in the KSL database that incorporate a location change, 0.49% of the database. These three signs introduce no novel locations as they use only locations already listed in (4-6). Two novel locations are used in compounds in KSL, as shown in (4-9).

(4-9) KSL: novel locations used only in compounds

| **location** | **sign** |
|---|---|
| ζ□ | SONAMOO *pine* |
| }φ | GGOBCHOO *hunchback* |

### 4.1.3 New Zealand Sign Language

The New Zealand Sign Language database uses 31 locations for non-hand body contact monomorphemic signs that do not incorporate a location change. They are shown in (4-10).

(4-10) NZSL: location inventory

<table>
<tr><td>σ</td><td>σ□</td><td>ρ</td><td></td><td></td></tr>
<tr><td></td><td>υ□</td><td>θ</td><td>θ□</td><td></td></tr>
<tr><td>ϖ</td><td>ϖ□</td><td>ω</td><td></td><td></td></tr>
<tr><td>ψ</td><td>ψ□</td><td>ξ</td><td></td><td></td></tr>
<tr><td>ζ</td><td>ζ□</td><td></td><td></td><td></td></tr>
<tr><td>{</td><td></td><td></td><td></td><td></td></tr>
<tr><td>|</td><td>|□</td><td>|φ</td><td></td><td>□</td></tr>
<tr><td></td><td>}□</td><td></td><td></td><td>□</td></tr>
<tr><td>~</td><td>~□</td><td>~□□</td><td></td><td>□</td></tr>
<tr><td>ο</td><td>ο□</td><td>ο□□</td><td>οφ</td><td>□</td></tr>
<tr><td>□</td><td></td><td></td><td></td><td></td></tr>
</table>

Including all signs, the NZSL database uses the sixteen additional locations shown in (4-11).

(4-11) NZSL: neutral space and hand locations

π   2   49   69   3   3B   78

|   29   :   38   3B8   ;

28   3≅B

3≅A

The eleven handshapes in (4-12) are used in Type 2 signs as duplicate handshapes.

(4-12) NZSL: duplicate handshapes used as the hand location | in Type 2 signs

29   49   <B   3   3B

28   4X   39   3A

38   3A8

There are only two signs in the NZSL database that incorporate a location change, 0.3% of the database. These two signs introduce no novel locations as they use only locations already listed in (4-10). Compound signs do not introduce any novel locations either.

### 4.1.4 Finnish Sign Language

The Finnish Sign Language database uses 26 locations for non-hand body contact monomorphemic signs that do not incorporate a location change. They are shown in (4-13).

(4-13) SVK: location inventory

<table>
<tr><td>σ</td><td>σ□</td><td></td><td></td></tr>
<tr><td>υ</td><td>υ□</td><td>θ</td><td></td></tr>
<tr><td>ϖ</td><td></td><td>ω</td><td></td></tr>
<tr><td>ψ</td><td>ψ□</td><td>ξ</td><td></td></tr>
<tr><td>ζ</td><td>ζ□</td><td></td><td></td></tr>
<tr><td>{</td><td></td><td></td><td></td></tr>
<tr><td>|</td><td>|□</td><td></td><td></td></tr>
<tr><td>}</td><td>}□</td><td>}φ</td><td>□</td></tr>
<tr><td>~</td><td>~□</td><td></td><td>□</td></tr>
<tr><td>ο</td><td>ο□</td><td></td><td>□</td></tr>
<tr><td>□</td><td></td><td></td><td></td></tr>
</table>

Including all signs, the SVK database uses the eight additional locations shown in (4-14).

(4-14) SVK: neutral space and hand locations

<table>
<tr><td>π</td><td>2</td><td>49</td><td>3</td></tr>
<tr><td>|</td><td>29</td><td></td><td>38</td></tr>
<tr><td></td><td></td><td></td><td>3≅</td></tr>
</table>

The sixteen handshapes in (4-15) are used in Type 2 signs as duplicate handshapes.

(4-15) SVK: duplicate handshapes used as the hand location | in Type 2 signs

<table>
<tr><td>2</td><td>49</td><td>:</td><td>69</td><td>3</td><td>78</td></tr>
<tr><td>29</td><td>4X</td><td>;</td><td>59</td><td>3B</td><td></td></tr>
<tr><td></td><td>49B</td><td></td><td>69B</td><td>3A</td><td></td></tr>
<tr><td></td><td>48□B</td><td></td><td></td><td></td><td></td></tr>
<tr><td></td><td>4≅B</td><td></td><td></td><td></td><td></td></tr>
</table>

There are six signs in the SVK database that incorporate a location change, 1.0% of the database. These signs introduce the novel location shown in (4-16), which is used in two signs. Compound signs do not introduce any new locations.

(4-16) SVK: novel location used only in location change signs

□ VIOMA *strength* and KANGAS *cloth*

#### 4.1.5 Old Finnish Sign Language

The Old Finnish Sign Language database uses 16 locations for non-hand body contact monomorphemic signs that do not incorporate a location change. They are shown in (4-17).

(4-17) VSVK: location inventory

| | | |
|---|---|---|
| σ | σ□ | |
| υ | θ | θ□ |
| ϖ | | ω |
| ψ | ψ□ | ξ |
| ζ | | |
| \| | \|□ | □ |
| ~ | ο□ | □ |

Including all signs, the VSVK database uses the ten additional locations shown in (4-18).

(4-18) VSVK: neutral space and hand locations

| | | | | | |
|---|---|---|---|---|---|
| π | 2 | 59 | 3 | 78 | ; |
| \| | 29 | | 38 | | |
| | 28 | | | | |

The six handshapes in (4-19) are used in Type 2 signs as duplicate handshapes.

(4-19) VSVK: duplicate handshapes used as the hand location | in Type 2 signs

| | | | |
|---|---|---|---|
| 2 | 49 | 59 | 3 |
| 29 | 4≅A | | |

There are two signs in the VSVK database that incorporate a location change, 0.78% of the database. These signs introduce no novel locations. Compound signs do not introduce any new locations either.

### 4.1.6 Comparison

The sizes of the location inventories of ASL, KSL, NZSL, SVK and VSVK are indicated in (4-20).

(4-20) Location inventory sizes

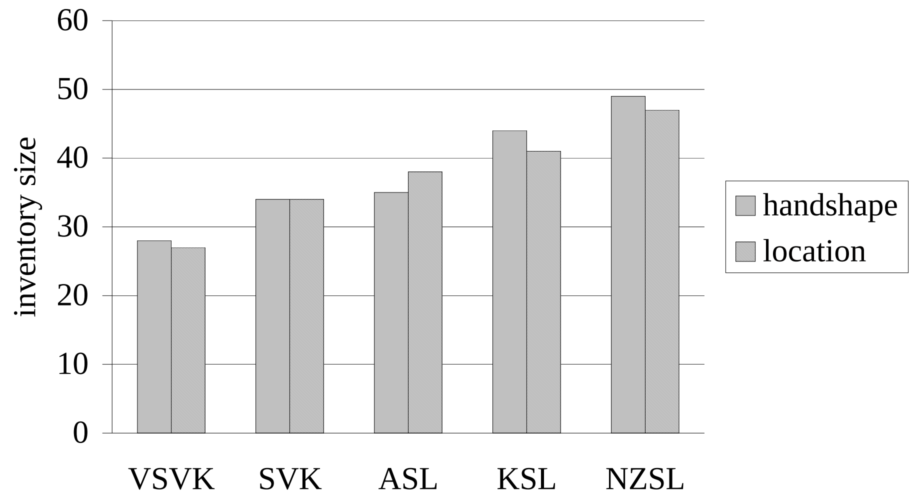
| | body | hand and $\pi$ | total |
|---|---|---|---|
| ASL | 27 | 11 | 38 |
| KSL | 27 | 15 | 42 |
| NZSL | 31 | 16 | 47 |
| SVK | 26 | 8 | 34 |
| VSVK | 17 | 10 | 27 |

The sizes of the handshape and location inventories are compared cross-linguistically in (4-21) and (4-22). Notice that it is not the case that a language compensates for a small inventory of one parameter by having a large inventory of another parameter (see section 1.4); instead, there is a strong direct correspondence between the size of the handshape inventory and the size of the location inventory, with correlation coefficient $r = 0.964$.

(4-21) Comparison of handshape and location inventory sizes

| | handshape | location | total |
|---|---|---|---|
| VSVK | 28 | 27 | 55 |
| SVK | 34 | 34 | 68 |
| ASL | 35 | 38 | 73 |
| KSL | 44 | 41 | 85 |
| NZSL | 49 | 47 | 96 |

(4-22) Graph of handshape and inventory sizes




The locations in (4-23) are shared by all four languages, excluding VSVK.

(4-23) Shared locations: locations used by all four languages

<table>
<tr><th colspan="3">Body</th><th></th><th colspan="2">Other</th></tr>
<tr><td>σ</td><td>σ□</td><td>θ</td><td></td><td>π</td><td>2</td></tr>
<tr><td></td><td>ω</td><td></td><td></td><td>|</td><td>29</td></tr>
<tr><td>ϖ</td><td>ξ</td><td></td><td></td><td></td><td>49</td></tr>
<tr><td>ψ</td><td>ψ□</td><td></td><td></td><td></td><td>3</td></tr>
<tr><td>ζ</td><td></td><td></td><td></td><td></td><td></td></tr>
<tr><td>{</td><td></td><td></td><td></td><td></td><td></td></tr>
<tr><td>|</td><td>|□</td><td>□</td><td></td><td></td><td></td></tr>
<tr><td></td><td>}□</td><td>□</td><td></td><td></td><td></td></tr>
<tr><td>~</td><td>~□</td><td></td><td></td><td></td><td></td></tr>
<tr><td></td><td>ο</td><td></td><td></td><td></td><td></td></tr>
</table>

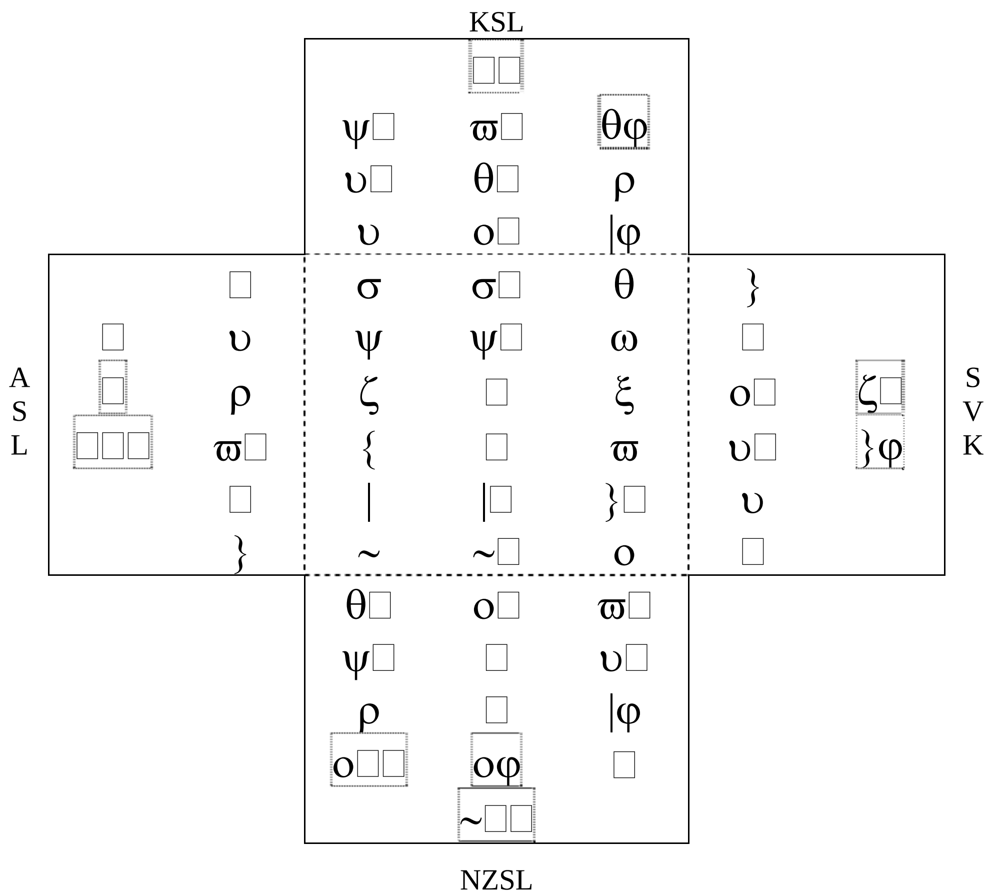
The locations in (4-24) are used by only one language.

(4-24) Locations occurring in only one language

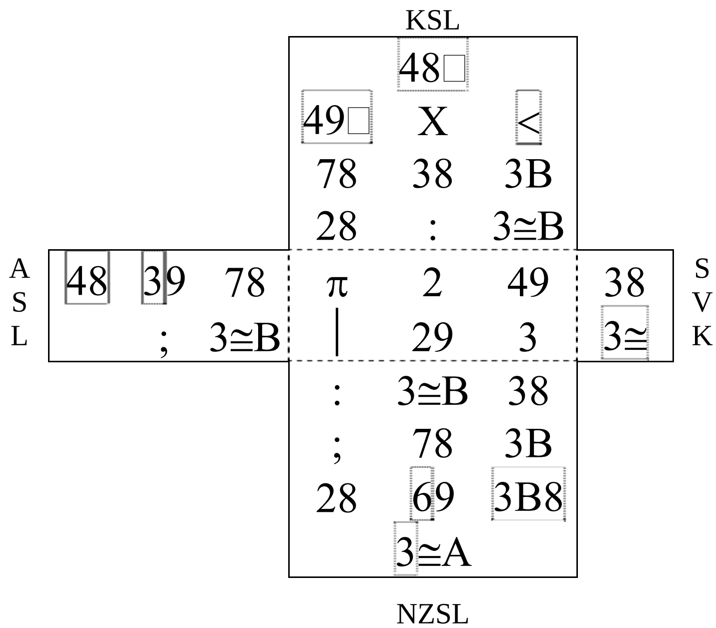
| Body | | Other | |
|---|---|---|---|
| □ | θφ | 48 | 39 |
| oφ | }φ | < | 3B8 |
| ~□□ | □□ | 48□X | 3≅A |
| o□□ | □□□ | 49□ | 69 |

The diagram in (4-25) shows the body location inventories of all four languages, the diagram in (4-26) shows the other location inventories, and the diagram in (4-27) shows all locations combined. The rectangle in the center of each diagram contains the shared locations. Note that there is considerable overlap in the location inventories. The locations unique to each language are outlined with dotted lines. Other locations in the peripheral rectangles for each language are shared with at least one other language.

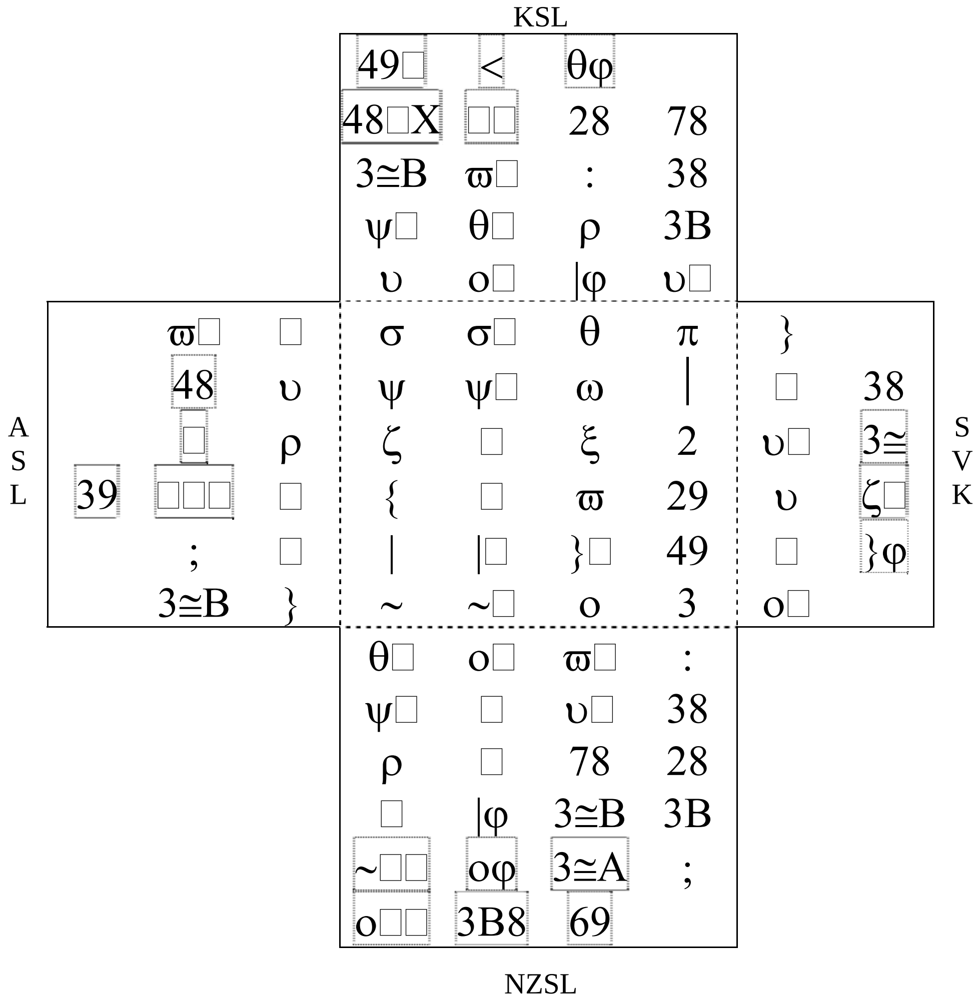
(4-25) Body location inventories of all four languages



(4-26) Non-body location inventories of all four languages



(4-27) Combined location inventories for all four languages




### 4.2 Distribution of locations in a lexicon

#### 4.2.1 Location rank

To determine the ranking of the locations, it is necessary to count how many signs use each location. As in the determination of inventory, only monomorphemic signs without location change were counted. Since the languages do not share the same set of locations, only the set of shared locations was ranked. Many of the shared locations are among the most frequently used locations in each language. The 18 shared body locations are used in 87% of the ASL database, 79% of the KSL database, 77% of the NSZL database, and 87% of the SVK database (while these 18 locations are only 64%, 67%, 58% and 69% of the inventories, respectively). However, it is not the case that the shared locations are necessarily the most common ones. For example, |□ is used in only one sign in ASL, KSL and SVK, and □ is used in one sign in NZSL. The rankings of the shared body locations are shown in (4-28).

(4-28) Ranking of shared body locations

| rank | ASL | KSL | NZSL | SVK |
|---|---|---|---|---|
| 1 | ψ | ψ | ζ | ψ |
| 2 | ~ | ~ | }□ | ~ |
| 3 | σ | ξ | ~ | σ |
| 4 | ξ | θ | ξ | ξ |
| 5 | σ□ | ϖ | ψ | ζ |
| 6 | ϖ | { | σ□ | ο |
| 7 | }□ | ζ | ϖ | ϖ |
| 8 | θ | σ□ | \|□ | θ |
| 9 | ζ | \| | ω | σ□ |
| 10 | ~□ | σ | ~□ | { |
| 11 | { | }□ | \| | }□ |
| 12 | ψ□ | ο | θ | ω |
| 13 | □ | □ | σ | \| |
| 14 | \| | ω | □ | □ |
| 15 | □ | □ | { | ~□ |
| 16 | ω | ψ□ | ο | □ |
| 17 | ο | \|□ | ψ□ | \|□ |
| 18 | \|□ | ~□ | □ | ψ□ |

The shared body locations appear to vary in their rankings somewhat more than the shared handshapes. For example, in each language the first and second most frequent handshapes were the same, 3 and 49. But for location, it is possible only to say that ψ, ~ and ξ are among the top five locations in each language. Is it the case, however, that there is basic similarity in the rankings across all four languages? Using the Spearman rank correlation test to determine whether the rankings are independent or whether they vary in the same or opposite directions, the rank correlation coefficient, $r_s$, was calculated for each pair of languages in (4-28). The results appear in (4-29). Note that it is necessary that $r_s > 0.468$ for the correlation to be significant at the 0.05 level, which is attained by all pairings, and that $r_s > 0.542$ for correlation to be significant at the 0.02 confidence level, which is almost but not quite attained by the NZSL / SVK comparison.[73] Thus, it is possible to assert that though the inventories are of different sizes and the shared locations are distributed throughout the entire inventory, the shared locations are ranked very similarly across these four languages. However, the similarity in rankings of the shared locations is not as strong as for the rankings of the shared handshapes, where all language pairings attained significance at the 0.01 confidence level.

(4-29) Spearman rank correlation test for body locations:

$$r_s = \sum_{i=1}^{n}(x_i - \overline{X})(y_i - \overline{Y})^2 \Big/ \sqrt{\sum_{i=1}^{n}(x_i - \overline{X})^2 \sum_{i=1}^{n}(y_i - \overline{Y})^2} \qquad \begin{array}{l} \overline{X} = \overline{Y} = 9.5 \\ n = 18 \end{array}$$

X = {$x_1$, $x_2$, …, $x_{18}$} and Y = {$y_1$, $y_2$, …, $y_{18}$}are sets of rankings of the 18 shared handshapes for the two languages being compared.

$r_s > 0.468$ for 0.05 significance level and $r_s > 0.542$ for 0.02 significance level.

| | ASL | KSL | NZSL | SVK |
|---|---|---|---|---|
| ASL | 1 | 0.745 | 0.554 | 0.751 |
| KSL | | 1 | 0.561 | 0.869 |
| NZSL | | | 1 | 0.541 |
| SVK | | | | 1 |

[73] The 0.05 level is the usual cut-off level for claiming significance. All pairings exceed 0.05.

When hand locations and neutral space are included in the location inventory, the ranking is as in (4-30). All four languages have π, 3, and | as the three most common locations. In addition, six of the eight most common locations are shared: {π 3 | ψ ~ ξ}.[74]

---

[74] Even if all locations are considered, not just the shared ones, six out of the top nine locations are shared.

(4-30) Ranking of all shared locations

| **rank** | **ASL** | **KSL** | **NZSL** | **SVK** |
|---|---|---|---|---|
| 1 | π | π | π | π |
| 2 | 3 | 3 | 3 | 3 |
| 3 | \| | \| | \| | \| |
| 4 | ψ | ψ | ζ | ψ |
| 5 | ~ | ~ | }□ | ~ |
| 6 | σ | 29 | ~ | σ |
| 7 | ξ | ξ | ξ | ξ |
| 8 | 29 | θ | ψ | ζ |
| 9 | σ□ | ϖ | σ□ | ο |
| 10 | 49 | { | ϖ | ϖ |
| 11 | ϖ | ζ | 49 | θ |
| 12 | }□ | σ□ | \|□ | σ□ |
| 13 | θ | \| | ω | { |
| 14 | ζ | σ | ~□ | }□ |
| 15 | ~□ | }□ | \| | ω |
| 16 | { | 49 | θ | 49 |
| 17 | ψ□ | ο | σ | \| |
| 18 | □ | □ | □ | □ |
| 19 | \| | 2 | { | ~□ |
| 20 | □ | ω | ο | □ |
| 21 | ω | □ | 29 | \|□ |
| 22 | ο | ψ□ | ψ□ | ψ□ |
| 23 | \|□ | \|□ | □ | 29 |
| 24 | 2 | ~□ | 2 | 2 |

Including signs articulated in neutral space as well as signs having the nondominant hand as the location, the overall similarity of the rankings of shared locations becomes even greater. In each language, neutral space is the most common location, followed by 3, then by |. According to the Spearman rank correlation test in (4-31), it is possible to assert at the 0.01 significance level that the shared locations are ranked very similarly across the four languages. The $r_s$ values obtained in (4-31) are comparable to those obtained in 3.2.1 for handshape rank.

(4-31) Spearman rank correlation test for all locations

It must be that $r_s > 0.537$ for correlation to be significant at the 0.01 level.

| | ASL | KSL | NZSL | SVK |
|---|---|---|---|---|
| ASL | 1 | 0.814 | 0.694 | 0.748 |
| KSL | | 1 | 0.605 | 0.785 |
| NZSL | | | 1 | 0.753 |
| SVK | | | | 1 |

### 4.2.2 Frequency distribution

#### 4.2.2.1 ASL: location rank-frequency

The distribution of body locations in monomorphemic, non-location-changing signs in the ASL database is shown in (4-32). The five most frequent locations, {ψ σ ~ ξ σ□ }, comprise 48% of the set of signs in the ASL database with a body location, while the five least frequent body locations, {ο □□□ □ □ |□ }, comprise only 2.8% of the sample. The six least frequent locations appear in only one monomorphemic, non-location-changing sign each, PREGNANT, WHEELCHAIR, PANTS, WATCH, and BROKE, respectively.

(4-32) ASL: body location frequency distribution

| **rank** | **loc** | **count** | **freq** | **rank** | **loc** | **count** | **freq** |
|---|---|---|---|---|---|---|---|
| 1 | ψ | 21 | 0.1160 | 16 | □ | 4 | 0.0221 |
| 2 | σ | 18 | 0.0994 | 17 | □ | 4 | 0.0221 |
| 3 | ~ | 18 | 0.0994 | 18 | \| | 4 | 0.0221 |
| 4 | ξ | 16 | 0.0884 | 19 | ω | 3 | 0.0166 |
| 5 | σ□ | 14 | 0.0773 | 20 | } | 2 | 0.0111 |
| 6 | ϖ | 9 | 0.0497 | 21 | □ | 2 | 0.0111 |
| 7 | }□ | 9 | 0.0497 | 22 | ϖ□ | 1 | 0.0055 |
| 8 | ζ | 8 | 0.0442 | 23 | \|□ | 1 | 0.0055 |
| 9 | θ | 8 | 0.0442 | 24 | ο | 1 | 0.0055 |
| 10 | ~□ | 7 | 0.0387 | 25 | □□□ | 1 | 0.0055 |
| 11 | { | 7 | 0.0387 | 26 | □ | 1 | 0.0055 |
| 12 | υ | 6 | 0.0331 | 27 | □ | 1 | 0.0055 |
| 13 | ψ□ | 5 | 0.0276 | sum | | 180 | 1 |
| 14 | □ | 5 | 0.0276 | | | | |
| 15 | ρ | 4 | 0.0221 | | | | |

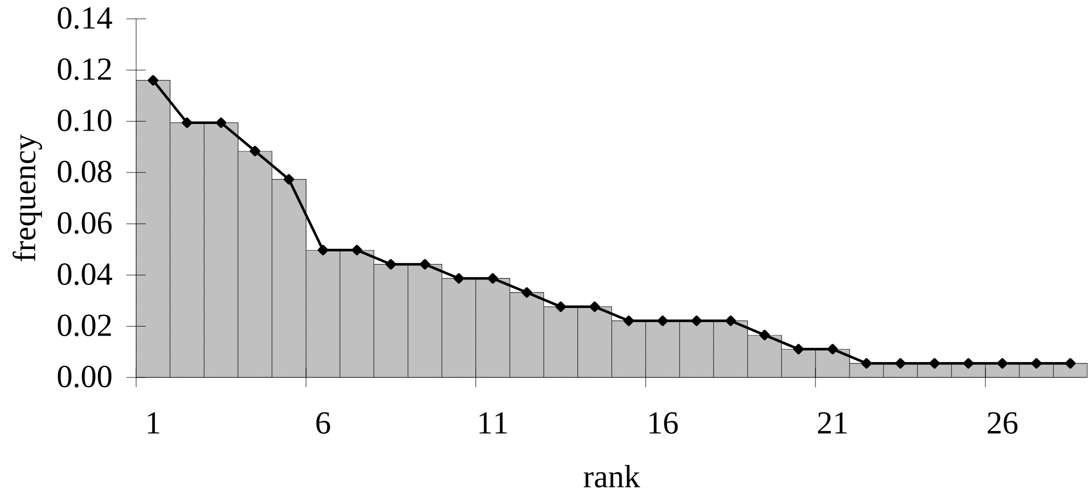
The distribution of all locations in monomorphemic, non-location-changing signs in the ASL database, including neutral space and hand locations, is shown in (4-33). Notice that 41% of the signs are articulated in neutral space and involve no body contact or proximity. The second most frequent location is 3, used as a location on the non-dominant hand in 14% of the signs. The five most common locations comprise 68% of the set of monomorphemic, non-location-changing signs in the ASL database. The five least common locations comprise 0.85% of the sample.

(4-33) ASL: combined location frequency

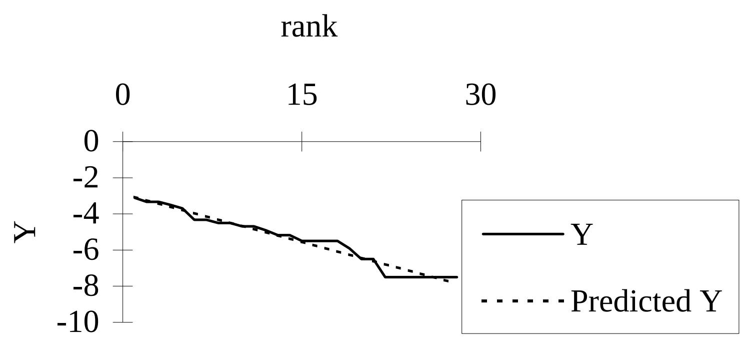
| rank | loc | count | freq |
|---|---|---|---|
| 1 | π | 236 | 0.4126 |
| 2 | 3 | 83 | 0.1451 |
| 3 | \| | 33 | 0.0577 |
| 4 | ψ | 21 | 0.0367 |
| 5 | ~ | 18 | 0.0315 |
| 6 | σ | 18 | 0.0315 |
| 7 | ξ | 17 | 0.0297 |
| 8 | σ□ | 14 | 0.0245 |
| 9 | 29 | 14 | 0.0245 |
| 10 | 49 | 12 | 0.0210 |
| 11 | }□ | 9 | 0.0157 |
| 12 | ϖ | 9 | 0.0157 |
| 13 | θ | 8 | 0.0140 |
| 14 | ζ | 8 | 0.0140 |
| 15 | { | 7 | 0.0122 |
| 16 | ~□ | 7 | 0.0122 |
| 17 | υ | 6 | 0.0105 |
| 18 | □ | 5 | 0.0087 |
| 19 | ψ□ | 5 | 0.0087 |
| 20 | \| | 4 | 0.0070 |
| 21 | □ | 4 | 0.0070 |
| 22 | □ | 4 | 0.0070 |
| 23 | ρ | 4 | 0.0070 |
| 24 | 39 | 3 | 0.0052 |
| 25 | 3≅B | 3 | 0.0052 |
| 26 | 78 | 3 | 0.0052 |
| 27 | ω | 3 | 0.0052 |
| 28 | ; | 2 | 0.0035 |
| 29 | □ | 2 | 0.0035 |
| 30 | } | 2 | 0.0035 |
| 31 | 2 | 1 | 0.0017 |
| 32 | \|□ | 1 | 0.0017 |
| 33 | □ | 1 | 0.0017 |
| 34 | □ | 1 | 0.0017 |
| 35 | □□□ | 1 | 0.0017 |
| 36 | ο | 1 | 0.0017 |
| 37 | ϖ□ | 1 | 0.0017 |
| 38 | 48 | 1 | 0.0017 |
| sum | | 572 | 1 |

The rank-frequency graph for ASL body location is shown in (4-34). This graph is an exponential decay curve, as shown by the linearity of the graph in (4-35). Linear regression applied to x = rank and Y = $\log_2$(frequency) results in an adjusted correlation coefficient R square value of 0.9629, indicating that body location in ASL, like handshape, declines with close fidelity to the exponential decay law, where $y = 0.13(2^{-0.18})$.

(4-34) ASL: body location rank-frequency graph




(4-35) ASL: Y = $\log_2 y$ where x = rank and y = frequency



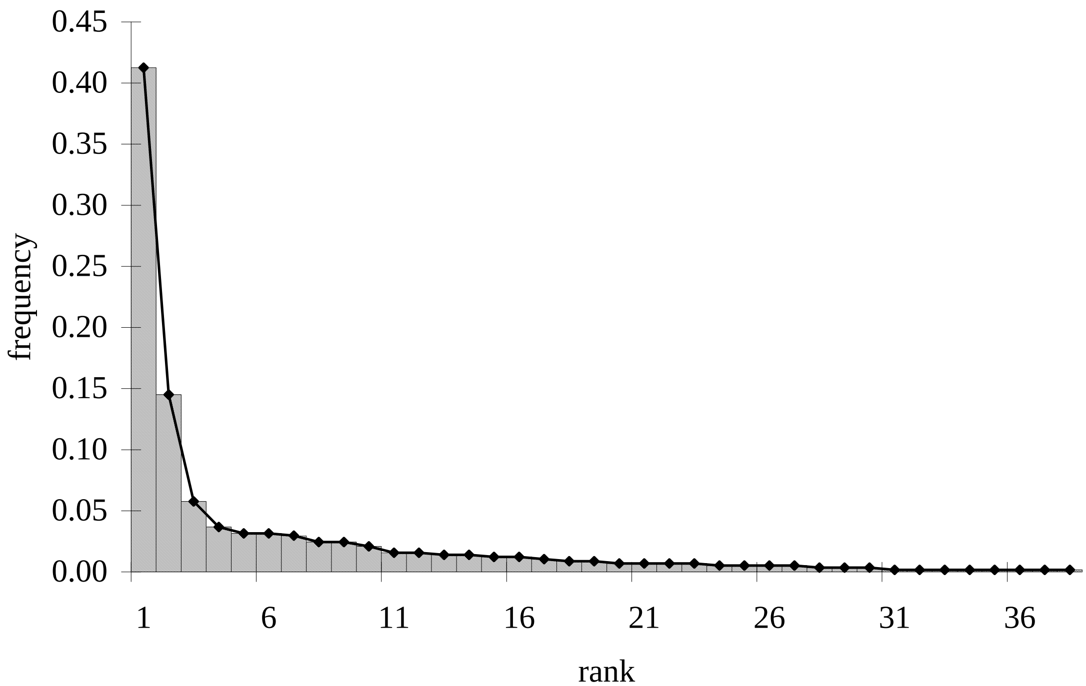

The rank-frequency graph for all ASL locations is shown in (4-36). This graph is not an exponential decay curve; instead, it is a hyperbolic curve, as shown by the linearity of the graph in (4-37)[75]. Linear regression applied to X = $\log_2$(rank) and Y = $\log_2$(frequency) results in an adjusted correlation coefficient R square value of 0.9391, indicating that the rank frequency of

---

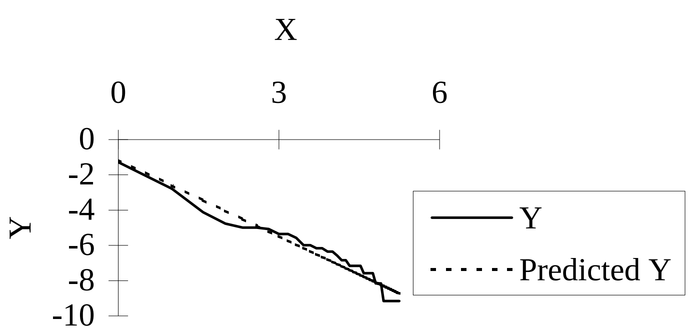
[75] $y = ax^c$ implies $\log y = \log (ax^c)$. Since $\log (ax^c) = \log a + \log (x^c) = \log a + c \log x$, it is the case that $\log y = \log a + c \log x$, which is a linear equation of the form Y = $m$X + $b$, where Y = $\log y$, X = $\log x$, the slope $m$ equals c, and the intercept $b$ equals $\log a$.

location in ASL, when neutral space and hand locations are included, declines with close fidelity to the power law, $y = 0.45x^{-1.4}$.

(4-36) ASL: rank-frequency graph of all locations




(4-37) ASL: X = $\log_2$x and Y = $\log_2$y where x = rank and y = frequency



#### 4.2.2.2 KSL: location rank-frequency

The distribution of body locations in monomorphemic, non-location-changing signs in the KSL database is shown in (4-38). The five most frequent locations, {ψ ξ ~ υ ϖ }, comprise 50% of the set of signs in the KSL database with a body location, while the five least frequent locations, { □□ □ ~□ |φ |□ }, comprise only 3.9% of the sample, with each of these five least frequent locations appearing in only one monomorphemic, non-location-changing sign each, CHEKGABANG *bag*, GYUDRANGI *armpit*, JUT *breast*, GISANG *rising*, and JA *let's*, respectively.

(4-38) KSL: body location frequency distribution

| rank | loc | count | freq |
|---|---|---|---|
| 1 | ψ | 22 | 0.1732 |
| 2 | ξ | 12 | 0.0945 |
| 3 | ~ | 12 | 0.0945 |
| 4 | υ | 9 | 0.0709 |
| 5 | ϖ | 8 | 0.0630 |
| 6 | θ | 8 | 0.0630 |
| 7 | ζ | 6 | 0.0472 |
| 8 | θ□ | 6 | 0.0472 |
| 9 | { | 6 | 0.0472 |
| 10 | σ□ | 5 | 0.0394 |
| 11 | σ | 4 | 0.0315 |
| 12 | \| | 4 | 0.0315 |
| 13 | ο | 3 | 0.0236 |
| 14 | }□ | 3 | 0.0236 |
| 15 | ω | 2 | 0.0157 |

| rank | loc | count | freq |
|---|---|---|---|
| 16 | ϖ□ | 2 | 0.0157 |
| 17 | υ□ | 2 | 0.0157 |
| 18 | ρ | 2 | 0.0157 |
| 19 | ο□ | 2 | 0.0157 |
| 20 | □ | 2 | 0.0157 |
| 21 | ψ□ | 1 | 0.0079 |
| 22 | θφ | 1 | 0.0079 |
| 23 | □□ | 1 | 0.0079 |
| 24 | □ | 1 | 0.0079 |
| 25 | ~□ | 1 | 0.0079 |
| 26 | \|φ | 1 | 0.0079 |
| 27 | \|□ | 1 | 0.0079 |
| sum | | 127 | 1 |

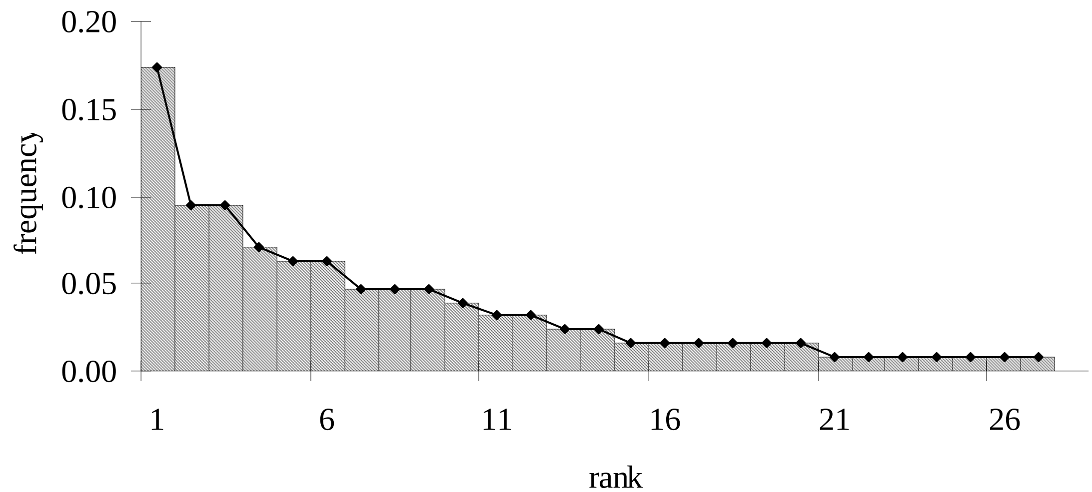
The distribution of all locations in monomorphemic, non-location-changing signs in the KSL database, including neutral space and hand locations, is shown in (4-39). Notice that 40% of the signs are articulated in neutral space and involve no body contact or proximity. The second most frequent location is 3, used as a location on the non-dominant hand in 14% of the signs. The five most common locations comprise 67% of the set of monomorphemic, non-location-changing signs in the KSL database. The five least common locations comprise 1.2% of the sample.

(4-39) KSL: combined location frequency

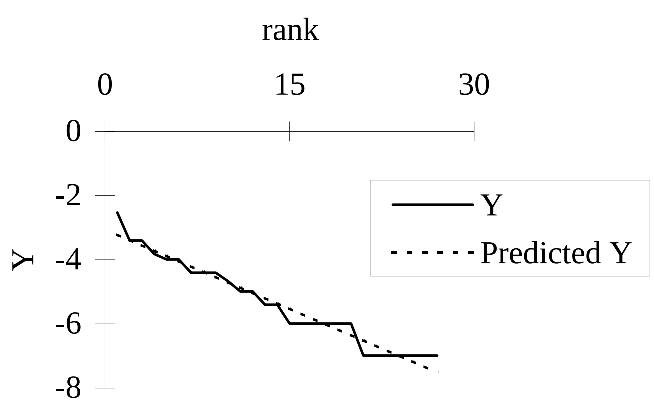
| rank | loc | count | freq | rank | loc | count | freq |
|---|---|---|---|---|---|---|---|
| 1 | π | 170 | 0.4009 | 26 | □ | 2 | 0.0047 |
| 2 | 3 | 58 | 0.1368 | 27 | ο□ | 2 | 0.0047 |
| 3 | \| | 23 | 0.0543 | 28 | ρ | 2 | 0.0047 |
| 4 | ψ | 22 | 0.0519 | 29 | υ□ | 2 | 0.0047 |
| 5 | ~ | 12 | 0.0284 | 30 | ϖ□ | 2 | 0.0047 |
| 6 | ξ | 12 | 0.0284 | 31 | ω | 2 | 0.0047 |
| 7 | 29 | 12 | 0.0284 | 32 | 78 | 2 | 0.0047 |
| 8 | 28 | 9 | 0.0213 | 33 | : | 1 | 0.0024 |
| 9 | υ | 9 | 0.0213 | 34 | \|□ | 1 | 0.0024 |
| 10 | θ | 8 | 0.0189 | 35 | \|φ | 1 | 0.0024 |
| 11 | ϖ | 8 | 0.0189 | 36 | ~□ | 1 | 0.0024 |
| 12 | { | 6 | 0.0142 | 37 | □□ | 1 | 0.0024 |
| 13 | θ□ | 6 | 0.0142 | 38 | 48□X | 1 | 0.0024 |
| 14 | ζ | 6 | 0.0142 | 39 | 49□ | 1 | 0.0024 |
| 15 | 3≅B | 5 | 0.0142 | 40 | □□ | 1 | 0.0024 |
| 16 | σ□ | 5 | 0.0118 | 41 | θφ | 1 | 0.0024 |
| 17 | \| | 4 | 0.0094 | 42 | ψ□ | 1 | 0.0024 |
| 18 | σ | 4 | 0.0094 | sum | | 424 | 1 |
| 19 | 2 | 3 | 0.0071 | | | | |
| 20 | 38 | 3 | 0.0071 | | | | |
| 21 | }□ | 3 | 0.0071 | | | | |
| 22 | < | 3 | 0.0071 | | | | |
| 23 | 3B | 3 | 0.0071 | | | | |
| 24 | ο | 3 | 0.0071 | | | | |
| 25 | 49 | 3 | 0.0071 | | | | |

The body location rank-frequency graph for KSL is shown in (4-40). This graph, like that for ASL, is also an exponential decay curve, as shown by the linearity of the graph in (4-41). Linear regression applied to x = rank and Y = $\log_2$(frequency) results in an adjusted correlation coefficient R square value of 0.9573, indicating that location in KSL declines with close fidelity to the exponential decay law, where $y = 0.12(2^{-0.16x})$.

(4-40) KSL: body location rank-frequency graph

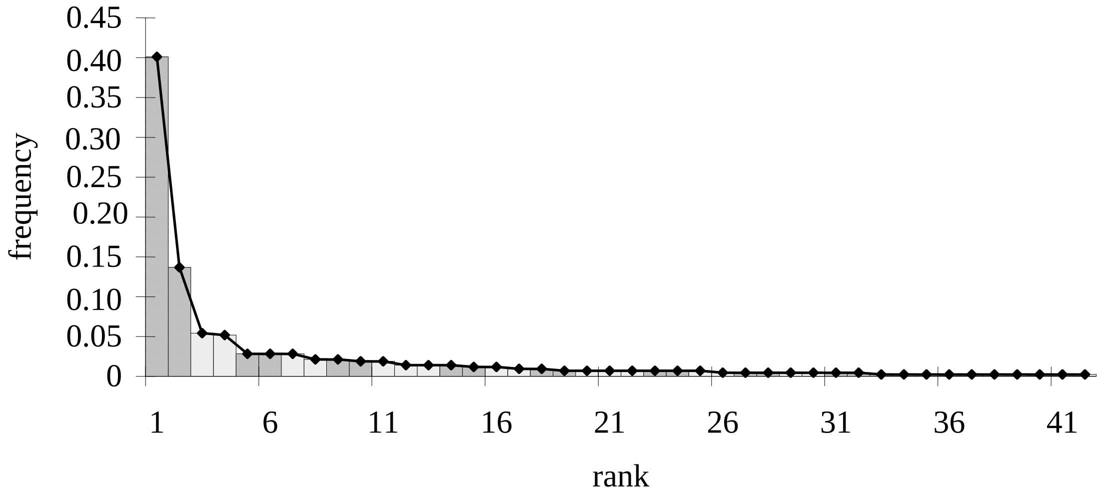



(4-41) KSL: $Y = \log_2 y$ where x = rank and y = frequency

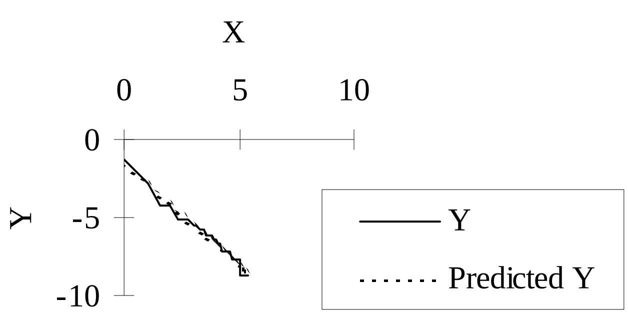



The rank-frequency graph for all KSL locations is shown in (4-42). This graph is not an exponential decay curve; instead, it is a hyperbolic curve, as shown by the linearity of the graph in (4-42). Linear regression applied to X = $\log_2$(rank) and Y = $\log_2$(frequency) results in an adjusted correlation coefficient R square value of 0.9625, indicating that the rank frequency of location in KSL, when neutral space and hand locations are included, declines with close fidelity to the power law, $y = 0.35x^{-1.3}$.

(4-42) KSL: rank-frequency graph for all locations




(4-43) KSL: X = $\log_2 x$ and Y = $\log_2 y$ where x = rank and y = frequency




#### 4.2.2.3 NZSL: location rank-frequency

The distribution of body locations in monomorphemic, non-location-changing signs in the NZSL database is shown in (4-44). The five most frequent locations, {ζ }□ ξ ~ σ□ }, comprise 40% of the set of signs in the NZSL database with a body location, while the five least frequent locations, { οφ □ □ ~□□ |φ }, comprise only 3.5% of the sample, with each of these five least frequent locations appearing in only one monomorphemic, non-location-changing sign each, UNDERHAND, STRONG, POOR, RUGBY UNION, and SPINAL MENINGITIS, respectively.

(4-44) NZSL: body location frequency distribution

| rank | loc | count | freq | rank | loc | count | freq |
|---|---|---|---|---|---|---|---|
| 1 | ζ | 15 | 0.1034 | 21 | { | 3 | 0.0207 |
| 2 | }□ | 12 | 0.0828 | 22 | ζ□ | 2 | 0.0138 |
| 3 | ξ | 11 | 0.0759 | 23 | ψ□ | 2 | 0.0138 |
| 4 | ~ | 11 | 0.0759 | 24 | ρ | 2 | 0.0138 |
| 5 | σ□ | 9 | 0.0621 | 25 | ο□□ | 2 | 0.0138 |
| 6 | ψ | 9 | 0.0621 | 26 | □ | 2 | 0.0138 |
| 7 | ϖ | 7 | 0.0483 | 27 | οφ | 1 | 0.0069 |
| 8 | ω | 6 | 0.0414 | 28 | □ | 1 | 0.0069 |
| 9 | \|□ | 6 | 0.0414 | 29 | □ | 1 | 0.0069 |
| 10 | ~□ | 5 | 0.0345 | 30 | ~□□ | 1 | 0.0069 |
| 11 | θ□ | 4 | 0.0276 | 31 | \|φ | 1 | 0.0069 |
| 12 | θ | 4 | 0.0276 | sum | | 145 | 1 |
| 13 | ο□ | 4 | 0.0276 | | | | |
| 14 | □ | 4 | 0.0276 | | | | |
| 15 | \| | 4 | 0.0276 | | | | |
| 16 | σ | 4 | 0.0276 | | | | |
| 17 | ϖ□ | 3 | 0.0207 | | | | |
| 18 | υ□ | 3 | 0.0207 | | | | |
| 19 | ο | 3 | 0.0207 | | | | |
| 20 | □ | 3 | 0.0207 | | | | |

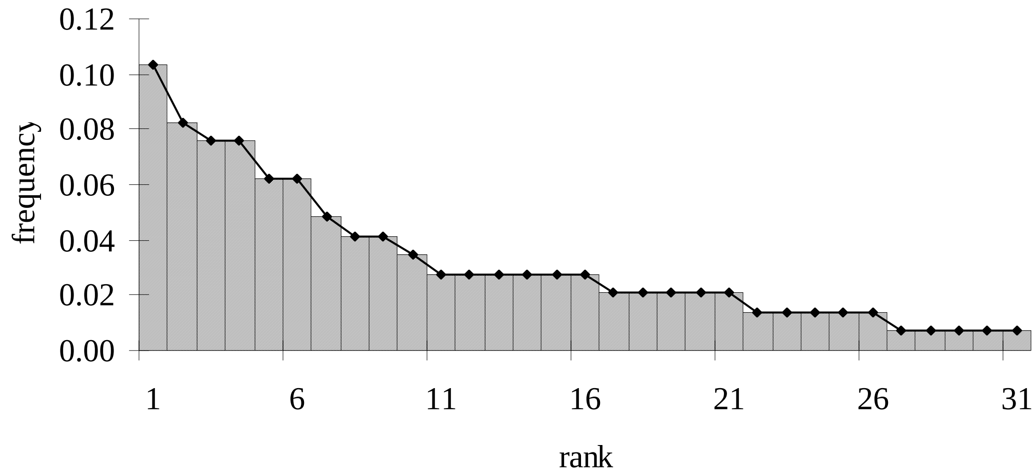
The distribution of all locations in monomorphemic, non-location-changing signs in the NZSL database, including neutral space and hand locations, is shown in (4-45). Notice that 61% of the signs are articulated in neutral space and involve no body contact or proximity. The second most frequent location is 3, used as a location on the non-dominant hand in 4% of the signs. The five most common locations comprise 75% of the set of monomorphemic, non-location-changing signs in the NZSL database. The five least common locations comprise 0.84% of the sample.

(4-45) NZSL: combined location frequency

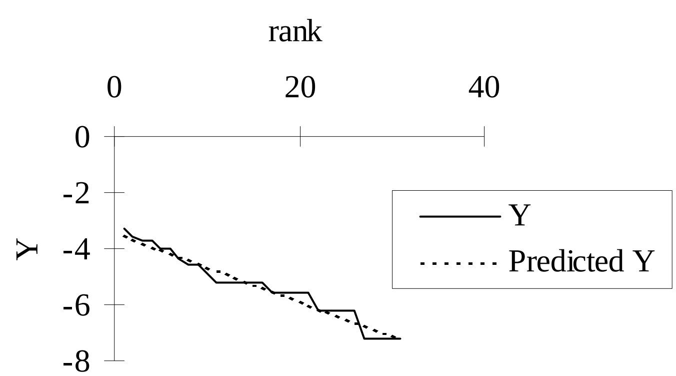
| rank | loc | count | freq |
|---|---|---|---|
| 1 | π | 364 | 0.6128 |
| 2 | 3 | 25 | 0.0421 |
| 3 | \| | 22 | 0.0370 |
| 4 | 38 | 17 | 0.0286 |
| 5 | ζ | 15 | 0.0253 |
| 6 | }□ | 12 | 0.0202 |
| 7 | ~ | 11 | 0.0185 |
| 8 | ξ | 11 | 0.0185 |
| 9 | σ□ | 9 | 0.0152 |
| 10 | ψ | 9 | 0.0152 |
| 11 | ϖ | 7 | 0.0118 |
| 12 | 49 | 7 | 0.0118 |
| 13 | \|□ | 6 | 0.0101 |
| 14 | ω | 6 | 0.0101 |
| 15 | ~□ | 5 | 0.0084 |
| 16 | \| | 4 | 0.0067 |
| 17 | □ | 4 | 0.0067 |
| 18 | ο□ | 4 | 0.0067 |
| 19 | θ | 4 | 0.0067 |
| 20 | θ□ | 4 | 0.0067 |
| 21 | σ | 4 | 0.0067 |
| 22 | { | 3 | 0.0051 |
| 23 | □ | 3 | 0.0051 |
| 24 | ο | 3 | 0.0051 |
| 25 | υ□ | 3 | 0.0051 |
| 26 | ϖ□ | 3 | 0.0051 |
| 27 | 29 | 2 | 0.0034 |
| 28 | 69 | 2 | 0.0034 |
| 29 | 78 | 2 | 0.0034 |
| 30 | ; | 2 | 0.0034 |
| 31 | □ | 2 | 0.0034 |
| 32 | 3≅B | 2 | 0.0034 |
| 33 | ο□□ | 2 | 0.0034 |
| 34 | ρ | 2 | 0.0034 |
| 35 | ψ□ | 2 | 0.0034 |
| 36 | ζ□ | 2 | 0.0034 |
| 37 | 28 | 1 | 0.0017 |
| 38 | : | 1 | 0.0017 |
| 39 | \|φ | 1 | 0.0017 |
| 40 | ~□□ | 1 | 0.0017 |
| 41 | □ | 1 | 0.0017 |
| 42 | 3≅A | 1 | 0.0017 |
| 45 | 3B | 1 | 0.0017 |
| 46 | 3B8 | 1 | 0.0017 |
| sum | | 594 | 1 |

The body location rank-frequency graph for NZSL is shown in (4-46). This graph is also an exponential decay curve, as shown by the linearity of the graph in (4-47). Linear regression applied to x = rank and Y = $\log_2$(frequency) results in an adjusted correlation coefficient R square value of 0.9603, indicating that location in NZSL declines with close fidelity to the exponential decay law, where $y = 0.09(2^{-0.12x})$.

(4-46) NZSL: body location rank-frequency graph

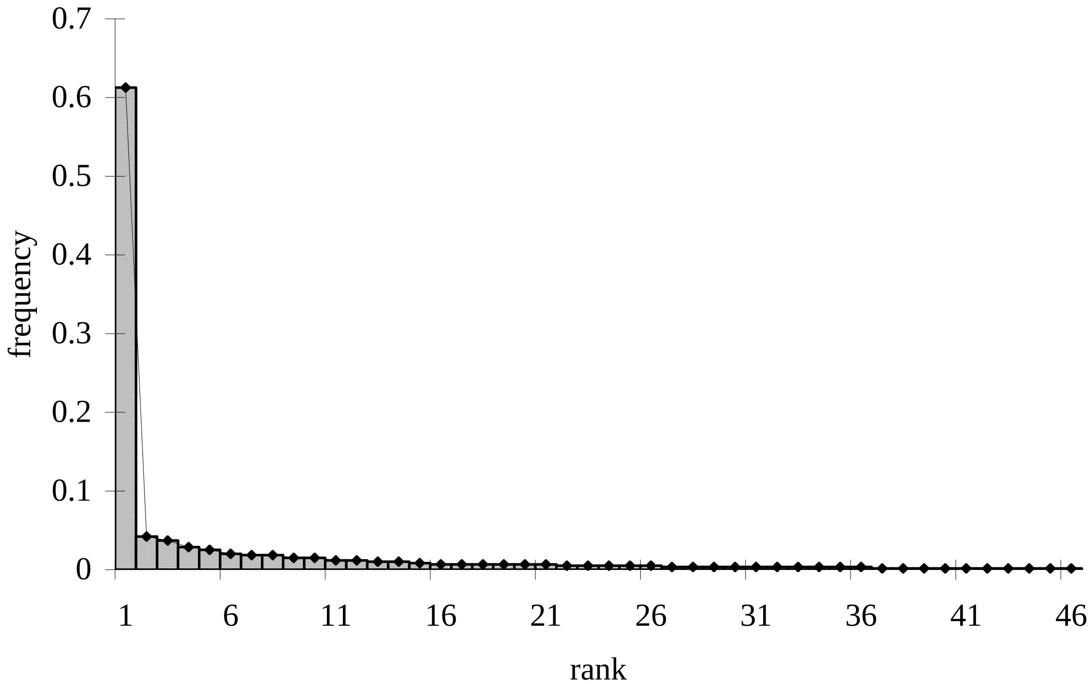



(4-47) NZSL: Y = $\log_2 y$ where x = rank and y = frequency

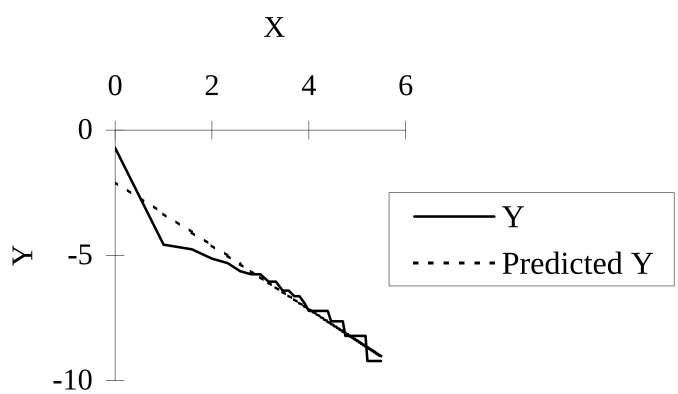



The rank-frequency graph for all NZSL locations is shown in (4-48). This graph is not an exponential decay curve; instead, it is a hyperbolic curve, as shown by the linearity of the graph in (4-49). Linear regression applied to X = $\log_2$(rank) and Y = $\log_2$(frequency) results in an adjusted correlation coefficient R square value of 0.9406, indicating that the rank frequency of location in NZSL, when neutral space and hand locations are included, declines with close fidelity to the power law, $y = 0.24x^{-1.3}$.

(4-48) NZSL: rank-frequency graph for all locations




(4-49) NZSL: X = $\log_2 x$ and Y = $\log_2 y$ where x = rank and y = frequency




#### 4.2.2.4 SVK: location rank-frequency

The distribution of body locations in monomorphemic, non-location-changing signs in the SVK database is shown in (4-50). The five most frequent locations, {ψ ~ σ ξ ζ}, comprise 56% of the set of signs in the SVK database with a body location, while the five least frequent locations, {}φ ο□ υ□ ψ□ ζ□}, comprise only 2.6% of the sample, with each of these five least frequent locations appearing in only one monomorphemic, non-location-changing sign each, SELKÄ *back*, VIERAS *guest*, SIPULI *onion*, KUKKA *flower*, and LEIKKI *joke*, respectively.

(4-50) SVK: body location frequency distribution

| **rank** | **loc** | **count** | **freq** |
|---|---|---|---|
| 1 | ψ | 26 | 0.1347 |
| 2 | ~ | 24 | 0.1244 |
| 3 | σ | 21 | 0.1088 |
| 4 | ξ | 21 | 0.1088 |
| 5 | ζ | 16 | 0.0829 |
| 6 | } | 10 | 0.0518 |
| 7 | ο | 9 | 0.0466 |
| 8 | ϖ | 9 | 0.0466 |
| 9 | υ | 8 | 0.0415 |
| 10 | θ | 7 | 0.0363 |
| 11 | σ□ | 7 | 0.0363 |
| 12 | { | 6 | 0.0311 |
| 13 | }□ | 5 | 0.0259 |
| 14 | ω | 4 | 0.0207 |
| 15 | \| | 3 | 0.0155 |
| 16 | □ | 3 | 0.0155 |
| 17 | ~□ | 2 | 0.0104 |
| 18 | □ | 2 | 0.0104 |
| 19 | □ | 2 | 0.0104 |
| 20 | □ | 2 | 0.0104 |
| 21 | \|□ | 1 | 0.0052 |
| 22 | }φ | 1 | 0.0052 |
| 23 | ο□ | 1 | 0.0052 |
| 24 | υ□ | 1 | 0.0052 |
| 25 | ψ□ | 1 | 0.0052 |
| 26 | ζ□ | 1 | 0.0052 |
| sum | | 193 | 1 |

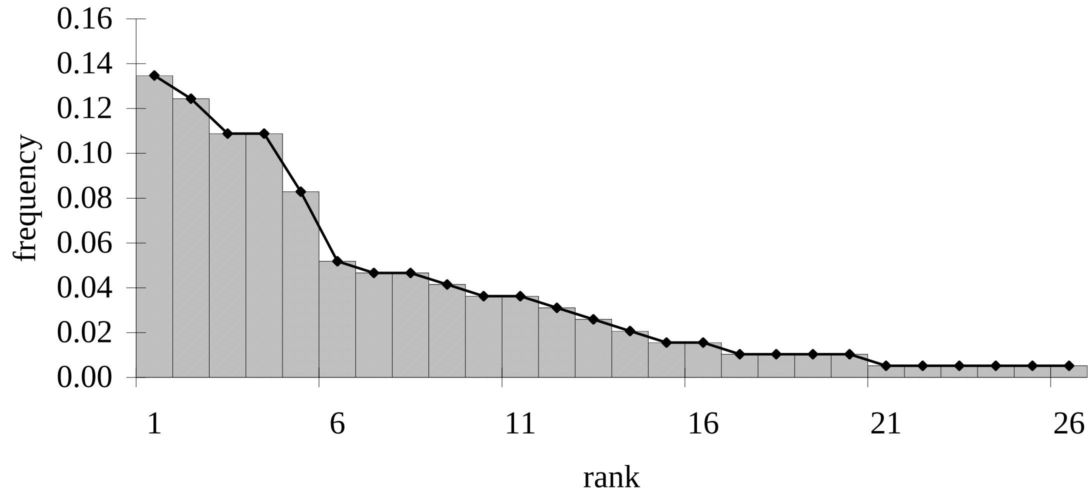
The distribution of all locations in monomorphemic, non-location-changing signs in the SVK database, including neutral space and hand locations, is shown in (4-51). Notice that 39% of the signs are articulated in neutral space and involve no body contact or proximity. The second most frequent location is 3, used as a location on the non-dominant hand in 14% of the signs. The five most common locations comprise 73% of the set of monomorphemic, non-location-changing signs in the SVK database. The five least common locations comprise 0.90% of the sample.

(4-51) SVK: combined location frequency

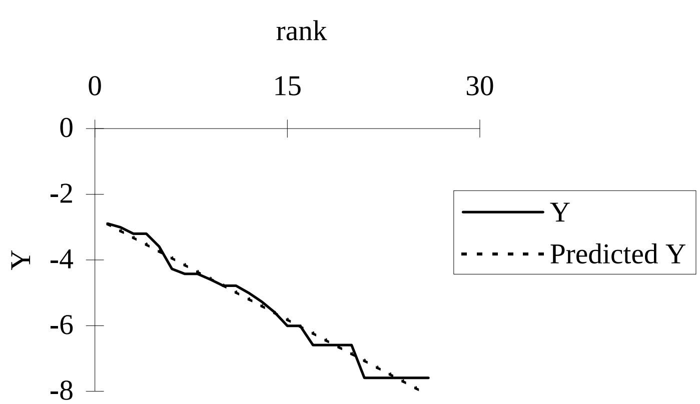
| **rank** | **loc** | **count** | **freq** |
|---|---|---|---|
| 1 | π | 214 | 0.3870 |
| 2 | 3 | 78 | 0.1410 |
| 3 | \| | 61 | 0.1103 |
| 4 | ψ | 26 | 0.0470 |
| 5 | ~ | 24 | 0.0434 |
| 6 | σ | 21 | 0.0380 |
| 7 | ξ | 21 | 0.0380 |
| 8 | ζ | 16 | 0.0289 |
| 9 | } | 10 | 0.0181 |
| 10 | ο | 9 | 0.0163 |
| 11 | ϖ | 9 | 0.0163 |
| 12 | υ | 8 | 0.0145 |
| 13 | θ | 7 | 0.0127 |
| 14 | σ□ | 7 | 0.0127 |
| 15 | { | 6 | 0.0108 |
| 16 | }□ | 5 | 0.0090 |
| 17 | ω | 4 | 0.0072 |
| 18 | 49 | 3 | 0.0054 |
| 19 | \| | 3 | 0.0054 |
| 20 | □ | 3 | 0.0054 |
| 21 | ~□ | 2 | 0.0036 |
| 22 | □ | 2 | 0.0036 |
| 23 | □ | 2 | 0.0036 |
| 24 | □ | 2 | 0.0036 |
| 25 | 2 | 1 | 0.0018 |
| 26 | 29 | 1 | 0.0018 |
| 27 | 38 | 1 | 0.0018 |
| 28 | \|□ | 1 | 0.0018 |
| 29 | }φ | 1 | 0.0018 |
| 30 | 3≅ | 1 | 0.0018 |
| 31 | ο□ | 1 | 0.0018 |
| 32 | υ□ | 1 | 0.0018 |
| 33 | ψ□ | 1 | 0.0018 |
| 34 | ζ□ | 1 | 0.0018 |
| sum | | 553 | 1 |

The body location rank-frequency graph for SVK is shown in (4-52). This graph is also an exponential decay curve, as shown by the linearity of the graph in (4-53). Linear regression applied to x = rank and Y = $\log_2$(frequency) results in an adjusted correlation coefficient R square value of 0.9767, indicating that location in SVK declines with close fidelity to the exponential decay law, where $y = 0.15(2^{-0.21x})$.

(4-52) SVK: body location rank-frequency graph

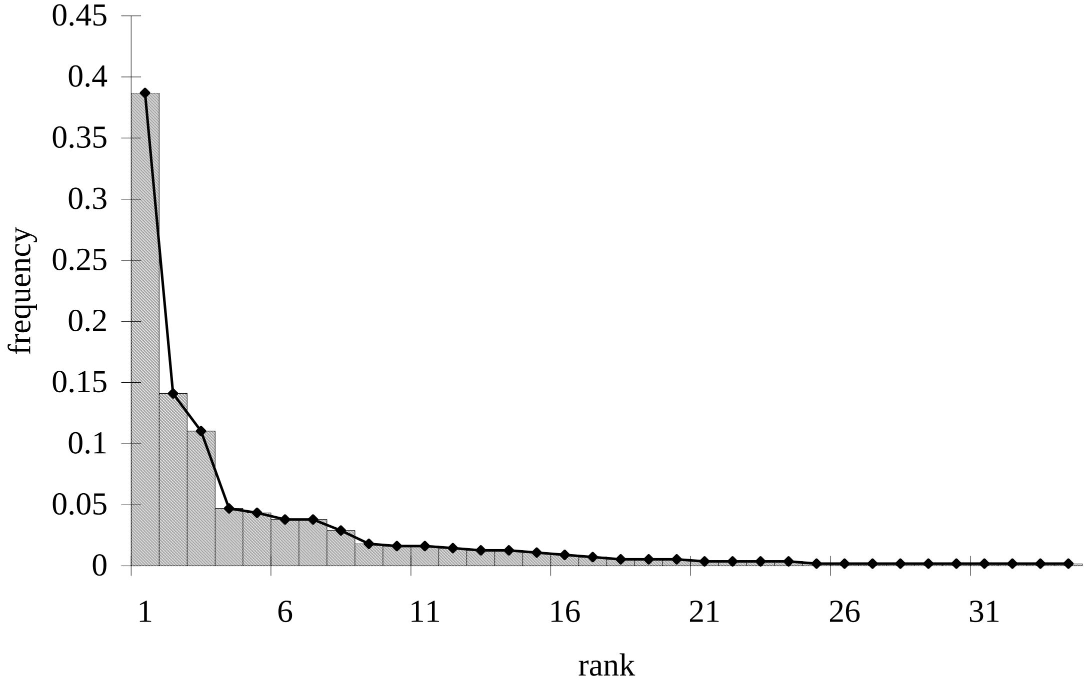



(4-53) SVK: Y = $\log_2 y$ where x = rank and y = frequency

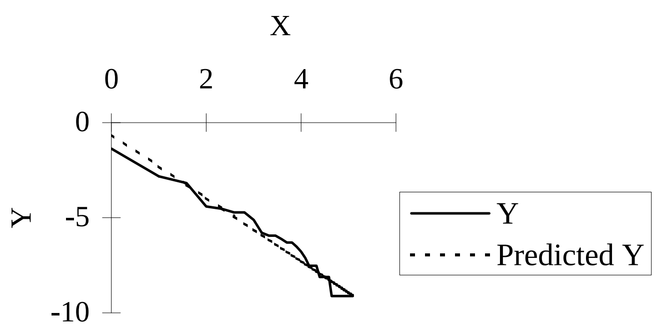



The rank-frequency graph for all SVK locations is shown in (4-54). This graph is not an exponential decay curve; instead, it is a hyperbolic curve, as shown by the linearity of the graph in (4-55). Linear regression applied to X = $\log_2$(rank) and Y = $\log_2$(frequency) results in an adjusted correlation coefficient R square value of 0.9601, indicating that the rank frequency of location in SVK, when neutral space and hand locations are included, declines with close fidelity to the power law, $y = 0.64x^{-1.7}$ .

(4-54) SVK: rank-frequency graph for all locations




(4-55) SVK: X = $\log_2 x$ and Y = $\log_2 y$ where x = rank and y = frequency




#### 4.2.2.5 VSVK: location rank-frequency

The distribution of body locations in monomorphemic, non-location-changing signs in the VSVK database is shown in (4-56). The five most frequent locations, {ψ σ ~ ζ □}, comprise 59% of the set of signs in the SVK database with a body location, while the five least frequent locations, { θ □ σ□ |□ ψ□}, comprise 10% of the sample, with the last four of these five locations appearing in only one monomorphemic, non-location-changing sign each, VERI *blood*, TALONPOIKA *peasant*, VALKOINEN *white*, and HUUTAA *shout*, respectively.

(4-56) VSVK: body location frequency distribution

| rank | loc | count | freq |
|---|---|---|---|
| 1 | ψ | 12 | 0.2034 |
| 2 | σ | 9 | 0.1525 |
| 3 | ~ | 5 | 0.0847 |
| 4 | ζ | 5 | 0.0847 |
| 5 | □ | 4 | 0.0678 |
| 6 | υ | 4 | 0.0678 |
| 7 | ξ | 4 | 0.0678 |
| 8 | \| | 2 | 0.0339 |
| 9 | ο□ | 2 | 0.0339 |
| 10 | θ□ | 2 | 0.0339 |
| 11 | ϖ | 2 | 0.0339 |
| 12 | ω | 2 | 0.0339 |
| 13 | θ | 2 | 0.0339 |
| 14 | □ | 1 | 0.0169 |
| 15 | σ□ | 1 | 0.0169 |
| 16 | \|□ | 1 | 0.0169 |
| 17 | ψ□ | 1 | 0.0169 |
| sum | | 59 | 1 |

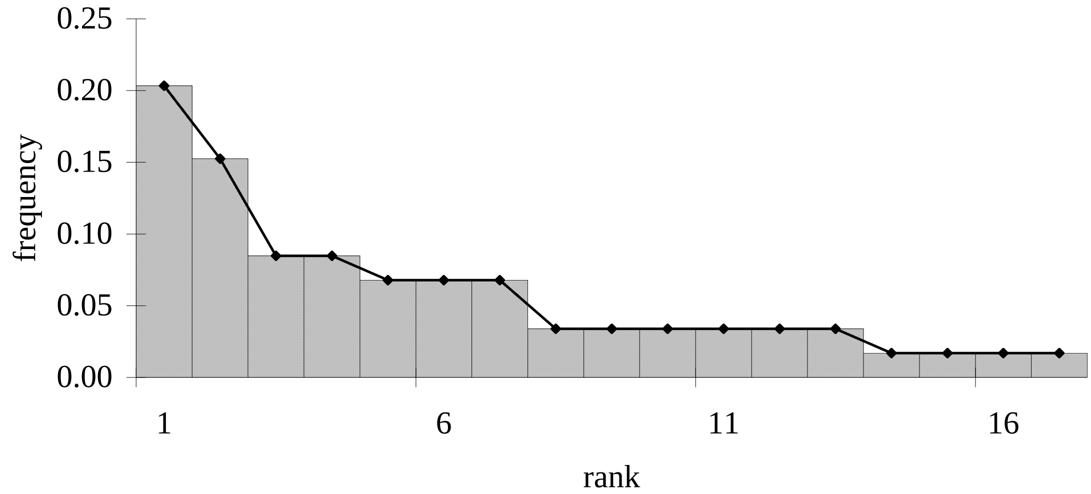
The distribution of all locations in monomorphemic, non-location-changing signs in the VSVK database, including neutral space and hand locations, is shown in (4-57). Notice that 52% of the signs are articulated in neutral space and involve no body contact or proximity. The second most frequent location is 3, used as a location on the non-dominant hand in 10% of the signs. The five most common locations comprise 80% of the set of monomorphemic, non-location-changing signs in the VSVK database. The five least common locations comprise 2.0% of the sample.

(4-57) VSVK: combined location frequency

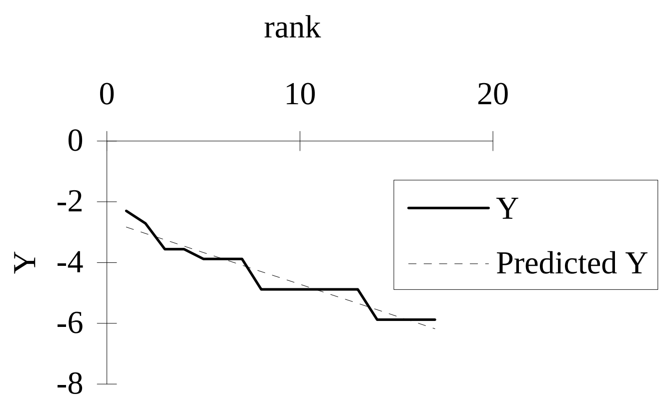
| rank | loc | count | freq |
|---|---|---|---|
| 1 | π | 129 | 0.5202 |
| 2 | 3 | 26 | 0.1048 |
| 3 | \| | 22 | 0.0887 |
| 4 | ψ | 12 | 0.0484 |
| 5 | σ | 9 | 0.0363 |
| 6 | ~ | 5 | 0.0202 |
| 7 | 29 | 5 | 0.0202 |
| 8 | ζ | 5 | 0.0202 |
| 9 | □ | 4 | 0.0161 |
| 10 | υ | 4 | 0.0161 |
| 11 | ξ | 4 | 0.0161 |
| 12 | \| | 3 | 0.0121 |
| 13 | ο□ | 2 | 0.0081 |
| 14 | θ | 2 | 0.0081 |
| 15 | θ□ | 2 | 0.0081 |
| 16 | ϖ | 2 | 0.0081 |
| 17 | ω | 2 | 0.0081 |
| 18 | \|□ | 1 | 0.0040 |
| 19 | ψ□ | 1 | 0.0040 |
| 20 | ; | 1 | 0.0040 |
| 21 | □ | 1 | 0.0040 |
| 22 | 2 | 1 | 0.0040 |
| 23 | 28 | 1 | 0.0040 |
| 24 | 38 | 1 | 0.0040 |
| 25 | 59 | 1 | 0.0040 |
| 26 | 78 | 1 | 0.0040 |
| 27 | σ□ | 1 | 0.0040 |
| sum | | 248 | 1 |

The body location rank-frequency graph for VSVK is shown in (4-58). This graph is also an exponential decay curve, as shown by the linearity of the graph in (4-59). Linear regression applied to x = rank and Y = $\log_2$(frequency) results in an adjusted correlation coefficient R square value of 0.9125, indicating that location in VSVK declines with close fidelity to the exponential decay law, where $y = 0.16(2^{-0.21x})$.

(4-58) VSVK: body location rank-frequency graph

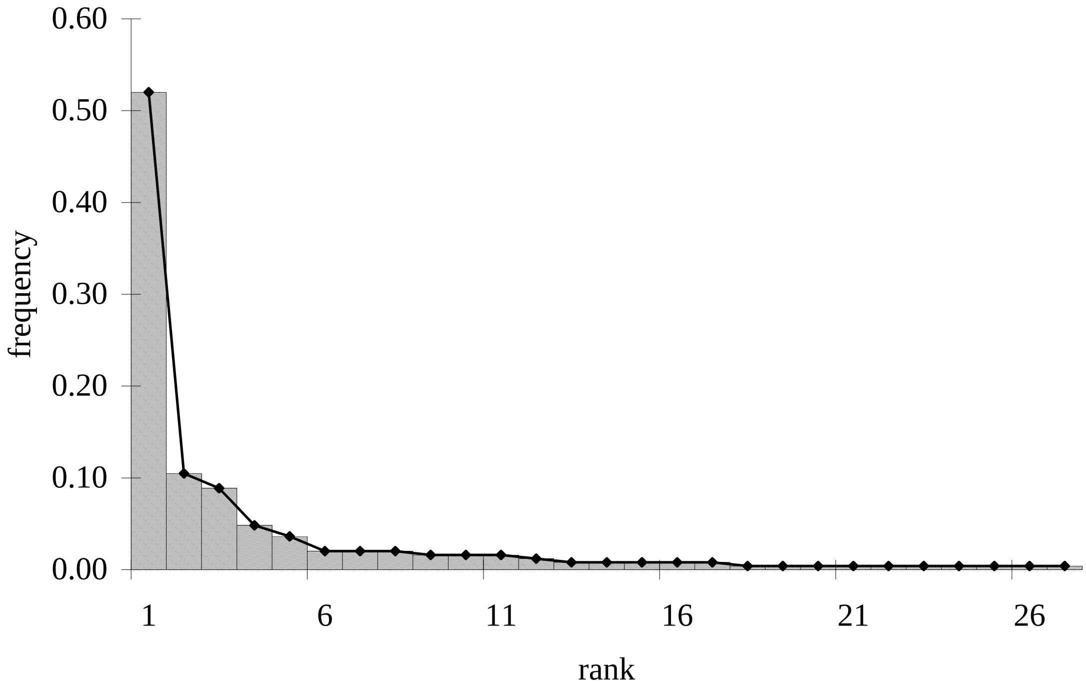



(4-59) VSVK: Y = $\log_2 y$ where x = rank and y = frequency

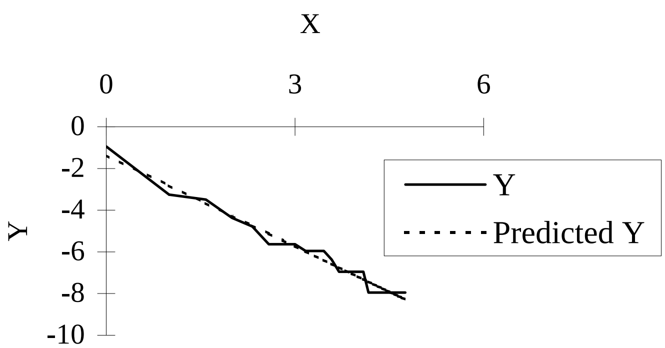



The rank-frequency graph for all VSVK locations is shown in (4-60). This graph is not an exponential decay curve; instead, it is a hyperbolic curve, as shown by the linearity of the graph in (4-61). Linear regression applied to X = $\log_2$(rank) and Y = $\log_2$(frequency) results in an adjusted correlation coefficient R square value of 0.9745, indicating that the rank frequency of location in VSVK, when neutral space and hand locations are included, declines with close fidelity to the power law, $y = 0.39x^{-1.5}$.

(4-60) VSVK: rank-frequency graph for all locations




(4-61) VSVK: $X = \log_2 x$ and $Y = \log_2 y$ where x = rank and y = frequency

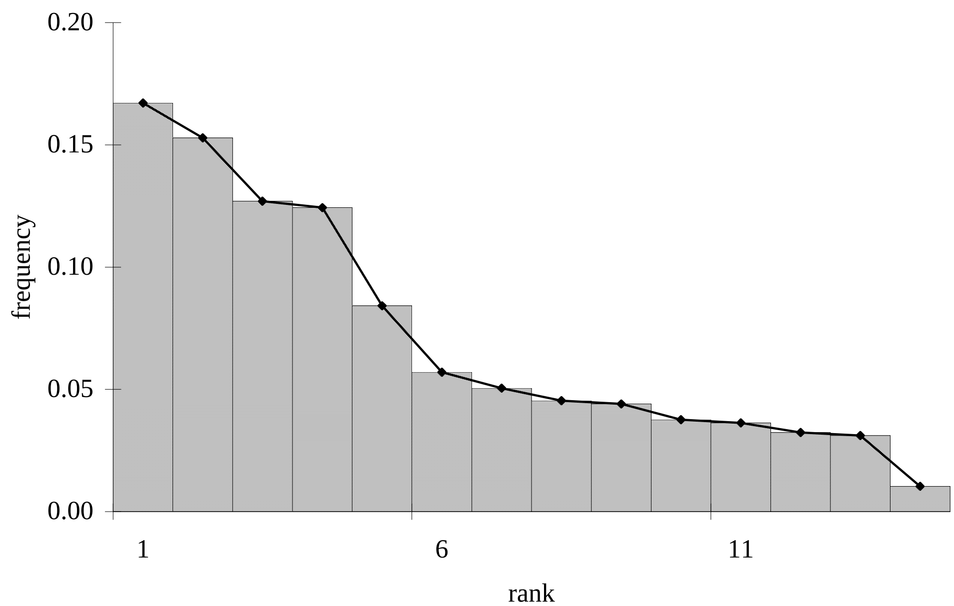



#### 4.2.2.6 LIS: location rank frequency

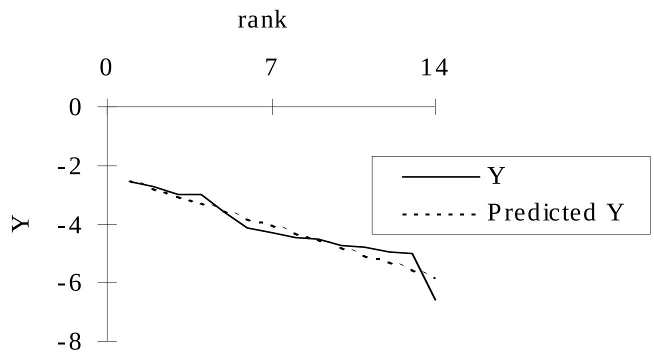
Because of the differences in the methods used in compiling the inventory data for Italian Sign Language (Lingua Italiana dei Segni: LIS) and the other four databases, no direct comparison of location inventory size and composition is made. The location categories used in

Pietrandrea (1998) are broader than the ones I used; for example, no distinction was notated between the center and the side of a location. Also, the nondominant hand was counted as only one location, and it is not clear whether the location of Type 2 signs is neutral space or the nondominant hand. For these reasons only body locations are examined. LIS still patterns like ASL, KSL, NZSL, and SVK, though the correspondence is not so exact, probably due to the differences in methodology. Neutral space is the most frequent location, used in 48% of the lexicon, and the nondominant hand is the second most frequent, used in 13% of the lexicon. The LIS body location rank-frequency graph is shown in (4-62). Like the other body location rank-frequency graphs, it is also an exponential decay curve, as shown by the linearity of the graph in (4-63). Linear regression applied to x = rank and Y = $\log_2$(frequency) results in an adjusted correlation coefficient R square value of 0.9029, indicating that handshape frequency declines with fairly close fidelity to the exponential decay law, where $y = 0.21(2^{-0.25x})$ for the LIS graph.

(4-62) LIS: rank-frequency graph for body locations



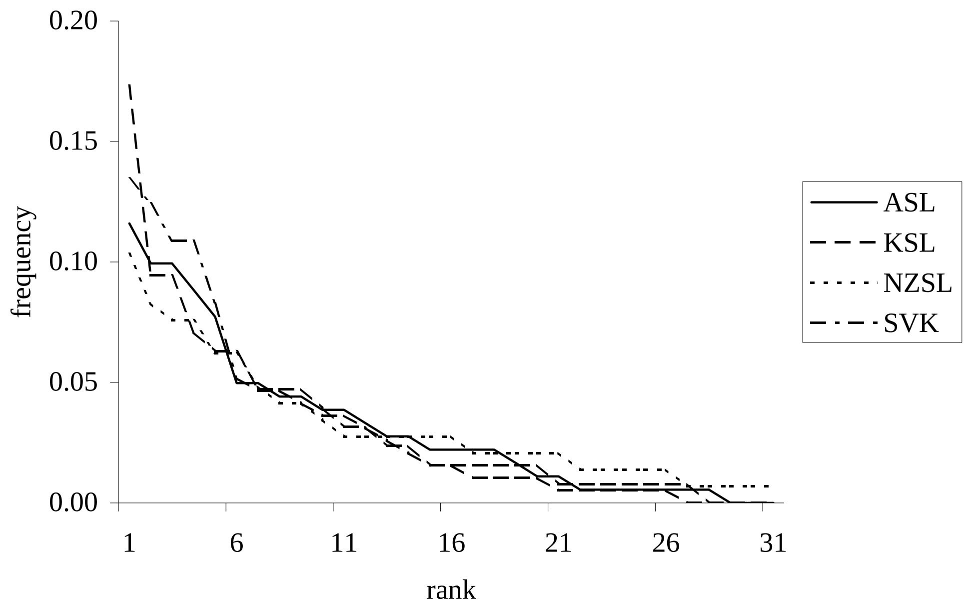
(4-63) LIS: $Y = \log_2 y$ where x = rank and y = frequency



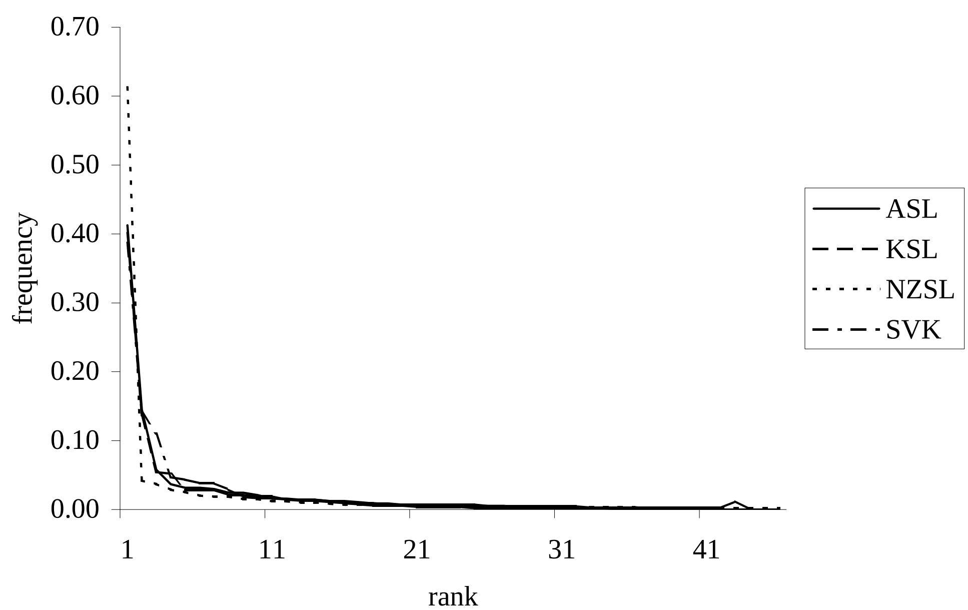

### 4.2.3 Comparison

In this section the location distributions are compared. The rank-frequency graphs of body location in ASL, KSL, NZSL, and SVK have been shown to be exponential decay curves, while the graphs for all locations are hyperbolic curves. How similar are these curves to each other? Is it the case that although all are the same type of curve, they are nevertheless significantly different instantiations of $y = a2^{-cx}$ and $y = ax^{-c}$, or is the variation attributable to sampling error? The rank-frequency graph for body location of all four languages is shown in (4-64), while the graph for all locations is shown in (4-65).

(4-64) All languages: rank-frequency graph for body locations




(4-65) All languages: rank-frequency graph for all locations




Examining two languages at a time, the frequencies associated with each rank for body locations were paired, and the signed differences between these frequencies were tested to ascertain whether they can be considered a random sample from a population with mean $\mu = 0$ by using the paired-sample *t* test.  The results of this pair-wise comparison show at the 0.001 level of significance that the differences between each pair of body location distributions is not significant.  This procedure was repeated for all languages with the same results.  Thus, all curves are very similar to each other.

### 4.2.4 Universal location pool

Pooling the locations used in ASL, KSL, SVK and NZSL provides a rough idea of the "International Location Alphabet," analogous to the International Phonetic Alphabet for spoken languages.  Shown in (4-66) is an approximation of the set of body locations available for use in the world's sign languages, while in (4-67) is an approximation of the set of other locations.

(4-66) "International Body Location Alphabet"

<table>
<tr><td>ρ</td><td></td><td></td><td></td><td></td><td></td></tr>
<tr><td>θ</td><td>θ□</td><td></td><td>θφ</td><td></td><td></td></tr>
<tr><td>σ</td><td>σ□</td><td></td><td></td><td></td><td></td></tr>
<tr><td>υ</td><td>υ□</td><td>ω</td><td></td><td></td><td></td></tr>
<tr><td>ϖ</td><td>ϖ□</td><td>ξ</td><td></td><td></td><td></td></tr>
<tr><td>ψ</td><td>ψ□</td><td></td><td></td><td></td><td></td></tr>
<tr><td>ζ</td><td>ζ□</td><td></td><td></td><td></td><td></td></tr>
<tr><td>{</td><td></td><td></td><td></td><td></td><td></td></tr>
<tr><td>|</td><td>|□</td><td></td><td>|φ</td><td></td><td></td></tr>
<tr><td>}</td><td>}□</td><td></td><td>}φ</td><td>□</td><td></td></tr>
<tr><td>~</td><td>~□</td><td>~□□</td><td></td><td>□</td><td></td></tr>
<tr><td>ο</td><td>ο□</td><td>ο□□</td><td>οφ</td><td>□</td><td></td></tr>
<tr><td>□</td><td>□□</td><td>□□□</td><td></td><td>□</td><td>□</td></tr>
</table>

(4-67) "International Other Location Alphabet"

<table>
<tr><td>π</td><td>2</td><td>49</td><td>:</td><td>49□</td><td>69</td><td>3</td><td>3B</td><td>78</td><td>;</td></tr>
<tr><td>|</td><td>29</td><td>48</td><td><</td><td></td><td></td><td>38</td><td>3B8</td><td></td><td></td></tr>
<tr><td></td><td>28</td><td>48□X</td><td></td><td></td><td></td><td>39</td><td>3≅B</td><td></td><td></td></tr>
<tr><td></td><td></td><td>4≅B</td><td></td><td></td><td></td><td>3≅</td><td>3≅A</td><td></td><td></td></tr>
</table>

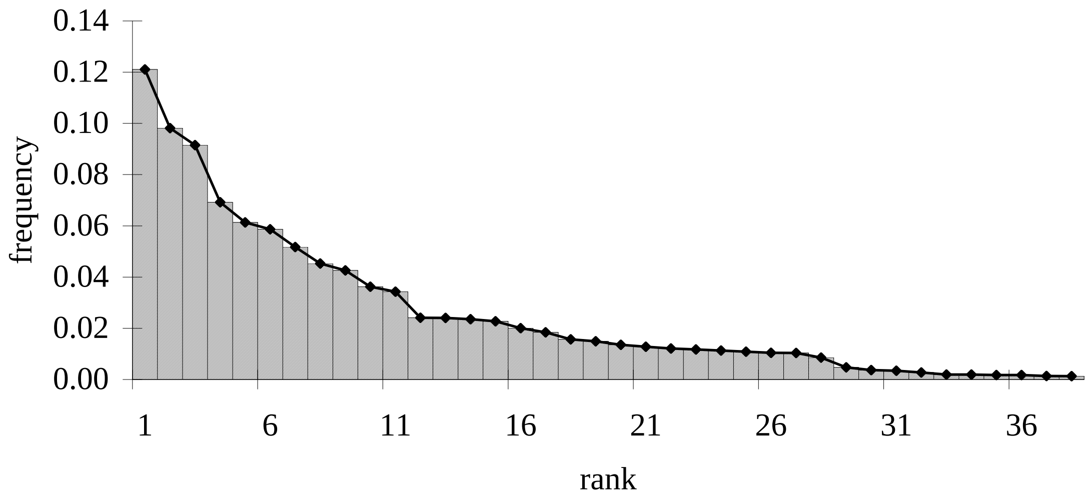
Moreover, pooling the lexicons of the four languages provides an approximation of the universal distribution of body location, as shown in (4-68).

(4-68) Pooled data: body location frequency distribution

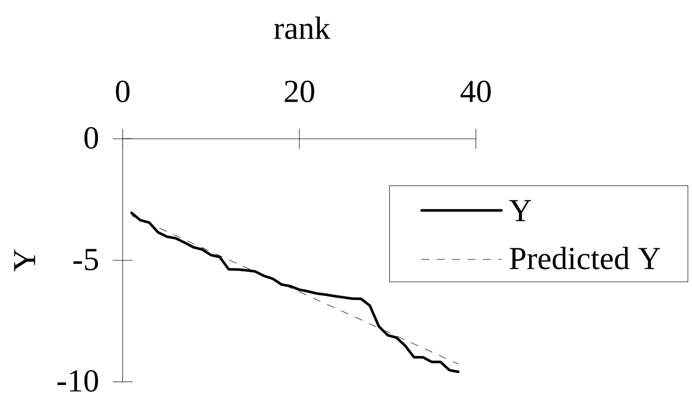
| rank | loc | count | freq |
|---|---|---|---|
| 1 | ψ | 78 | 0.1211 |
| 2 | ~ | 64 | 0.0982 |
| 3 | ξ | 59 | 0.0915 |
| 4 | ζ | 45 | 0.0692 |
| 5 | σ | 40 | 0.0613 |
| 6 | σ□ | 38 | 0.0587 |
| 7 | ϖ | 34 | 0.0517 |
| 8 | }□ | 29 | 0.0453 |
| 9 | θ | 28 | 0.0426 |
| 10 | υ | 23 | 0.0362 |
| 11 | { | 22 | 0.0343 |
| 12 | \| | 16 | 0.0241 |
| 13 | ο | 16 | 0.0241 |
| 14 | ω | 15 | 0.0235 |
| 15 | ~□ | 15 | 0.0227 |
| 16 | θ□ | 13 | 0.0201 |
| 17 | □ | 12 | 0.0184 |
| 18 | } | 10 | 0.0157 |
| 19 | \|□ | 10 | 0.015 |
| 20 | ψ□ | 9 | 0.0135 |
| 21 | ρ | 8 | 0.0128 |
| 22 | ο□ | 8 | 0.0121 |
| 23 | □ | 8 | 0.0117 |
| 24 | □ | 7 | 0.0113 |
| 25 | □ | 7 | 0.0108 |
| 26 | ϖ□ | 7 | 0.0105 |
| 27 | υ□ | 7 | 0.0104 |
| 28 | □ | 6 | 0.0085 |
| 29 | ζ□ | 3 | 0.0047 |
| 30 | \|φ | 2 | 0.0037 |
| 31 | ο□□ | 2 | 0.0034 |
| 32 | □ | 2 | 0.0027 |
| 33 | □□ | 1 | 0.002 |
| 34 | θφ | 1 | 0.002 |
| 35 | ~□□ | 1 | 0.0017 |
| 36 | οφ | 1 | 0.0017 |
| 37 | □□□ | 1 | 0.0014 |
| 38 | }φ | 1 | 0.0013 |
| sum | | 648 | 1 |

The body location rank-frequency graph for this pooled data, a sample of 648 signs from all four languages, is given in (4-69). This graph is also an exponential decay curve, as shown by the linearity of the graph in (4-70). Linear regression applied to x = rank and Y = $\log_2$(frequency) results in an adjusted correlation coefficient R square value of 0.9673, indicating that body location frequency declines with close fidelity to the exponential decay law, where $y = 0.13(2^{-0.17x})$ for the pooled data.

(4-69) Pooled data: body location rank-frequency graph




(4-70) Pooled data: Y= $\log_2 y$ where x = rank and y = frequency



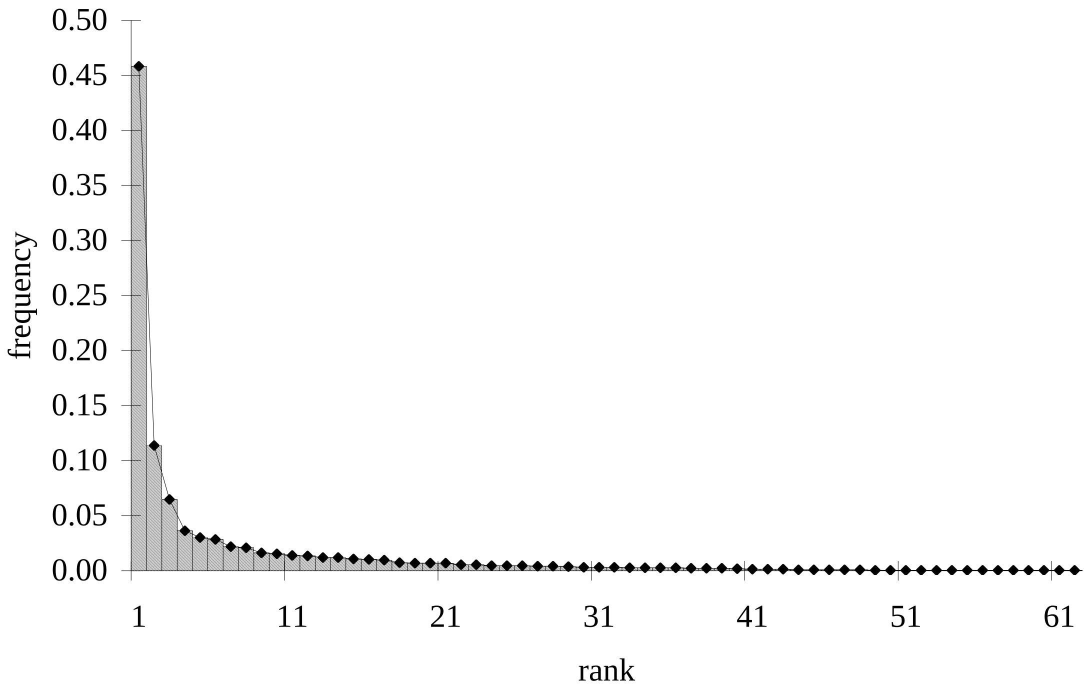

Pooling the lexicons of the four languages also provides an approximation of the universal distribution of all locations combined, as shown in (4-71).

(4-71) Pooled data: frequency distribution for all locations

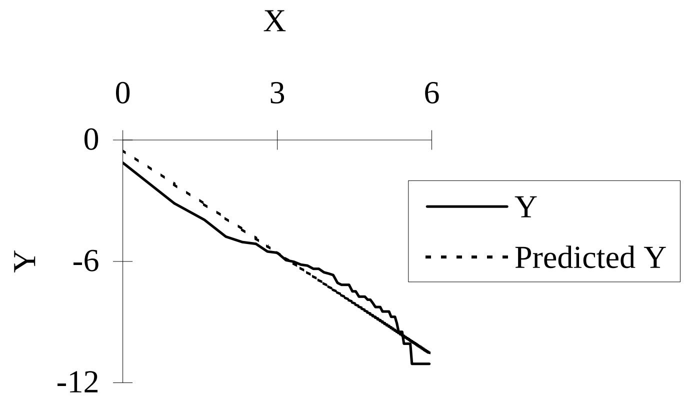
| **rank** | **loc** | **freq** | **rank** | **loc** | **freq** | **rank** | **loc** | freq |
|---|---|---|---|---|---|---|---|---|
| 1 | π | 0.4583 | 26 | θ□ | 0.0047 | 51 | ~□□ | 0.0005 |
| 2 | 3 | 0.1138 | 27 | \|□ | 0.0042 | 52 | □□ | 0.0005 |
| 3 | □ | 0.0648 | 28 | ψ□ | 0.0042 | 53 | □□□ | 0.0005 |
| 4 | ψ | 0.0364 | 29 | ρ | 0.0037 | 54 | 3≅ | 0.0005 |
| 5 | ~ | 0.0303 | 30 | □ | 0.0033 | 55 | 3≅A | 0.0005 |
| 6 | ξ | 0.0284 | 31 | □ | 0.0033 | 56 | 3B8 | 0.0005 |
| 7 | σ | 0.0219 | 32 | o□ | 0.0033 | 57 | 4≅B | 0.0005 |
| 8 | ζ | 0.021 | 33 | 78 | 0.0028 | 58 | 48□X | 0.0005 |
| 9 | σ□ | 0.0163 | 34 | □ | 0.0028 | 59 | 49□ | 0.0005 |
| 10 | ϖ | 0.0154 | 35 | υ□ | 0.0028 | 60 | □□ | 0.0005 |
| 11 | 29 | 0.014 | 36 | ϖ□ | 0.0028 | 61 | θ | 0.0005 |
| 12 | }□ | 0.0135 | 37 | 2 | 0.0023 | 62 | θφ | 0.0005 |
| 13 | 49 | 0.0121 | 38 | ; | 0.0023 | sum | | 1 |
| 14 | θ | 0.0121 | 39 | □ | 0.0023 | | | |
| 15 | υ | 0.0107 | 40 | 3B | 0.0019 | | | |
| 16 | { | 0.0103 | 41 | 39 | 0.0014 | | | |
| 17 | 38 | 0.0098 | 42 | < | 0.0014 | | | |
| 18 | o | 0.0075 | 43 | ζ□ | 0.0014 | | | |
| 19 | \| | 0.007 | 44 | 69 | 0.0009 | | | |
| 20 | ~□ | 0.007 | 45 | : | 0.0009 | | | |
| 21 | ω | 0.007 | 46 | □ | 0.0009 | | | |
| 22 | } | 0.0056 | 47 | \|φ | 0.0009 | | | |
| 23 | □ | 0.0056 | 48 | o□□ | 0.0009 | | | |
| 24 | 28 | 0.0047 | 49 | 48 | 0.0005 | | | |
| 25 | 3≅B | 0.0047 | 50 | }φ | 0.0005 | | | |

The rank-frequency graph of all locations for this pooled data, a sample of 2,145 signs from all four languages, is given in (4-72). This graph is a hyperbolic curve, as shown by the linearity of the graph in (4-73). Linear regression applied to X = $\log_2$(rank) and Y=$\log_2$(frequency) results in an adjusted correlation coefficient R square value of 0.9251, indicating that location frequency declines with close fidelity to the power law, where $y = 0.71x^{-1.7}$ for the pooled data.

(4-72) Pooled data: rank-frequency graph for all locations



(4-73) Pooled data: $X = \log_2 x$ and $Y = \log_2 y$ where x = rank and y = frequency




## 4.3 Location by number of hands

This section investigates the distribution of locations across one-handed and two-handed signs. For this analysis signs that are articulated on the hand, such as Type 2 and Type 3 signs, are excluded; clearly these signs cannot be two-handed. Signs that have the arm as their location are also excluded from the analysis since it is the case that all such signs are one-handed. While it is physically possible to articulate, for example, a two-handed sign with the elbow as the location, no such signs exist in any of the four databases.[76]

### 4.3.1 Neutral space versus contact

The first question addressed is whether a one-handed Type 0 sign is more likely to contact the body or to be articulated in neutral space. Contact entails either actual contact between an articulator and a location or proximity to this location (see footnote 5). Overwhelmingly, in all four languages, chi-square tests indicate that more one-handed signs contact the body than expected and more two-handed signs are articulated in neutral space than expected (p=0.000 for all four languages), as can be seen in (4-74). (Since each language has different grand totals of signs, frequencies rather than counts are reported.) NZSL differs slightly in that it has more one-handed signs articulated in neutral space than contacting the body, but it also has a much larger proportion of signs articulated in neutral space than do the other languages.

---

[76] See section A.2 of Appendix A for categorizations of locations, such as arm locations, face locations, and torso locations.

(4-74) One- and two-handed signs articulated with contact and in neutral space

ASL

| obs | one | two | sum |
|---|---|---|---|
| c | **0.28** | 0.13 | 0.41 |
| ns | 0.14 | **0.45** | 0.59 |
| sum | 0.42 | 0.58 | 1.00 |

| exp | one | two |
|---|---|---|
| c | 0.17 | 0.24 |
| ns | 0.25 | 0.34 |

KSL

| obs | one | two | sum |
|---|---|---|---|
| c | **0.31** | 0.12 | 0.43 |
| ns | 0.19 | **0.38** | 0.57 |
| sum | 0.50 | 0.50 | 1.00 |

| exp | one | two |
|---|---|---|
| c | 0.21 | 0.22 |
| ns | 0.28 | 0.29 |

NZSL

| obs | one | two | sum |
|---|---|---|---|
| c | **0.26** | 0.03 | 0.29 |
| ns | 0.34 | **0.37** | 0.71 |
| sum | 0.60 | 0.40 | 1.00 |

| exp | one | two |
|---|---|---|
| c | 0.17 | 0.12 |
| ns | 0.43 | 0.29 |

SVK

| obs | one | two | sum |
|---|---|---|---|
| c | **0.39** | 0.08 | 0.47 |
| ns | 0.15 | **0.38** | 0.53 |
| sum | 0.54 | 0.46 | 1.00 |

| exp | one | two |
|---|---|---|
| c | 0.26 | 0.21 |
| ns | 0.29 | 0.24 |

### 4.3.2 Contact locations versus number of hands

The second question is whether certain locations prefer one-handed or two-handed signs when neutral space, in addition to hand and arm locations, is excluded. For each language, one-handed and two-handed signs made in contact with or proximity to the body were included in a data set that consisted of ordered pairs of the form (location, number of hands). For example, the ordered pair for the ASL sign MOTHER is (ζ, 1). The mutual information value of this data set was compared to the mutual information values of the 1,000 scrambled data sets. In all four cases, the mutual information was high enough to signal dependence between location and number of hands (ASL: p=0.000; KSL: p=0.019; NZSL: p=0.024; SVK: p=0.000).

Examining the four databases shows that in general if a two-handed sign occurs at a particular location, a one-handed sign will occur at that location. There are two exceptions in KSL: □ has three two-handed signs, and τ has one two-handed sign; there are no one-handed signs at these two locations. There are three exceptions in ASL: □ has one two-handed sign, □□□ has one two-handed sign, and ~□ has six two-handed signs; there are no one-handed signs at these three locations. These exceptions suggest that an attraction between the torso and two-handed signs might be part of this dependence.

Frishberg (1975) found the following two historical changes in ASL: two-handed signs at the face become one-handed, and one-handed signs below the face become two-handed. Another distinction Frishberg (1975) finds relevant in historical change is that of center versus perimeter location. She reports that the location of signs at the face move from the middle to the side, while the location of signs below the face move from the side to the middle. Similarly, Siple (1973; 1978; 1980) explains that human vision is most accurate in the area surrounding the viewer's focal point, the point on which the viewer's eyes focus, which for sign language viewers is the middle of the face. Vision is less accurate farther out from this center point. Based on this property of the human visual system, she predicts that signs that are made close to the perceiver's focal point, that is, the center of the face, are more likely to be two-handed, while signs made farther away from the face are more likely to be two-handed.[77] An examination of the ASL database shows that more one-handed signs occur on the face and more two-handed signs occur on the body, as described in detail below.[78] Are these generalizations true of the other three languages? Is the middle / side distinction relevant to whether a sign is one-handed or two-handed?

#### 4.3.2.1 ASL: number of hands versus contact location

For the first analysis, each ASL sign was labeled as one-handed or two-handed and as having a face location or a torso location. A chi-square analysis indicates with p=0.001 that number of hands is dependent on the face / torso distinction. More one-handed signs than expected have the face as their location, while more two-handed signs than expected have the torso as their location. For the second analysis, the location of each sign was labeled as middle or side.[79] A chi-square analysis indicates with p=0.000 that the number of hands is dependent on the middle / side distinction. More one-handed signs than expected have center locations, while more two-handed signs than expected have side locations. As the tables in (4-75) indicate, these effects

---

[77] See section 5.2.5 for further discussion of Siple (1973; 1978; 1980).

[78] Frishberg (1975) defines the face strictly; she notes that signs made on the periphery of the face, such as DEER, have one-handed and two-handed variants.

[79] I use *middle* as a neutral term to avoid the spatial, visual and phonetic connotations that the more specific terms *perimeter* or *peripheral* might carry.

are additive, so that far more than expected one-handed signs are articulated at the middle of the face, while far more than expected two-handed signs are articulated at the side of the torso. For those locations where the demands of these two tendencies compete, that is, signs articulated at the side of the face and the middle of the torso, there is independence between the variables of location and hand number. I(loc; hand number)=0.137, which indicates that there is independence, with p=0.490, where only those signs articulated at the side of the face and the middle of the torso are included in the data set.

(4-75) ASL: number of hands versus location

| observed | one | two | sum |
|---|---|---|---|
| face: middle | **56** | 5 | 61 |
| side | 21 | 19 | 40 |
| torso: middle | 15 | 8 | 23 |
| side | 2 | **12** | 14 |
| sum | 94 | 44 | 138 |

| expected | one | two |
|---|---|---|
| face: middle | **42** | 19 |
| side | 27 | 13 |
| torso: middle | 16 | 7 |
| side | 10 | **4** |

### 4.3.2.2 KSL: number of hands versus contact location

KSL was analyzed the same way as ASL, but with different results. To determine whether the dependence between the variables of location and number of hands can be accounted for by the same location distinctions relevant in ASL, each KSL sign was labeled as one-handed or two-handed and as having a face location or a torso location. A chi-square analysis indicates with p=0.489 that number of hands is independent of the face / torso distinction. Next, the location of each sign was labeled as middle or side. A chi-square analysis indicates with p=0.743 that the number of hands is independent of the middle / side distinction.

(4-76) KSL: number of hands versus location

| observed | one | two | sum |
|---|---|---|---|
| face: middle | 41 | 11 | 52 |
| side | 29 | 13 | 42 |
| torso: middle | 10 | 9 | 19 |
| side | 9 | 0 | 9 |
| sum | 89 | 33 | 122 |

| expected | one | two |
|---|---|---|
| face: middle | 38 | 14 |
| side | 31 | 11 |
| torso: middle | 14 | 5 |
| side | 7 | 2 |

The hand number is independent of location when only locations in the middle of the face and the side of the torso are included ( I(loc; hand number)=0.210, p=0.138). Signs articulated at the middle of the face have more one-handed signs than expected, as was the case with ASL, and as will be seen to be the case with NZSL and SVK. However, signs articulated at the side of the torso do also. Middle face and side torso does not seem to be an externally motivated grouping, nor does side face and middle torso. Hence, it appears that the face / torso and middle / side are not relevant distinctions in KSL as they were in ASL. I have examined the mutual information calculation for location versus hand number and am unable to determine a natural description of the locations that have predominantly one-handed signs versus those that are predominantly two-handed.

#### 4.3.2.3 NZSL: number of hands versus contact location

NZSL was analyzed the same way as ASL, with similar results. To determine whether the dependence between the variables of location and hand number can be accounted for by the same location distinctions relevant in ASL, each NZSL sign was labeled as one-handed or two-handed and as having a face location or a torso location. Because there are cells with counts below five, the chi-square test cannot be used. The mutual information significance program indicates with p=0.016 that number of hands and the face / torso distinction are dependent variables. Next, the location of each sign was labeled as middle or side. The mutual information significance program indicates with p=0.000 that the number of hands and the middle / side distinction are dependent variables.

(4-77) NZSL: number of hands versus location

| observed | | one | two | sum |
|---|---|---|---|---|
| face: | middle | **35** | 3 | 38 |
| | side | 43 | 5 | 48 |
| torso: | middle | 21 | 0 | 21 |
| | side | 21 | **9** | 30 |
| sum | | 120 | 17 | 137 |

| expected | | one | two |
|---|---|---|---|
| face: | middle | 30 | 4 |
| | side | 42 | 5 |
| torso: | middle | 18 | 2 |
| | side | 26 | 3 |

Notice that as in ASL, these two tendencies are additive, in that there are more one-handed signs at the middle of the face and more two-handed signs at the side of the torso than expected. For those locations where the demands of these two tendencies compete, that is, signs articulated at the side of the face and the middle of the torso, there is independence between the variables of location and hand number: I(loc; hand number)=0.114 and p=0.482. Yet, what is most notable about the NZSL data is not so much the face / torso distinction or the middle / side distinction, but the tiny numbers of two-handed signs. In section 4.3.1 it was shown that only 3% of NZSL signs are two-handed signs with a body location. Thus, the dependency on hand number induced by these body location distinctions, though existing, play only a small role in the lexicon.

#### 4.3.2.4 SVK: number of hands versus contact location

SVK was analyzed the same way as ASL, with somewhat similar results; the face / torso distinction is relevant, but the middle / side is not. To determine whether the dependence between location and hand number can be accounted for by the same location distinctions relevant in ASL, each SVK sign was labeled as one-handed or two-handed and as having a face location or a torso location. A chi-square analysis indicates with p=0.000 that the number of hands and the face / torso distinction are dependent variables. Next, the location of each sign was labeled as middle or side. A chi-square analysis indicates with p=0.262 that the number of hands and the middle / side distinction are independent variables.

(4-78) SVK: number of hands versus location

| observed | one | two | sum |
|---|---|---|---|
| face: middle | **78** | 8 | 86 |
| side | 36 | 5 | 41 |
| torso: middle | 31 | 17 | 48 |
| side | 9 | 1 | 10 |
| sum | 154 | 31 | |

| expected | one | two |
|---|---|---|
| face: middle | **72** | 14 |
| side | 34 | 7 |
| torso: middle | 40 | 8 |
| side | 8 | 2 |

When only face locations are included, there is no dependence between location and hand number; I(face; hand number)=0.176, p=0.186. When only torso locations are included, there is no dependence between location and hand number; I(torso; hand number)=0.116, p=0.060. Thus, it is the case that for SVK, the face / torso distinction influences whether a sign is one-handed or two-handed, but within these categories, there is not further relevant distinction.

## 4.4 Summary

The first part of this chapter has presented an overview of the location inventories of four sign languages. First just the body locations were examined, then all locations, including neutral space and nondominant hand locations. As with handshape, despite differences in size and content of the inventories, the manner in which the location resource is used is uniform cross-linguistically. When only shared locations were considered, it was seen that they are ordered very similarly across all four languages. Even when the entire location inventories were considered, the use of location resources is similar cross-linguistically. For every language, the exponential decay law, $y = a(2^{-cx})$, is a better fit than the power law, $y = ax^{-c}$, to the rank-frequency graph for body locations alone, while the reverse is true when all locations are included: the power law is a better fit than the exponential decay law. I propose that the following is a property of all natural sign languages: the body location rank-frequency distribution is modeled by an exponential decay equation, and the all location rank-frequency distribution is modeled by a power law.

The second part of this chapter investigated the question of whether the number of hands used in a sign is dependent upon the location at which the sign is articulated. To summarize, hand number and neutral space versus body location are dependent variables in all four languages, with more two-handed signs than expected occurring in neutral space. The face / torso distinction and the middle / side distinction that are operative diachronically and synchronically in ASL are less important or even non-existent in other languages. Caution is therefore necessary in applying even robust distinctions in ASL to other languages before their relevance in those languages is understood.

# Chapter 5: Duets: handshape and location pairs

A sign comprises both a handshape and a location, along with other phonological parameters, both manual and non-manual. The co-occurrence of a handshape and a location in a sign will be called a *duet*, to underscore the fact that only two parameters are considered and that these two parameters occur simultaneously. Duets also play a special role in the task of sign recognition. In the first step of the two step process, the handshape and location duet is recognized, and a cohort of lexical entries sharing this duet is activated; then the movement parameter is recognized, allowing the identification of the sign. (Grosjean, 1981; Corina and Emmorey, 1990; Emmorey, 2002).

Having determined the inventory and distribution of handshape and location in Chapters 3 and 4, this chapter first examines the inventory and distribution of duets in the lexicons of the four languages under study. Secondly, this chapter analyzes the composition of duets with respect to dependence between the variable of handshape and location. Is it the case that handshapes and locations freely co-occur? Or do certain combinations of handshape and location occur more frequently than expected in the lexicon?

## 5.1 Duet tables

### 5.1.1 ASL duets

The set of ASL duets of handshape and body location is shown in (5-1), and the set of duets of handshape and other locations, such as neutral space and the non-dominant hand is shown in (5-2). The data are presented in two arrays because a combined array featuring all handshape and location duets is too large to display in the format of this dissertation and is unwieldy to peruse. In all arrays of handshape by location, each variable is ordered from most to least frequent; location is ordered left to right, and handshape is ordered from top to bottom. In 405 signs, 36 handshapes and 36 locations are used, yielding 36 x 36 = 1296 possible duets, of which 192, or 15%, are actually attested. Thus, 85% of the cells in the array are empty. The question of whether these are accidental gaps or the result of structural restrictions is addressed in section 5.2.

(5-1) ASL: duets of handshape and body location

| | ψ | ~ | ξ | σ | σ | }□ | ~□ | ϖ | Ψ | ζ | { | \| | □ | □ | θ | ρ | υ | □ | □ | □□ | □ | ο | ϖ | ω | sum |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 49 | 3 | 1 | 3 | | 1 | | | 1 | 2 | 2 | | 1 | | 1 | | | | | | 1 | 1 | | 1 | | 18 |
| 3 | 2 | 3 | 1 | | 1 | 2 | | 1 | | | | | | 1 | 2 | | | | | | | | | | 13 |
| 78 | | | 2 | | 1 | 1 | | | | 1 | 1 | | 1 | | | | | | | | | 1 | | | 8 |
| 49B | | | 3 | 1 | | | | 1 | 1 | | | | 1 | | | | 1 | | | | | | | | 8 |
| 28 | 2 | 1 | | 1 | | | 1 | | | 1 | | | | | | | | | | | | | | | 6 |
| 3≅B | 1 | 1 | | | 1 | | | | | | 1 | | | 1 | | | 1 | | | | | | | | 6 |
| 39 | 1 | 1 | | 1 | | | | | | | 1 | | 1 | | 1 | | | | | | | | | | 6 |
| 78□ | | 3 | | 2 | | 1 | | | | | | | | | | | | | | | | | | | 6 |
| < | | | 1 | | | 1 | | 1 | | | | 1 | | | | | | | 1 | | | | | | 5 |
| 29 | | 1 | | | | | 1 | 1 | 1 | | | | | | | | | | | | | | | | 4 |
| 3A | | | | | | 1 | 1 | | | | | 1 | | | | | | 1 | | | | | | | 4 |
| 69 | 1 | | | 1 | 1 | | | | | | | 1 | | | | | | | | | | | | | 4 |
| 2 | | 2 | | | | | 1 | | | | | | 1 | | | | | | | | | | | | 4 |
| 7B8 | 1 | | 1 | | | | | | | | 1 | | | | | 1 | | | | | | | | | 4 |
| ;A | 2 | | 1 | | | | | | | | | | | | | | | | | | | | | | 3 |
| 48□ | | | 1 | | 1 | | | | | 1 | | | | | | | | | | | | | | | 3 |
| 59 | 1 | | | | | | | | 1 | | | | | | | | | | | | | | | | 2 |
| 69□ | | | | | | | | 1 | | | | | | | | | 1 | | | | | | | | 2 |
| 79 | 1 | | | 1 | | | | | | | | | | | | | | | | | | | | | 2 |
| 49□ | | | | | | | 1 | | | | | | | | | 1 | | | | | | | | | 2 |
| 68 | 1 | | | 1 | | | | | | | | | | | | | | | | | | | | | 2 |
| 69B | | | 1 | | | | | | | | | | | | | | | | | | | | | | 1 |
| 7≅B | | | | | | | | | | | | | | | | | | | | | | | | 1 | 1 |
| 4B8 | | | | | 1 | | | | | | | | | | | | | | | | | | | | 1 |
| 38 | | 1 | | | | | | | | | | | | | | | | | | | | | | | 1 |
| 48 | | | | | | | | | | | | | | | | 1 | | | | | | | | | 1 |
| 78B | | | | | | | 1 | | | | | | | | | | | | | | | | | | 1 |
| 3B≅ | | | | 1 | | | | | | | | | | | | | | | | | | | | | 1 |
| sum | 16 | 14 | 14 | 9 | 7 | 6 | 6 | 6 | 5 | 5 | 4 | 4 | 4 | 3 | 3 | 3 | 3 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 119 |

(5-2) ASL: duets of handshape and other location

| | π | 3 | \| | 29 | 49 | 2 | ; | 3≅B | 39 | 78 | 48 | sum |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 3 | 28 | 4 | 15 | 1 | 1 | | | | | | | 49 |
| 49 | 21 | 12 | 7 | | | | | | | 2 | 1 | 43 |
| 78 | 15 | 1 | | 1 | 1 | | | | | | | 18 |
| 29 | 11 | 1 | 5 | | | | | | | | | 17 |
| 28 | 8 | 1 | 4 | | | 3 | | | | | | 16 |
| 3A | 8 | 6 | | | | | | | | | | 14 |
| 69 | 4 | 4 | 1 | 1 | 2 | | | | 1 | | | 13 |
| < | 7 | 1 | 1 | | | | 1 | | 1 | | | 11 |
| 59 | 3 | 3 | 3 | 1 | | | 1 | | | | | 11 |
| 69B | 7 | 1 | 1 | 1 | | | | 1 | | | | 11 |
| 49B | 6 | 1 | 1 | | 1 | | | | | | | 9 |
| 2 | 5 | 1 | 1 | | 1 | | | | | | | 8 |
| ;A | 4 | 2 | 1 | | | | | | | | | 7 |
| 3≅B | 4 | 2 | 1 | | | | | | | | | 7 |
| 7≅B | 4 | 3 | | | | | | | | | | 7 |
| 48□ | 3 | 2 | | | 1 | | | | | | | 6 |
| 7B8 | 4 | 1 | | | | | | 1 | | | | 6 |
| 4B8 | 3 | 1 | | | | | | | | | | 4 |
| :□ | 1 | 1 | 1 | | | | | | | | | 3 |
| 6□□□ | 2 | 1 | | | | | | | | | | 3 |
| 69□□ | 2 | | | 1 | | | | | | | | 3 |
| 79 | 3 | | | | | | | | | | | 3 |
| ; | 2 | | | | | | | | | | | 2 |
| 38 | | | 2 | | | | | | | | | 2 |
| 39 | 1 | 1 | | | | | | | | | | 2 |
| 48 | 1 | 1 | | | | | | | | | | 2 |
| 49□ | 2 | | | | | | | | | | | 2 |
| 78□A | 1 | 1 | | | | | | | | | | 2 |
| ?□ | | | | 1 | | | | | | | | 1 |
| 3≅A | | 1 | | | | | | | | | | 1 |
| 3B | 1 | | | | | | | | | | | 1 |
| 68 | 1 | | | | | | | | | | | 1 |
| 78B | 1 | | | | | | | | | | | 1 |
| sum | 163 | 53 | 44 | 7 | 7 | 3 | 2 | 2 | 2 | 2 | 1 | 286 |

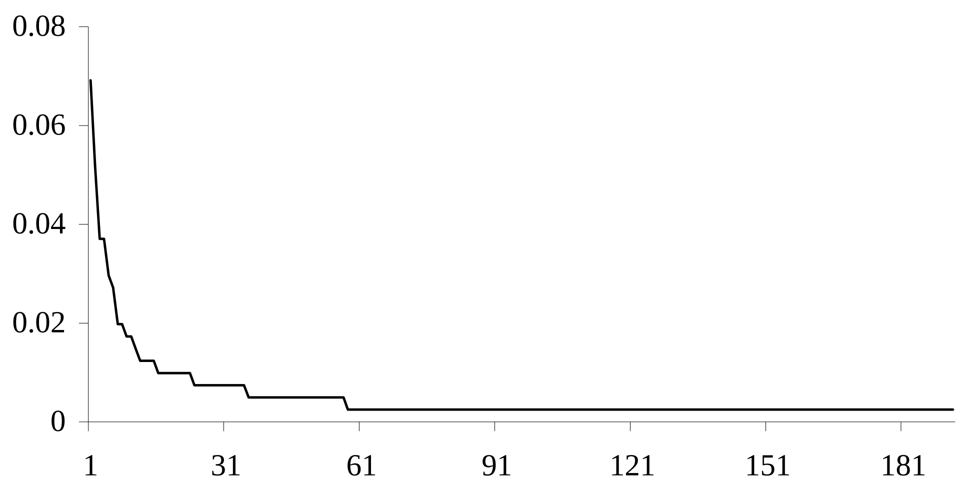
The five most common ASL duets and their percent of the ASL sample are listed in (5-3). These duets are used in 22.5% of the monomorphemic, non-handshape-changing, non-location changing signs in the ASL database.

(5-3) ASL: most common duets

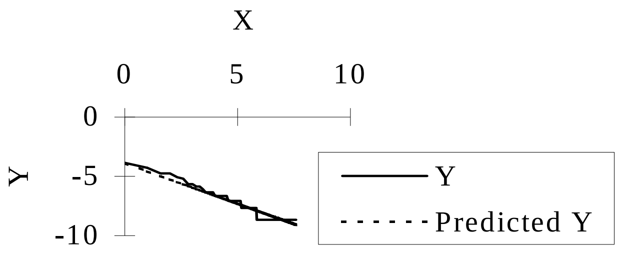
| rank | duet | count | percent | example |
|---|---|---|---|---|
| 1 | 3 π | 28 | 6.9 | THING |
| 2 | 49 π | 21 | 5.2 | WHERE |
| 3 | 78 π | 15 | 3.7 | FINISH |
| 4 | 3 \| | 15 | 3.7 | SCHOOL |
| 5 | 49 3 | 12 | 3.0 | HOUR |
| sum | | 91 | 22.5 | |

The duet rank-frequency graph for ASL is shown in (5-4). This graph is a hyperbolic curve, as shown by the linearity of the graph in (5-5). Linear regression applied to $X = \log_2(\text{rank})$ and $Y = \log_2(\text{frequency})$ results in an adjusted correlation coefficient R square value of 0.9053, indicating that the rank frequency of duets in ASL declines with close fidelity to the power law, $y = 0.067x^{-0.69}$.

(5-4) ASL: duet rank-frequency graph



(5-5) ASL: $X = \log_2 x$ and $Y = \log_2 y$ where x = rank and y = frequency




### 5.1.2 KSL duets

The set of KSL duets of handshape and body location is shown in (5-6), and the set of duets of handshape and other locations, such as neutral space and the non-dominant hand, is shown in (5-7). Note that in 384 signs, 41 handshapes and 32 locations are used, yielding 41 x 32 = 1312 possible duets, of which 189, or 14%, are actually attested. Thus, 86% of the cells in the array are empty.

(5-6) KSL: duets of handshape and body location

| | ψ | ξ | ~ | υ | ϖ | { | θ | ζ | \| | σ | υ□ | }□ | ο | □ | θ□ | ρ | σ□ | ϖ□ | ω | \|□ | \|φ | ~□ | □□□ | □□ | ο□ | θφ | ψ□ | sum |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 49 | 5 | 1 | | 1 | 3 | 1 | | | 2 | | 1 | | | | | | 1 | 1 | | | | | | 1 | | | | 17 |
| 3 | 1 | 1 | 4 | | 1 | | 3 | 1 | 1 | 1 | | 1 | | | 1 | | | | | | | | | | | | | 15 |
| 59 | 2 | 1 | | 1 | 2 | | | | 1 | | 1 | | | 1 | | | | 1 | | | | | | | | | | 10 |
| 28 | 2 | | 3 | | | | | 1 | | | | | 1 | | | 1 | | | | 1 | | | | | | | | 9 |
| 29 | 1 | | 1 | | 1 | 1 | | | | | 1 | | 1 | | | | | | | | 1 | | | | 1 | 1 | | 9 |
| < | 1 | | 1 | 2 | | | 1 | 1 | | | | | | | | | | | | | | | | | | | | 6 |
| 49B | | 1 | 1 | 1 | | 1 | | | | | | | | | | | | | 2 | | | | | | | | | 6 |
| 3A | | | | | | 2 | | | | | | | 1 | | | | 1 | | | | | | | | | | | 4 |
| 3B | 1 | 1 | | | | | | | | 1 | | | | 1 | | | | | | | | | | | | | | 4 |
| 3≅B | 1 | | | 1 | | | | | | | | | | | | | | | | | | 1 | | | | | | 3 |
| 48 | 1 | | 1 | | | | 1 | | | | | | | | | | | | | | | | | | | | | 3 |
| 48□ | | 1 | | | | | | 2 | | | | | | | | | | | | | | | | | | | | 3 |
| : | 1 | 1 | | | | | | | | | | | | | | | | | | | | | | | | | | 2 |
| :A | | 1 | | | | | 1 | | | | | | | | | | | | | | | | | | | | | 2 |
| ;A | | 1 | | | | | | | | 1 | | | | | | | | | | | | | | | | | | 2 |
| 3≅A | 1 | 1 | | | | | | | | | | | | | | | | | | | | | | | | | | 2 |
| 38 | | | | | 1 | 1 | | | | | | | | | | | | | | | | | | | | | | 2 |
| 4≅A | | 1 | | | | | | | | | | 1 | | | | | | | | | | | | | | | | 2 |
| 49□ | | | | | | | | 1 | | | 1 | | | | | | | | | | | | | | | | | 2 |
| 69 | | | | 1 | | | | | | | | | | | 1 | | | | | | | | | | | | | 2 |
| ; | | | | | | | | | | | | | | | | | | | | | | | | | | | 1 | 1 |
| 2 | | 1 | | | | | | | | | | | | | | | | | | | | | | | | | | 1 |
| 39 | | | | | | | | | | | | | | | | | | | | | | | 1 | | | | | 1 |
| 48□X□X | | | | 1 | | | | | | | | | | | | | | | | | | | | | | | | 1 |
| 49A | 1 | | | | | | | | | | | | | | | | | | | | | | | | | | | 1 |
| 58 | | | | | | | | | | 1 | | | | | | | | | | | | | | | | | | 1 |
| 69B | | | | | | | | | | | | 1 | | | | | | | | | | | | | | | | 1 |
| 7≅B | | | | | | | | | | | | | | | | 1 | | | | | | | | | | | | 1 |
| sum | 18 | 12 | 11 | 8 | 8 | 6 | 6 | 6 | 4 | 4 | 4 | 3 | 3 | 2 | 2 | 2 | 2 | 2 | 2 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 113 |

(5-7) KSL: duets of handshape and other location

| | π | 3 | \| | 29 | 28 | 3≅B | < | 38 | 3B | 2 | 49 | 78 | : | 49□ | sum |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 3 | 41 | 11 | 7 | | 1 | | | 1 | | | | | | 1 | 62 |
| 49 | 17 | 7 | | 5 | 3 | | 2 | 1 | | | 1 | 1 | 1 | | 38 |
| 29 | 16 | 3 | 2 | 1 | 1 | 2 | | | | | | | | | 25 |
| 28 | 7 | 5 | 1 | | | 1 | | | 2 | | | | | | 16 |
| < | 6 | 6 | 1 | | | | | | | | | 1 | | | 14 |
| 59 | 1 | 6 | 5 | 1 | | | | | | 1 | | | | | 14 |
| 3B | 8 | | 1 | 2 | | | | 1 | | | | | | | 12 |
| : | 4 | 1 | 2 | 1 | 1 | | 1 | | | | | | | | 10 |
| :A | 4 | | 1 | 1 | | | | | | | | | | | 6 |
| 2 | 3 | 2 | | | 1 | | | | | | | | | | 6 |
| 7≅B | 3 | 2 | | | 1 | | | | | | | | | | 6 |
| ; | 3 | 1 | 1 | | | | | | | | | | | | 5 |
| 49B | 3 | 2 | | | | | | | | | | | | | 5 |
| 78 | 4 | | 1 | | | | | | | | | | | | 5 |
| ;A | 3 | | | | | | | | | | 1 | | | | 4 |
| 3A | 2 | 1 | | | | 1 | | | | | | | | | 4 |
| 4 | 2 | | | 1 | | | | | 1 | | | | | | 4 |
| 48□ | 2 | 2 | | | | | | | | | | | | | 4 |
| 3≅B | 2 | 1 | | | | | | | | | | | | | 3 |
| 4≅B | 2 | 1 | | | | | | | | | | | | | 3 |
| 59B | 2 | 1 | | | | | | | | | | | | | 3 |
| 38 | 1 | 1 | | | | | | | | | | | | | 2 |
| 39 | 2 | | | | | | | | | | | | | | 2 |
| 4B≅ | 2 | | | | | | | | | | | | | | 2 |
| 69 | 1 | | | | 1 | | | | | | | | | | 2 |
| <□ | 1 | | | | | | | | | | | | | | 1 |
| 38B | 1 | | | | | | | | | | | | | | 1 |
| 4≅ | 1 | | | | | | | | | | | | | | 1 |
| 4≅A | | 1 | | | | | | | | | | | | | 1 |
| 48 | 1 | | | | | | | | | | | | | | 1 |
| 49□ | 1 | | | | | | | | | | | | | | 1 |
| 49A | | | | | | 1 | | | | | | | | | 1 |
| 49A□ | 1 | | | | | | | | | | | | | | 1 |
| 4X | | 1 | | | | | | | | | | | | | 1 |
| 5A9 | | 1 | | | | | | | | | | | | | 1 |
| 6≅A | 1 | | | | | | | | | | | | | | 1 |
| 68 | 1 | | | | | | | | | | | | | | 1 |
| 69B | | 1 | | | | | | | | | | | | | 1 |
| 79 | | | | | | | | | | 1 | | | | | 1 |
| sum | 149 | 57 | 22 | 12 | 9 | 5 | 3 | 3 | 3 | 2 | 2 | 2 | 1 | 1 | 271 |

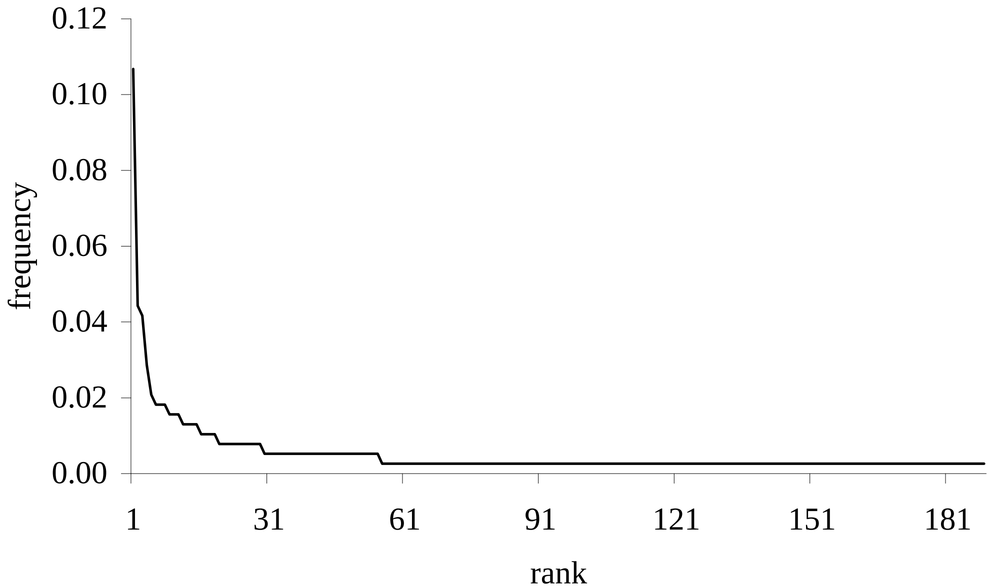
The five most frequent KSL duets and their percent of the KSL sample are listed in (5-8). There are three duets tied for fifth place. These duets are used in 24% of the monomorphemic, non-handshape-changing, non-location-changing signs in the KSL database.

(5-8) KSL: most common duets

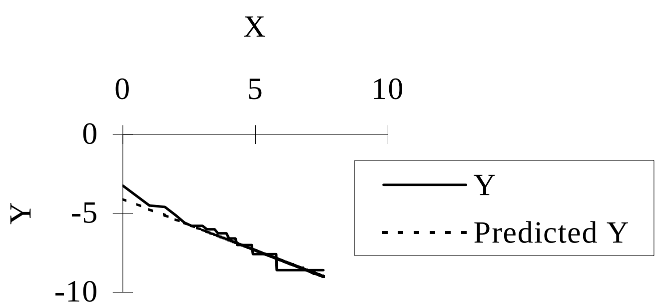
| rank | duet | count | percent | example |
|---|---|---|---|---|
| 1 | 3 π | 41 | 10.7 | UNJAE *when* |
| 2 | 49 π | 17 | 4.4 | GAKGAK *each* |
| 3 | 29 π | 16 | 4.2 | CHUNGSO *cleaning* |
| 4 | 3 3 | 11 | 2.9 | NANOODA *divide* |
| 5 | 3B π | 8 | 1.8 | CHAMGA *participation* |
| sum | | 92 | 24.0 | |

The duet rank-frequency graph for KSL is shown in (5-9). This graph is a hyperbolic curve, as shown by the linearity of the graph in (5-10). Linear regression applied to $X = \log_2(\text{rank})$ and $Y = \log_2(\text{frequency})$ results in an adjusted correlation coefficient R square value of 0.8988, indicating that the rank frequency of duets in ASL declines with reasonable fidelity to the power law, $y = 0.057x^{-0.65}$.

(5-9) KSL: duet rank-frequency graph




(5-10) KSL: $X = \log_2 x$ and $Y = \log_2 y$ where x = rank and y = frequency




### 5.1.3 NZSL duets

The set of NZSL duets of handshape and body location is shown in (5-11), and the set of duets of handshape and other locations, such as neutral space and the non-dominant hand is shown in (5-12). Note that in 461 signs, 50 handshapes and 45 handshapes are used, yielding 50 x 45 = 2250 possible duets, of which 188, or 8.4%, are actually attested. Thus, 91.6% of the cells in the array are empty.

(5-11) NZSL: duets of handshape and body location

| | ζ | ~ | }□ | σ□ | ψ | ω | ξ | \|□ | ϖ | \| | ~□ | □ | { | □ | ο□ | ϖ□ | □ | ο | θ | θ□ | ρ | υ□ | ψ□ | \|φ | ~□□ | □ | □ | ο□□ | σ | ζ□ | sum |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 3 | 1 | 1 | 1 | 1 | | 1 | 1 | | 1 | | | | 1 | 1 | | | | | | 2 | | | 1 | 1 | | | | 1 | | | 14 |
| 49 | 4 | | 1 | 2 | | | | 1 | 1 | | | | | | | 1 | | | | | | | | | | | 1 | | | | 11 |
| 78 | | 2 | 2 | | | 1 | | | | 1 | 2 | | | | 1 | 1 | | | | | | | | | | | | | | | 10 |
| 49B | 2 | | | 3 | | | | | | | | | | | | 1 | 1 | | | | | 2 | | | | | | | | | 9 |
| 39 | 1 | 2 | | | 1 | | | | | | | | | | 1 | | | | 2 | | | | | | | | | | | | 7 |
| 2 | 1 | 1 | | 1 | | 1 | 1 | | | | 1 | | | | | | | | | | | | | | | | | | | | 6 |
| 28 | | 1 | | | | | 1 | 1 | | 1 | | | 1 | | | | | | | | | | | | | | | | 1 | | 6 |
| 7≅B | | 1 | | | 1 | 1 | | | | | | 1 | | | | | | | | | 2 | | | | | | | | | | 6 |
| 3B | 1 | 1 | | 1 | | | | | | | | | 1 | | | | | | | | | | | | 1 | | | | | | 5 |
| 4 | 2 | | | | | 1 | | | | | | | | 1 | | | | | | | | | | | | | | | | | 4 |
| 4X | 1 | | 1 | | | | 1 | 1 | | | | | | | | | | | | | | | | | | | | | | | 4 |
| ; | | | 1 | | 1 | | | | 1 | | | | | | | | | | | | | | | | | | | | | | 3 |
| < | | | | | 1 | | 1 | | | | | 1 | | | | | | | | | | | | | | | | | | | 3 |
| 29 | | | | | 2 | | | | | | | | | | | | | 1 | | | | | | | | | | | | | 3 |
| 4B8 | | | 1 | | | | 1 | | | | | | | | | | | | | | | | | | | | | | | 1 | 3 |
| 69 | | | | 1 | 1 | | | | | 1 | | | | | | | | | | | | | | | | | | | | | 3 |
| 38 | | | | | | | | | | 1 | | | | 1 | | | | | | | | | | | | | | | | | 2 |
| 49A | | | | | | 1 | | | | | | | | | | | | | | | | | 1 | | | | | | | | 2 |
| 69B | | | | | | | | 1 | 1 | | | | | | | | | | | | | | | | | | | | | | 2 |
| 7B | | | | | | | | | | | | | | | | | 1 | 1 | | | | | | | | | | | | | 2 |
| <B | | | | | | | | | 1 | | | | | | | | | | | | | | | | | | | | | | 1 |
| 38B | | | | | | | | | | | 1 | | | | | | | | | | | | | | | | | | | | 1 |
| 3A | | | 1 | | | | | | | | | | | | | | | | | | | | | | | | | | | | 1 |
| 4≅B | | | | | | | | | | | | | | | 1 | | | | | | | | | | | | | | | | 1 |
| 48 | | | | | | | | 1 | | | | | | | | | | | | | | | | | | | | | | | 1 |
| 4A | | | | | | | | | | | | 1 | | | | | | | | | | | | | | | | | | | 1 |
| 4B9 | | | 1 | | | | | | | | | | | | | | | | | | | | | | | | | | | | 1 |
| 6□□□ | | 1 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | 1 |
| 79 | | | | | | | | | | | | 1 | | | | | | | | | | | | | | | | | | | 1 |
| 7B8 | | | | | | | | | | | | | | | | | | | | | | | | | | 1 | | | | | 1 |
| sum | 13 | 10 | 9 | 9 | 7 | 6 | 6 | 5 | 5 | 4 | 4 | 4 | 3 | 3 | 3 | 3 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 115 |

(5-12) NZSL: duets of handshape and other locations

| | π | 3 | \| | 38 | 49 | 3≅B | :□ | ; | 2 | 29 | 3B | 3B8 | 69 | 69□□□ | 78 | sum |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 3 | 34 | 5 | 5 | | | | | | | | | | | | | 44 |
| 49 | 31 | 2 | 3 | 1 | 3 | | | 1 | | | | 1 | | | | 42 |
| 78 | 27 | | | 1 | | 1 | | | | | | | | | | 29 |
| 2 | 21 | | | 1 | 1 | | | | | | | | | | | 23 |
| 38 | 12 | | 2 | 4 | | | | | | | | | | 1 | | 19 |
| 29 | 13 | 2 | 1 | | | | | | | | | | | | | 16 |
| 4X | 11 | 1 | 1 | | | | 1 | | | | | | | | | 14 |
| 28 | 11 | 1 | | | | | | | 1 | | | | | | | 13 |
| 3A | 8 | 3 | 1 | | | | | | | | | | | | | 12 |
| 7≅B | 10 | | | 1 | | 1 | | | | | | | | | | 12 |
| 3≅B | 10 | | | | | | | | | | | | | | | 10 |
| < | 9 | | | | | | | | | | | | | | | 9 |
| 4 | 7 | 1 | | | | | | | | | | | 1 | | | 9 |
| 69 | 6 | 1 | | 1 | | | | | | | | | | | | 8 |
| : | 3 | 2 | | | 1 | | | | | | 1 | | | | | 7 |
| 4≅B | 7 | | | | | | | | | | | | | | | 7 |
| <B | 4 | | 2 | | | | | | | | | | | | | 6 |
| 49□ | 4 | 1 | | 1 | | | | | | | | | | | | 6 |
| 7B | 5 | | | | | | | | | | | | | | | 5 |
| :A | 3 | | | 1 | | | | | | | | | | | | 4 |
| ; | 3 | | | | | | | | | 1 | | | | | | 4 |
| 39 | 3 | | 1 | | | | | | | | | | | | | 4 |
| 49A | 4 | | | | | | | | | | | | | | | 4 |
| 59 | 3 | 1 | | | | | | | | | | | | | | 4 |
| 3B | 2 | | 1 | | | | | | | | | | | | | 3 |
| 48□ | 3 | | | | | | | | | | | | | | | 3 |
| 78□A | 2 | | | 1 | | | | | | | | | | | | 3 |
| 3≅ | 2 | | | | | | | | | | | | | | | 2 |
| 3A8 | 1 | | 1 | | | | | | | | | | | | | 2 |
| 4≅A | 2 | | | | | | | | | | | | | | | 2 |
| 48 | 2 | | | | | | | | | | | | | | | 2 |
| 4B8 | 2 | | | | | | | | | | | | | | | 2 |
| 69□□ | 2 | | | | | | | | | | | | | | | 2 |
| 69B | | 1 | 1 | | | | | | | | | | | | | 2 |
| ;B | 1 | | | | | | | | | | | | | | | 1 |
| ? | | | | | | | | | | | | | | | 1 | 1 |
| 3≅A | 1 | | | | | | | | | | | | | | | 1 |
| 3X | 1 | | | | | | | | | | | | | | | 1 |
| 3X8 | 1 | | | | | | | | | | | | | | | 1 |
| 4□ | 1 | | | | | | | | | | | | | | | 1 |
| 49□ | 1 | | | | | | | | | | | | | | | 1 |
| 5≅A | | 1 | | | | | | | | | | | | | | 1 |
| 6□□□ | 1 | | | | | | | | | | | | | | | 1 |
| 79 | 1 | | | | | | | | | | | | | | | 1 |
| 79□X | 1 | | | | | | | | | | | | | | | 1 |
| 7B8 | 1 | | | | | | | | | | | | | | | 1 |
| sum | 277 | 22 | 19 | 12 | 5 | 2 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 346 |

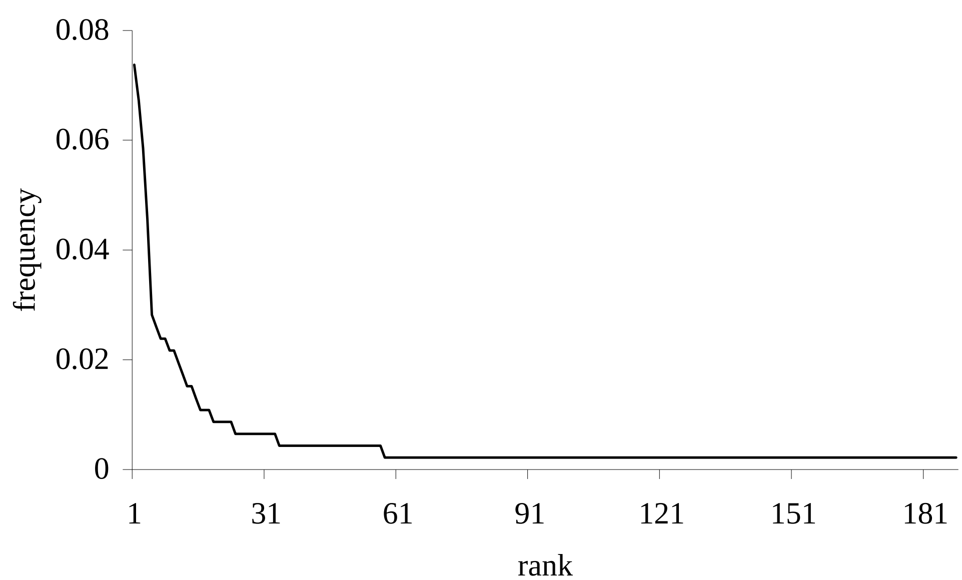
The five most common NZSL duets and their percent of the NZSL sample are listed in (5-13). These duets are used in 27.3% of the monomorphemic, non-handshape-changing, non-location-changing signs in the NZSL database.

(5-13) NZSL: most common duets

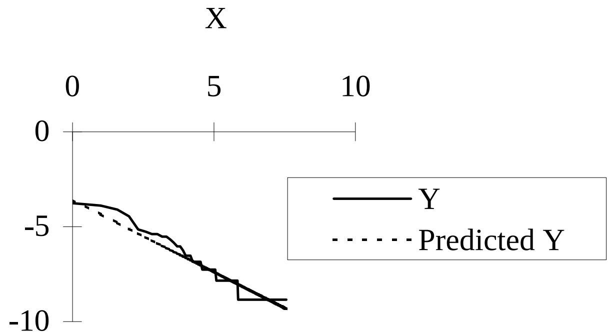
| rank | duet | count | percent | example |
|---|---|---|---|---|
| 1 | 3 π | 34 | 7.4 | READ |
| 2 | 49 π | 31 | 6.7 | WHO |
| 3 | 78 π | 27 | 5.9 | NONSENSE |
| 4 | 2 π | 21 | 4.6 | USE |
| 5 | 29 π | 13 | 2.8 | FIRM |
| sum | | 126 | 27.3 | |

The duet rank-frequency graph for NZSL is shown in (5-14). This graph is a hyperbolic curve, as shown by the linearity of the graph in (5-15). Linear regression applied to $X = \log_2(\text{rank})$ and $Y = \log_2(\text{frequency})$ results in an adjusted correlation coefficient R square value of 0.9033, indicating that the rank frequency of duets in ASL declines with reasonable fidelity to the power law, $y = 0.083x^{-0.76}$.

(5-14) NZSL: duet rank-frequency graph



(5-15) NZSL: $X = \log_2 x$ and $Y = \log_2 y$ where x = rank and y = frequency




### 5.1.4 SVK duets

The set of SVK duets of handshape and body location is shown in (5-16), and the set of duets of handshape and other locations, such as neutral space and the non-dominant hand is shown in (5-17). Note that in 482 signs, 34 handshapes and 25 locations are used, yielding 34 x 32 = 1088 possible duets, of which 189, or 17%, are actually attested. Thus, 82.6% of the cells in the array are empty.

(5-16) SVK: duets of handshape and body location

| | ~ | ψ | σ | ξ | ζ | } | ο | ϖ | θ | υ | { | }□ | σ□ | □ | ω | \| | ~□ | □ | □ | □ | \|□ | }□φ | ο□ | υ□ | ζ□ | sum |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 3 | 3 | 3 | | 5 | | 2 | 2 | | 4 | 2 | | | 2 | | | | 1 | 2 | 2 | | | | 1 | | | 29 |
| 49 | 1 | 5 | 1 | 2 | 3 | 1 | 1 | 1 | 1 | 2 | 1 | | 3 | | 1 | | | | | | | | | | | 23 |
| 78 | 6 | 1 | 3 | | | 2 | 1 | 1 | | | 1 | | | | | | | | | | | | | | | 15 |
| 28 | 4 | | | | 3 | | | | | | | 1 | | | | | | | | | | | | | | 8 |
| 7≅B | 1 | 1 | 1 | 1 | | 1 | 1 | 1 | | | | | | | | | | | | | | | | | | 7 |
| 39 | 2 | | 2 | | | | | | | | 1 | | | 1 | | | | | | | | | | | | 6 |
| 3A | | 1 | 1 | 1 | | | 1 | | | | | | | | | | | | | 1 | 1 | | | | | 6 |
| 3B | | 2 | 1 | | | | 1 | | | | | | | 1 | | | | | | | | 1 | | | | 6 |
| 49B | | 2 | 1 | | | | | | | 1 | 1 | | | | 1 | | | | | | | | | | | 6 |
| 69 | | | 1 | 2 | | 1 | | 2 | | | | | | | | | | | | | | | | | | 6 |
| 2 | 3 | | | | | | | | | | | | | 1 | | | | | | | | | | | 1 | 5 |
| 59 | | 1 | 1 | | | | | 1 | | | | 1 | | | | | | | | 1 | | | | | | 5 |
| 68 | | | 1 | | 1 | 1 | 1 | | | | | | | | | | 1 | | | | | | | | | 5 |
| 69B | | 2 | 1 | 1 | | | | | 1 | | | | | | | | | | | | | | | | | 5 |
| ;A | 1 | 1 | 2 | | | | | | | | | | | | | | | | | | | | | | | 4 |
| 48□ | | | 1 | | 1 | | | 1 | | | | | | | | 1 | | | | | | | | | | 4 |
| 4X | | 1 | 1 | | | | | | | | | 1 | | | | | | | | | | | | 1 | | 4 |
| 29 | 1 | | | | 1 | | | | | | | 1 | | | | | | | | | | | | | | 3 |
| 4B8□ | | | | 2 | | | 1 | | | | | | | | | | | | | | | | | | | 3 |
| 79 | | | | 1 | | | | 1 | | | 1 | | | | | | | | | | | | | | | 3 |
| : | | | | 1 | | 1 | | | | | | | | | | | | | | | | | | | | 2 |
| 3≅B | | | | | 2 | | | | | | | | | | | | | | | | | | | | | 2 |
| 49A | | 1 | | | | | | | | | | | | | 1 | | | | | | | | | | | 2 |
| 4B≅□ | | 1 | | | | | | | | | | | | | | 1 | | | | | | | | | | 2 |
| 7B | | | | 1 | | | | | | | | 1 | | | | | | | | | | | | | | 2 |
| 4≅B | | | | | | | | | | 1 | | | | | | | | | | | | | | | | 1 |
| 4A | | | | | 1 | | | | | | | | | | | | | | | | | | | | | 1 |
| 7≅ | 1 | | | | | | | | | | | | | | | | | | | | | | | | | 1 |
| sum | 23 | 22 | 18 | 17 | 12 | 9 | 9 | 8 | 6 | 6 | 5 | 5 | 5 | 3 | 3 | 2 | 2 | 2 | 2 | 2 | 1 | 1 | 1 | 1 | 1 | 166 |

(5-17) SVK: duets of handshape and other locations

| | π | \| | 3 | 49 | 29 | 3≅ | 38 | sum |
|---|---|---|---|---|---|---|---|---|
| 3 | 36 | 35 | | | | | | 71 |
| 49 | 21 | 9 | 8 | | | | 1 | 39 |
| 78 | 20 | 4 | 4 | | | | | 28 |
| 29 | 13 | 4 | 2 | | | | | 19 |
| 69 | 9 | 3 | 6 | | | | | 18 |
| 3B | 11 | 3 | 2 | | | | | 16 |
| 7≅B | 11 | | 3 | | | | | 14 |
| 28 | 9 | | 2 | | | | | 11 |
| 4X | 7 | 2 | 1 | 1 | | | | 11 |
| 2 | 6 | 2 | 1 | | | | | 9 |
| 69B | 5 | 2 | 2 | | | | | 9 |
| 59 | 3 | 3 | 2 | | | | | 8 |
| < | 4 | 1 | 1 | 1 | | | | 7 |
| 49B | 3 | 2 | 1 | | 1 | | | 7 |
| ;A | 4 | | 1 | | | | | 5 |
| 39 | 3 | | 2 | | | | | 5 |
| 3A | | 1 | 4 | | | | | 5 |
| : | 1 | 1 | 2 | | | | | 4 |
| 4B8□ | 2 | 2 | | | | | | 4 |
| 3≅B | 2 | | 1 | | | | | 3 |
| 49□ | 2 | | 1 | | | | | 3 |
| 7B | 3 | | | | | | | 3 |
| ; | | | | 1 | | 1 | | 2 |
| 4≅B | 1 | 1 | | | | | | 2 |
| 48□ | 2 | | | | | | | 2 |
| 49A | 1 | | 1 | | | | | 2 |
| 4B≅□ | 1 | | 1 | | | | | 2 |
| 7≅ | 1 | | 1 | | | | | 2 |
| 3≅ | 1 | | | | | | | 1 |
| 3≅A | | | 1 | | | | | 1 |
| 48 | 1 | | | | | | | 1 |
| 68 | 1 | | | | | | | 1 |
| 79 | 1 | | | | | | | 1 |
| sum | 185 | 75 | 50 | 3 | 1 | 1 | 1 | 316 |

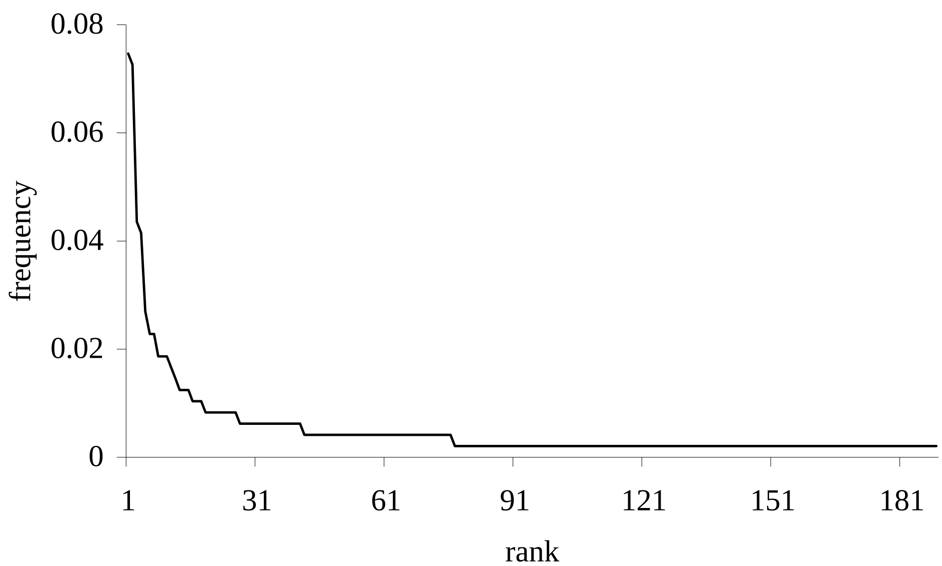
The five most common SVK duets and their percent of the SVK sample are listed in (5-18). These duets are used in 26.1% of the monomorphemic, non-handshape-changing, non-location-changing signs in the SVK database.

(5-18) SVK: most common duets

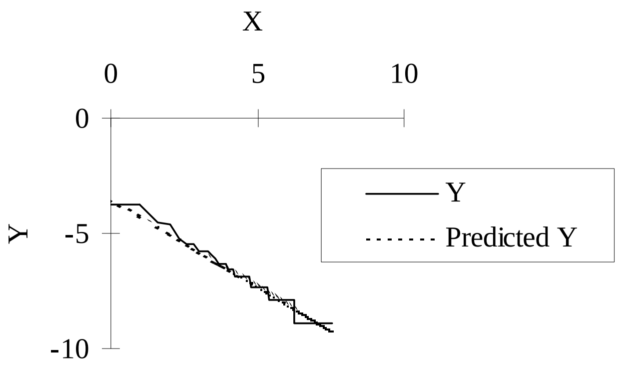
| rank | duet | count | percent | example |
|---|---|---|---|---|
| 1 | 3 π | 36 | 7.5 | TAIVAS *sky* |
| 2 | 3 \| | 35 | 7.3 | SIISTI *clean* |
| 3 | 49 π | 21 | 4.4 | KÄYDÄ *visit* |
| 4 | 78 π | 20 | 4.2 | VIITTOA *sign* |
| 5 | 29 π | 13 | 2.7 | MYRSKY *storm* |
| sum | | 126 | 26.1 | |

The duet rank-frequency graph for SVK is shown in (5-19). This graph is a hyperbolic curve, as shown by the linearity of the graph in (5-20). Linear regression applied to $X = \log_2(\text{rank})$ and $Y = \log_2(\text{frequency})$ results in an adjusted correlation coefficient R square value of 0.9450, indicating that the rank frequency of duets in ASL declines with close fidelity to the power law, $y = 0.089x^{-0.77}$.

(5-19) SVK: duet rank-frequency graph



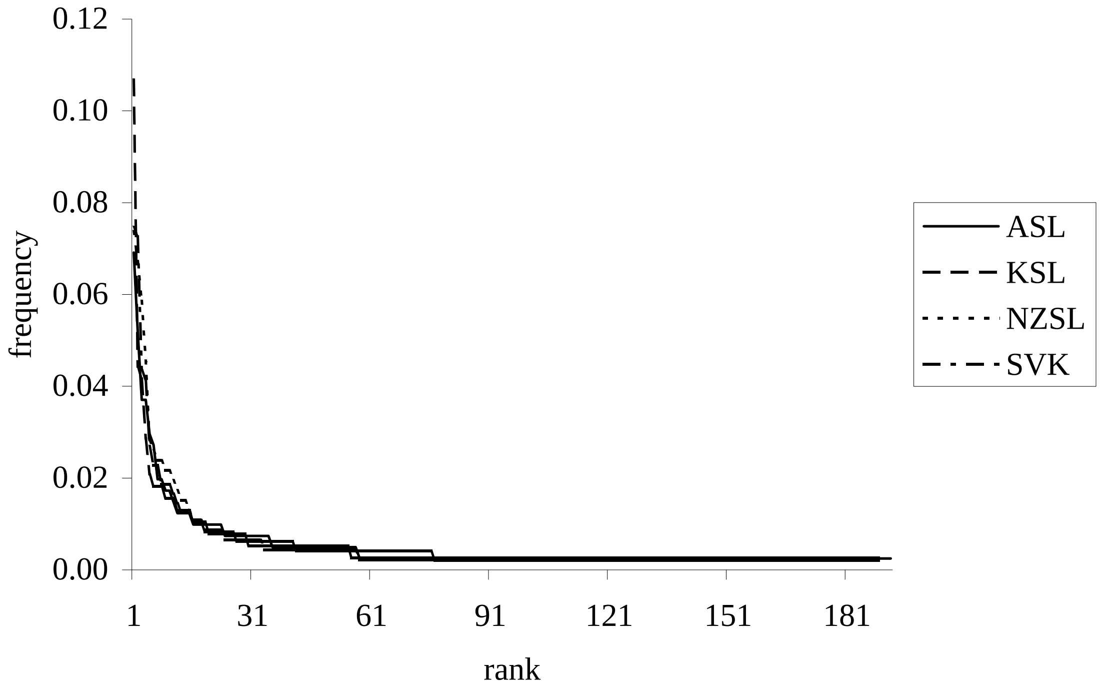
(5-20) SVK: X = $\log_2 x$ and Y = $\log_2 y$ where x = rank and y = frequency




### 5.1.5 Comparison

The next question to consider is the comparison of the duet distributions. The rank-frequency graphs of duets in ASL, KSL, NZSL, and SVK have been shown to be hyperbolic curves. How similar are these curves to each other? Is it the case that although all are the same type of curve, they are nevertheless significantly different instantiations of $y = ax^{-c}$, or is the variation attributable to sampling error? The rank-frequency graph for body location of all four languages is shown in (5-21).

(5-21) All languages: rank-frequency graph for duets




Examining two languages at a time, the frequencies associated with each rank for duets were paired, and the signed differences between these frequencies were tested to ascertain whether they can be considered a random sample from a population with mean $\mu = 0$ by using the paired-sample *t* test. The results of this pair-wise comparison show at the 0.001 level of significance that the differences between each pair of duet distributions is not significant. Thus, all the curves are very similar to each other.

### 5.1.6 Universal duet pool

Pooling the data from ASL, KSL, SVK and NZSL produces the set of handshape and body locations shown in (5-22), and the set of duets of handshape and other locations, such as neutral space and the non-dominant hand shown in (5-23). Note that in 1732 signs, 68 handshapes and 60 locations are used, yielding 68 x 60 = 4080 possible duets, of which 504, or 29%, are actually attested. Thus, 71% of the cells in the array are empty. The pooled duet data is a step toward identifying the set of potentially available duets and their distributions.

(5-22) Pooled data: duets of handshape and body location

| | ψ | ~ | ξ | ζ | σ | ϖ | }□ | σ□ | { | θ | υ | ο | \| | ~□ | ω | □ | } | ψ□ | \|□ | ρ | υ□ | □ | ϖ□ | □ | □ | ο□ | □ | θ□ | \|φ | □□□ | ζ□ | □ | }□φ | ~□□ | 3 | □□ | ο□□ | θφ | sum |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 3 | 6 | 11 | 8 | 2 | 1 | 3 | 4 | 4 | 1 | 9 | 2 | 2 | 1 | 1 | 1 | 2 | 2 | 1 | | | | 2 | | | 2 | 1 | | 3 | 1 | | | | | | | | 1 | | 71 |
| 49 | 13 | 2 | 6 | 9 | 1 | 6 | 1 | 7 | 2 | 1 | 3 | 1 | 3 | | 1 | 1 | 1 | 2 | 1 | | 1 | | 3 | | | | 2 | | | 1 | | | | | | 1 | | | 69 |
| 78 | 1 | 8 | 2 | 1 | 3 | 1 | 3 | 1 | 2 | | | 2 | 1 | 2 | 1 | | 2 | | | | | | 1 | 1 | | 1 | | | | | | | | | | | | | 33 |
| 28 | 4 | 9 | 1 | 5 | 2 | | 1 | | 1 | | | 1 | 1 | 1 | | | | | 2 | 1 | | | | | | | | | | | | | | | 1 | | | | 30 |
| 49B | 2 | 1 | 4 | 2 | 2 | 1 | | 3 | 2 | | 3 | | | | 3 | | | 1 | | | 2 | | 1 | 1 | 1 | | | | | | | | | | | | | | 29 |
| 39 | 2 | 5 | | 1 | 3 | | | | 2 | 3 | | | | | | 1 | | | | | | | | 1 | | 1 | | | | 1 | | | | | | | | | 20 |
| 29 | 3 | 3 | | 1 | | 2 | 1 | | 1 | | | 2 | | 1 | | | | 1 | | | 1 | | | | | 1 | | | 1 | | | | | | | | | 1 | 19 |
| 59 | 4 | | 1 | | 1 | 3 | 1 | | | | 1 | | 1 | | | 1 | | 1 | | | 1 | | 1 | | | | 1 | | | | | | | | | | | | 17 |
| 2 | | 6 | 2 | 1 | | | | 1 | | | | | | 2 | 1 | 1 | | | | | | | | 1 | | | | | | | 1 | | | | | | | | 16 |
| 3A | 1 | | 1 | | 1 | | 2 | 1 | 2 | | | 2 | 1 | 1 | | | | | 1 | | | | | | | | 1 | | | | | 1 | | | | | | | 15 |
| 3B | 3 | 1 | 1 | 1 | 2 | | | 1 | 1 | | | 1 | | | | 2 | | | | | | | | | | | | | | | | | 1 | 1 | | | | | 15 |
| 69 | 2 | | 2 | | 2 | 2 | | 2 | | | 1 | | 2 | | | | 1 | | | | | | | | | | | 1 | | | | | | | | | | | 15 |
| 7≅B | 2 | 2 | 1 | | 1 | 1 | | | | | | 1 | | | 2 | | 1 | | | 3 | | 1 | | | | | | | | | | | | | | | | | 15 |
| < | 2 | 1 | 2 | 1 | | 1 | 1 | | | 1 | 2 | | 1 | | | | | | | | | 1 | | | 1 | | | | | | | | | | | | | | 14 |
| 3≅B | 2 | 1 | | 2 | | | | 1 | 1 | | 2 | | | 1 | | 1 | | | | | | | | | | | | | | | | | | | | | | | 11 |
| 48□ | | | 2 | 4 | 1 | 1 | | 1 | | | | | 1 | | | | | | | | | | | | | | | | | | | | | | | | | | 10 |
| ;A | 3 | 1 | 2 | | 3 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | 9 |
| 69B | 2 | | 2 | | 1 | 1 | 1 | | | 1 | | | | | | | | | 1 | | | | | | | | | | | | | | | | | | | | 9 |
| 4X | 1 | | 1 | 1 | 1 | | 2 | | | | | | | | | | | | 1 | | 1 | | | | | | | | | | | | | | | | | | 8 |
| 68 | 1 | | | 1 | 2 | | | | | | | 1 | | 1 | | | 1 | | | | | | | | | | | | | | | | | | | | | | 7 |
| 78□A | | 3 | | | 2 | | 1 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | 6 |
| 79 | 1 | | 1 | | 1 | 1 | | | 1 | | | | | | | | | | | | | 1 | | | | | | | | | | | | | | | | | 6 |
| 38 | | 1 | | | | 1 | | | 1 | | | | 1 | | | 1 | | | | | | | | | | | | | | | | | | | | | | | 5 |
| 48 | 1 | 1 | | | | | | | | 1 | | | | | | | | | 1 | 1 | | | | | | | | | | | | | | | | | | | 5 |
| 49A | 2 | | | | | | | | | | | | | | 2 | | | 1 | | | | | | | | | | | | | | | | | | | | | 5 |
| 7B8 | 1 | | 1 | | | | | | 1 | | | | | | | | | | | 1 | | | | 1 | | | | | | | | | | | | | | | 5 |
| : | 1 | | 2 | | | | | | | | | | | | | | 1 | | | | | | | | | | | | | | | | | | | | | | 4 |
| ; | 1 | | | | | 1 | 1 | | | | | | | | | | | 1 | | | | | | | | | | | | | | | | | | | | | 4 |

| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 4 | | | | 2 | | | | | | | | | | | 1 | 1 | | | | | | | | | | | | | | | | | | | | | | 4 |

(5-22) continued

| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 49□ | | | | 1 | | | | | | | | | 1 | | | | | | 1 | 1 | | | | | | | | | | | | | | | | | | 4 |
| 4B8 | | | 1 | | | | 1 | 1 | | | | | | | | | | | | | | | | | | | | | | 1 | | | | | | | | 4 |
| 7B | | | 1 | | | | 1 | | | | | 1 | | | | | | | | | | | | 1 | | | | | | | | | | | | | | 4 |
| 4B8□ | | | 2 | | | | | | | | | 1 | | | | | | | | | | | | | | | | | | | | | | | | | | 3 |
| :A | | | 1 | | | | | | | 1 | | | | | | | | | | | | | | | | | | | | | | | | | | | | 2 |
| 3≅A | 1 | | 1 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | 2 |
| 4≅A | | | 1 | | | | 1 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | 2 |
| 4≅B | | | | | | | | | | | 1 | | | | | | | | | | | | | | | 1 | | | | | | | | | | | | 2 |
| 4A | | | | 1 | | | | | | | | | | | | | | | | | | 1 | | | | | | | | | | | | | | | | 2 |
| 4B≅□ | 1 | | | | | | | | | | | | 1 | | | | | | | | | | | | | | | | | | | | | | | | | 2 |
| 69□□ | | | | | | 1 | | | | | 1 | | | | | | | | | | | | | | | | | | | | | | | | | | | 2 |
| <B | | | | | | 1 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | 1 |
| 38B | | | | | | | | | | | | | | 1 | | | | | | | | | | | | | | | | | | | | | | | | 1 |
| 3B≅ | | | | | 1 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | 1 |
| 48□X□X | | | | | | | | | | | 1 | | | | | | | | | | | | | | | | | | | | | | | | | | | 1 |
| 4B9 | | | | | | | 1 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | 1 |
| 58 | | | | | 1 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | 1 |
| 6□□□ | | 1 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | 1 |
| 7≅ | | 1 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | 1 |
| 78B | | | | | | | | | | | | | | 1 | | | | | | | | | | | | | | | | | | | | | | | | 1 |
| sum | 63 | 58 | 49 | 36 | 32 | 27 | 23 | 23 | 18 | 17 | 17 | 15 | 14 | 13 | 12 | 11 | 9 | 8 | 7 | 7 | 7 | 6 | 6 | 5 | 5 | 5 | 4 | 4 | 2 | 2 | 2 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 514 |

(5-23) Pooled data: duets of handshape and other locations

| | π | 3 | | 29 | 49 | 38 | 28 | 3≅B | 3 | 2 | 78 | 3B | ; | < | 39 | 48 | : | :□ | 3≅ | 3B8 | 49□ | 69 | 69□□□ | sum |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 3 | 139 | 16 | 62 | 1 | 1 | 1 | 1 | | 4 | | | | | | | | | | | | 1 | | | 226 |
| 49 | 90 | 29 | 19 | 5 | 4 | 3 | 3 | | | | 3 | | 1 | 2 | | 1 | 1 | | | 1 | | | | 162 |
| 78 | 66 | 5 | 5 | 1 | 1 | 1 | | 1 | | | | | | | | | | | | | | | | 80 |
| 29 | 53 | 8 | 12 | 1 | | | 1 | 2 | | | | | | | | | | | | | | | | 77 |

| | | | | | | | | | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 28 | 35 | 8 | 5 | | | | | 1 | | 4 | | 2 | | | | | | | | | | | 55 |
| 2 | 35 | 4 | 3 | | 2 | 1 | 1 | | | | | | | | | | | | | | | | 46 |
| < | 26 | 8 | 3 | | 1 | | | | | | 1 | | 1 | | 1 | | | | | | | | 41 |
| 69 | 20 | 11 | 4 | 1 | 2 | 1 | 1 | | | | | | | | 1 | | | | | | | | 41 |
| 7≅B | 28 | 8 | | | | 1 | 1 | 1 | | | | | | | | | | | | | | | 39 |
| 59 | 10 | 12 | 11 | 2 | | | | | | 1 | | | 1 | | | | | | | | | | 37 |
| 3A | 18 | 13 | 2 | | | | | 1 | 1 | | | | | | | | | | | | | | 35 |
| 3B | 22 | 2 | 5 | 2 | | 1 | | | | | | | | | | | | | | | | | 32 |
| 4X | 18 | 3 | 3 | | 1 | | | | | | | | | | | | | 1 | | | | | 26 |
| 3≅B | 18 | 4 | 1 | | | | | | | | | | | | | | | | | | | | 23 |
| 38 | 13 | 1 | 4 | | | 4 | | | | | | | | | | | | | | | | 1 | 23 |
| 69B | 12 | 5 | 4 | 1 | | | | 1 | | | | | | | | | | | | | | | 23 |
| : | 8 | 5 | 3 | 1 | 1 | | 1 | | | | | 1 | | 1 | | | | | | | | | 21 |
| 49B | 12 | 4 | 3 | 1 | 1 | | | | | | | | | | | | | | | | | | 21 |
| ;A | 11 | 3 | 1 | | 1 | | | | | | | | | | | | | | | | | | 16 |
| 48□ | 10 | 4 | | | 1 | | | | | | | | | | | | | | | | | | 15 |
| ; | 8 | 1 | 1 | 1 | 1 | | | | | | | | | | | | | | 1 | | | | 13 |
| 39 | 9 | 2 | 1 | | | | | | 1 | | | | | | | | | | | | | | 13 |
| 4 | 9 | 1 | | 1 | | | | | | | | 1 | | | | | | | | | 1 | | 13 |
| 4≅B | 10 | 1 | 1 | | | | | | | | | | | | | | | | | | | | 12 |
| 49□ | 9 | 2 | | | | 1 | | | | | | | | | | | | | | | | | 12 |
| :A | 7 | | 1 | 1 | | 1 | | | | | | | | | | | | | | | | | 10 |
| 7B | 8 | | | | | | | | | | | | | | | | | | | | | | 8 |
| 49A | 5 | 1 | | | | | | 1 | | | | | | | | | | | | | | | 7 |
| 7B8 | 5 | 1 | | | | | | 1 | | | | | | | | | | | | | | | 7 |
| <B | 4 | | 2 | | | | | | | | | | | | | | | | | | | | 6 |
| 48 | 5 | 1 | | | | | | | | | | | | | | | | | | | | | 6 |
| 4B8 | 5 | 1 | | | | | | | | | | | | | | | | | | | | | 6 |

(5-23) continued

| | | | | | | | | | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 79 | 5 | | | | | | | | | 1 | | | | | | | | | | | | | 6 |
| 69□□ | 4 | | | 1 | | | | | | | | | | | | | | | | | | | 5 |
| 78□A | 3 | 1 | | | | 1 | | | | | | | | | | | | | | | | | 5 |

| | | | | | | | | | | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 4B8□ | 2 | | 2 | | | | | | | | | | | | | | | | | | | | | 4 |
| 6□□□ | 3 | 1 | | | | | | | | | | | | | | | | | | | | | | 4 |
| :□ | 1 | 1 | 1 | | | | | | | | | | | | | | | | | | | | | 3 |
| 3≅ | 3 | | | | | | | | | | | | | | | | | | | | | | | 3 |
| 3≅A | 1 | 2 | | | | | | | | | | | | | | | | | | | | | | 3 |
| 4≅A | 2 | 1 | | | | | | | | | | | | | | | | | | | | | | 3 |
| 59B | 2 | 1 | | | | | | | | | | | | | | | | | | | | | | 3 |
| 68 | 3 | | | | | | | | | | | | | | | | | | | | | | | 3 |
| 3A8 | 1 | | 1 | | | | | | | | | | | | | | | | | | | | | 2 |
| 4B≅ | 2 | | | | | | | | | | | | | | | | | | | | | | | 2 |
| 4B≅□ | 1 | 1 | | | | | | | | | | | | | | | | | | | | | | 2 |
| 7≅ | 1 | 1 | | | | | | | | | | | | | | | | | | | | | | 2 |
| ;B | 1 | | | | | | | | | | | | | | | | | | | | | | | 1 |
| ? | | | | | | | | | | | 1 | | | | | | | | | | | | | 1 |
| ?□ | | | | 1 | | | | | | | | | | | | | | | | | | | | 1 |
| <□ | 1 | | | | | | | | | | | | | | | | | | | | | | | 1 |
| 38B | 1 | | | | | | | | | | | | | | | | | | | | | | | 1 |
| 3X | 1 | | | | | | | | | | | | | | | | | | | | | | | 1 |
| 3X8 | 1 | | | | | | | | | | | | | | | | | | | | | | | 1 |
| 4□ | 1 | | | | | | | | | | | | | | | | | | | | | | | 1 |
| 4≅ | 1 | | | | | | | | | | | | | | | | | | | | | | | 1 |
| 49A□ | 1 | | | | | | | | | | | | | | | | | | | | | | | 1 |
| 49□ | 1 | | | | | | | | | | | | | | | | | | | | | | | 1 |
| 5≅A | | 1 | | | | | | | | | | | | | | | | | | | | | | 1 |
| 5A9 | | 1 | | | | | | | | | | | | | | | | | | | | | | 1 |
| 6≅A | 1 | | | | | | | | | | | | | | | | | | | | | | | 1 |
| 78B | 1 | | | | | | | | | | | | | | | | | | | | | | | 1 |
| 79□X | 1 | | | | | | | | | | | | | | | | | | | | | | | 1 |
| sum | 774 | 175 | 160 | 21 | 17 | 16 | 9 | 9 | 6 | 6 | 5 | 4 | 3 | 3 | 2 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1218 |

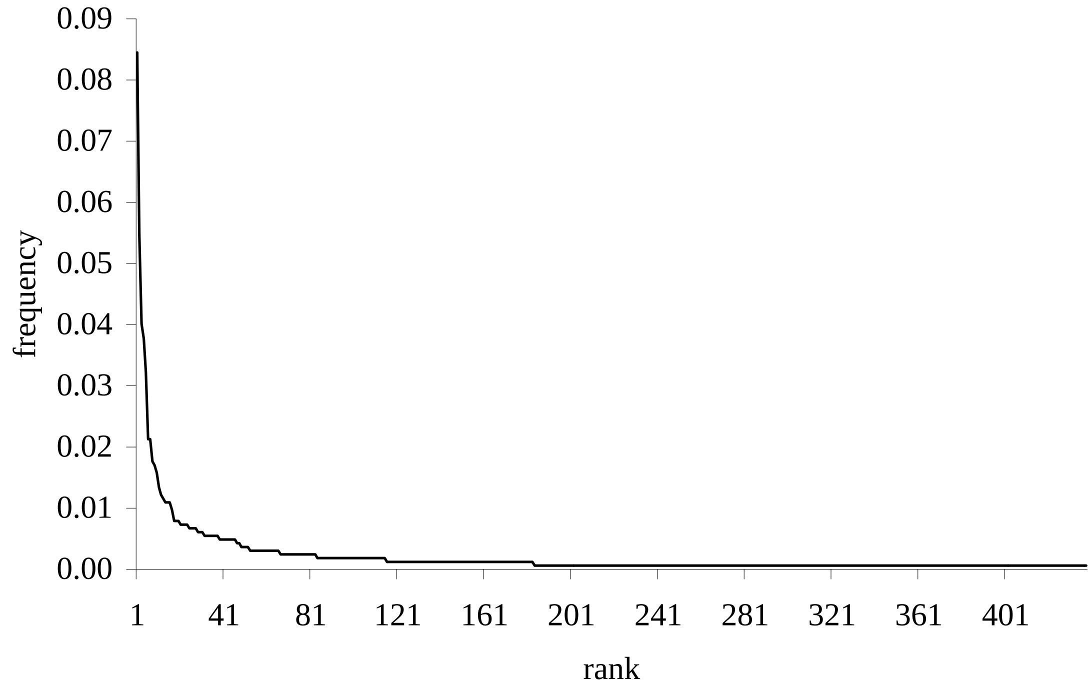
The five most common duets in the pooled data and their percents are listed in (5-24). These duets are used in 25.9% of the monomorphemic, non-handshape-changing, non-location-changing signs in the pooled database.

(5-24) Pooled data: most common duets

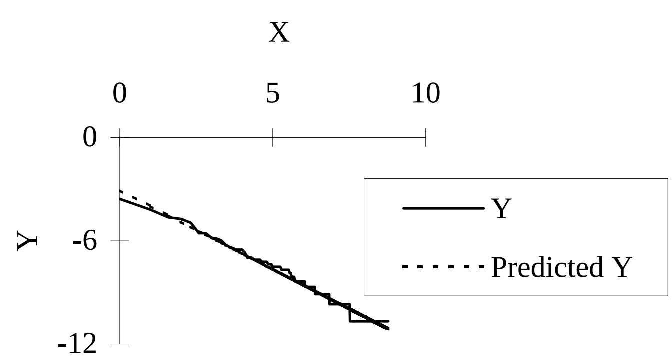
| rank | duet | count | percent |
|---|---|---|---|
| 1 | 3 π | 139 | 8.4 |
| 2 | 49 π | 90 | 5.5 |
| 3 | 78 π | 66 | 4.0 |
| 4 | 3 \| | 62 | 3.8 |
| 5 | 29 π | 53 | 3.2 |
| sum | | 410 | 24.9 |

The duet rank-frequency graph for this pooled data, a sample of 1645 signs from all four languages, is given in (5-25). Like the other duet graphs, this is also a hyperbolic curve, as shown by the linearity of the graph in (5-26). Linear regression applied to X = $\log_2$(rank) and Y = $\log_2$(frequency) results in an adjusted correlation coefficient R square value of 0.9553, indicating that the rank frequency of duets in the pooled data declines with close fidelity to the power law, $y = 0.12x^{-0.92}$.

(5-25) Pooled data: duet rank-frequency graph




(5-26) Pooled data: Y= log $_2$ y where x = rank and y = frequency




## 5.2 Handshape distributions across locations

This section investigates how handshapes are distributed across locations, or equivalently, how locations are distributed across handshapes. We know from chapters 3 and 4 that handshapes and locations themselves are not distributed uniformly in the lexicon, but given this knowledge of handshape and location frequency distribution, is the distribution of duets in each language as expected, or do certain locations attract certain handshapes and repel others? For example, one well-known statement about duet distribution is Siple (1973; 1978; 1980), which predicts that signs with finer articulatory distinctions will be made on the face and neck, while signs with grosser distinctions will be made outside this area. This statement has been interpreted to as “signs on the face and neck use more marked handshapes than signs on the torso.” Is this interpretation of the prediction true of the languages in this sample? Are there other groupings of handshapes and groupings of locations that show affinities for each other? A part of this question has already been looked at in section 3.3 which investigated the distribution of handshape across sign types, in that Type 2 and Type 3 signs have a special type of location, the hand. As in section 3.3, the two handshape markedness groupings based on the universal, frequency criterion and on the language-specific, phonological criterion are used (see section 1.5.3). Locations are grouped by the contact/neutral space distinction and by the hand / arm / face / torso distinction.

### 5.2.1 ASL: handshape by location

The first test was for general dependence between the handshape and location variables, without any grouping. Ordered pairs consisted of a handshape and a location. The mutual information for this data set was I(hs; loc)=0.999, with p=0.054.[80] Since this value is on the border-line of significance, I investigated further. The first distinction tested was the neutral space / contact location distinction. Is it the case that different distributions of handshapes are used in neutral space than with contact locations? The location of each sign was labeled as “contact” or as “neutral space.” The data set to be tested was the set of ordered pairs, each consisting of a handshape and either “contact” or “neutral space”. For example, the ordered pairs for the ASL signs MOTHER and YES are (78, contact) and (29, neutral space), respectively. There was independence between the handshape variable and the contact / neutral space variable; p=0.813. Thus, knowing that a sign is articulated in neutral space conveys no information about what the handshape might be.

Handshape versus contact locations (that is, excluding signs made in neutral space) was tested next. Signs articulated in neutral space were excluded from the data set, which consisted of

---

80 This test was run five times. Out of the five sets of 1,000 scrambled data sets, 44, 53, 56, 59 and 60 sets had higher mutual information values (average is 57), suggesting that dependence between handshape and location in ASL is weak.

ordered pairs of the form (handshape, contact location); for example, the ASL sign MOTHER was represented as (78,ζ). The mutual information significance program tested this data set and found it to be dependent, with p=0.017. Thus, handshapes are not used as expected across all possible contact locations.

To ascertain the source of this dependence, contact locations were classified as occurring on the arm, hand, face or torso. Four data sets were created, excluding each of these body areas in turn, and the dependence of these sets was tested by the mutual information significance program. The results are shown in (5-27). The only set for which handshape and location are independent is the set that includes arm, face and torso locations. When two areas at a time are tested, the three data sets that include the hand are dependent (arm and hand: p=0.033; face and hand: p=0.013; torso and hand: p=0.024). The three data sets that exclude the hand are independent (arm and face: p=0.696; face and torso: p=0.398; arm and torso: p=0.579).

(5-27) ASL: handshape by contact location

| areas included | p value | Are hs and loc dependent variables? |
|---|---|---|
| arm, face, hand | 0.015 | dependent |
| **arm, face, torso** | **0.516** | **independent** |
| arm, hand, torso | 0.023 | dependent |
| face, hand, torso | 0.008 | dependent |

Recall from section 3.3 that in ASL, handshape was distributed differently in Type 2 signs than in signs of other types, in that a larger than expected number of unmarked handshapes were used in Type 2 signs. When Type 2 signs, that is, signs with location |, are excluded from the data set, but all other signs are included, handshape and location are independent, with p=0.309. In fact, the set that includes all signs except those with handshape 3 and location | has no dependence between the handshape and location variables, with p=0.097. Thus, the weak dependence between the handshape and location variable is accounted for by the larger than expected number of unmarked handshapes with the location |, especially those with 3 as the handshape.

### 5.2.2 KSL: handshape by location

The first test was for general dependence between the handshape and location variables, without any grouping. Ordered pairs consisted of a handshape and a location. The mutual information for this data set was I(hs; loc)=1.175, with p=0.066.[81] Since this value is not very much above the minimum for attaining significance, I investigated further. The first test was on the contact / neutral space distinction. Each sign was labeled as "contact" or as "neutral space", and the set of ordered pairs consist of a handshape and either "contact" or "neutral space". There was dependence between the handshape variable and the contact / neutral space variable; p=0.010. This result implies that handshapes are used with different distributions in neutral space than in contact locations.

In addition, handshape versus contact locations (that is, excluding signs made in neutral space) was tested and found to be independent, with p=0.093. As in ASL, contact locations were classified as occurring on the arm, hand, face or torso. Four data sets were created, excluding each of these body areas in turn, and the dependence of these sets was tested by the mutual information significance program. Not surprisingly, the results show independence for these groupings ( arm, face, hand: p=0.159; arm, face, torso: p=0.088; arm, hand, torso: p=0.219; face, hand, torso: p=0.080).

Thus, the weak dependence between handshape and location seems to depend on the use of handshape across the contact / neutral space distinction. What is different about the handshapes used in signs with contact locations versus those used in signs in neutral space? To determine whether markedness in involved, the unmarked and marked sets defined by the universal frequency criterion and the language-specific phonological criterion were tested against the contact / neutral space distinction. As in section 3.3, neither markedness criterion produced dependency; p=0.073 for the universal unmarked set and p=0.914 for the language-specific unmarked set, as shown in (5-28). However, the tendency, though not statistically significant, is for more marked handshapes to be used with contact locations and more unmarked handshapes to be used in neutral space.

---

81 This test was run five times. Out of the five sets of 1,000 scrambled data sets, 58, 62, 63, 73, and 76 sets had higher mutual information values (average is 66), suggesting that dependence between handshape and location in KSL is weak, as it was in ASL.

(5-28) KSL: marked/unmarked handshape versus contact/neutral space location

**universal unmarked set**

| observed | contact | π | sum |
|---|---|---|---|
| Marked | 118 | 60 | 178 |
| Unmarked | 117 | 87 | 204 |
| sum | 237 | 143 | 382 |

| expected | contact | π |
|---|---|---|
| Marked | 110 | 68 |
| Unmarked | 126 | 79 |

p=0.073 independent

**language-specific unmarked set**

| observed | contact | π | sum |
|---|---|---|---|
| marked | 58 | 37 | 95 |
| unmarked | 177 | 110 | 287 |
| sum | 235 | 147 | 382 |

| expected | contact | π |
|---|---|---|
| marked | 58 | 37 |
| unmarked | 177 | 110 |

p=0.914 independent

An examination of the mutual information calculation for contact / neutral space versus handshape shows that one handshape, 59, contributes far more to the mutual information value, I(hs; contact / π)=0.133, than any other handshape. There are 23 signs using 59 that contact a location, but there is only one sign using 59 in neutral space; the expected values are 15 and 9. Of the signs that have 59 for the handshape and use contact locations, 10 are Type 0 signs (that is, one-handed), 4 are Type 2, and 9 are Type 3. MAJOON *meeting* is the only sign using 59 articulated in neutral space, and it is the only Type 1, that is, two-handed, sign with this handshape. When this handshape is excluded, the mutual information between the remaining handshapes and the binary contact / neutral space variable is I(hs\59; contact / π)=0.106, with p=0.282, indicating that the two variables are now independent. Thus, the weak dependence between handshape and location appears to be due to the exceptional distribution of 59.

### 5.2.3 NZSL: handshape by location

NZSL was difficult to analyze; while handshape and location are dependent, I was unable to find any natural groupings of either handshape or location that could account for this dependency. The first test was for general dependence between the handshape and location variables, without any grouping. Ordered pairs consisted of a handshape and a location. The mutual information for this data set was I(hs; loc)=1.059, with p=0.013, which indicates that handshape and location are not randomly distributed with respect to each other.

To determine the source of this dependence, the contact / neutral space distinction was first tested. Each sign was labeled as “contact” or as “neutral space”, and the set of ordered pairs consisted of a handshape and either “contact” or “neutral space”. There was dependence between the handshape variable and the contact / neutral space variable; p=0.015. However, when handshapes are labeled as “marked” or “unmarked” and locations as “contact” or “neutral space”, a chi-square analysis indicates that these two variables are independent for both markedness criteria, with p=0.090 for the universal criterion and p=0.134 for the language-specific criterion. Though these results do not attain dependence, for either markedness criteria, the tendency is for more marked handshapes to occur with contact locations and more unmarked handshapes to occur in neutral space, as in KSL. This is the expected result, since section 3.3.3 showed that more marked handshapes occur in one-handed signs, and section 4.3.2.3 showed that more one-handed signs have contact locations.

Then handshapes were tested against contact locations alone, and these two variables were found to be dependent, also, with p=0.004. Other location groupings were tested, but no pattern emerged. When tested against handshape, the location distinctions side / middle, hand / non-hand and face / non-face all tested as independent (p=0.648, p=0.191, p=0.077). Handshape was also tested against the variable with the four values “arm”, “face”, “hand” and “torso”, and was found to be dependent (p=0.007). However, when groups of locations were excluded and the remaining two or three groups were tested, there was no pattern to the results. Some location subsets are independent relative to handshape and some are dependent.

Next, handshapes were labeled as “marked” and “unmarked” according to the universal and language-specific criteria. The binary handshape variable was tested against all locations and against just contact locations, but all tests indicate independence between the handshape and location variables, as shown in (5-29).

(5-29) NZSL: marked/unmarked handshape versus all locations and contact locations

| | universal criteria | language-specific criteria |
|---|---|---|
| all locations | p=0.920 | p=0.477 |
| contact locations | p=0.919 | p=0.485 |

Finally, the mutual information calculation was examined. One variable value contributes generously to the total mutual information, I(hs; loc) = 1.059. The handshape 49B contributed 0.08. This handshape is used on the face eight times and on the arm one time. It is never used in neutral space, on the nondominant hand, or on the torso. To test the impact of this unusual distribution on the mutual information, signs with 49B as the handshape were excluded. Using this subset of signs, handshape versus location was tested and found to be independent, with p=0.132. Then, the locations in this subset were labeled as “contact” or “neutral space”; this binary location variable was tested against handshape, and found to be independent, with p=0.230. Finally, only signs with contact locations and without handshape 49B were considered, and handshape versus location was tested and found to be independent, with p=0.050. No other handshape or location has such a great impact on the mutual information value. For example, 3 as a nondominant hand location is a big contributor to mutual information; however, removal of signs with 3 as nondominant hand location does not produce a set in which handshape and location are independent. I suggest that the unusual distribution of 49B across locations is at least partly responsible for the dependency between handshape and location.

### 5.2.4 SVK: handshape by location

The first test was for general dependence between the handshape and location variables, without any grouping. Ordered pairs consisted of a handshape and a location. The mutual information for this data set was I(hs; loc)=0.812, with p=0.001, which indicates that handshape and location are not randomly distributed with respect to each other. To determine the source of this dependence, the contact / neutral space distinction was first tested. Each sign was labeled as “contact” or as “neutral space”, and the set of ordered pairs consisted of a handshape and either “contact” or “neutral space”. There was independence between the handshape variable and the contact / neutral space variable; p=0.457. Thus, the dependency is between handshape and contact locations, that is, excluding signs made in neutral space.

Handshape versus contact location was tested and found to be dependent, with p=0.002. Contact locations were classified as occurring on the arm, hand, face or torso. Four data sets were created, excluding each of these body areas in turn, and the dependence of these sets was tested by the mutual information significance program. The results show dependence for all groupings of these four location areas (arm, face, hand: p=0.009; arm, face, torso: p=0.010; arm, hand, torso: p=0.011; face, hand, torso: p=0.003). Next, two data sets were created by labeling handshapes "marked" or "unmarked" by both the universal and the language-specific markedness criteria. Both sets were tested against contact locations, and both were found to be dependent (universal: p=0.000; language-specific: p=0.000). Next, contact locations were divided into two groups, face and non-face (that is, hand, arm and torso), and handshapes were divided by markedness. Chi-square tests using both markedness criteria indicate dependence for both markedness criteria, as shown in (5-30). More marked handshapes than expected appear on the face for both markedness criteria. In SVK, the face, therefore, is notable as an attractor of unmarked handshapes, as well as of one-handed signs, as was seen in section 4.3.2.4.

(5-30) SVK: marked/unmarked handshape versus face /non-face contact locations

**universal unmarked set**

| observed | face | non-face | sum |
|---|---|---|---|
| Marked | **62** | 80 | 142 |
| Unmarked | 46 | **108** | 154 |
| sum | 108 | 188 | 296 |

| expected | face | non-face |
|---|---|---|
| Marked | 52 | 90 |
| Unmarked | 56 | 98 |

p=0.014 dependent

**language-specific unmarked set**

| observed | face | non-face | sum |
|---|---|---|---|
| marked | **65** | 87 | 152 |
| unmarked | 43 | **101** | 144 |
| sum | 108 | 188 | 296 |

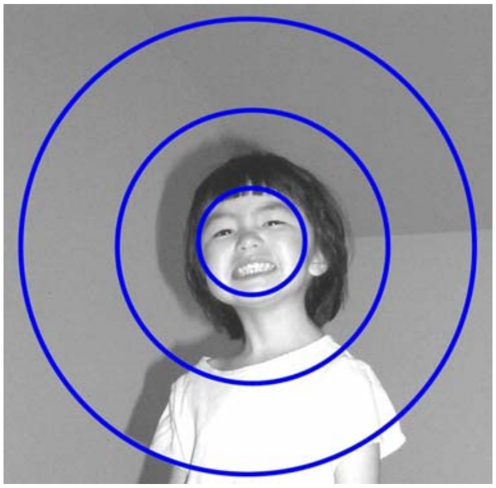
| expected | face | non-face |
|---|---|---|
| marked | 55 | 97 |
| unmarked | 53 | 91 |

p=0.021 dependent

### 5.2.5 Visual constraints on sign language

Siple (1973; 1978; 1980) discusses constraints imposed upon sign language by the visual modality. The human visual field is not uniformly accurate. Accuracy is greatest in the area around the focal point, in which a great deal of detail can be perceived, because the image of this area falls on the fovea, the most sensitive part of the retina. Outside of this area, accuracy diminishes, and less detail can be distinguished. Siple claims that since the sign language perceiver focuses on the signer's face, not the hands, the focal point is the center of the face. The zone of greater acuity is the inner circle centered on the face, and the zone of lesser acuity is the outermost circle, as shown in (5-31).

(5-31) Zones of acuity: visual acuity diminishes as distance from center increases



Siple states, "We would expect to find pairs of signs made in the areas about the face or upper chest to be visually more similar, i.e., to differ in a less detailed way, than signs made in the areas of lower acuity." For example, she suggests that a minimal pair in which the handshapes differ only in the number of fingers extended would occur on the face but not on the torso. The ASL minimal pair RED and SWEET, both articulated on the lips with initial handshapes 49 and 59, is such an example. Siple claims, "The set of signs made in the regions of lower acuity should not contain such pairs."[82] She also predicts that one-handed signs occur more frequently at locations on the head and neck, while two-handed signs occur more frequently outside this area, because the redundancy of the handshape signal when carried on two hands aids perception in this area of lesser visual acuity. Siple comments, "A look at the Dictionary of ASL (Stokoe et al., 1965) will show that these predictions are confirmed," but unfortunately she does not actually verify these interesting predictions with a quantitative analysis.

Indeed, a quantitative analysis of this prediction is a major undertaking. First, the database in use must be extensive enough to be able to verify the claim that certain types of signs

---

[82] But consider the SVK minimal pair NIMI *name* and OSOITE *address*, both Type 3 signs with the nondominant 3 hand as location. NIMI has handshape 49 and OSOITE has handshape 59. These signs are semantically related, a point which will be referred to in section 6.2.4.

do not exist. Second, a metric of phonological similarity must be developed in order to be able to quantify how similar two signs are. Third, the phonological similarity of every pair of signs in each of the two acuity zones must be calculated. Finally, the average phonological similarity over all pairs from each zone must be compared to ascertain if signs with face and neck locations are actually more alike.

Battison (1995:195) analyzes a simpler version of Siple's prediction. Locations in neutral space and on the hand and arm are excluded; the remaining locations are divided into two groups: head and neck locations and trunk locations. Handshapes are divided into marked and unmarked sets. As discussed in section 2.1.2.2.3, Battison identifies {3 2 29 3≅B ; 49 78} as the set of unmarked handshapes. Based on a count of signs from DASL, he finds that unmarked handshapes predominate on the torso and marked handshapes predominate on the face and neck, with p=0.043. I attempted unsuccessfully to reproduce this result for the four languages in this study. I preformed three tests using the universal unmarked set, the language-specific unmarked set, and the unmarked handshape set identified by Battison.[83] The results are shown in (5-32); the p values that indicate a dependence between handshape and location are in boldface type.

[83] Because my transcription system is different from that used in DASL, there were a number of handshapes whose markedness was uncertain. In addition to BASCO15 = {3 2 29 3≅B ; 49 78}, I included {39 38 28}as unmarked handshapes.

(5-32) All languages: marked/unmarked handshape by head and neck/torso

| | universal | language-specific | Battison | number of tokens |
|---|---|---|---|---|
| ASL | p=0.128 | p=0.466 | p=0.213 | 107 |
| KSL | p=0.138 | p=0.946* | p=0.065 | 111 |
| NZSL | p=0.913* | p=0.808* | p=0.762 | 109 |
| SVK | **p=0.018** | p=0.115 | **p=0.023** | 158 |

Dependence between the two variables was observed only in SVK and only for the unmarked sets defined by the universal criterion and by Battison's criterion. In most of the cases, the tendency was for there to be more than expected marked handshapes on the face and neck, although this effect was not significant. An asterisk indicates the particular cases in which there were fewer than expected marked handshapes on the face and neck.[84]

Recall from section 4.3.2 that in ASL, NZSL, and SVK whether a sign was one-handed or two-handed was dependent on the face / torso distinction, with more one-handed signs having face locations and more two-handed signs having torso locations. This result is consistent with Siple's prediction. However, in the analysis in section 4.3.2, the neck was considered part of the torso rather than being grouped with face locations. New chi-square analyses in which the neck is regrouped with the head do not change the results. It is still the case that in ASL, NZSL and SVK, there is dependence between the hand number and location variables, with p=0.002 for

[84] Note that the computation in Battison (1995) included 606 tokens, as compared to the 107 to 158 tokens. Perhaps a larger sample size would yield different results. These tests were repeated with all locations, including contact locations on the arm and hand as well as neutral space locations. The location grouping was face and neck versus all other locations. The three sets of marked / unmarked handshapes were the same. The results are shown below. I do not know whether the differences are due to the additional locations that were included or to the difference in sample size.

| | universal | language-specific | Battison | number of tokens |
|---|---|---|---|---|
| ASL | **p=0.047** | p=0.217 | p=0.212 | 394 |
| KSL | **p=0.018** | p=0.690 | **p=0.022** | 382 |
| NZSL | p=0.758 | p=0.466 | p=0.573 | 460 |
| SVK | **p=0.007** | **p=0.014** | **p=0.043** | 482 |

ASL, p=0.005 for NZSL, and p=0.001 for SVK. The hand number and location variables are still independent for KSL, with p=0.106.

## 5.3 Summary

The first part of this chapter has presented an overview of the duet inventories of four sign languages. Differing handshape and location inventories and distributions produce different duet inventories and distributions. Yet, the manner in which duet resources are used is uniform cross-linguistically. Roughly, each language uses a few duets very frequently, about five duets in about 25% of the signs, and it uses a lot of duets very rarely. More precisely, the rank-frequency graphs of duets follow the power law, $y = ax^{-c}$. I propose that the following is a property of all natural sign languages: the duet rank-frequency distribution is modeled by a power law.

The second part of this chapter has presented an analysis of the distribution of handshape relative to location. In contrast to the first part of the chapter, which found robust cross-linguistic similarities, language-specific differences emerge. In ASL and KSL, handshape and location influence each other only slightly. Thus, in these two languages, knowing the location of a sign conveys little information about what the handshape might be. What little dependence that exists between the variables of handshape and location can be accounted for by the predominance of the 3 handshape in Type 2 sign in ASL and by the exceptional distribution of 59 in KSL. In contrast, NZSL and SVK show strong dependence between handshape and location. I cannot determine the source of this dependence in NZSL, although signs with contact locations favor marked handshapes while signs in neutral space favor unmarked handshapes, though not significantly. The one significant source of dependence was the exceptional distribution of 49B. In SVK, the dependence between handshape and location is accounted for by the coincidence of marked handshapes and face locations. One notable cross-linguistically similarity was the fact that neither definition of handshape markedness played an important role in the dependency between handshape and location in any language except SVK. In particular, the affinity of the face and neck for marked handshapes that was predicted to occur in ASL was not confirmed.

# Chapter 6: Conclusion

## 6.1 Summary

### 6.1.1 Inventory and rank frequency

Handshape, location and duet inventories were determined for each of the four languages under investigation. Handshape and location inventories vary in both size and composition, but each language has approximately the same number of handshapes as locations. In contrast, duet inventories are remarkably similar in size, as summarized in (6-1).

(6-1) All languages: comparison of inventory sizes

| | handshapes | locations | total | duets |
|---|---|---|---|---|
| SVK | 34 | 34 | 68 | 189 |
| ASL | 35 | 38 | 73 | 192 |
| KSL | 44 | 41 | 85 | 189 |
| NZSL | 49 | 47 | 96 | 188 |

Although a sample of four is small, it does not appear that phonological simplicity in one parameter requires compensatory complexity in another, at least within the handshape and location inventories. It is possible that movement or orientation inventories are more abundant in a language with a smaller handshape and location inventory, or that there are more syllable types (see section 1.3.3 and footnote 12). I have not explicitly investigated these possibilities, but based on my knowledge of ASL and SVK, the languages with the smallest inventories, along with casual observation throughout the construction of the four databases, I do not think that KSL and NZSL are deficient in movement, orientation, or syllable types.[85] Another possibility not

---

[85] This is not to say that there is no diversity in these constituents. For example, the syllable type ML (movement to a location) is common in ASL; in SVK, LM (movement from a location) is common. HEAD in ASL and PÄÄ *head* in SVK are both formed with the 3 handshape with the head as location. But in ASL, the movement is toward the head, while in SVK, the movement is away from the head. The SVK sign SYY

explored in this dissertation is that languages with simpler inventories use the compounding process to a greater extent for compensatory complexity. Also, non-manual gestures, and in particular, mouth movements, could provide another means of adding complexity to a small handshape and location inventory. However, ASL and SVK, the languages with the smallest handshape and location inventories, are distinct in their use of mouth movements. SVK uses mouth movements to a great extent, to disambiguate homophonous forms, to nuance the meaning of a sign, to distinguish noun-verb pairs, and so on. In contrast, mouth movements are infrequently in ASL. Finally, notice the remarkable similarity cross-linguistically in the size of the duet inventories. All languages, even SVK with the smallest handshape and location inventories, are able to produce the same number of distinct duets. There would appear to be no need to resort to an elaborated set of other parameters in order to produce a lexicon of sufficiently distinct signs.[86]

The four languages share 22 handshapes. When these handshapes are ranked from most frequent to least frequent in each language, it was shown in section 3.2.1 that although the rankings appear to differ somewhat, they are similar enough to be considered samples drawn from a common source, with $p<0.01$ for all languages. Similarly, the four languages share 24 locations, 18 of which are body locations. Whether all locations are considered or just body locations, it was shown in section 4.2.1 that the rankings of location are similar enough to be considered samples drawn from a common source, with $p<0.01$ for all locations and $p<0.05$ for body locations.

According to one theory of spoken language inventory structure (see McCarthy (1999), for example), the consonant inventory is structured so that if an articulatory resource is available, maximum use is made of this resource. For example, if a language uses the alveolar place of articulation, this place combines with all available manners of articulation. This structural generalization is not true of the duet inventory, which comprises combinations of handshapes and locations. While some handshapes occur at almost all locations and some locations host almost all handshapes, most do not. Another theory of spoken language inventory structure is that the

*cause* is another example. The location is the nondominant 3; the fingertips of the dominant ; hand contact the palm of the 3 hand, then move away from this location opening to 78B.

[86] Of course, SVK must actually use a greater percentage of its theoretically possible duets than does NZSL. This was shown to be true in section 5.1; attestation of potential duets is as follows: ASL 15%; KSL 14%; NZSL 8.4%; SVK 17%.

vowel inventory is structured so as to maximize the phonetic distance between the set of vowels that occupy a language's vowel space (Liljencrants and Lindblom, 1972; Crothers, 1978). Neither is this structural generalization true of the duet inventory. If the entire space of potential duets is considered, that is, the whole handshape by location table, it is not the case that actually occurring duets disperse themselves evenly throughout this space.

What does sign language inventory size tell us about the optimum inventory size range of languages (section 1.4)? Regardless of what is considered segmental in sign language, the inventories are large. If only movement and location are segmental (Sandler, 1986; Sandler, 1987; Perlmutter, 1988; Sandler, 1989; Perlmutter, 1992; Perlmutter, 1993), there are approximately 44 to 57 segments.[87] (Recall that the inventory size of seventy percent of the spoken languages in UPSID (Maddieson, 1984) lies between 20 and 37 segments.) If handshape and location are considered segmental, there are 68 to 96 segments, and if handshape, location and movement are all considered segmental, there are 78 to 102 segments. Alternatively, if duets are considered segmental, the inventories are even larger, with about 200 duet segments in each language. Regarding the entire simultaneously occurring bundle of handshape, location and movement as a segment produces approximately 2000 segments. With such great combinatorial potential, it is not surprising that most signs comprise only one phonological constituent, whether that constituent is called a segment or a syllable.

Van der Hulst (1985) comments on this situation.

> "This leads us to a remarkable and potentially confusing conclusion. Since most morphemes in ASL (and other sign languages as well) have just one value for the attributes [of handshape, location, and movement], we must conclude that most morphemes in sign languages are monosegmental. But there is no contradiction here, since there is no principled reason why morphemes should consist of more than one segment. That this is typically so in spoken languages follows from the fact that the class of segments is many times smaller than the class of morphemes that languages seem to have."

Next, the distribution of handshapes, locations and duets in the lexicon was investigated. It is not the case that either the handshape resource or the location resource is used uniformly. Neither are duets used uniformly. Out of 405 ASL signs, 35 or 8.6% of them feature unique

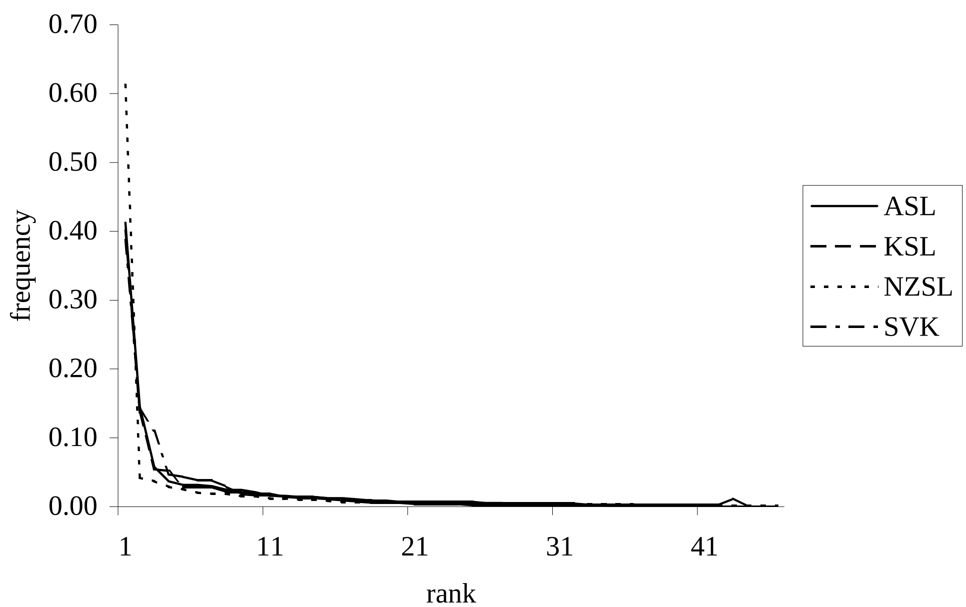
---

[87] The number of locations in (6-1) was added to ten, which is a rough estimate of the number of movement types gleaned from Stokoe (1965), Sandler (1989), and Brentari (1998).

duets, while 91 or 23% of them use the top five most frequent duets, {3π 49π 78π 3| 493}. In all four languages, a few resources are used very frequently, while many resources are used very rarely. There is remarkable cross-linguistic agreement. The handshape and body location rank-frequency graphs are best approximated by exponential decay equations, while duet and all location rank-frequency graphs are best approximated by power law equations. These statistical models are proposed to be universals of sign language lexicon structure. This proposal will be investigated further in section 6.2.

In addition to the theoretical arguments for the nondominant hand as the location in Type 2 and Type 3 signs presented in section 2.1.2.2.3, a quantitative argument can now be offered as well. Recall that in Chapter 4, the locations of signs articulated anywhere in neutral space are classified as π. The locations of all Type 2 signs are classified as |. The locations of Type 3 signs are differentiated by the handshape of the nondominant hand, 3, 49, etc., just as locations on the face are differentiated, ψ, υ, etc. With this classification of locations, the rank-frequency graphs for all four languages, plus VSVK and the pooled data, exhibited the same shape, that of a power-law distribution, repeated here as (6-2).

(6-2) All languages: rank-frequency graph for all locations



There are, of course, other ways to classify the locations, some of which are shown in (6-3).[88] For example, it is possible to consider the location of a Type 2 or Type 3 sign to be neutral space, the first possibility listed in (6-3). The classification actually used in Chapter 4 was for a Type 2 sign to have $\vert$ as its location, and for Type 3 signs to have an individually specified handshape, $h_1$, $h_2$, … , as its location; it is indicated in (6-3) by a dotted line.

(6-3) Possibilities for locations of Type 2 and Type 3 signs

| Type 2 | Type 3 | Type 2 | Type 3 | Type 2 | Type 3 |
|---|---|---|---|---|---|
| $\pi$ | $\pi$ | $\vert$ | $\pi$ | $h_1, h_2, \ldots$ | $\pi$ |
| $\pi$ | $\vert$ | $\vert$ | $\vert$ | $h_1, h_2, \ldots$ | $\vert$ |
| $\pi$ | $h_1, h_2, \ldots$ | $\vert$ | $h_1, h_2, \ldots$ | $h_1, h_2, \ldots$ | $h_1, h_2, \ldots$ |

The location rank-frequency graphs for all of the possibilities in (6-3) were drawn, and a linear regression was done on X = log x and Y = log y for each language. Every possibility investigated except the accepted one gave inconsistent results across ASL, KSL, NZSL, SVK, VSVK, and the pooled data. The location rank-frequency graph for the other possibilities did not have a consistent form, such as linear, exponential, or hyperbolic, for all four languages. The only classification that gave consistent results cross-linguistically was the one which is also supported theoretically: the nondominant hand is the location in Type 2 and Type 3 signs; the

---

88 One possibility not explored is that of a complex location, with the non-dominant hand as the location for the dominant hand, and neutral space as the location of the non-dominant hand or of the two-handed composite. This possibility is not so interesting for a combination in neutral space, as neutral space is usually represented with an empty location node. However, for ASL signs such as NOSE-TO-THE-GRINDSTONE or SURGERY-ON-BODY-PART, in which the dominant hand contacts the non-dominant hand while the non-dominant hand contracts a part of the body, this analysis of location is reasonable. Many such forms violate the Symmetry and Dominance conditions, and, are in fact morphologically complex or derived historically from polymorphemic forms. Such forms do not occur as monomorphemic lexical items; hence, the complexity of their location ought to be represented on another linguistic level. (One exception is the monomorphemic sign INTERNALIZE, which has a complex location involving the nondominant hand and the chest.) See Rozelle (1998) for further discussion.

location is an unspecified “hand” node, | , in Type 2 signs; the location is the particular handshape of the nondominant hand in Type 3 signs.

### 6.1.2 Distributions

In the second parts of Chapters 4, 5 and 6, the distribution of one sign property with respect to another property was investigated. Chapter 4 investigated the distribution of handshapes across all sign types, Chapter 5, the distribution of locations across sign Types 0 and 1 (one-handed and two-handed signs), and Chapter 6, the distribution of handshapes across locations. The analyses tested for dependence by using chi-square tests on r x c tables when all cell counts exceeded five and by using the mutual information significance program when there was a cell count five or less. The results are summarized in (6-4) for handshape by type, in (6-5) for location by number of hands, and in (6-6) for handshape by location. When a result is of borderline significance, or just a tendency based on observed versus expected values, it is parenthesized.

(6-4) Summary: handshape by Type 0/1/2/3

| | ASL | KSL | NZSL | SVK |
|---|---|---|---|---|
| hs x T 0/1/2/3 | *dependent*<br>T2: more<br>unmarked hs | (*dependent?*<br>T1: few 49 ) | *dependent*<br>T1: more<br>unmarked hs | **independent**<br>(T2: more<br>unmarked hs) |
| hs x one/two-handed | **independent** | *dependent*<br>T1: few 49 | *dependent*<br>2-handed: more<br>unmarked hs | **independent** |
| universal<br>marked/unmarked x<br>T 0/1/2/3 | *dependent*<br>T2: more<br>unmarked hs | **independent** | *dependent*<br>T1: more<br>unmarked hs | (*dependent*<br>T2: more<br>unmarked hs) |
| language-specific<br>marked/unmarked x<br>T 0/1/2/3 | **independent** | **independent** | **independent** | (*dependent*<br>T2: more<br>unmarked hs) |

(6-5) Summary: location by Type 0/1 (T0: one-handed; T1: two-handed)

| | ASL | KSL | NZSL | SVK |
|---|---|---|---|---|
| loc x T0/1 | *dependent* | *dependent* | *dependent* | *dependent* |
| contact/ $\pi$ x T0/1 | *dependent*<br>T1: more $\pi$ | *dependent*<br>T1: more $\pi$ | *dependent*<br>T1: more $\pi$ | *dependent*<br>T1: more $\pi$ |
| contact locs x T0/1 | *dependent* | *dependent* | *dependent* | *dependent* |
| face/torso x T0/1 | *dependent*<br>T1: more torso | **independent** | *dependent*<br>T1: more torso | *dependent*<br>T1: more torso |
| middle/side x T0/1 | *dependent*<br>T1: more side | **independent** | *dependent*<br>T1: more side | **independent** |

(6-6) Summary: handshape by location

| | ASL | KSL | NZSL | SVK |
|---|---|---|---|---|
| hs x loc | borderline? | **(independent?)** | *dependent* | *dependent* |
| hs x contact/ $\pi$ | **independent** | (*dependent*<br>more 59 in $\pi$) | *dependent*<br>( $\pi$ : more<br>unmarked hs) | **independent** |
| hs x contact locs | *dependent*<br>│ : more<br>unmarked hs | **independent** | *dependent*<br>49B skewed<br>distribution | *dependent*<br>face: more<br>marked hs |
| universal marked/unmarked x contact locs | *dependent*<br>│ : more<br>unmarked hs | **independent** | **independent** | *dependent*<br>face: more<br>marked hs |
| language-specific marked/unmarked x contact locs | **independent** | **independent** | **independent** | *dependent*<br>face: more<br>marked hs |

There are two notable observations about the preceding three summaries. First, there is a great deal of cross-linguistic variation in these distributions, unlike the rank-frequency distributions.  A natural grouping, such as face versus non-face, which strongly influences whether a sign will be one-handed or two-handed in one language, such as in ASL (p=0.001), has no effect in another language, such as in KSL (p=0.498).  Second, the concept of markedness, be it defined by the universal, frequency-based criterion or by the language-specific, phonology-based criterion, is not especially useful in capturing patterns of dependencies between type or location and handshape.

Section 5.2.5 investigated a well-known dependence that is supposed to exist in ASL between handshape and location.  It is claimed that marked handshapes occur on the face and neck, and unmarked handshapes occur on the torso (Siple, 1973; Siple, 1978; Siple, 1980; Battison, 1995).  Note that this theory of lexical structure promotes perceptual ease over productive ease.  This proposed dependence could not be confirmed for ASL, nor for KSL or NZSL, by using either using Battison's markedness criterion or the two markedness criteria tested in this dissertation.  However, SVK did show the proposed dependence between handshape and location, which is consistent with the results in (6-4), (6-5), and (6-6).  These results indicate that two relevant groupings in SVK are the face/neck versus torso location grouping and the marked versus unmarked handshape grouping, where markedness is defined by either Battison's criterion or the universal criterion.

Another dependence between handshape and location that is supposed to exist in the children's acquisition of handshape is proposed in Boyes Braem (1990).  According to this research, the two most important factors influencing acquisition of handshape and location duets by children are visual feedback and tactile feedback (kinesthetic feedback is the third factor), in this order.  In particular, if a sign is made in a location that cannot be visually monitored, the child language learner simplifies the handshape.  Note that this theory promotes productive ease over perceptual ease.  Is this pattern for child language acquisition true of the adult lexicon?  Face locations provide tactile feedback but no visual feedback.  Neutral space provides visual feedback but not necessarily tactile feedback.  However, signs with the nondominant hand as location provide both types of feedback.  Is it the case in the adult lexicon that Type 2 and Type 3 signs allow the most diverse set of handshapes, while face locations allow the least diverse set?  For all four languages, the subset of signs with face location was compared to the subset with hand locations.  Handshapes were divided into marked and unmarked sets by both the universal and language-specific criteria.  For all languages and for both markedness criteria, the observed value of marked handshapes on the face was equal to or greater than the expected value.  Thus, although Boyes Braem's observation for the acquisition of ASL might be true, it does not appear to affect the structure of the adult lexicon.

### 6.2 Discussion of rank-frequency

The striking cross-linguistic uniformity of the sign language rank-frequency distributions is interesting as a property of sign language distributions. But are these distributions characterized by statistical models alone or can they also be explained by some linguistic concept? Recall that the rank-frequency distributions of handshapes and body locations are graphs of exponential decay functions while the rank-frequency distributions of duets and combined locations are graphs of power law functions. There cannot be a single statistical model since two different distributions must be accounted for.

In section 6.2.1, the rank-frequency distributions of segments in spoken language are presented and examined to see if they show the same uniformity and behavior as the sign language distribution. In section 6.2.2, Pietrandrea's (1998; 2000) account of handshape and location rank-frequency distributions in LIS is discussed. Next, in section 6.2.3, another field of research in which rank-frequency distributions are important, ecological diversity, is considered to see if insights about species distributions can be applied to linguistics. Finally, in section 6.2.4, I hypothesize that the sign language lexicon is a self-organizing, complex system, as evidenced by the power law distribution of duets, and I propose a method for substantiating this suggestion.

#### 6.2.1 Spoken language rank-frequency

There is great agreement in the form of the rank-frequency curves for sign languages. Do spoken languages agree in their rank-frequency curves? If they do, do they exhibit exponential decay or the power law or neither? The phonological constituent that was measured was the segment, because segmental counts were most widely available; however, it is possible that another constituent would yield more satisfactory results.

For this inquiry, nine spoken languages were surveyed. The linguists who investigated these nine languages used different methods for counting segments. In some cases, tokens are counted and in others types. The sample size varies. Often, when frequency data is given, it is grouped in a way that is relevant to the phenomenon the researcher is investigating, such as syllable-initial versus syllable-final consonants. The languages found to have useable segment frequency counts available are Chamorro (Seiden, 1960), Czech (Kucera and Monroe, 1968), English (Roberts, 1965), Greek (Modern Dhimotiki) (Householder et al., 1964), German (Kucera and Monroe, 1968), Japanese (Bloch, 1950), Russian (Kucera and Monroe, 1968), Kwadacha (Fort Ware Sekani) roots (Hargus, In preparation)[89], and Tulu (Bhat, 1967). Their inventories and frequencies are listed in (6-7) and (6-8), together with whether the count is of segment types or tokens, in those cases where this information is available. Each language is notated in the original transcription system used by the author, and all and only the distinctions present in the published segment list are exhibited. I did not attempt to uniformly represent all segments in IPA, since it is the frequency distribution that is under investigation. Note that the numerals in the Japanese data represent tones, while the Sekani data group vowels that differ in tone and nasality.

89 Thanks to Sharon Hargus for letting me use her dictionary of Kwadacha (Fort Ware Sekani) while it was being prepared.

(6-7) Spoken language segment inventories and frequencies

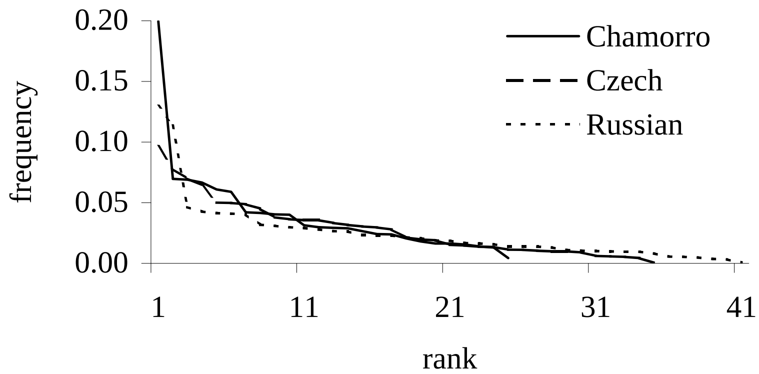
| **Chamorro** | ? | **Czech** | token | **English** | type | **German** | token | **Greek** | ? |
|---|---|---|---|---|---|---|---|---|---|
| ◊ | 0.1995 | ε | 0.0965 | ™ | 0.1182 | ε | 0.1193 | ◊ | 0.1303 |
| ι | 0.0696 | ο | 0.0779 | ι | 0.0929 | ν | 0.1007 | ι | 0.1191 |
| υ | 0.0690 | α | 0.0699 | τ | 0.0695 | τ | 0.0887 | ε | 0.0946 |
| ν | 0.0666 | ι | 0.0643 | ψ | 0.0677 | ρ | 0.0753 | ο | 0.0876 |
| ο | 0.0608 | σ | 0.0499 | ρ | 0.0658 | ι | 0.0699 | σ | 0.0761 |
| τ | 0.0589 | λ | 0.0498 | ν | 0.0629 | α | 0.0634 | τ | 0.0760 |
| λ | 0.0420 | τ | 0.0488 | ε | 0.0474 | σ | 0.0461 | ν | 0.0644 |
| γ | 0.0415 | ν | 0.0449 | ◊ | 0.0463 | δ | 0.0412 | κ | 0.0452 |
| σ | 0.0404 | μ | 0.0379 | ω | 0.0452 | λ | 0.0350 | μ | 0.0450 |
| μ | 0.0401 | ϖ | 0.0363 | σ | 0.0390 | υ | 0.0315 | ρ | 0.0436 |
| κ | 0.0314 | ι: | 0.0356 | λ | 0.0330 | μ | 0.0284 | π | 0.0375 |
| π | 0.0297 | κ | 0.0356 | δ | 0.0304 | ξ | 0.0266 | λ | 0.0289 |
| δ | 0.0292 | φ | 0.0333 | η | 0.0263 | ε: | 0.0257 | υ | 0.0284 |
| β | 0.0289 | υ | 0.0317 | μ | 0.0261 | ϕ | 0.0245 | ψ | 0.0265 |
| ? | 0.0265 | π | 0.0304 | κ | 0.0245 | ι: | 0.0214 | Δ | 0.0158 |
| ε | 0.0240 | ρ | 0.0297 | Δ | 0.0225 | ϖ | 0.0209 | ξ | 0.0143 |
| η | 0.0237 | δ | 0.0278 | ζ | 0.0200 | γ | 0.0200 | ζ | 0.0125 |
| ϕ | 0.0205 | α: | 0.0212 | υ | 0.0191 | κ | 0.0192 | ϕ | 0.0122 |
| │ | 0.0180 | ν□ | 0.0197 | ϖ | 0.0188 | ζ | 0.0182 | Τ | 0.0113 |
| χ | 0.0164 | ζ | 0.0190 | ϕ | 0.0170 | β | 0.0175 | Γ | 0.0091 |
| ω | 0.0164 | β | 0.0152 | β | 0.0163 | σʃ | 0.0173 | ϖ | 0.0077 |
| ζ | 0.0156 | η | 0.0147 | π | 0.0161 | α: | 0.0171 | Δ | 0.0059 |
| ® | 0.0139 | χ | 0.0137 | ® | 0.0154 | ο | 0.0170 | β | 0.0046 |
| ψ | 0.0131 | σʃ | 0.0135 | ο | 0.0154 | η | 0.0113 | ϕ | 0.0026 |
| ρ | 0.0044 | ε: | 0.0113 | │ | 0.0092 | π | 0.0103 | χ | 0.0005 |
| | | ρʃ | 0.0111 | γ | 0.0087 | ο: | 0.0083 | ζ | 0.0001 |
| | | ξ | 0.0104 | σʃ | 0.0072 | υ: | 0.0071 | | |
| | | τ□ | 0.0099 | ⌊ | 0.0064 | │ | 0.0068 | | |
| | | χʃ | 0.0099 | χʃ | 0.0046 | υ⌋: | 0.0044 | | |
| | | ζʃ | 0.0090 | Τ | 0.0042 | ο⌋: | 0.0018 | | |
| | | ϕ | 0.0060 | φ | 0.0036 | ο⌋ | 0.0017 | | |
| | | υ: | 0.0056 | ζʃ | 0.0003 | υ⌋ | 0.0004 | | |
| | | δ□ | 0.0053 | | | ζʃ | 0.0001 | | |
| | | γ | 0.0043 | | | | | | |
| | | ο: | 0.0002 | | | | | | |

(6-8) Spoken language segment inventories and frequencies, continued

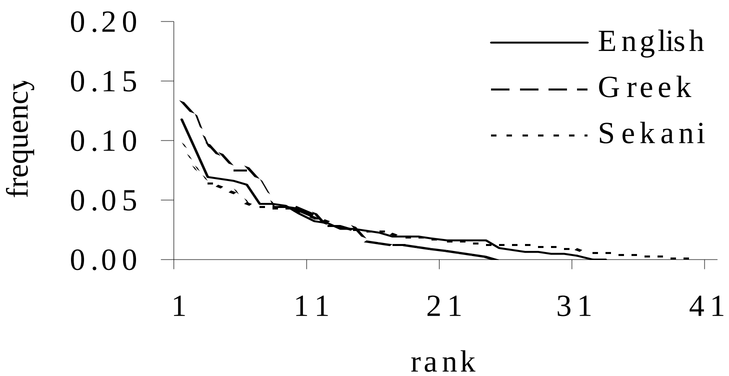
| **Japanese** | token | **Russian** | token | **Sekani** | type | **Tulu** | token |
|---|---|---|---|---|---|---|---|
| α | 0.1338 | α | 0.1296 | α | 0.0944 | α | 0.1432 |
| ο | 0.1085 | ι | 0.1135 | ε | 0.0764 | τ | 0.0727 |
| 3 | 0.0777 | α⌡ | 0.0463 | ι | 0.0647 | υ | 0.0630 |
| ι | 0.0737 | τ | 0.0427 | η | 0.0622 | ν | 0.0580 |
| 2 | 0.0732 | φ | 0.0414 | ™ | 0.0560 | ⎞ | 0.0566 |
| τ | 0.0619 | ν | 0.0410 | τΗ | 0.0486 | ε | 0.0559 |
| # | 0.0583 | ο⌡ | 0.0403 | ο | 0.0454 | ρ | 0.0471 |
| 4 | 0.0561 | κ | 0.0318 | τσΗ | 0.0428 | π | 0.0468 |
| ε | 0.0466 | σ | 0.0309 | τ | 0.0428 | ι | 0.0407 |
| κ | 0.0420 | ϖ | 0.0298 | σ | 0.0421 | κ | 0.0406 |
| μ | 0.0352 | ρ | 0.0291 | ν | 0.0351 | λ ¯ | 0.0333 |
| ν | 0.0348 | ι⌡ | 0.0278 | υ | 0.0340 | α⌉ | 0.0329 |
| υ | 0.0298 | λ | 0.0266 | λ | 0.0267 | δ ¯ | 0.0316 |
| ρ | 0.0266 | ε⌡ | 0.0261 | ζ | 0.0256 | μ | 0.0308 |
| σ | 0.0253 | μ | 0.0232 | κΗ | 0.0238 | σ | 0.0236 |
| σ⌠ | 0.0190 | π | 0.0231 | τσэ | 0.0234 | ν ¯ | 0.0219 |
| ν⌡ | 0.0140 | νэ | 0.0230 | ℜ | 0.0230 | ψ | 0.0219 |
| δ | 0.0130 | υ | 0.0213 | ⊗ | 0.0194 | λ | 0.0213 |
| ǀ | 0.0130 | λэ | 0.0208 | κэ | 0.0190 | ο | 0.0207 |
| η | 0.0094 | τэ | 0.0189 | τΣΗ | 0.0176 | ο⌉ | 0.0188 |
| β | 0.0094 | σэ | 0.0186 | κ | 0.0168 | β | 0.0187 |
| ω | 0.0094 | δ | 0.0167 | φ | 0.0154 | φ | 0.0148 |
| ψ | 0.0072 | χ⌠ | 0.0163 | τℜΗ | 0.0146 | γ | 0.0128 |
| φ | 0.0059 | σ⌠ | 0.0156 | τℜэ | 0.0132 | ε⌉ | 0.0102 |
| χ⌠ | 0.0054 | ρэ | 0.0138 | τэ | 0.0132 | Ε | 0.0094 |
| 1 | 0.0027 | ζ | 0.0137 | ξ | 0.0128 | υ⌉ | 0.0087 |
| χ | 0.0022 | υ⌡ | 0.0136 | τσ | 0.0128 | ϖ | 0.0079 |
| ζ | 0.0022 | γ | 0.0130 | π | 0.0121 | νϒ | 0.0077 |
| ǀ | 0.0013 | β | 0.0110 | Υ | 0.0113 | ν) | 0.0060 |
| γ | 0.0013 | δэ | 0.0104 | / | 0.0102 | τ | 0.0059 |
| ξ | 0.0009 | ϖэ | 0.0102 | ω | 0.0099 | χ | 0.0057 |
| | | ξ | 0.0099 | Χ | 0.0066 | ι⌉ | 0.0045 |
| | | ζ⌠ | 0.0095 | τΣ | 0.0062 | η | 0.0032 |
| | | ϕ | 0.0095 | τℜ | 0.0055 | σ ¯ | 0.0026 |
| | | μэ | 0.0081 | τΣэ | 0.0048 | σ⌠ | 0.0005 |
| | | χ | 0.0056 | Σ | 0.0037 | | |
| | | κэ | 0.0053 | μ | 0.0037 | | |
| | | πэ | 0.0048 | β | 0.0022 | | |
| | | βэ | 0.0037 | κΩΗ | 0.0011 | | |
| | | ζэ | 0.0032 | ξΩ | 0.0007 | | |
| | | ϕэ | 0.0006 | Ζ | 0.0004 | | |

The rank-frequency curves are shown in (6-9), (6-10), and (6-11). Although they bear some similarity to each other, they vary more than the handshape rank-frequency curves.

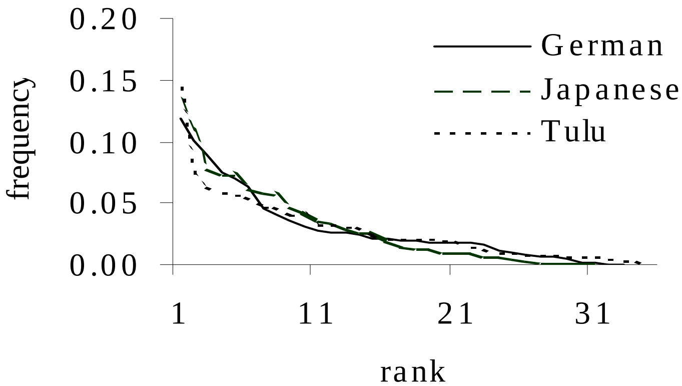
(6-9) Chamorro, Czech, Russian: rank-frequency graphs




(6-10) English, Greek, Sekani: rank-frequency graphs



(6-11) German, Japanese, Tulu: rank-frequency graphs




Linear regression was applied to x = rank and y = log (frequency), in the same manner as it was for handshape rank-frequency. In addition, linear regression was applied to the unlogged data, x = rank and y = frequency. The results are mixed. For some languages, the exponential decay graph is a better approximation to the rank-frequency curve, while for other languages, the linear graph is better. Except for Japanese, which is fit well by the exponential decay curve, none of the languages are fit as well by either method as the sign language data are. For sign language handshape rank-frequency, the fit with the exponential decay curve is very close, but when linear regression was applied to the unlogged handshape data, the fit is poor. The results are shown in (6-12).

(6-12) All languages: $R^2$ values for linear and exponential approximations

| | all segments | | consonants only |
|---|---|---|---|
| | exponential | linear | linear |
| Chamorro | 0.881 | 0.530 | 0.899 |
| Czech | 0.781 | 0.860 | 0.948 |
| English | 0.827 | 0.809 | 0.887 |
| German | 0.738 | 0.764 | 0.737 |
| Greek | 0.850 | 0.860 | 0.880 |
| Japanese | 0.965 | 0.815 | 0.823 |
| Russian | 0.846 | 0.558 | 0.937 |
| Sekani | 0.876 | 0.828 | 0.860 |
| Tulu | 0.901 | 0.719 | 0.895 |

| | handshapes | |
|---|---|---|
| | exponential | linear |
| ASL | 0.963 | 0.596 |
| KSL | 0.934 | 0.448 |
| NZSL | 0.974 | 0.589 |
| SVK | 0.961 | 0.484 |

The different methodologies used in the determination of the spoken language segment frequency data make it difficult to base a conclusion on these data. Nevertheless, I would tentatively hypothesize that spoken languages vary more in their segmental frequency distributions than do sign languages, and that they are in general not well approximated by an exponential decay curve.

Can a comparison of the distributions of handshapes, locations and duets with the distribution of spoken language constituents shed light on the analogy between sign and spoken language constituents? As discussed in section 1.3.3, phonological patterning does not provide an unambiguous answer. The distributions of segments in spoken language are fit only moderately well by either linear or exponential decay graphs, so it is not clear whether the distributional patterns provide support for analogizing segments with handshapes or locations. Segmental rank-frequency distributions are definitely not described by a power law. What about other spoken language phonological units? The distribution of consonants in the nine spoken languages was also investigated. The results in (6-12) show that the consonant distributions are more linear than the handshape or location distributions, and thus are not at all like duet inventories. The distribution of words in languages has been well-studied (Baayen, 2001); cross-linguistically, this distribution is in accord with Zipf's Law, which is a power law distribution.[90] These results suggest that a subset of sign language phonological parameters, such as handshape or location, do not pattern distributionally like consonants, which are a subset of spoken language segments. The distribution of handshapes is more like the distribution of segments than either consonants or words. The distribution of duets, however, is similar to that of words. Indeed, many signs can be uniquely identified by their handshape and location alone.[91]

### 6.2.2 Pietrandrea's account of LIS rank-frequency distributions

As discussed in sections 3.2.2.7 and 4.2.2.6, the handshape and location rank-frequency distributions in Italian Sign Language (Lingua Italiana dei Segni: LIS) is very similar to the distributions in ASL, KSL, NZSL and SVK. All handshape rank-frequency distributions are well-fit by an exponential decay curve of the form $y = a2^{-bx}$. Pietrandrea (1998; 2000) suggests

---

[90] Word distributions, of course, are based on token rather than type counts. It is likely that a count of duet tokens would also exhibit a power law distribution.

[91] Perhaps as the informational content of a phonological unit increases, its rank-frequency distribution evolves into a power-law.

that the form of the handshape rank-frequency curve can be explained by redundancy. Pietrandrea (1998:76) writes:

> "The results show that, although LIS signers have the choice of a wide range of alternative units in the creation of lexemes, they tend to use just a small group of these units in the citation lexicon. This phenomenon, that is characteristic of vocal languages as well, can be considered an example of the REDUNDANCY that is so prevalent in languages. Languages tend to consistently reuse the same material or to repeat the same information. The repetition of the same information (usually a syntactic phenomenon, such as agreement) allows users (especially the receiver) to better process language."

As was seen in section 6.2.1, spoken language rank-frequency graphs do not exhibit the same behavior as sign language graphs. Although they do tend to use a small number of phonemes frequently, their rank-frequency graphs are more variable and have different shapes than the signs language graphs.

In linguistics, "redundancy" refers to the repetition of information, not the reuse of material. For example, in spoken Finnish, the case is redundantly indicated on both the noun and its adjective, and in spoken Spanish, gender is redundantly indicated on the noun and both its adjective and determiner, as shown in (6-13). For example, the inessive case ending, -*ssa*, on the adjective is redundant information, since this information is already given by the inessive case ending of the noun.

(6-13) Redundancy in Finnish and Spanish

| **Finnish** | **Spanish** |
|---|---|
| *musta-ssa talo-ssa*<br>black-INESSIVE house-INESSIVE<br>'in the black house' | *la casa negra*<br>the-FEM house-FEM black-FEM<br>'the black house' |
| *musta-lla talo-lla*<br>black-ADESSIVE house-ADESSIVE<br>'on the black house' | *el sombrero negro*<br>the-MASC hat-MASC black-MASC<br>'the black hat' |

Another example of redundancy is in phonetics, when two or more cues are used in distinguishing a speech sound. For example, in English, the sounds [p] and [b] are distinguished syllable-initially not only by voicing, but also by aspiration: [p] is voiceless and aspirated, while [b] is voiced and unaspirated. The aspiration of [p] is predictable from its voicelessness, thus it is redundant information. I cannot see how the frequent use of a certain set of phonetic material, is related to redundancy. Just because a resource is used frequently does not imply that it is redundant. For example, handshape is not predictable from other information, such as location, as was discussed in section 5.2.

### 6.2.3 Ecological diversity distributions

Biologists seeking generalizations about the composition and structure of ecological communities often study rank-frequency distributions like the ones presented in this dissertation. In a given region, the number of different species and the relative abundance of each species, where abundance is quantified by an appropriate measure, such as number of individuals or biomass, is plotted as a rank-frequency distribution, termed a rank-abundance distribution.[92] This distribution is an important means of characterizing the diversity of an ecosystem, and ecologists endeavor to extract meaningful descriptions from these distributions. This task is parallel to the task presented by the rank-frequency distributions in Chapters 3, 4 and 5.

There are two approaches to analyzing rank-abundance data. The first approach, called the *mathematical model*, is descriptive and appears in the statistical literature: an appropriate statistical model is developed to fit the observed data. Unfortunately, the data available on ecosystems are often sparse and of questionable quality (Büssenschütt, 1997), so that the second approach is favored by ecologists.[93] I will call this approach the *conceptual model*, though it is sometimes called the *resource apportioning model* (Pielou, 1975). In this approach, a biologically-motivated model for a distribution is postulated, for which a mathematical expression of this model is sought. Sometimes these two approaches converge, and a purely descriptive mathematical model duplicates the results of a conceptual model (Pielou, 1975).[94]

---

[92] There are other types of distributions used by ecologists to compare the diversity of species (May, 1975; Pielou, 1975; Büssenschütt, 1997). Another important distribution is the *species-abundance* distribution, in which $f(r)$ is the frequency of species (or handshapes) that contain exactly r individuals (or signs), for $r$=1, 2, ….

[93] Pielou (1975) dismisses statistical versions as "ignorance in manageable form." The ultimate goal of linguistics, and presumably ecology, is explanation; however, knowing what needs to be explained, and putting this ignorance in manageable form, for example, by identifying a paradigm or a distribution, must occur first.

[94] One example of the convergence of mathematical and conceptual models is a distribution ubiquitous in ecosystems but not occurring in the sign language data (Preston, 1962a; 1962b). The conceptual model for this distribution is similar to the broken stick model discussed below, except that the stick is broken

Chapters 3, 4 and 5 presented well-fitting mathematical models of rank-frequency distributions without offering a conceptual model as a linguistic explanation for their occurrence. This section discusses two ecosystem models and their analogies to the handshape distributions, as well as a possible explanation for the duet power law distribution.

One conceptual model of distribution of species within an ecosystem is the *niche preemption model* (May, 1975). In this model, the dominant species, $n = 1$, appropriates a fraction, $k$, of available resources, the second strongest species, $n = 2$, appropriates the same fraction, $k$, of the remaining resources, and so on, so that the fraction of resources appropriated by the $n^{th}$ strongest species is $f(n) = k(1-k)^{n-1}$, for $n = 1, 2, \ldots s$.[95] Since this series forms a geometric progression, this model is also called the *geometric model*. Translating into linguistic terms, each handshape is a species, and the resource that is being apportioned is the lexicon. The "strongest" species, which is the most common handshape, claims a portion of the lexicon; in the case of ASL, $k = 0.152$. The geometric model is a good fit for the handshape data, but not for the duet data, which is not surprising since $f(x) = k(1-k)^{x-1}$ is an exponential decay function ($0 < 1-k < 1$ and $x-1 \geq 0$).

If this model has sufficient descriptive accuracy, what is the linguistic motivation behind it? I will suggest three possibilities, but I will not explore them in this dissertation. One possibility is that the frequency ranking corresponds to a language-specific markedness hierarchy. The "strongest" handshape is the least marked, and due to some property it maximizes (such as perceptual salience or articulatory ease), it is used by the largest number of signs in the lexicon. A second possibility is that a phonological representation of handshape employs the simplest representation for the "strongest" handshape, and increasingly more complex representations for the subsequent handshapes. Furthermore, perhaps the geometric progression in the distribution of handshape in the lexicon is inversely proportional to the number of features (or to some other measurement of representational complexity) needed in the phonological representation of the

---

sequentially rather than simultaneously into the needed number of pieces. When the number of pieces (species) is large, this model approaches the lognormal distribution, which is the mathematical model. However, the handshape data are not normally distributed even when logged, and the common handshapes are commoner and the rare handshapes are rarer than this model predicts, so the lognormal distribution does not fit.

95 To ensure that the sum of the frequencies is one, the final term in the series can be adjusted to be $f(s)=(1-k)^{s-1}$ (Pielou, 1975).

handshapes.[96] The problem with this second possibility is that it entails language-specific handshape representations in order to account for the differing handshape rankings in each language. A third possibility is that the ranking of the handshapes is the output of a language-specific Optimality Theory ranking of (sign) language universal constraints. This last suggestion is the most promising, because it exploits OT's ability to account for cross-linguistic variation by isolating language differences to the ranking of constraints while retaining linguistic universals in the form of the constraints themselves.

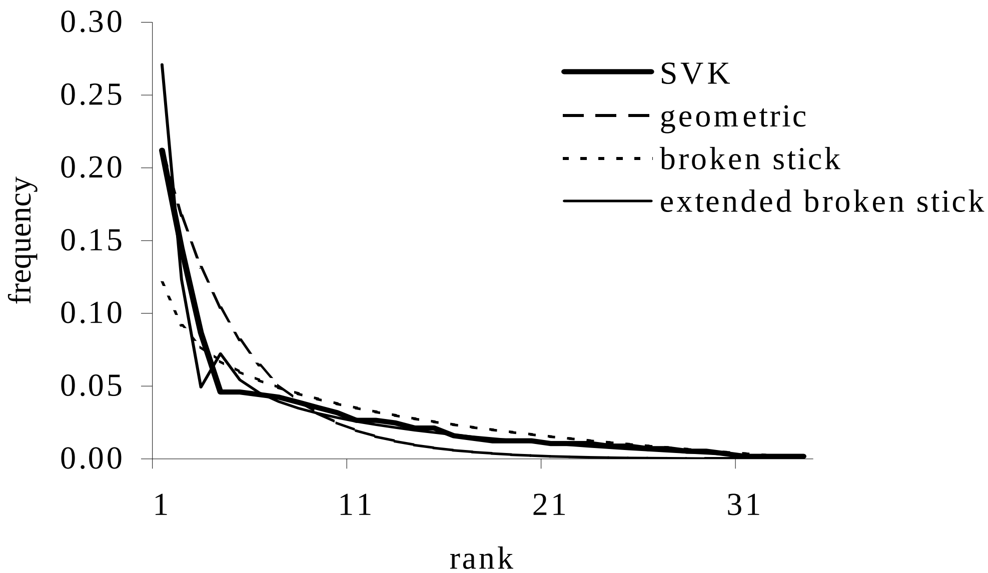
Another conceptual model of species distribution within an ecosystem is the *nonoverlapping niche* or *broken stick* model (MacArthur, 1957). In this model the environment is a limited resource shared as each species simultaneously appropriates a portion. The limited resource is thought of as a stick that is randomly broken into *s* disjoint pieces, which are then ordered from largest to smallest. The abundance of the species is proportional to the length of the pieces. The expected size of the $i^{\text{th}}$ largest piece is $E(i) = \frac{1}{s}\sum_{x=i}^{s}\frac{1}{x}$. The rank-abundance distributions of bird species in tropical forests and many temperate regions agree with this model.

It is sometimes the case that in a rank-abundance distribution, the common species are commoner and the rare species are rarer than the broken stick model predicts. This is, in fact, true of the handshape distributions, as well. To accommodate these two slopes, the broken stick model is modified so that the stick is first broken non-randomly into two pieces, and then the two pieces are separately broken into pieces of random lengths, so that the total number of pieces is *s* as required. This new model is called the *extended broken stick* model, and it is not as unmotivated it appears. Sometimes the set of species populating a certain ecosystem is heterogeneous and can be divided into two groups; there is competition within the two groups, but not between them. As can be seen in (6-14) for the best exemplar of this phenomenon, SVK, there is an inflection point at which the slope changes. The graph also shows the SVK handshape rank-frequency distribution modeled by the geometric, broken stick and extended broken stick models.

---

[96] Thanks to Sharon Hargus for suggesting that the ranking is in accord with increasing complexity in the phonological representation of handshape, and to Setsuko Shirai for suggesting the possibility that the geometric series arises from a geometrically increasing number of features needed for handshape representation.

(6-14) SVK: handshape rank-frequency graph with three models




It might be the case that the position of the inflection point defines a small, special set of handshapes that compete amongst themselves for approximately 50% of the lexicon, while the remaining handshapes are used in the rest of the lexicon. Perhaps this method produces a language-particular criterion for the set of "unmarked" handshapes. It would be interesting to test whether the set of handshapes below the inflection point have special status in the language, perhaps by using this subdivision of the set of handshapes to test for dependence between handshape and type or location, as in sections 3.3 and 5.2.

### 6.2.4 Duet rank-frequency distributions and complexity

The mathematical and conceptual models of the previous section apply to the handshape rank-frequency distributions. Is there a conceptual model that applies to the duet distributions? As quoted in Chapter 1, Miller, in his 1965 introduction to Zipf (1935:vi) observes that, "Faced with this massive statistical regularity, you have two alternatives. Either you can assume that it reflects some universal property of human mind, or you can assume that it represents some necessary consequence of the laws of probabilities." Li (1992) shows that symbols generated independently with equal probabilities (including the blank space symbol) yield a distribution quite similar to Zipf's Law for word frequencies, and implies that Zipf's law is a purely mathematical model requiring no conceptual model. However, Günther et al. (1996) shows that an underlying assumption to all statistical explanations, such as Li's, is that the symbols are independent and do not interact, and asserts that the fact that simple, structureless systems exhibit Zipf's Law does not preclude Zipf's Law from reflecting characteristics of complex, interacting systems.

This is a third possibility that had not yet been developed in 1965 when Miller wrote: Zipf's Law, and other such power laws, are a characteristic of complex systems. Bak et al. (1988) goes so far as to assert that power law behavior is the "fingerprint" of self-organized, complex systems. A "system" is an arrangement of related items that show a plan. A "complex" system has many independent but interacting items. A "self-organizing" complex system organizes the whole of itself spontaneously, through local interactions of the items (Waldrop, 1992).

I propose that the sign language lexicon is a structured system of dependent, interacting elements, and not a collection of randomly assembled elements. In particular, duets interact with each other in order to maximize phonological distance between themselves. This is essentially the claim of Siple (1973; 1978; 1980). Battison (1995) stated , "Siple also proposed that in the areas in the outer reaches of sign space, in areas of low visual acuity, … there [should] be signs with simpler handshapes (i.e., more unmarked handshapes)." However, the actual claim of Siple (1980:325) is different:

> "We would expect to find pairs of signs made in the areas about the face or upper chest to be visually more similar, i.e., *to differ in a less detailed way* [emphasis added], than signs made in the areas of lower acuity. … The set of signs made in the regions of lower acuity should not contain such pairs [as the ASL minimal pair RED and SWEET, both articulated on the lips with initial handshapes 49 and 59 [97]]."

Notice that this claim does not necessarily concern markedness. Instead it is about the interaction of duets; the existence of one duet precludes the existence of another duet if the two duets are not sufficiently different in form.

I suggest that phonological distance is operational in the structure of sign language lexicons. Phonological distance can be expressed as a suitably defined metric function that maps a pair of signs to a nonnegative numerical value; the smaller the value, the more similar the signs are. The fundamental constraint is that the phonological distance between a pair of signs must be sufficiently large. In support of this hypothesis is the much reported dearth of minimal pairs in

---

[97] This is the example given by Siple (1980:325).

ASL (Brentari, 1998:4).[98] Following Siple, when the phonological distance between a pair of signs is less than the minimum allowed, the signs are articulated on the face or neck. In addition, I propose that when the phonological distance between a pair of signs not articulated on the face or neck is less than the minimum allowed, the signs must be semantically related. Consider the SVK minimal pair NIMI *name* and OSOITE *address*, both Type 3 signs with the nondominant 3 hand as location; NIMI has handshape 49 and OSOITE has handshape 59. SVK signers say the sign is derived from the writing on, for example, an envelope, where the first line is the name and the following lines are the address.

In contrast to the well-populated core locations of the face and neck, contrasting points in neutral space are sparsely used. A minimal pair distinguished only by separate locations in neutral space must comprise two signs that are morphologically related, for example, by the use of space for personal pronoun reference or for verbs of location. Likewise, movement is a sparsely used phonological distinction. A minimal pair distinguished only by different movements must comprise signs that are morphologically related, for example, by aspectual or distributional marking.[99]

The phonological similarity metric proposed here, together with restrictions on the minimum allowable distance between signs could provide the key for explaining the self-organizing, complex structure of the sign language lexicon.

### 6.3 Conclusion

This dissertation has taken a step toward the ultimate goal of predicting possible forms of naturally occurring sign language lexicons. The handshape and location inventories of four sign languages were established in a uniform manner that allows cross-linguistic comparison. Both the size and content of the handshape and location inventories vary. In contrast, the contents of the duet inventories vary, but their sizes do not. The rank-frequency distributions of these

---

[98] I have not been able to verify that the sign languages in my sample have fewer minimal pairs than spoken languages do, nor do I know how to quantitatively verify this observation.

[99] The morphological use of neutral space locations and movement types that are not available for phonological use is in contradiction to a tenet of lexical phonology that requires morphology to use only contrastive resources; hence, only phonetic material available at the phonological level can be used in the morphological level.

languages also show uniformity and diversity.  While the different types of distributions vary, there is remarkable cross-linguistic uniformity in the form of the handshape distributions, location distributions and duet distributions, showing that sign languages use these phonetic resources in similar ways.  I hypothesize that the duet power law distribution is evidence that the sign language lexicon is a self-organizing complex system.  I propose that a metric of phonological similarity together with lower bounds on the value this metric is permitted to take might be the key to this self-organization.  In order to analyze dependence between sign parameters, I adapted a method from information theory and applied it to the parameters of handshape and type, handshape and location, and location and number of hands.  Results vary greatly across the four languages, distinctions highly relevant in one language, such as contact versus neutral space, can be completely irrelevant in another language.  In particular, neither of the markedness criteria that were used were especially useful in accounting for dependencies between sign parameters.

# Bibliography


Adams, C. M. 1969. A survey of Australian English intonation. *Phonetica* 20:81-130.

Ann, Jean. 1991. Constraining sign language handshapes: Towards a phonetically grounded account of handshapes in Taiwan Sign Language and American Sign Language. Paper presented at *WECOL*, Simon Fraser University, Vancouver, BC.

Ann, Jean. 1993. *A linguistic investigation of the relationship between physiology and handshape*. University of Arizona: PhD dissertation.

Ann, Jean. 1998. Contact between a Sign Language and a Written Language: Character Signs in Taiwan Sign Language. In *Pinky Extension and Eye Gaze: Language Use in Deaf Communities*, ed. Ceil Lucas, 59-99. Washington, D. C.: Gallaudet University Press.

Baayen, R. Harald. 2001. *Word Frequency Distributions*. Dordrecht: Kluwer Academic Publishers.

Bak, Per, Chao Tang, and Kurt Wiesenfeld. 1988. Self-organized criticality. *Physical Review A* 38:364-375.

Bak, Per. 1996. *How nature works: the science of self-organized criticality*. New York: Copernicus Springer-Verlag.

Battison, Robbin. 1974. Phonological deletion in American Sign Language. *Sign Language Studies* 5:1-19.

Battison, Robbin. 1978. *Lexical borrowing in American Sign Language*. Silver Spring, MD: Linstok Press.

Battison, Robbin. 1995. Analyzing Signs. In *Linguistics of American Sign Language*, eds. Clayton Valli and Ceil Lucas. Washington D.C.: Gallaudet University Press.

Baucom, K.L. 1974. Proto-Central Khoisan. In *Third Annual Conference on African Linguistics*, ed. E. Voeltz, 3-38. Bloomington: Indiana University Publications.

Bhat, D. N. S. 1967. *Descriptive analysis of Tulu*. Building Centenary and Silver Jubilee Series, 15. Poona-Deccan College Postgraduate and Research Institute.

Bloch, B. 1950. Studies in colloquial Japanese, IV: Phonemics. *Language* 26:86-125.

Bloomfield, Leonard. 1933. *Language*. New York: Holt, Rinehart and Winston.

Bonvillian, John D., and Theodore Jr. Siedlecki. 1996. Young children's acquisition of the location aspect of American Sign Language signs: Parental report findings. *Journal of Communication Disorders* 29:13-35.

Boyes Braem, Penny. 1973. A study of the acquisition of dez in American Sign Language. Ms. La Jolla, CA: Salk Institute.

Boyes Braem, Penny. 1990. Acquisition of the handshape in American sign language: a preliminary analysis. In *From gesture to language in hearing and deaf children*, eds. Virginia Volterra and Carol J. Erting, 107-127. New York: Springer.

Brentari, Diane. 1990a. Licensing in ASL handshape change. In *Sign language research. Theoretical issues*, ed. Ceil Lucas, 57-70. Washington, D. C.: Gallaudet University Press.

Brentari, Diane. 1990b. *Theoretical foundations of American Sign Language phonology*. University of Chicago: PhD dissertation.

Brentari, Diane. 1992. Phonological representation in American Sign Language. *Language* 68:359-374.

Brentari, Diane. 1993. Establishing a sonority hierarchy in American Sign Language: The use of simultaneous structure in phonology. *Phonology* 10:281-306.

Brentari, Diane, and Howard Poizner. 1994. A phonological analysis of a deaf Parkinsonian signer. *Language and Cognitive Processes* 9:69-99.

Brentari, Diane, Howard Poizner, and Judy Kegl. 1995. Aphasic and Parkinsonian Signing: Differences in Phonological Disruption. *Brain and Language* 48:69-105.

Brentari, Diane. 1998. *A prosodic model of sign language phonology*. Cambridge, MA: MIT Press.

Brentari, Diane, ed. 2001. *Foreign Vocabulary in Sign Language*. Mahwah, NJ: Lawrence Erlbaum Associates, Inc.

Brentari, Diane, and Carol A. Padden. 2001. Native and Foreign Vocabulary in American Sign Language. In *Foreign Vocabulary in Sign Languages*, ed. Diane Brentari, 87-119. Mahwah: Lawrence Erlbaum Associates, Inc.

Büssenschütt, Martin. 1997. *Pattern and Performance: On the Modeling of Self-Organized Behavior in Ecosystems*. Swiss Federal Institute of Technology: PhD dissertation.

Butterworth, G. 1992. Origins of self-perception in infancy. *Psychological Inquiry* 3:103-111.

Channon, Rachel. 2002. Beads on a string? Representations of repetition in spoken and signed languages. In *Modality and Structure in Signed and Spoken Languages*, eds. Richard P. Meier, Kearsy Cormier and David Quinto-Pozos. Cambridge: Cambridge University Press.

Chinchor, Nancy. 1978. The syllable in ASL: Simultaneous and sequential phonology. Ms. Brown University.

Cline, Melissa. 2000. *Protein Sequence Alignment Reliability: Prediction and Measurement*. University of California at Santa Cruz: PhD dissertation.

Collins-Ahlgren, Marianne. 1990. Word Formation in New Zealand Sign Language. In *Theoretical Issues in Sign Language Research, Vol. 1: Linguistics*, eds. Susan Fischer and Patricia Siple. Chicago: University of Chicago Press.

Conlin, Kimberly E., Gene R. Mirus, Claude Mauk, and Richard P. Meier. 2000. The Acquisition of First Signs: Place, Handshape, and Movement. In *Language Acquisition by Eye*, eds. Charlene Chamberlain, Jill P. Morford and Rachel I. Mayberry. Mahwah, NJ: Lawrence Erlbaum Associates.

Corina, David. 1993. To branch or not to branch: Underspecification in ASL handshape contours. In *Phonetics and phonology, Vol. 3, Current issues in ASL phonology*, ed. G. Coulter. San Diego: Academic Press.

Corina, David P., and Elizabeth Sagey. 1989. Predictability in ASL handshapes and handshape sequences, with implications for features and feature geometry. Ms. La Jolla, CA: Salk Institute.

Corina, David P. 1990. Handshape assimilations in hierarchical phonological representation. In *Sign language research: Theoretical issues*, ed. Ceil Lucas, 27-49.

Corina, David P., and Karen Emmorey. 1990. Lexical recognition in sign language: Effects of phonetic structure and morphology. *Perceptual and Motor Skills* 71:1227-1252.

Corina, David P. 1996. ASL syllables and prosodic constraints. *Lingua* 98:73-102.

Corina, David P. 1998. Aphasia in users of sign language. In *Aphasia in atypical populations*, eds. Patrick Coppens, Yvan Lebrum and Anna Baddo, 261-309. Hillsdale, NJ: Lawrence Erlbaum Associates.

Corina, David P., Susan L. McBurney, Carl Dodrill, Kevin Hinshaw, Jim Brinkley, and George Ojemann. 1999. Functional roles of Broca's area and SMG: Evidence from cortical stimulation mapping in a deaf signer. *NeuroImage* 10:570-581.

Corina, David P. 2000a. Some Observations Regarding Paraphasia in American Sign Language. In *The Signs of Language Revisited: An Anthology to Honor Ursula Bellugi and Edward Klima*, eds. Karen Emmorey and Harlan Lane, 493-508. Mahwah, NJ: Lawrence Erlbaum Associates.

Corina, David P. 2000b. Is ASL phonology psychologically real? Paper presented at *Texas Linguistic Society Meeting: The Effects of Modality on Language and Linguistic Theory*, Austin.
Corina, David P., and Ursula C. Hildebrandt. 2002. Psycholinguistic Investigations of Phonological Structure in American Sign Language. In *Modality and Structure in Signed and Spoken Languages*, eds. Richard P. Meier, David Quinto-Pozos and Kearsy Cormier. Cambridge: Cambridge University Press.
Crasborn, Onno. 1998. SignPhon Software (v. 2.20). http://www.leidenuniv.nl/hil/sign-lang/signphon2.html. Leiden University.
Crasborn, Onno, Harry van der Hulst, and Els van der Kooij. 1998. SignPhon 2 manual. Ms. Leiden University.
Crothers, John. 1978. Typology and Universals of Vowel Systems. In *Universals of Human Language*, ed. Joseph Greenberg, 93-152. Stanford: Stanford University Press.
Davis, J. 1989. Distinguishing language contact phenomena in ASL interpretation. In *The sociolinguistics of the Deaf community*, ed. Ceil Lucas, 85-102. New York: Academic Press.
Davis, S. 1994. Loanwords: Phonological treatment. In *The encyclopedia of language and linguistics*, ed. R. E. Asher, 2273-2276. New York and Oxford: Pergamon Press.
Denes, P. B. 1963. On the statistics of spoken English. *Journal of the Acoustical Society of America* 35:892-904.
Edmondson, William. 1987. Segments in signed languages: Do they exist and does it matter? In *Proceedings of the Fourth International Symposium on Sign Language Research, Lappeenranta, Finland*, eds. W.H. Edmondson and F. Karlsson, 66-74: Signum Press, Hamburg.
Emmorey, Karen, and David P. Corina. 1990. Lexical recognition in sign language: Effects of phonetic structure and morphology. *Perceptual and Motor Skills* 71:1227-1252.
Emmorey, Karen, and Harlan Lane, eds. 2000. *The Signs of Language Revisited*. Mahwah, NJ: Lawrence Erlbaum Associates.
Emmorey, Karen. 2002. *Language, Cognition and the Brain: Insights from Sign Language Research*. Mahwah, NJ: Lawrence Erlbaum Associates.
Friedman, Lynn A. 1976. *Phonology in a Soundless Language: Phonological Structure of the American Sign Language*. U.C. Berkeley: PhD dissertation.
Friedman, Lynn A. 1977. Formational Properties of American Sign Language. In *On the Other Hand: New Perspectives on American Sign Language*, ed. L. Friedman. New York: Academic Press.
Frisch, Stephan A., Michael Broe, and Janet B. Pierrehumbert. 1997. Similarity and phonotactics in Arabic. Ms., *ROA-223 Rutgers Optimality Archive, http://ruccs.rutgers.edu/roa.html*. Bloomington, IN and Evanston, IL: Indiana University and Northwestern University.
Frisch, Stephan A., and Bushra Adnan Zawaydeh. 2001. The Psychological Reality of OCP-Place in Arabic. *Language* 77.
Frishberg, Nancy. 1975. Arbitrariness and iconicity: historical change in American Sign Language. *Language* 51:696-719.
Goldsmith, John A. 1976. *Autosegmental phonology*. University of Southern California: PhD dissertation.
Goldsmith, John A. 1998. On information theory, entropy, and phonology in the 20th century. Paper presented at *Royaumont CTIP II Round Table on Phonology in the 20th Century*.
Grace, G.W. 1959. *The position of the Polynesian languages within the Austronesian (Malayo-Polynesian) language family*. vol. 11: IJAL Memoir. Bloomington: Indiana University.
Greenberg, Joseph. 1957. *Essays in linguistics*. Chicago: University of Chicago Press.

Grimes, Barbara, et al., ed. 1996. *Ethnologue [electronic source]: Languages of the World*. Dallas, TX: SIL. http://www.ethnologue.com.

Groce, Nora Ellen. 1985. *Everyone Here Spoke Sign Language: Hereditary Deafness on Martha's Vineyard*. Cambridge, MA: Harvard University Press.

Grosjean, François. 1981. Sign and word recognition: A first comparison. *Sign Language Studies* 32:195-219.

Guerra Currie, Anne-Marie, Richard P. Meier, and Keith Walters. 2002. A crosslinguistic examination of the lexicons of four signed languages. In *Modality and Structure in Signed and Spoken Languages*, eds. Richard P. Meier, Kearsy Cormier and David Quinto-Pozos. Cambridge: Cambridge University Press.

Günther, R., L. Levitin, B. Schapiro, and P. Wagner. 1996. Zipf's Law and the Effect of Ranking on Probability Distributions. *International Journal of Theoretical Physics* 35:395-417.

Hamilton, Harley. 1986. Perception of sign features by deaf children. *Sign Language Studies* 50:73-77.

Hargus, Sharon. In preparation. *Dictionary of Kwadacha (Fort Ware Sekani)*.

Hayden, Rebecca. 1950. The Relative Frequency of Phonemes in General-American English. *Word* VI:217-223.

Hayes, Bruce. 1993. Against Movement. In *Phonetics and phonology: Current Issues in ASL Phonology, Vol. 3*, ed. Geoffrey R. Coulter, 213-226. New York: Academic Press.

Herdan, Gustav. 1960. *Type-Token Mathematics*. The Hague: Mouton.

Herdan, Gustav. 1966. *The Advanced Theory of Language as Choice and Chance*. New York: Springer-Verlag.

Hirn, D. F. 1910. *Kuuromykkäin Viittomakieli*. Helsinki: Osakeyhtiö Weilin and Göös Aktiebolag.

Horgan, John. 1995. From Complexity to Perplexity. *Scientific American* 272:104-109.

Householder, Fred, Kostas Kazazis, and Andreas Koutsoudas. 1964. Reference Grammar of Literary Dhimotiki. *International Journal of American Linguistics* 30.

Hulst, Harry van der. 1985. Acquisitional evidence for the phonological composition of handshapes. Ms. Holland Institute of Generative Linguistics, Leiden University.

Hulst, Harry van der. 1993. Units in the analysis of signs. *Phonology* 10:209-241.

Hulst, Harry van der. 1995. Dependency relations in the phonological representation of signs. In *Sign Language Research 1994: Proceedings of the Fourth European Congress on Sign Language Research, Munich, September 1-3, 1994*, eds. Heleen F. Bos and Gertrude M. Schermer, 11-38. Hamburg: Signum.

Hulst, Harry van der. 1996a. Introduction. Issues in sign linguistics: Phonetics, phonology and morpho-syntax. *Lingua. Special issue on "Sign linguistics: Phonetics, phonology and morpho-syntax"* 98:3-18.

Hulst, Harry van der. 1996b. On the other hand. *Lingua* 98:121-143.

Jakobson, Roman, Gunnar Fant, and Morris Halle. 1952. *Preliminaries to speech analysis: the distinctive features and their correlates*. MIT Acoustics Laboratory. Cambridge: MIT Press.

Kantor, R. 1980. The acquisition of classifiers in American Sign Language. *Sign Language Studies* 28:193-208.

Karplus, Kevin. 1995. Regularizers for estimating distributions of amino acids from small samples. In *Intelligent Systems for Molecular Biology-95*. Cambridge, England.

Kennedy, Graeme, et al. 1999. *A dictionary of New Zealand Sign Language*. Auckland, New Zealand: Auckland University Press.

Kim, Jong-Sung, Won Jang, and Zeungnam Bien. 1996. A Dynamic Gesture Recognition System for the Korean Sign Language (KSL). *IEEE Transactions on Systems, Man, and Cybernetics -- Part B: Cybernetics* 26:354-359.
Kirchner, Robert. 1998. *An Effort-Based Approach to Consonant Lenition*. University of California, Los Angeles: PhD dissertation.
Klima, Edward S., and Ursula Bellugi, eds. 1979. *The Signs of Language*. Cambridge, MA: Harvard University Press.
Kooij, Els van der. 1997. *Contact: A Phonological or a Phonetic Feature of Signs?* vol. 14: Linguistics in the Netherlands 1997. Amsterdam: John Benjamins Publishing Company.
Krámský, J. 1976a. A Quantitative Typology of Language. In *Papers in General Linguistics*. The Hague: Mouton.
Krámský, J. 1976b. The Frequency of Occurrence of Vowel Phonemes in Languages Possessing Vowel Systems of Identical Structure. In *Papers in General Linguistics*. The Hague: Mouton.
Kucera, Henry, and George K. Monroe. 1968. *A comparative quantitative phonology of Russian, Czech and German*. New York: American Elsevier Publishing Company.
Kuurojen Liitto. 1973. *Viittomakielen kuvasanakirja*. Helsinki, Finland: Kuurojen Liitto r.y.
Kuurojen Liitto. 1998. *Suomalaisen viittomakielen perussanakirja*. Helsinki, Finland: Kuurojen Liitto.
Kyle, James G., and Bencie Woll. 1985. *Sign language: The study of deaf people and their language*. Cambridge: Cambridge University Press.
Lane, Harlan, Penny Boyes Braem, and Ursula Bellugi. 1980. A featural analysis of handshapes. In *The Signs of Language,* eds. Edward S. Klima and Ursula Bellugi, 164-180. Cambridge, MA: Harvard University Press.
Lane, Harlan. 1984. *When the Mind Hears*. New York: Random House, Inc.
Lee, Chan-Su, Gyu-tae Park, Jong-Sung Kim, Zeungnam Bien, Won Jang, and Sung-Kwon Kim. 1997. Real-time Recognition System of Korean Sign Language based on Elementary Components. In *IEEE International Conference on Fuzzy Systems Proceedings*, 1463-1468. Piscataway, NJ: IEEE.
Li, W. 1992. Random texts exhibit Zipf's-law-like word frequency distributions. *IEEE Transactions on Information Theory* 38:1842.
Liddell, Scott K. 1980. *American Sign Language Syntax*. vol. 52: Approaches to Semiotics. The Hague: Mouton.
Liddell, Scott K. 1983. Compound formation rules in American Sign Language. In *Proceedings of the III International Symposium on Sign Language Research*, eds. William C. Stokoe and Virginia Volterra. Silver Spring, MD: Linstok Press.
Liddell, Scott K. 1984a. THINK and BELIEVE: Sequentiality in ASL. *Language* 60:372-399.
Liddell, Scott K. 1984b. Unrealized-Inceptive Aspect in American Sign Language: Feature Insertion in Syllabic Frames. In *Papers from the Twentieth Regional Meeting, Chicago Linguistic Society*. Chicago: Chicago Linguistic Society, University of Chicago.
Liddell, Scott K., and Robert E. Johnson. 1985. American Sign Language compounds: Implications for the structure of the lexicon. In *Georgetown University Round Table on Languages and Linguistics*. Washington, DC: Georgetown University.
Liddell, Scott K., and Robert E. Johnson. 1986. American Sign Language Compound Formation Processes, Lexicalization, and Phonological Remnants. *Natural Language and Linguistic Theory* 4:445-513.
Liddell, Scott K., and Robert E. Johnson. 1989. American Sign Language: The phonological base. *Sign Language Studies* 64:197-277.

Liddell, Scott K. 1990. Structures for representing handshape and local movement at the phonemic level. In *Theoretical issues in Sign Language Research. Vol. 1: Linguistics*, eds. Susan D. Fischer and Patricia Siple, 37-66. Chicago, London: University of Chicago Press.

Liddell, Scott K. 1995. Real, Surrogate, and Token Space: Grammatical Consequences in ASL. In *Language, Gesture, and Space*, eds. Karen Emmorey and Judy Reilly, 19-42. Hillsdale, NJ: Lawrence Erlbaum Associates.

Liljencrants, J., and B. Lindblom. 1972. Numerical simulation of vowel quality contrasts: the role of perceptual contrast. *Language* 48:839-862.

Long, J. Schuyler. 1918. *The Sign Language: A Manual of Signs*. Washington DC: Gallaudet College.

MacArthur, R. H. 1957. On the relative abundance of bird species. *Proceedings of the National Academy of Sciences of the United States of America* 43:293-295.

Maddieson, Ian. 1984. *Patterns of sounds*. Cambridge Studies in speech science and communication. Cambridge: Cambridge University Press.

Maddieson, Ian, and Karen Precoda. 1989. Updating UPSID. *UCLA Working Papers in Phonetics* 74:104-111.

Mandel, Mark A. 1981. *Phonotactics and morphophonology in American Sign Language*. University of California, Berkeley: PhD dissertation.

Mandel, Mark A. 1982. Major information from a "minor parameter": Point of Contact in sign language phonology. Paper presented at *Berkeley Linguistic Society*, University of California, Berkeley.

Mandelbrot, Benoit B. 1953. *The Fractal Geometry of Nature*. New York: Freeman.

Mandelbrot, Benoit B. 1977. *Fractals: Form, Chance, and Dimension*. San Francisco: W. H. Freeman and Company.

Manning, Christopher D., and Hinrich Schütze. 1999. *Foundations of Statistical Natural Language Processing*. Cambridge, MA: The MIT Press.

Marentette, Paula F., and Rachel Mayberry, I. 2000. Principles for an Emerging Phonological System: A Case Study of Early ASL Acquisition. In *Language Acquisition by Eye*, eds. Charlene Chamberlain, Jill P. Morford and Rachel I. Mayberry. Mahwah, NJ: Lawrence Erlbaum Associates.

May, Robert M. 1975. Patterns of Species Abundance and Diversity. In *Ecology and Evolution of Communities*, eds. Martin L. Cody and Jared M. Diamond. Cambridge, MA: Belknap Press of Harvard University.

McCarthy, John J. 1994. The phonetics and phonology of Semitic pharyngeals. In *Papers in laboratory phonology 3: Phonological structure and phonetic form*, ed. Patricia Keating, 191-283. Cambridge: Cambridge University Press.

McCarthy, John J., and Alan Prince. 1993. Prosodic Morphology I: Constraint Interaction and Satisfaction. Ms. University of Massachusetts, Amherst, and Rutgers University, New Brunswick, NJ.

McCarthy, John J. 1999. Distinctive Features. In *The MIT Encyclopedia of the Cognitive Sciences*, eds. Robert A. Wilson and Frank C. Keil. Cambridge, MA: MIT Press.

McIntire, Marina L. 1977. The acquisition of American Sign Language hand configurations. *Sign Language Studies* 6:247-266.

McIntire, Marina L. 1980. Acquisition of ASL hand configurations. In *Sign and culture. A reader for students of American Sign Language*, ed. William C. Stokoe, 186-205. Silver Spring, MD: Linstok Press.

McKee, David, and Graeme Kennedy. 2000. Lexical comparison of signs from American, Australian, British and New Zealand Sign Languages. In *The signs of language revisited:*

*An anthology to honor Ursula Bellugi and Edward Klima*, eds. Karen Emmorey and Harlan Lane, 49-76. Mahwah, NJ: Lawrence Erlbaum Associates.
Miyamoto, Ichiro. 2000. JSL (Japanese Sign Language) and KSL (Korean Sign Language). Paper presented at *Seventh International Conference on Theoretical Issues in Sign Language Research*, Amsterdam.
Neisser, U. 1991. Two perceptually given aspects of the self and their development. *Developmental Review* 11:197-209.
Nettle, Daniel. 1995. Segmental inventory size, word length, and communicative efficiency. *Linguistics* 33:359-367.
Pareto, V. 1897. *Cour d'Economie Politique*. Lausanne and Paris [reprinted in Oevre Completes, Genf Droz].
Perlmutter, David. 1988. A moraic theory of American Sign Language syllable structure. Paper presented at *the second conference on Theoretical Issues in Sign Language Research*, Gallaudet University.
Perlmutter, David. 1991. Feature geometry in a language with two active articulators. Paper presented at *Conference on Phonological Feature Organization*, University of California, Santa Cruz.
Perlmutter, David. 1992. Sonority and Syllable Structure in American Sign Language. *Linguistic Inquiry* 23:407-442.
Perlmutter, David. 1993. Sonority and syllable structure in American Sign Language. In *Phonetics and Phonology: Current Issues in ASL Phonology*, *Vol. 3*, ed. G. Coulter, 227-259. New York: Academic Press.
Pielou, E. C. 1975. *Ecological Diversity*. New York: John Wiley and Sons.
Pietrandrea, Paola. 1998. Sublexical Regularities and Linguistic Economy in Italian Sign Language: A Quantitative Analysis. *Proceedings of the First Annual High Desert Linguistics Society Conference* 1:63-80.
Pietrandrea, Paola. 2000. Economy and Redundancy in Italian Sign Language Lexicon. Paper presented at *Seventh International Conference on Theoretical Issues in Sign Language Research*, Amsterdam.
Poizner, Howard, and Harlan Lane. 1978. Discrimination of location in American Sign Language. In *Understanding language through sign language research*, ed. Patricia Siple, 271-287. New York: Academic Press.
Poizner, Howard, Edward S. Klima, and Ursula Bellugi. 1987. *What the hands reveal about the brain*. Cambridge, MA: MIT Press.
Preston, F. W. 1962a. The Canonical Distribution of Commonness and Rarity: Part I. *Ecology* 43:185-215.
Preston, F. W. 1962b. The Canonical Distribution of Commonness and Rarity: Part II. *Ecology* 43:410-432.
Prillwitz, Siegmund et al. 1987. HamNoSys. Version 2.0; Hamburg Notation System for Sign Language. An Introductory guide. In *International Studies on Sign Language and Communication of the Deaf, 5*. Hamburg: Signum.
Prince, Alan, and Paul Smolensky. 1993. *Optimality Theory: Constraint Interaction and Generative Grammar*. ROA 537-0802, Rutgers Optimality Archive, http://roa.rutgers.edu.
Rissanen, Terhi. 1987. Finnish Sign Language. In *Gallaudet encyclopedia of deaf people and deafness*, ed. John V. Van Cleve. New York: McGraw-Hill.
Roberts, A. Hood. 1965. *A statistical linguistic analysis of American English*. The Hague: Mouton.

Rozelle, Lorna. 1992. A partial feature geometry for American Sign Language. Paper presented at *Conference on Theoretical Issues in Sign Language Research*, University of California, San Diego.

Rozelle, Lorna. 1996a. Assimilation in compounds in Finnish Sign Language. Paper presented at *Theoretical Issues in Sign Language Research*, McGill University, Montreal.

Rozelle, Lorna. 1996b. Optimality and the Hand Number feature in ASL and SVK. Paper presented at *North West Linguistics Conference*, University of Washington.

Rozelle, Lorna. 1998. Two Hands are Better than One: The Representation of the Non-Dominant Hand in American Sign Language and Finnish Sign Language. Ms. University of Washington.

Sandler, Wendy. 1986. The spreading hand autosegment of ASL. *Sign Language Studies* 15:1-28.

Sandler, Wendy. 1987. *Sequentiality and simultaneity in American Sign Language phonology*. University of Texas: PhD dissertation.

Sandler, Wendy. 1989. *Phonological Representation of the Sign: Linearity and Nonlinearity in American Sign Language*. Dordrecht: Foris Publications.

Sandler, Wendy. 1993a. Hand in hand: The role of the nondominant hand in ASL. *The Linguistic Review* 10:337-390.

Sandler, Wendy. 1993b. A sonority cycle in American Sign Language. *Phonology* 10:243-279.

Sandler, Wendy. 1996. Representing handshapes. In *International Review of Sign Linguistics*, eds. William H. Edmondson and Ronnie B. Wilbur, 115-158. Hillsdale, N.J: Erlbaum.

Sapir, Edward. 1921. *Language: An Introduction to the Study of Speech*. New York: Harcourt Brace Jovanovich.

Sasaki, Daisuke. 2001. Comparing the lexicon of sign languages in East Asia: The influence of Japanese Sign Language over Taiwan Sign Language and Korean Sign Language. Paper presented at *Conference on Sign Linguistics Deaf Education and Deaf Culture in Asia*.

Seiden, William. 1960. Chamorro phonemes. *Anthropological Linguistics* 2:6-35.

Shankweiler, D., and K. Harris. 1966. An experimental approach to the problem of articulation in aphasia. *Cortex* 2:277-292.

Shannon, Claude E., and W. W. Weaver. 1949. *The mathematical theory of communication*. Urbana, IL: University of Illinois Press.

Siedlecki, Theodore, and John D. Bonvillian. 1993. Location, handshape & movement: Young children's acquisition of the formational aspects of American Sign Language. *Sign Language Studies* 22:31-51.

Siedlecki, Theodore, and John D. Bonvillian. 1997. Young children's acquisition of the handshape aspect of American Sign Language signs: Parental report findings. *Applied Psycholinguistics* 18:17-39.

Siple, Patricia. 1973. Visual Constraints for Sign Language Communication. Ms. La Jolla, CA: Salk Institute.

Siple, Patricia. 1978. Visual Constraints for Sign Language Communication. *Sign Language Studies* 19:97-112 (Working paper 1973, Salk Institute).

Siple, Patricia. 1980. Visual Constraints for Sign Language Communication. In *Sign and culture. A reader for students of American Sign Language*, ed. William C. Stokoe, 319-333. Silver Spring, MD: Linstok Press.

Snedecor, George W., and William G. Cochran. 1982. *Statistical Methods*. Ames: Iowa State University Press.

Son, Won-Jae. 1988. *Korean Sign Language For the Guide*. Seoul: Chang-Rok Press.

Stack, Kelly. 1988. *Tiers and syllable structure: Evidence from phonotactics*. University of California, Los Angeles: MA thesis.

Stokoe, William C. 1960. *Sign Language Structure: An Outline of the Visual Communication Systems of the American Deaf.* vol. 8: Studies in Linguistics, Occasional Papers. Silver Spring, MD: Linstock Press.
Stokoe, William C., D. C. Casterine, and C. G. Croneberg. 1965. *A dictionary of American Sign Language on linguistic principles*. Washington, DC: Gallaudet College Press.
Stungis, Jim. 1981. Identification and discrimination of handshape in American Sign Language. *Perception and Psychophysics* 29:261-276.
Sutton, Valerie. 2003. Signwriting. http://www.signwriting.org.
Swadesh, M. 1955. Toward greater accuracy in lexicostatistic dating. *International Journal of American Linguistics* 21:121-137.
Takkinen, Ritva. 1990. Sign articulation in the age of 2.3-3-4 years. A case study. In *Studies in Logopedics and Phonetics*, eds. M. Leiwo and R. Aulanko, 233-244: Department of Phonetics, University of Helsinki.
Tennant, Richard A., and Marianne Gluszak Brown. 1998. *The American Sign Language handshape dictionary*. Washington, DC: Gallaudet University Press.
Tervoort, B. T. M. 1953. *Structurele analyse van visueel taalgebruik binnen een groep dove kinderen*. Amsterdam: Noord-Hollandsche Uitgevers Maatschappij: PhD dissertation. (English summary available at http://www.leidenuniv.nl/hil/sign-lang/tervoort1953.html).
Têšitelová, Marie. 1992. *Quantitative Linguistics*. Amsterdam: John Benjamins.
Traill, A. 1978. Research on the Non-Bantu African languages. In *Language and Communication Studies in South Africa*, eds. L.W. Lanhan and K.P. Prinsloo, 117-137. Cape Town: Oxford University Press.
Trubetzkoy, N. S. 1967 [1939]. *Principes de Phonologie*. Paris: Klincksieck. [University of California Press: Berkeley. English translation *Principles of Phonology* by J. Cantineau published 1969].
Uyechi, Linda. 1994. *The Geometry of Visual Phonology*. Stanford University: PhD dissertation [Published 1996 by CSLI Publications, Stanford].
Waldrop, M. Mitchell. 1992. *Complexity*. New York: Touchstone.
Whittemore, Gregory Lynn. 1987. *The Production of ASL Signs*. University of Texas at Austin: PhD dissertation.
Wilbur, Ronnie B. 1987. *American Sign Language, Linguistic and Applied Dimensions*. Boston: Little, Brown.
Wilbur, Ronnie B. 1990. Why Syllables?  What the Notion Means for ASL Research. In *Theoretical Issues in Sign Language Research*, eds. Susan Fischer and Patricia Siple, 81-108. Chicago: University of Chicago Press.
Wilbur, Ronnie B. 1993. Syllables and Segments: Hold the Movement and Move the Holds! In *Phonetics and Phonology: Current Issues in ASL Phonology, Vol. 3*, ed. G. Coulter, 135-168. New York: Academic Press.
Woodward, James. 1978. Historical Bases of American Sign Language. In *Understanding language through sign language research*, ed. Patricia Siple. New York: Academic Press.
Woodward, James. 1982. Single finger extension: For a theory of naturalness in sign language phonology. *Sign Language Studies* 37:289-304.
Woodward, James. 1985. Universal constraints on two-finger extension across sign languages. *Sign Language Studies* 46:53-72.
Woodward, James. 1987. Universal constraints across sign languages: single finger contact handshapes. *Sign Language Studies* 16:375-385.
Woodward, James. 1991. Sign language varieties on Costa Rica. *Sign Language Studies* 73:329-346.

Woodward, James. 1993. The relationship of sign language varieties in India, Pakistan, and Nepal. *Sign Language Studies* 78:15-22.

Woodward, James. 1996. Modern Standard Thai Sign Language: Influence from American Sign Language and its relationship to original Thai Sign Language varieties. *Sign Language Studies* 92.

Woodward, James. 2000. Sign Languages and Sign Language Families in Thailand and Viet Nam. In *The signs of language revisited: An anthology to honor Ursula Bellugi and Edward Klima*, eds. Karen Emmorey and Harlan Lane, 23-47. Mahwah, NJ: Lawrence Erlbaum Associates.

Zipf, George Kingsley. 1935. *The Psycho-Biology of Language*. Cambridge: Cambridge University Press.

Zipf, George Kingsley. 1949. *Human Behavior and the Principle of Least Effort. An Introduction to Human Ecology*. Cambridge, MA: Addison-Wesley Press, Inc.

# Appendix A: HamNoSys handshape and location directory

## A.1 Handshape directory

In this section are photographs as well as the HamNoSys notation for all handshapes that have been mentioned. The organization of this section is shown in (A-1).

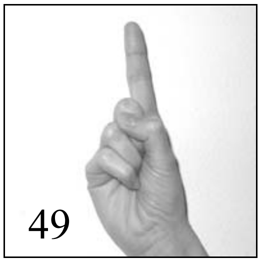
(A-1) Handshape directory organization

- A.1.1 No fingers selected
- A.1.2 One finger selected
  - A.1.2.1 Selected finger open
    - A.1.2.1.1 Index finger selected
    - A.1.2.1.2 Non-index finger selected
  - A.1.2.2 Selected finger closed
- A.1.3 Two fingers selected
  - A.1.3.1 Selected fingers open
  - A.1.3.2 Selected fingers closed
- A.1.4 Three fingers selected
- A.1.5 All fingers selected
  - A.1.5.1 Fingers open
    - A.1.5.1.1 Fingers touching
    - A.1.5.1.2 Fingers spread
  - A.1.5.2 Fingers closed

To find a handshape it is necessary to determine the number of selected fingers. The thumb is not included in this count. A selected finger (Mandel, 1981) is a finger that can contact the location or can change positions when a handshape change is involved in the articulation of a sign. Compare the handshapes in (A-2). The handshape in (A-2 a) has one selected finger, the index. This finger can flex, as in the ASL sign QUESTION, and can contact a location, as in the ASL sign THINK, in which the tip of the index finger contacts the forehead. The handshape in (A-2 b) has three selected fingers, the thumb and all fingers except the index. In ASL it is usually used in initialized signs involving the English letter “d”, such as DORM, in which the tips of the selected fingers contact the cheek.

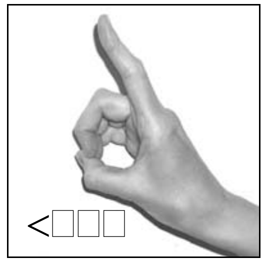
(A-2) Similar handshapes with different selected fingers that are open or closed

a.



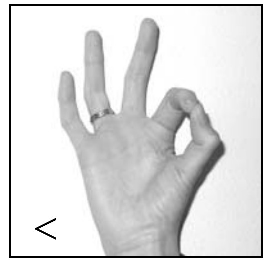

index selected;
selected finger open

b.




middle, ring and
pinky selected and
closed

c.



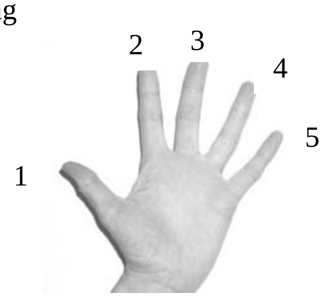

index selected;
selected finger closed

In addition, the selected fingers can be open or closed. In (A-2 a) the selected finger, the index, is open. In (A-2 c) the same selected finger is closed. The tips of the index finger and thumb can contact a location, as in ASL CAT, while the middle, ring and pinky cannot contact a location.[100] (In (A-2 b) the selected fingers are closed.)

The HamNoSys notation system labels the digits one through five beginning with the thumb, as shown in (A-03).

(A-03) HamNoSys digit labeling

1 thumb
2 index finger
3 middle finger
4 ring finger
5 pinky



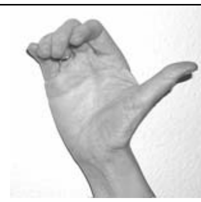

If a diacritic is used, its domain of application is the part of the handshape it appears over, usually all selected fingers, such as 38B, in which all five fingers are curved. Occasionally, the diacritic appears over and applies to only some of the selected fingers, as in 3B8, in which the

---

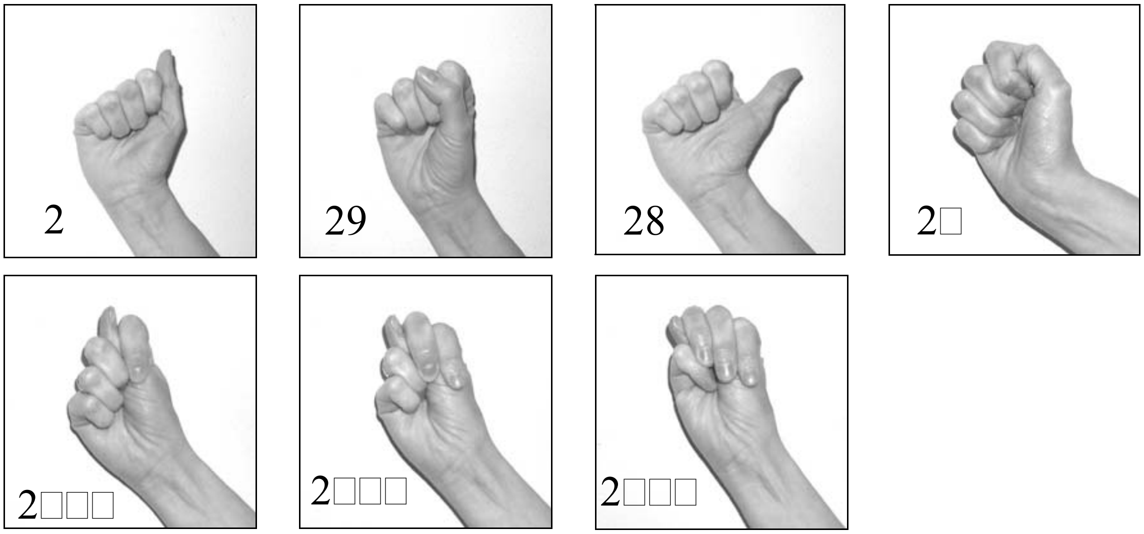
[100] It is apparently possible for two physically identical handshapes to have different selected fingers in different languages. In ASL the handshape < in (A-2 c) only contacts a location with the tips of the thumb and index finger. However, in Chinese Sign Language, this same handshape can contact a location with the middle, ring and index fingers. In this case these three fingers would be considered the selected fingers.

diacritic does not appear over the thumb, and so the thumb is straight. If the domain of application of a diacritic is not all selected fingers, and if this distinction is not clearly discernible from the notation, further specification is indicated. For example, 48B and 4B8 are difficult to disambiguate by examining the notation, and 4≅B and 4B≅ are impossible. Thus, in these cases, the convention is for the diacritic to refer to all selected fingers. To notate handshapes with different specifications for different selected fingers, that is, handshapes in which the diacritic does not apply to all selected fingers, the notation in (A-4) is used.

(A-4) Notation convention for diacritics

| | | | |
|---|---|---|---|
| 3B8<br>not 38 □BB□B□B□B |  | 38B<br>not 38 □B□BB□B□B□B |  |
| 48□B<br>not 48B |  | 48□B<br>not 4B8 |  |
| 4≅B<br>not 48□B□B |  | 4≅□B<br>not 4B≅ |  |

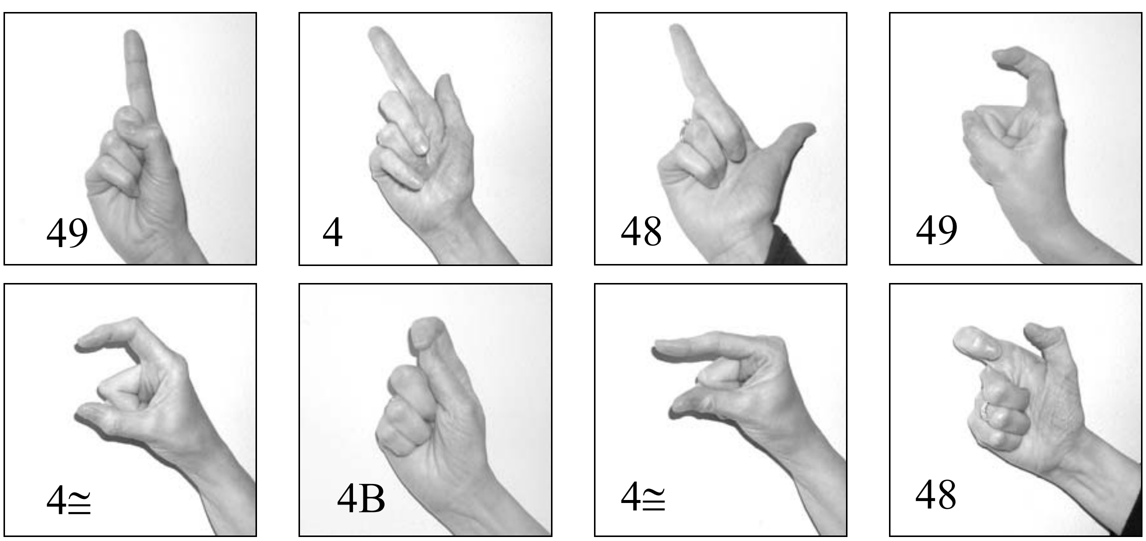
**A.1.1 No fingers selected**

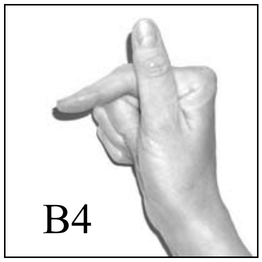


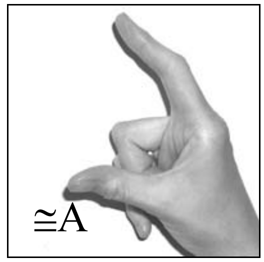

**A.1.2 One finger selected**

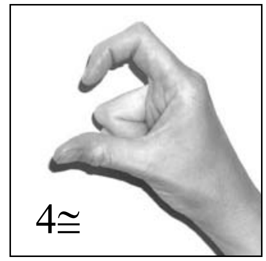
**A.1.2.1 Selected finger open**

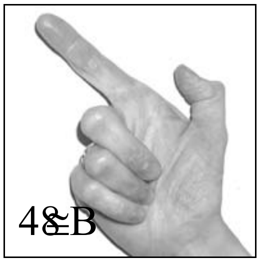
**A.1.2.1.1 Index finger selected**

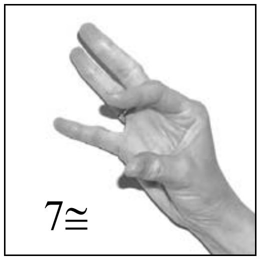


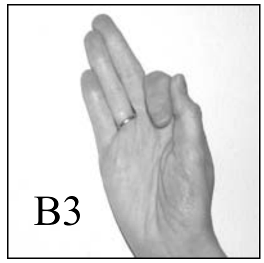



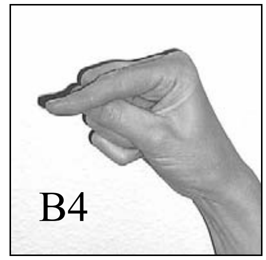



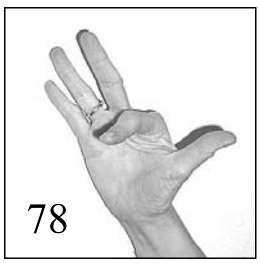



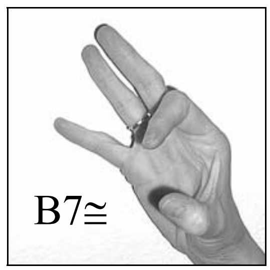



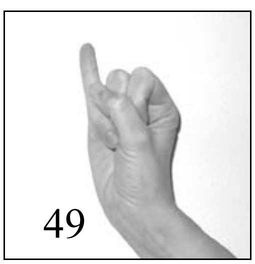



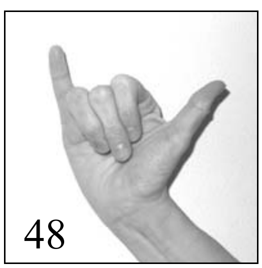


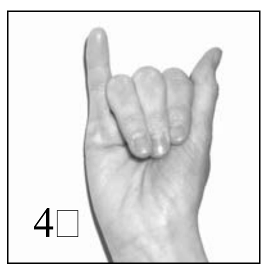


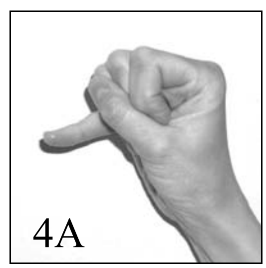



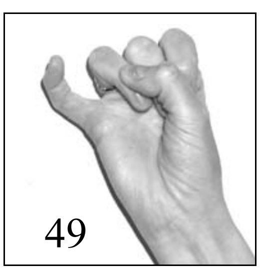

**A.1.2.1.2 Non-index finger selected**



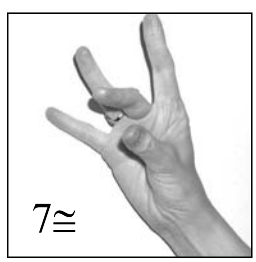



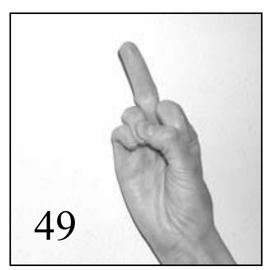



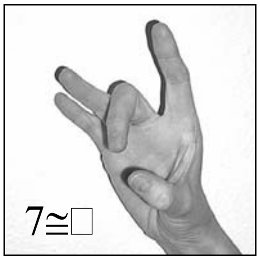





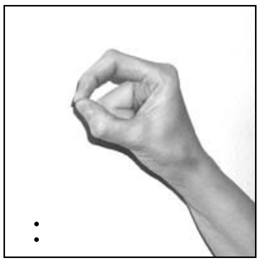


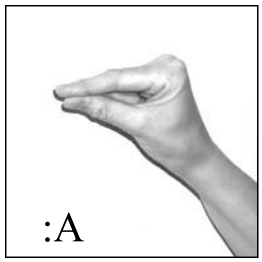


7≅□

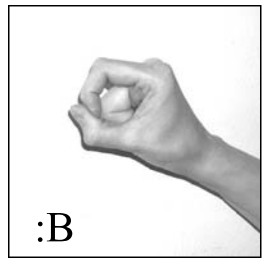
### A.1.2.2 Selected finger closed



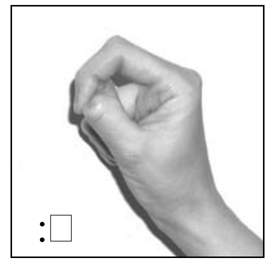
:



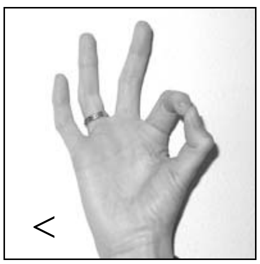
:A



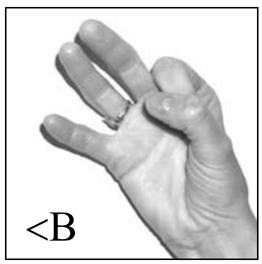
:B



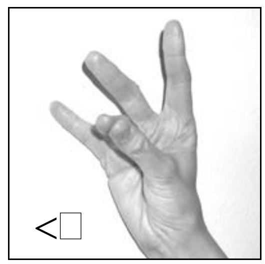
:□



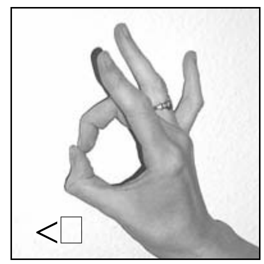
<



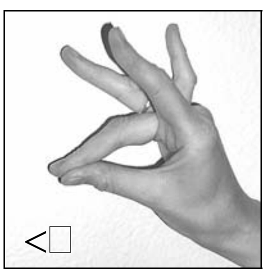
<B



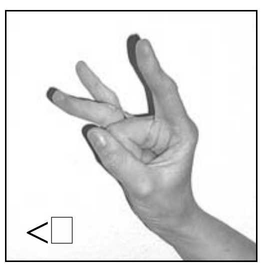
<□



<□



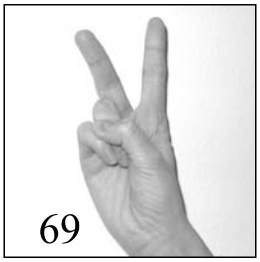
<□



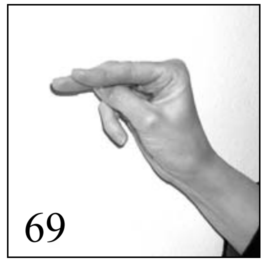
<□

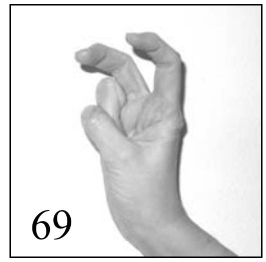
### A.1.3 Two fingers selected

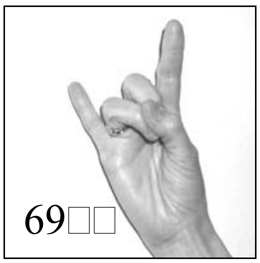
#### A.1.3.1 Selected fingers open

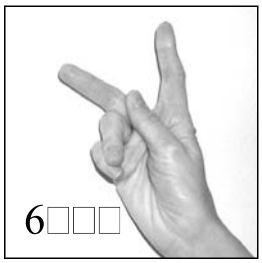

69

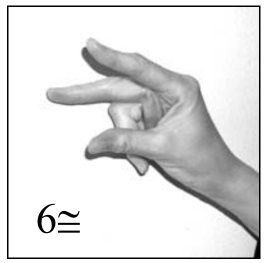

68

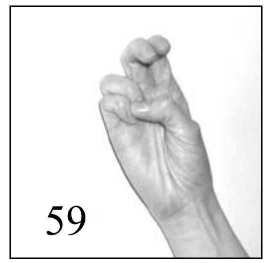

69

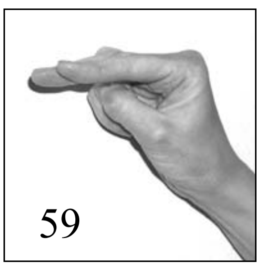

69

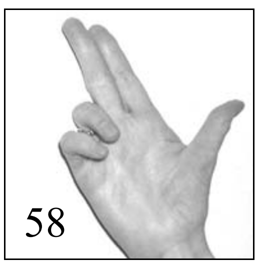

69□□

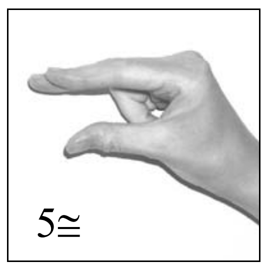

6□□□

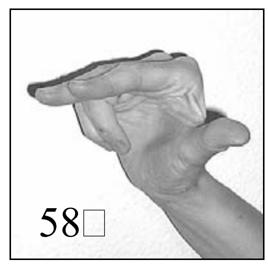

6≅

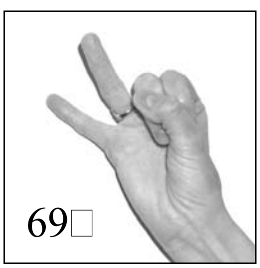

59

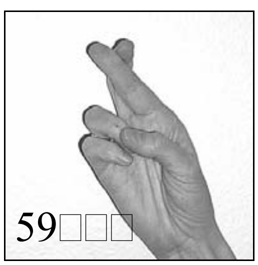

59


58

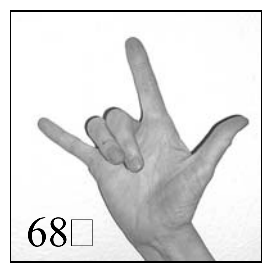

5≅

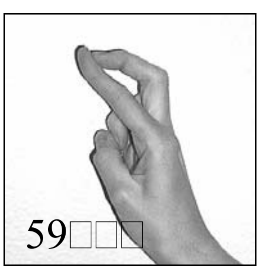

58□

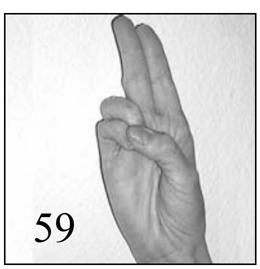

69□

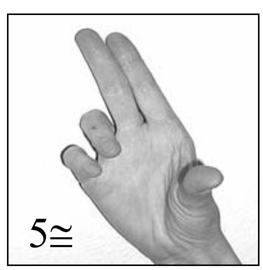

59□□□


68

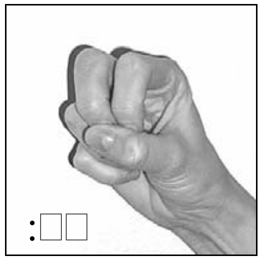

68□

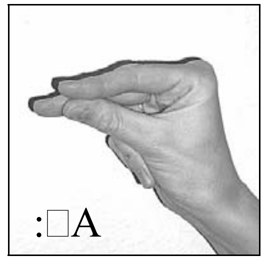

59□□□


59

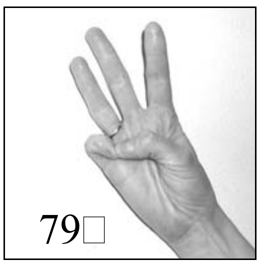

5≅

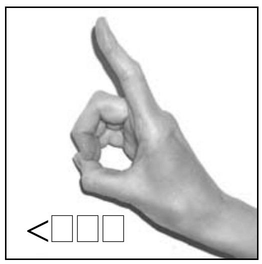
**A.1.3.2 Selected fingers closed**

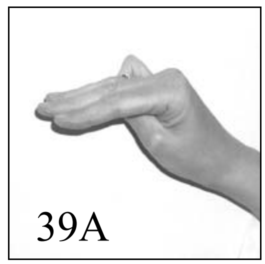



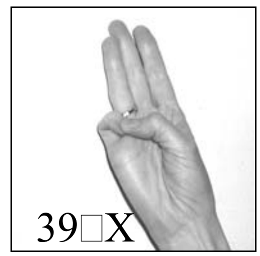



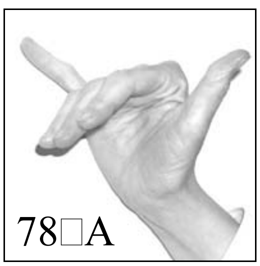

**A.1.4 Three fingers selected**



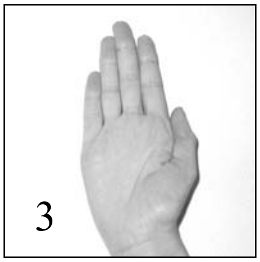



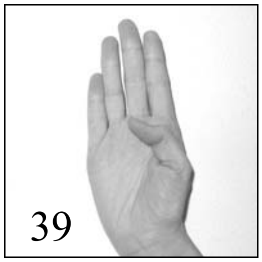



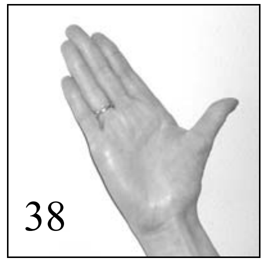

## A.1.5 All fingers selected

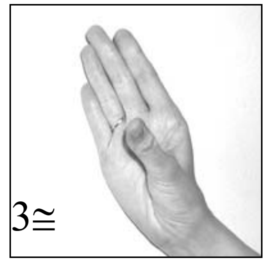
### A.1.5.1 Fingers open

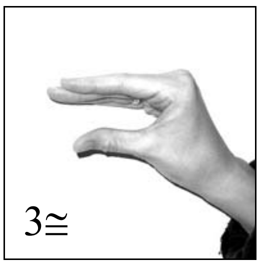
#### A.1.5.1.1 Fingers touching

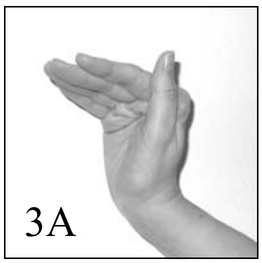

3

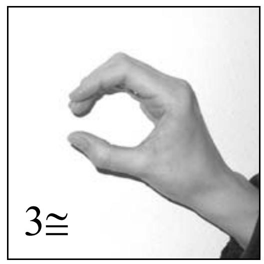

39

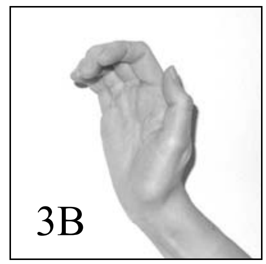

38

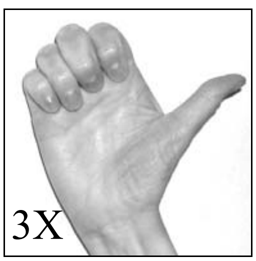

3≅

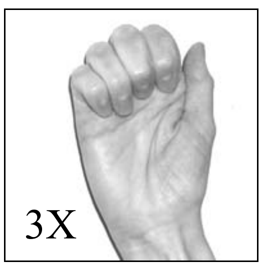

3≅

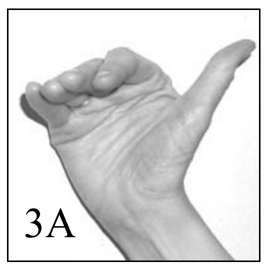

3A

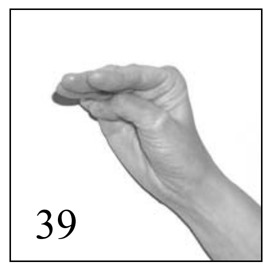

3≅

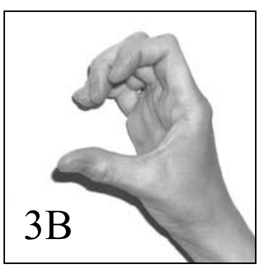

3B

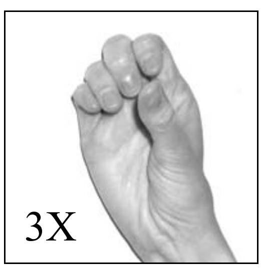

3X

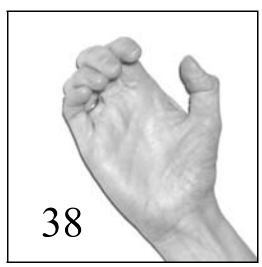

3X

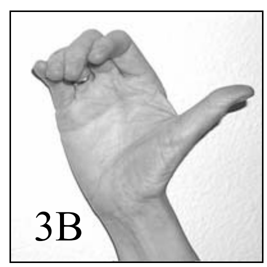

3A

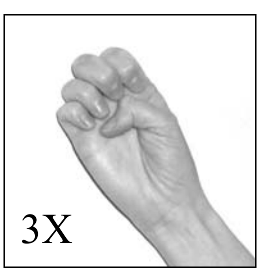

39


3B

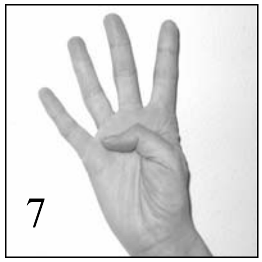

3X

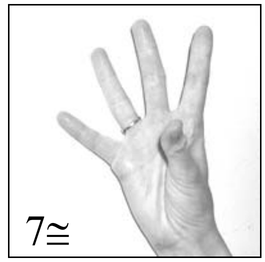

38

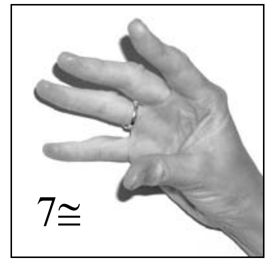

3B


3X

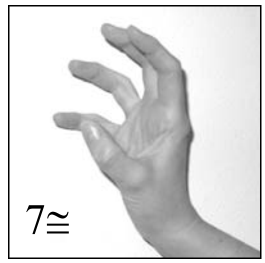
#### A.1.5.1.2 Fingers spread

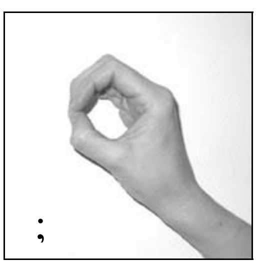



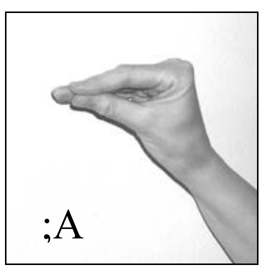


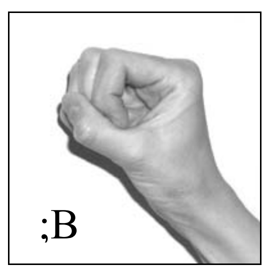






#### A.1.5.2 Fingers closed



## A.2 Location directory

| symbol | location |
|---|---|
| ρ | top of head |
| θ | whole head or face |
| σ | forehead |
| υ | eye or eyes |
| ϖ | nose |
| ψ | lips |
| ζ | chin |
| { | under chin |
| \| | neck |
| ω | ear |
| ξ | cheek |
| □ | upper arm |
| □ | inside of elbow |
| □ | outside of elbow |
| □ | forearm |
| □ | wrist pronated |
| □ | wrist supinated |
| } | upper torso |
| ~ | middle torso |
| o | lower torso |
| □ | below waist |

| diacritic | meaning |
|---|---|
| X φ | behind area X |
| X □ | to the right or left side of area X |
| X □□ | at the side of the body near area X |

| “face” locations |
|---|
| ρ θ σ υ ϖ ψ ζ { ω ξ |

| “torso” locations |
|---|
| \| } ~ o □ |

| “arm” locations |
|---|
| □ □ □ □ □ □ |

| “middle” locations |
|---|
| σ υ ϖ ψ ζ { \| } ~ o □ |

| “side” locations |
|---|
| ρ θ ω ξ |

and any location of the form X □

# Appendix B: Source code for mutual information program

```
#ifndef _UNICODE
#define _UNICODE
#endif

#include <tchar.h>
#include <wchar.h>
#include <stdio.h>
#include <memory.h>
#include <math.h>
#include <stdlib.h>
#include <string.h>
#include <time.h>
#include <assert.h>
#define double float
/////////////////////////////////////////////////////////////////////////
FILE  *hfileOut;        // unicode output file
FILE  *hfileHeadOut;    // unicode output file: header information
wchar_t     szOut[1024];

int cTest = 1000;
const int cBucket = 102400;
/////////////////////////////////////////////////////////////////////////
#define min(x,y)  ( (x < y) ? x : y )
#define max(x,y)  ( (x > y) ? x : y )
/////////////////////////////////////////////////////////////////////////
typedef struct tCharSet
{
      int         cch;
      wchar_t     rgch[1024][16];
}     CharSet;

CharSet     rgchX;
CharSet rgchY;
/////////////////////////////////////////////////////////////////////////
typedef struct tRawData
{
      int         x;                // index into rgX array
      int         y;                // index into rgY array
      int         count;
}     RawData;

typedef struct tInputData
{
      int         cData;            // how many data points
      RawData     rgRawData[10240]; // raw input data
}     InputData;

InputData   data;
/////////////////////////////////////////////////////////////////////////
#define     xMax rgchX.cch                /* maximum Column */
#define     yMax rgchY.cch                /* maximum Row    */
```

```
double              counts[1024][1024];     // input data
double              countX[1024];           // sum of columns
double              countY[1024];           // sum of rows

double              rgProb[1024][1024];     //probability of a data point
double              rgProbX[1024];    // marginal probability of columns
double              rgProbY[1024];    // marginal probability of rows

double              rgLogOdds[1024][1024];  //probability of a data point
double              rgProbLogOdds[1024][1024];    // Prob * LogOdds

double              pseudoCounts = 0.00;//if used avoids log(0) issues
///////////////////////////////////////////////////////////////////////
double              mutualInfo;
double              entropyX;
double              entropyY;

int                 rgMutualInfo[ cBucket ];
double              mutualInfoAvg = 0;
double              mutualInfoMin = 1000.0;
double              mutualInfoMax = -1000.0;
///////////////////////////////////////////////////////////////////////
//
void Help()
{
      printf(
"\n\nThis program calculates the \"mutual information\" between two data
sets."
"\n\nUsage : MutuInfo file1.txt [output=output.txt] [test=1000]
[resolution=0.01]"
"\n\nOR      MutuInfo file1.txt [output=output.txt] [test=1000]
[bucketsize=100] > output.txt"
"\n\nfile1.txt is a unicode text file "
"\n(usually from the Excel spreadsheet)"
"\n\n[test=1000] is an optional parameter that sets the number of "
"\nrandom tests to execute to generate histograms."
"\n\noutput.txt is a text file in tab separated variable format (for
Excel)."
"\nthe character \">\" before output.txt tells the computer to put the "
"\ndata into a file, not to the screen."
"\n\n[resolution] and [bucketsize] are mutually exclusive optional
parameters."
"\n - [resolution=0.01] sets the width of each histogram data bucket"
"\n - [bucketsize=100] sets the number of buckets between 0 and 1"
"\n - the maximum for bucketsize is 100000"
"\n - the minimum for resolution is 0.0001"
 );

      exit(1);
}

///////////////////////////////////////////////////////////////////////
void  MyPutWS( wchar_t *sz )
{
```

```
	while ( *sz )
	{
		fputwc( *sz++, hfileOut );
	}
}
///////////////////////////////////////////////////////////////////////
//	step zero : read in unicode file and translate into indices
void	ReadRawData( wchar_t *szFile )
{
	FILE	*hfile;// = _wfopen( szFile, _T("rb") );
	wchar_t		szIn[1024];
	wchar_t		*pchOne;
	wchar_t		*pchTwo;
	wchar_t		*pchTab;		// comma pointer
	wchar_t		*pchLF;			// line feed pointer
	wchar_t		*pch = szFile;	// temp var
	int			i, x,y;
	int			cLine = 0;
	char	szAscii[1024];
	char	*pchAscii = szAscii;

	if ( *pch == 0xfeff )	pch++;
	while ( *pchAscii++ = (char)*pch++ );
	hfile = fopen( szAscii, "rb" );

	rgchX.cch = 0;
	rgchY.cch = 0;

	if ( !hfile )
	{
		fprintf( stderr, "\n\n\tfile %s not found\n\n", szFile );
		exit(1);
	}

	wchar_t ch = fgetwc( hfile );		// check for unicode flag
	if( ch != 0xfeff )
	{
		ungetwc( ch, hfile );
		printf( "\n\nThis is not a unicode file\n\n" );
		exit(1);
	}


	wprintf( _T("\n\nSource file = %s\n\n"), szFile );
	pchOne = fgetws( szIn, 1024, hfile ); // read in the header line

	pchLF = wcsstr( szIn, _T("\r") );
	if ( pchLF )	*pchLF = 0;			// remove line feed

	wprintf( _T("\nThrowing away text : %s"), pchOne );

	while ( pchOne = fgetws( szIn, 1024, hfile ) )
	{
//		find the first two character strings
		pchLF = wcsstr( szIn, _T("\r") );
```

```
            if ( pchLF )      *pchLF = 0;       // remove line feed

            pchTab = wcsstr( pchOne, _T("\t") );
            if ( pchTab )
            {
                  pchTwo = pchTab + 1;
                  *pchTab = 0;            // mark end of first char
                  pchTab = wcsstr( pchTwo, _T("\t") );
                  if ( pchTab )*pchTab = 0;// mark end of second char
            }
            else
            {
                  printf( "\n\nstopping after %s", szIn );
                  break;
            }

            if ( ( wcslen( pchOne ) == 0 ) || ( wcslen( pchOne ) == 0 )
)
            {
                  wprintf( _T("\nthrowing away ..%s..%s.."), pchOne,
pchTwo );
                  continue;                     // bad input text
            }

//          fwprintf( hfileOut, _T("\n%s\t%s"), pchOne, pchTwo );
            fputwc( '\n', hfileOut );
            MyPutWS( pchOne );
            fputwc( '\t', hfileOut );
            MyPutWS( pchTwo );
            fputwc( 0x0d, hfileOut );
//          fputwc( 0x0a, hfileOut );


            wprintf( _T("\n%s\t%s"), pchOne, pchTwo );
            wprintf( _T("\t %d"), ++cLine );

//          find X index
            x = 0;
            y = 0;

            for ( i = 0; i < rgchX.cch; i++ )
            {
                  if ( wcscmp( pchOne, (wchar_t *)&rgchX.rgch[i] ) == 0
)
                  {
                        x = i;
                        break;
                  }
            }

            if ( wcscmp( pchOne, (wchar_t *)&rgchX.rgch[x] ) )
            {
                  x = rgchX.cch;
                  wcscpy( (wchar_t *)&rgchX.rgch[rgchX.cch++], pchOne );
            }
```

```
////  Find Y index
            for ( i = 0; i < rgchY.cch; i++ )
            {
                  if ( wcscmp( pchTwo, (wchar_t *)&rgchY.rgch[i] ) == 0
)
                  {
                        y = i;
                        break;
                  }
            }

            if ( wcscmp( pchTwo, (wchar_t *)&rgchY.rgch[y] ) )
            {
                  y = rgchY.cch;
                  wcscpy( (wchar_t *)&rgchY.rgch[rgchY.cch++], pchTwo );
            }

            data.rgRawData[data.cData].x = x;
            data.rgRawData[data.cData].y = y;
            data.cData++;

//          printf( "\t%s,%s\t%d,%d", rgchX.rgch[x], rgchY.rgch[y], x, y
);
      }

      fclose( hfile );
}

///////////////////////////////////////////////////////////////////////
//    step one : convert raw data into counts of data points
void  CountDataPoints()
{
      int   i;

      memset( &counts, 0, sizeof( counts ) );

      for ( i = 0; i < data.cData; i++ )
      {
            counts[ data.rgRawData[i].x ][ data.rgRawData[i].y ] += 1.0;

            countX[ data.rgRawData[i].x ] += 1.0;
            countY[ data.rgRawData[i].y ] += 1.0;
      }

}

///////////////////////////////////////////////////////////////////////
//    step two ; find probabilities of each data point
void  SetProbabilities()
{
      int x, y;

      memset( &rgProb, 0, sizeof( rgProb ) );
      memset( &rgProbX, 0, sizeof( rgProbX ) );
```

```
      memset( &rgProbY, 0, sizeof( rgProbY ) );

      for ( x = 0; x < xMax; x++ )
      {
            for ( y = 0; y < yMax; y++ )
            {
                  rgProb[x][y] = counts[x][y] / (double)data.cData;
            }
      }

      for ( x = 0; x < xMax; x++ )
      {
            for ( y = 0; y < yMax; y++ )
            {
                  rgProbX[x] += rgProb[x][y];
                  rgProbY[y] += rgProb[x][y];
            }
      }
}

////////////////////////////////////////////////////////////////////////
#define Log2(x)   (log(x) / log(2.0) )
///////////////////////////////////////////////////////////////////////
//    step three : find Log Odds Table
void  SetLikelyhood()
{
      int         ix, iy;
      double      denom;            // denominator of the division

      memset( &rgLogOdds, 0, sizeof( rgLogOdds ) );

      for ( ix = 0; ix < xMax; ix++ )
      {
            for ( iy = 0; iy < yMax; iy++ )
            {
                  denom = rgProbX[ix] * rgProbY[iy];
                  if ( ( denom != 0.0 ) && ( rgProb[ix][iy] != 0 ) )
                  {
                        rgLogOdds[ix][iy] = (double)Log2( rgProb[ix][iy]
/ denom );
                  }
                  else
                        rgLogOdds[ix][iy] = 0.0;
            }
      }
}

///////////////////////////////////////////////////////////////////////
//    step four : calculate Mutual Info and entropy
void  CalculateMutualInfo()
{
      int         x, y;

//    calculate Mutual Info
      mutualInfo = 0.0;
```

```
	for ( x = 0; x < xMax; x++ )
	{
		for ( y = 0; y < yMax; y++ )
		{
			rgProbLogOdds[x][y] = rgProb[x][y] * rgLogOdds[x][y];
			mutualInfo += rgProb[x][y] * rgLogOdds[x][y];
		}

	}

//	Calculate Entropy for X, Y
	entropyX = 0.0;
	entropyY = 0.0;

	for ( x = 0; x < xMax; x++ )
		entropyX += rgProbX[x] * (double)Log2( rgProbX[x] );

	for ( y = 0; y < yMax; y++ )
		entropyY += rgProbY[y] * (double)Log2( rgProbY[y] );

	entropyX = -entropyX;
	entropyY = -entropyY;
}

/////////////////////////////////////////////////////////////////////
void	MutualInfo()
{
	CountDataPoints();
	SetProbabilities();
	SetLikelyhood();
	CalculateMutualInfo();
}

/////////////////////////////////////////////////////////////////////
void	ShowCountTable()
{
	int	ix, iy;

	fwprintf( hfileOut,
_T("\n\n=========================================================") );
	fwprintf( hfileOut, _T("\nCount Table\n\n") );

	fwprintf( hfileOut, _T("\n") );
	for ( ix = 0; ix < xMax; ix++ )
	{
		fwprintf( hfileOut, _T("\t%4s"), rgchX.rgch[ix] );
	}

	for ( iy = 0; iy < yMax; iy++ )
	{
		fwprintf( hfileOut, _T("\n%4s "), rgchY.rgch[iy] );
		for ( ix = 0; ix < xMax; ix++ )
		{
			fwprintf( hfileOut, _T("\t") );
```

```
                  if ( counts[ix][iy] )   fwprintf( hfileOut,
_T("%5.0f"), counts[ix][iy] );
            }
            fwprintf( hfileOut, _T("\t%5.0f"), countY[iy] );
      }

      fwprintf( hfileOut, _T("\n") );
      for ( ix = 0; ix < xMax; ix++ )
      {
            fwprintf( hfileOut, _T("\t%5.0f"), countX[ix] );
      }
}

//////////////////////////////////////////////////////////////////////
void  ShowProbTable()
{
      int   ix, iy;

      fwprintf( hfileOut,
_T("\n\n=========================================================") );
      fwprintf( hfileOut, _T("\nProbability Table\n\n") );

      fwprintf( hfileOut, _T("\n") );
      for ( ix = 0; ix < xMax; ix++ )
      {
            fwprintf( hfileOut, _T("\t%4s"), rgchX.rgch[ix] );
      }

      for ( iy = 0; iy < yMax; iy++ )
      {
            fwprintf( hfileOut, _T("\n%4s "), rgchY.rgch[iy] );
            for ( ix = 0; ix < xMax; ix++ )
            {
                  fwprintf( hfileOut, _T("\t") );
                  if ( rgProb[ix][iy] )   fwprintf( hfileOut, _T("%f"),
rgProb[ix][iy] );
            }
            fwprintf( hfileOut, _T("\t%f"), rgProbY[iy] );
      }

      fwprintf( hfileOut, _T("\n") );
      for ( ix = 0; ix < xMax; ix++ )
      {
            fwprintf( hfileOut, _T("\t%f"), rgProbX[ix] );
      }
}

//////////////////////////////////////////////////////////////////////
void  ShowLogOdds()
{
      int   ix, iy;

      fwprintf( hfileOut,
_T("\n\n=========================================================") );
      fwprintf( hfileOut, _T("\nLog Odds Table\n\n") );
```

```
	fwprintf( hfileOut, _T("\n") );
	for ( ix = 0; ix < xMax; ix++ )
	{
		fwprintf( hfileOut, _T("\t%4s"), rgchX.rgch[ix] );
	}

	for ( iy = 0; iy < yMax; iy++ )
	{
		fwprintf( hfileOut, _T("\n%4s "), rgchY.rgch[iy] );
		for ( ix = 0; ix < xMax; ix++ )
		{
			fwprintf( hfileOut, _T("\t") );
			if ( rgLogOdds[ix][iy] )	fwprintf( hfileOut,
_T("%f"), rgLogOdds[ix][iy] );
		}
	}

	fwprintf( hfileOut, _T("\n") );
}

///////////////////////////////////////////////////////////////////////
void	ShowProbLogOdds()
{
	int	ix, iy;

	fwprintf( hfileHeadOut,
_T("\n\n===========================================================") );
	fwprintf( hfileHeadOut, _T("\nProbability * Log Odds Table\n\n")
);

	fwprintf( hfileHeadOut, _T("\n") );
	for ( ix = 0; ix < xMax; ix++ )
	{
		fwprintf( hfileHeadOut, _T("\t%4s"), rgchX.rgch[ix] );
	}

	for ( iy = 0; iy < yMax; iy++ )
	{
		fwprintf( hfileHeadOut, _T("\n%4s "), rgchY.rgch[iy] );
		for ( ix = 0; ix < xMax; ix++ )
		{
			fwprintf( hfileHeadOut, _T("\t") );
			if ( rgProbLogOdds[ix][iy] )	fwprintf( hfileHeadOut,
_T("%f"), rgProbLogOdds[ix][iy] );
		}
	}

	fwprintf( hfileHeadOut, _T("\n") );
	fwprintf( hfileHeadOut,
_T("\n\n==========================================================") );
}
///////////////////////////////////////////////////////////////////////
#define ShowVar(x) fwprintf( hfileOut, _T("\n") _T(#x) _T(" = %f"), x );
```

```
#define ShowVarHead(x)  fwprintf( hfileHeadOut, _T("\n") _T(#x) _T(" =
%f"), x );
///////////////////////////////////////////////////////////////////////
void  ShowResults()
{
      ShowCountTable();
      ShowProbTable();
      ShowLogOdds();

      fwprintf( hfileOut, _T("\n\n") );
      ShowVar( mutualInfo );
      ShowVar( entropyX );
      ShowVar( entropyY );

      ShowVarHead( mutualInfo );

}

///////////////////////////////////////////////////////////////////////
void RandomizeData()
{
      int   i;
      int   j;
      int   oldY;       // temp var

      for ( i = 0; i < data.cData; i++ )
      {
            j = rand() % data.cData;
            oldY = data.rgRawData[i].y;
            data.rgRawData[i].y = data.rgRawData[j].y;
            data.rgRawData[j].y = oldY;
      }

}


///////////////////////////////////////////////////////////////////////
void  CloseMergeFiles()
{
      char szAscii[1024];                        // new file name
      int   i = 0;

      while ( szOut[i] )
      {
            szAscii[i] = (char)szOut[i];
            i++;
      }
      szAscii[i] = 0;

      int iName = 0;
      while ( rename("Head.out", szAscii ) && (iName < 50) )
      {
            char sz[1024];
            sprintf( sz, "%s.%d", szAscii, iName );
            printf( "\ntrying to rename to file %s", sz );
```

```
            rename( szAscii, sz );
            iName++;
      }

      if ( iName >= 50 )
      {
            char sz[1024];
            iName = 0;

            sprintf( sz, "%s.%d", szAscii, iName );
            while ( rename( sz, "Head.out" ) && (iName < 50) )
            {
                  iName++;
                  sprintf( sz, "%s.%d", szAscii, iName );
            }

            printf( "\nfile %s cannot be renamed.  Output file is %s",
                  szAscii, sz );
            strcpy( szAscii, sz );
      }

      hfileOut = fopen( szAscii, "ab" );
      FILE *hfileTemp = fopen( "temp.out", "rb" );

      // copy the temp data
      wchar_t     rgch[1024];
      while ( fgetws(rgch, 1024, hfileTemp ) )
      {
            fputws( rgch, hfileOut );
      }

      fclose( hfileOut );
      fclose( hfileTemp );
      remove( "temp.out" );
}


double            rgProbLogOddsPrimary[1024][1024];   // the interesting
Probability Log Odds Table
///////////////////////////////////////////////////////////////////////
int _cdecl wmain( int argc, wchar_t *argv[] )
{
      double      bucketSize = 100.0;
      double      resolution = (double)1.0 / bucketSize;
      double      mutualInfoOrig;
      int         cGreaterThan      = 0;
      int         cLessThan         = 0;
      int         iarg;
      double      d;
      wchar_t     sz[100];
      wchar_t     *pch;

      srand( (unsigned)time( NULL ) );//seed random number generator

      if ( argc == 1 )
```

```
	{
		Help();
	}

	wcscpy( szOut, argv[1] );
	pch = wcsstr( szOut, _T("." ) );
	if ( pch )
	{
		pch++;
		*pch++ = 'o';
		*pch++ = 'u';
		*pch++ = 't';
	}

	for ( iarg = 2; iarg < argc; iarg++ )
	{
		wcscpy( sz, argv[iarg] );
		wcslwr( sz );
	// force argument to lower case

		pch = wcsstr( argv[iarg], (wchar_t *)"=" ); // find numeric
value of argument
		if ( pch )  d = (double)_wtof( pch+1 );
	// in the form "arg=###.##"
		else  d = 0.0;

		if ( wcsstr( argv[iarg], (wchar_t *)"test" ) )
	// number of tests to do
		{
			if ( d > 1 )		cTest = (int)d;
		}
		else if ( wcsstr( argv[iarg], (wchar_t *)"resolution" ) )
		{
			resolution = d;
			bucketSize = (double)1.0 / resolution;
		}
		else if ( wcsstr( argv[iarg], (wchar_t *)"bucketsize" ) )
		{
			bucketSize = d;
			resolution = (double)1.0 / bucketSize;
		}
		else if ( wcsstr( argv[iarg], (wchar_t *)"output" ) )
		{
			wcscpy( szOut, pch+1 );
		}
	}

//	hfileOut = _wfopen( szOut, _T("wb") );
	hfileOut = fopen( "temp.out", "wb" );
	hfileHeadOut = fopen( "Head.out", "wb" ); //fileheader information

	if ( hfileOut == NULL )
	{
		wprintf( _T("\n\ncannot open output file %s.  Program
terminated\n\n"),
```

```
                szOut );
            return(1);
        }
        fputwc( 0xfeff, hfileOut );
        fputwc( 0xfeff, hfileHeadOut );
        fwprintf( hfileOut, _T("\n\nThe Unicode Version\tFeb 2003\n\n") );

        printf( "\nExecuting %d random data sets", cTest );
        printf( "\nResolution = %8.4f", resolution );

        ReadRawData( argv[1] );

        MutualInfo();
        mutualInfoOrig = mutualInfo;
        ShowResults();
        _flushall();

        memset( rgMutualInfo, 0, sizeof( rgMutualInfo ) );
        fprintf( stderr, "\n\n\n\n" );
        memcpy( rgProbLogOddsPrimary, &rgProbLogOdds, sizeof(
rgProbLogOdds ) );

        for ( int i = 0; i < cTest; i++ )
        {
            RandomizeData();
            MutualInfo();
            mutualInfoAvg += mutualInfo;
            mutualInfoMin = min( mutualInfoMin, mutualInfo );
            mutualInfoMax = max( mutualInfoMax, mutualInfo );

            if ( mutualInfoOrig > mutualInfo )  cGreaterThan++;
            if ( mutualInfoOrig < mutualInfo )  cLessThan++;

            int   j = (int)(mutualInfo  * bucketSize);
            rgMutualInfo[j]++;
            if ( j > cBucket )
            {
                printf( "\n\nproblem in line %d", __LINE__ );
                printf( "\n\nplease try a larger bucket size\n\n" );
                exit(1);
            }

            if ( ( i & 15 ) == 0 )
            {
                fprintf( stderr, "\r tests executed = %d   (%8.2f
sec)", i, (double)clock() / (double)CLOCKS_PER_SEC);
            }
        }

        mutualInfoAvg = mutualInfoAvg / cTest;

        fwprintf( hfileOut, _T("\n\n") );
        fwprintf( hfileOut, _T("\nMin Mutual Info = %g"), mutualInfoMin );
        fwprintf( hfileOut, _T("\nMax Mutual Info = %g"), mutualInfoMax );
```

```
	fwprintf( hfileOut, _T("\n\nAverage Mutual Info = %g of %d
tests"), mutualInfoAvg, cTest );
	fwprintf( hfileOut, _T("\n\nOrigData > %d (%g%) tests"),
		cGreaterThan, (double)cGreaterThan / (double)cTest );
	fwprintf( hfileOut, _T("\nOrigData < %d (%g%) tests"),
		cLessThan, (double)cLessThan / (double)cTest );
	fwprintf( hfileHeadOut, _T("\nOrigData < %d (%g%) tests"),
		cLessThan, (double)cLessThan / (double)cTest );
	fwprintf( hfileOut, _T("\n\nDistance of mutualInfo from Average =
%g"), fabs( mutualInfoOrig - mutualInfoAvg ) );
	fwprintf( hfileOut, _T("\n\n") );

	memcpy( &rgProbLogOdds, rgProbLogOddsPrimary, sizeof(
rgProbLogOdds ) );
	ShowProbLogOdds();

	for ( i = (int)(mutualInfoMin * bucketSize); i < (mutualInfoMax *
bucketSize); i++ )
	{
		fwprintf( hfileOut, _T("\n%8.3f\t%d"), i / bucketSize,
rgMutualInfo[i] );
	}

	fwprintf( hfileOut, _T("\n\nElapsed time = %8.2f"),
(double)clock() / (double)CLOCKS_PER_SEC );
	fclose( hfileOut );
	fclose( hfileHeadOut );

	CloseMergeFiles();

	printf( "\n\n" );
	for (int iArg = 0; iArg < argc; iArg++ )
	{
		wprintf( _T("%s "), argv[iArg] );
	}
	printf( "\n\n\a" );

	return 0;
}
```

# Curriculum Vitae

**Lorna Rozelle**

Department of Linguistics
Box 354340
University of Washington
Seattle, WA 98105

rozelle@u.washington.edu

**Education**

Ph.D. Linguistics, University of Washington, April 2003
Dissertation: "The Structure of Sign Language Lexicons:
Inventory and Distribution of Handshape and Location"
Ph.C. Linguistics, University of Washington, October 1998
M.A. Linguistics, University of Washington, June 1992
Thesis: "Constraining Radical Underspecification"
M.A. Mathematics, University of Washington, December 1986
A.B. Mathematics, Cornell University, May 1983

**Research Interests**

| | | |
|---|---|---|
| Linguistic theory | Signed languages | Morphology |
| Phonology | American Sign Language | Quantitative linguistics |
| Acoustic phonetics | Finnish Sign Language | Finnish |

**Experience**

Teaching Assistant
- Department of Linguistics, University of Washington
  - Developed course and shared teaching with another graduate student:
    - Linguistics 403: The Structure of American Sign Language (Spring 1999)
  - Full teaching responsibility:
    - Linguistics 200: Introduction to Linguistic Thought (Autumn 1997)
    - Linguistics 452: Phonology I (Autumn 1996)
    - Linguistics 451: Phonology II (Winter 1994)
  - Partial teaching responsibility:
    - Linguistics 200: Introduction to Linguistic Thought (4 terms, 1992 - 1995)
- Department of Mathematics, University of Washington, 1984 - 1986
- Department of Mathematics, Columbia University, 1983 - 1984
- Department of Mathematics, Cornell University, 1983

Lecturer: full teaching responsibility
- Department of Mathematics, University of Washington, 1986 - 1991
- Department of Mathematics, North Seattle Community College, 1987 -1990
- Department of Mathematics, Seattle Central Community College, 1988

Research Assistant
Department of Space Sciences, Cornell University, 1978 – 1983

**Publications and Presentations**

Rozelle, Lorna. 2002. The inventory and distribution of handshape and location in four sign languages. Poster presented at Linguistic Society of America Annual Meeting, San Francisco.

Taff, Alice, Rozelle, Lorna, Cho, Taehong, Ladefoged, Peter, Dirk, Moses and Wegelin, Jacob. 2001. Phonetic Structures of Aleut. *Journal of Phonetics*, 29: 231-271.

Rozelle, Lorna. 1997. The effect of stress on vowel length in Aleut. In *Fieldwork Studies of Targeted Languages, UCLA Working Papers in Phonetics*, 95, UCLA.

Rozelle, Lorna. 1996. Assimilation in compounds in Finnish Sign Language. Poster presented at Theoretical Issues in Sign Language Research, McGill University, Montreal.

Rozelle, Lorna. 1996. Optimality and the Hand Number feature in ASL and SVK. Paper presented at North West Linguistics Conference, University of Washington.

Rozelle, Lorna. 1992. A feature geometry for American Sign Language. *University of Washington Working Papers in Linguistics,* 10.

Rozelle, Lorna. 1992. ASL feature geometry. Paper presented at Theoretical Issues in Sign Language Research, UCSD.

Rozelle, Lorna, Schoenberg, Beth, Hargus, Sharon and McKee, Cecile. 1992. Simultaneity and the classification of ASL morphology. Poster presented at Theoretical Issues in Sign Language Research, UCSD.

Wong, Lorna, and Gradie, Jonathan. 1983. *Operating Manual and Technical Description of the Cornell Photogoniometer* CRSR Report. Cornell University.

Thomas, Peter, Veverka, Joseph, Gingeris, D., and Wong, Lorna. 1984. " 'Dust' Streaks on Mars." *Icarus,* 60: 161-179.

**Professional Memberships**

Linguistic Society of America

**Awards**

2000 University of Washington Graduate School Arts and Humanities and Social Sciences and Social Professions Dissertation Fellowship

1999 University of Washington, Arts and Sciences Curriculum Development Grant (with Susan McBurney for the development and teaching of Linguistics 403: The Structure of American Sign Language)

1999 University of Washington, Huckabay Fellowship (with Susan McBurney for the development and teaching of Linguistics 403: The Structure of American Sign Language)

1995 Linguistic Society of American Summer Institute fellowship, Cornell University

1992 Fulbright Fellowship, Finland

1991 Linguistic Society of American Summer Institute fellowship, University of New Mexico

**Service**

President, Linguistic Students at the University of Washington, 1996 - 1997.

Volunteer, Deaf-Blind Service Center, 1991 - 1993
Washington State Deaf Blind Citizens, 1991 - 1993
Deaf-Blind Camp, 1989