Exploring Reduced Feature Sets for American Sign Language Dictionaries

Ben Kosa
School of Computer, Data, & Information Sciences
University of Wisconsin?Madison
Madison, Wisconsin, USA
bkosa@cs.wisc.edu

Aashaka Desai
Paul G. Allen School of Computer Science and Engineering
University of Washington
Seattle, Washington, USA
aashakad@cs.washington.edu

Alex X Lu
Microsoft Research
Cambridge, Massachusetts, USA
lualex@microsoft.com

Richard E. Ladner
Paul G. Allen School of Computer Science & Engineering
University of Washington
Seattle, Washington, USA
ladner@cs.washington.edu

Danielle Bragg
Microsoft Research
Cambridge, Massachusetts, USA
danielle.bragg@microsoft.com

Abstract

There is currently no easy way to look up signs in sign language. Feature-based dictionaries help overcome this challenge by enabling users to look up a sign by inputting descriptive visual features, such as handshape and movement. However, feature-based dictionaries are typically cumbersome, including large numbers of complex features that the user must sort through. In this work, we explore simplifying the set of features used in feature-based American Sign Language (ASL) dictionaries. We present two studies: 1) a simulation study focused on lookup accuracy for various reduced feature sets, and 2) a user study focused on understanding human preferences between feature sets. Our results suggest that it is possible to dramatically reduce the number of features needed to search for signs without significantly impacting the accuracy of search results, and that smaller feature sets can improve the user experience in some cases.

CCS Concepts

• **Human-centered computing** → **Accessibility**.

Keywords

American Sign Language (ASL), Dictionary, Search, Education

ACM Reference Format:
Ben Kosa, Aashaka Desai, Alex X Lu, Richard E. Ladner, and Danielle Bragg. 2025. Exploring Reduced Feature Sets for American Sign Language Dictionaries. In *CHI Conference on Human Factors in Computing Systems (CHI '25), April 26–May 01, 2025, Yokohama, Japan.* ACM, New York, NY, USA, 14 pages. https://doi.org/10.1145/3706598.3714118

CHI '25, Yokohama, Japan

ACM ISBN 979-8-4007-1394-1/25/04
https://doi.org/10.1145/3706598.3714118

1 Introduction

Across the world, there are roughly 70 million Deaf and Hard of Hearing (DHH) people who use one of at least 158 different sign languages (there are many more that are undocumented) [2, 14]. In the U.S. alone, there are roughly 500,000 people who are DHH and use ASL as their primary form of communication [24, 25]. Due to its widespread use and importance to the Deaf community, ASL is the third-most studied language within the U.S. [23]. To support language learning and documentation, bidirectional dictionaries (i.e. ASL-to-English and English-to-ASL lookup) are essential. There are many English-to-ASL dictionaries, but very few ASL-to-English dictionaries–which we also refer to as reverse sign language dictionaries throughout this paper.

Querying ASL-to-English dictionaries is significantly more challenging than English-to-ASL. Unlike English, there is no commonly-adopted written form of ASL, and thus no standardized system to describe a sign for look-up in a dictionary. Digital sign language input is commonly accomplished using two different input strategies: (1) feature-based input [7, 10, 15, 18, 22, 30, 39] and (2) video or example-based input [1, 11, 38]. Feature-based input allows users to describe a sign by inputting common features such as the shape of the hand (handshape), location, palm orientation, and movement. Video or example-based input allows users to physically demonstrate a sign as input, but the performance of these example-based systems is still lacking for effective sign lookup [6, 13].

In this work, we focus on improving feature-based dictionaries. While numerous ASL-to-English feature-based dictionaries exist, most of these use phonological systems designed by trained linguists that require users to choose between a large set of features when searching for a sign, and it is unclear if these systems are accessible to everyday sign language users. Motivated by these challenges, we seek to explore simpler and more usable feature sets for feature-based dictionaries. To do this, we present two studies: 1) a simulation study focused on understanding the tradeoff between feature set size and dictionary retrieval accuracy, and 2) a user study to understand user preferences for feature sets proposed in our simulation study. In our simulation study, we investigated different approaches to reducing the feature sets typically used in feature-based sign language search. We found that effective search could be maintained using only reduced handshape and movement features, and dropping all other feature types (e.g. location). In our user study, we evaluated user preferences for these reduced feature sets and observed that, in certain cases, reduced feature sets can enhance the user experience. While the work this paper focuses on ASL and English, the methods we explored for reducing feature sets could be easily be extended to other sign languages.

2 Background and Related Work

In this section, we provide a brief background on sign languages and their linguistic features, describe existing approaches to sign language dictionaries, and outline current approaches to feature-based sign language search.

2.1 Sign Languages and Linguistic Features

Sign languages are the primary languages used in many DHH communities, with over 158 different sign languages wordwide. Our work focuses on ASL, which is used predominantly in North America. A sign in sign language can be described in terms of phonological features, similar to how words in spoken languages can be described in terms of voicing, place, and manner of articulation. Currently, the field has largely adopted features grounded in decades of sign language linguistic theory pioneered by linguist William Stokoe and later refined by other linguists [4, 16, 32, 33], which we refer to as *Stokoe features* for short[1].

Stokoe features document five different phonological parameters of signs: handshape, movement, location (relative to the body), (palm) orientation, and non-manual markers [17, 40]. Handshapes are used to describe the configuration of hands and fingers in a sign. In ASL, they are often described in relation to a more granular version of the ASL alphabet (e.g. A-handshape, Open-A handshape, B-handshape, Bent-B handshape, etc), the ASL number system (numbers 1-9), and classifiers (a set of handshapes used to represent specific categories of objects, actions, or characteristics in space). Movement refers to how the hands and arms change position while signing (e.g., up-and-down, away-from-the-body, etc.). Location describes the position of the hand relative to the body, and orientation describes the direction the signer's palm is facing. Non-manual markers refer to facial expressions, head movements, and other body movements outside of the hands and arms that convey grammatical or emotional information and are an essential part of the meaning in signed languages. These different features make it possible to document and search for signs, which we discuss below.

To describe a sign entirely, we need to describe handshape, movement, location, and palm orientation for both hands. Since some signers are left-handed and others are right-handed, features are described in terms of *dominant* and *non-dominant hand*. Some signs are one-handed (e.g., APPLE, CANDY), and some signs are two-handed (e.g., CLASS, HOW). Two handed-signs can be *symmetric*– where both hands have the same handshape and same or opposite direction, orientation, and location–or *asymmetric*–where both hands have different directions, orientation, and location but may have the same handshape in both hands (e.g. AGAIN or PAPER) [5]. In our work, we draw upon linguistic theories on the *Dominance Condition* and the *Symmetry Condition* [4, 5] as well. The Symmetry Condition states that if a sign is two-handed sign where both hands have the same handshape, then both hands must have the same or opposite movements, locations, and orientations. The Dominance Condition states that in two-handed signs where both hands have different handshapes, the non-dominant hand becomes constrained to a small subset of possible handshapes and should not have movement. To our knowledge, we are the first to use assumptions from Dominance and Symmetry conditions to propose smaller and more usable sets of phonological features for

There are other types of linguistic features that can be used to describe a sign, such as morphological or lexical semantic features [29] as well. However, these are often accessible only to linguists due to their requirement of prior linguistic knowledge. Thus for our work on digital sign language dictionary search, we focus on the the above phonological parameters.

2.2 Sign Language Dictionaries

Dictionaries are valuable resources for any language. Bilingual and bidirectional (i.e., ASL-to-English and English-ASL) dictionaries are valuable resources for both novices and experts in sign language. For sign language learners, dictionaries support looking up the meaning of unfamiliar signs as well how to sign new concepts. This is useful to both hearing learners as well as DHH individuals who learn sign language later in life. For those already fluent in a sign language, dictionaries allow them to search for specific jargon [12, 41] or different dialectal variations of a certain sign along with any unfamiliar signs. Overall, dictionaries support greater access to learning resources and contribute to the documentation and preservation of sign languages.

Querying dictionaries for sign languages is complicated compared to languages with both written and spoken forms (e.g. English). While approaches to transcribe linguistic features of signs, such as Stokoe Notation or HamNoSys [19] exist, these systems are typically geared towards linguistic or computational use. Other sign writing systems (such as SignWriting [36], si5s [3], SignFont [26], ASL-phabet [35], ASL Orthography, SLIPA [27], and ASLSJ [34]) are designed for more everyday use, but still require a steep learning curve. Thus, there is a need for different computational approaches to sign language dictionary search.

Printed ASL-to-English dictionaries use the original Stokoe features, or their modern adaptations, to organize signs, often organizing them by handshape [37]. Digital dictionaries use a variety of input methods: videos [1, 11, 38], features [7, 10, 15, 18, 22, 30, 39], or a combination of the two [20]. Outside of dictionaries designed for study in research settings only, there are only several digital dictionaries that exist for ASL which are feature-based [15, 22, 30]. To our knowledge, there also exists at least one digital sign language dictionary that supports searching in multiple sign languages[2] by using Sutton's SignWriting system as input [36].

In contrast to feature-based retrieval, video-based retrieval methods would in principle allow users to search by example, or by demonstrating a sign. For example, recent research has investigated

[1]This is done merely for simplicity. We are fully aware that the current set of handshapes, movements, location, and orientations used today are heavily iterated versions of linguistic work done by William Stokoe, Diane Brentari, Lynn A. Friedman, and other sign language linguists.

[2]https://www.signbank.org/

hybrid systems that start with the user demonstrating a sign, and then further refining their search with feature-based input [20]. Video-based approaches may be more usable for those who are less familiar with using Stokoe features to describe signs. However, implementing the underlying retrieval algorithm requires large amounts of training data and computational resources – both of which have been bottlenecks for improving performance of these example-based dictionaries [6].

In our study, we chose to focus on feature-based retrieval, not only because these are the only retrieval methods currently adopted by digital dictionaries in practice, but also because they have other benefits over video-based retrieval. As retrieval reduces to identifying signs in the dictionary with similar features (e.g. by vector similarity), computational complexity of feature-based retrieval is low. They also allow users to search for signs discreetly, while video-based methods require users to film themselves demonstrating a sign. In comparison to video-based methods (e.g., deep learning models) which require large amounts of recordings, feature-based methods approach with simpler user queries might generalize better to other lower-resource sign languages.

Because of these advantages, we focus on feature-based dictionaries in our work. In particular, we build on the work of [7], leveraging their method for building a dictionary from past queries and using their feature set (which we refer to as *Stokoe features* in this paper) as a point of comparison for smaller feature sets.

2.3 Feature-Based Dictionary Search

The features used in existing feature based dictionaries vary widely, ranging from handshapes only [22, 30] to some combination of handshape, location, orientation, and movement [7, 10, 15, 18, 39].

Existing approaches to feature-based dictionaries have several limitations relating to their underlying retrieval algorithms and user interfaces. These include poor matching of features to signs that are not robust to a large vocabulary, an inability for users to omit features, cumbersome search interfaces, and/or requirements that the user specify between a large number of features [10, 18, 22, 30]. Bragg et al. [7] addresses some of these limitations, but still requires the user to choose between a large set of features when searching for a sign–a result of using the Stokoe-based phonological features that were developed mainly for use by academically trained Linguists, not signers generally. In this work, we explore both smaller and more easily understandable features in the context of dictionary search while maintaining algorithm performance to help address this gap.

3 Simulation Study of Reduced Feature Set Accuracy

In our simulation study, our aim was to identify smaller sets of features that could still preserve accuracy in retrieving signs from ASL-to-English dictionaries, under the hypothesis that the large number of features in current dictionary systems is a usability barrier. To do this, we start with a version of the Stokoe-inspired features used in an ASL-to-English dictionary search system employed by [7]. With 177 phonological features in total, this system consists of 41 handshapes, 17 movements, 11 locations, and 11 orientations for each hand (dominant and non-dominant) along with two additional feature categories: *Relative Position* and *Relative Movement*. We call these the *Stokoe features* in our study. We then systematically ablate categories of features from the baseline features and measure the impact on dictionary retrieval performance on a vocabulary of 1145 signs, to identify which features were important for retrieval performance. We found that some feature categories barely impact retrieval performance, while others were critical for performance. To design reduced feature sets, we prune feature categories that were not impactful for retrieval. For the feature categories important for retrieval, we experimented with clustering them based on different levels of phonological similarity and find some of these reduced featured sets to be similarly effective to the *Stokoe features* in retrieval. In subsequent sections, we expand on each of these results.

3.1 Query Data

To simulate dictionary query data, we collected a dataset of feature evaluations over a comprehensive ASL vocabulary. For a particular sign (e.g., APPLE) , each query consists of a set of feature selections for each of the 177 *Stokoe features* (e.g., 'A-handshape' for handshape feature, 'away from the body' for movement). Our vocabulary spans 1,145 signs, representing the first 1,145 signs taught in a standard American Sign Language (ASL) curriculum. Since different signers may describe signs differently, and dialectal variations exist, each sign was represented as an aggregate of annotations by at least four annotators using the *Stokoe features* for a total of 6,180 annotations. Each annotator was a second-year ASL student who annotated signs based on their own knowledge or video demonstration from a fluent ASL signer. Since ASL signs can be performed using either the left hand or the right hand depending on the signer's dominant hand, we simulate this by duplicating all annotations where hands are asymmetrical, and swapping all feature values for the hands. After this augmentation, this led to a total of 9,838 annotations– for the purpose of our study, we consider these annotations as simulating queries, as they reflect how human annotators record signs as features.

3.2 Dictionary Simulation Process

To simulate dictionary lookup, we adopt the process proposed in Bragg et al [7], where past lookup queries are aggregated to form a dictionary, and a held-out set simulates new incoming dictionary queries. We simulate dictionary retrieval by indexing annotations using a topic modeling method called Latent Semantic Analysis (LSA)[3]. We split our annotations collected from ASL learners into a 80% - 20% train-test split, retaining this split across all evaluated feature sets in our study. Using the training dataset alone, we ran a grid search to find the optimal number of components for LSA, which consist of k–the dimension of the latent space we reduce all of our annotations to for indexing. For k, we searched over 10 evenly spaced k-values ranging from 5 to 177 (the full number of baseline features). For each k-value, we conducted a 10-fold cross-validation on the training dataset, measuring retrieval performance using Discounted Cumulative Gain (DCG), a standard metric for evaluating search engine performance by assessing the relevance of results ranked in a search query [21]. The DCG score accounts for

[3] More details regarding our exact simulation process in [7].

both the relevance of the returned results and their position in the ranking, rewarding higher-ranked relevant results more than those ranked lower. We averaged the DCG scores across folds to obtain a single performance score for each *k*-value. We then evaluated each annotation in our held-out test dataset, reporting the DCG for LSA indexing based on the *k*-value that achieved the best performance during training as our final retrieval performance metric.

3.3 Feature Category Ablation

Towards the aim of reducing the number of features users need to input to query signs, we first sought to understand if all features contribute equally to dictionary retrieval performance, or if some features were disproportionately important. To do this, we performed an ablation study by removing different categories from the full 177 *Stokoe features* –which as a reminder are grouped into six categories (handshape, movement, location, orientation, relative position, and relative movement). Due to the large number of features, we opted to remove entire categories at once, rather than individual features.

Features Dropped	Number of Features	DCG Scores (Avg)	Percent Drop in Features From Stokoe	Difference in DCG From Stokoe
None	177	0.689		
Handshape (Both Hands)	95	0.453	46%	0.236
Drop Movement (Both Hands)	135	0.593	23%	0.096
Orientation (Both Hands)	155	0.674	11%	0.015
Location (Both Hands)	155	0.667	11%	0.022
Drop Rel Position	165	0.683	6%	0.007
Drop Rel Movement	165	0.686	6%	0.003
Orientation, Location, Rel Position, Rel Movement	115	0.633	34%	0.057
All Features From Hand 2	95	0.442	46%	0.247
Handshape (NDH Only)	135	0.576	23%	0.113
Movement (NDH Only)	155	0.644	11%	0.045
Orientation (NDH Only)	165	0.679	6%	0.01
Location (NDH Only)	165	0.676	6%	0.013

Table 1: The results of the Ablation in tabular form, which show average Discounted Cumulative Gain (DCG) performance for feature sets when each feature category was removed, when hand 2 and all of its encoded features were removed, and when the orientation, location, relative position, and relative movement features were removed

Figure 1 shows the effect of removing each of the six feature categories on dictionary retrieval performance in our test set. We report results removing feature categories from both hands, and for the non-dominant hand only. We observe that relative to the existing *Stokoe Handshapes* baseline features (DCG = 0.689), removing the handshape (DCG = 0.453) or the movement category (0.593) induces large drops in DCG when removing features from both hands. Dropping any of the other four categories for both hands only induces a negligible reduction in DCG performance (>0.03 DCG score). We observed similar trends even when dropping these features from the non-dominant hand only, although the drop overall is less pronounced across all feature categories. Overall, this suggests that search performance primarily relies upon handshape and movement features.

3.4 Reducing Handshape Features

Next, following our observation that handshape and movement features largely drive retrieval performance, we sought to design reduced feature sets for each of these categories. First, we focused on handshape features. We experimented with three different methods to reduce the number of handshape features from the 41 handshapes in the *Stokoe features* –or the *Stokoe Handshapes* –while trying to maintain search accuracy and linguistic intuitiveness: (1) clustering different variations of an alphabetic handshape together, (2) clustering features by visual similarity, and (3) using a smaller set of combinatorial features instead of disjoint ones.

No Variation. The baseline handshape features include multiple versions of handshapes from the manual alphabet but with different flexions and spreads (e.g. 'b', 'open b', 'bent b'). We decided to group different variations of the same manual alphabet handshape together into one feature (e.g. 'b', 'open b', and 'bent b' are grouped into just 'b'), reducing 41 handshapes down to 26. Throughout this paper, we refer to this set of reduced handshape features as *No Variation* .

Clustering by Visual Similarity. Drawing inspiration from previous feature sets that group handshapes by visual similarity [28], we designed a reduced feature set where handshapes were clustered by visual similarity. An author who is a native ASL signer proposed an initial clustering, and these clusters were reviewed by several other fluent signers and ASL learners for feedback. After several iterations of feedback, we finalized nine clusters: fist (zero fingers), one finger, two fingers, three fingers, four fingers, five fingers, pinching, and curved/round. After adding an additional "other" feature in the case a handshape wasn't described, this reduced the original 41 handshapes down to ten. Throughout this paper, we refer to this set of reduced handshape features as *Clustered Handshapes* .

Clustering by Number of Fingers. During the pilot of our user study, some participants gave feedback that the pinching and curved/round features were confusing. Following this feedback, we additionally evaluated the *Clustered Handshapes* features excluding the pinching and curved/round features, reducing the ten handshapes in *Clustered Handshapes* to six. Throughout this paper, we refer to this set of reduced handshape features as *Number of Fingers Handshapes* .

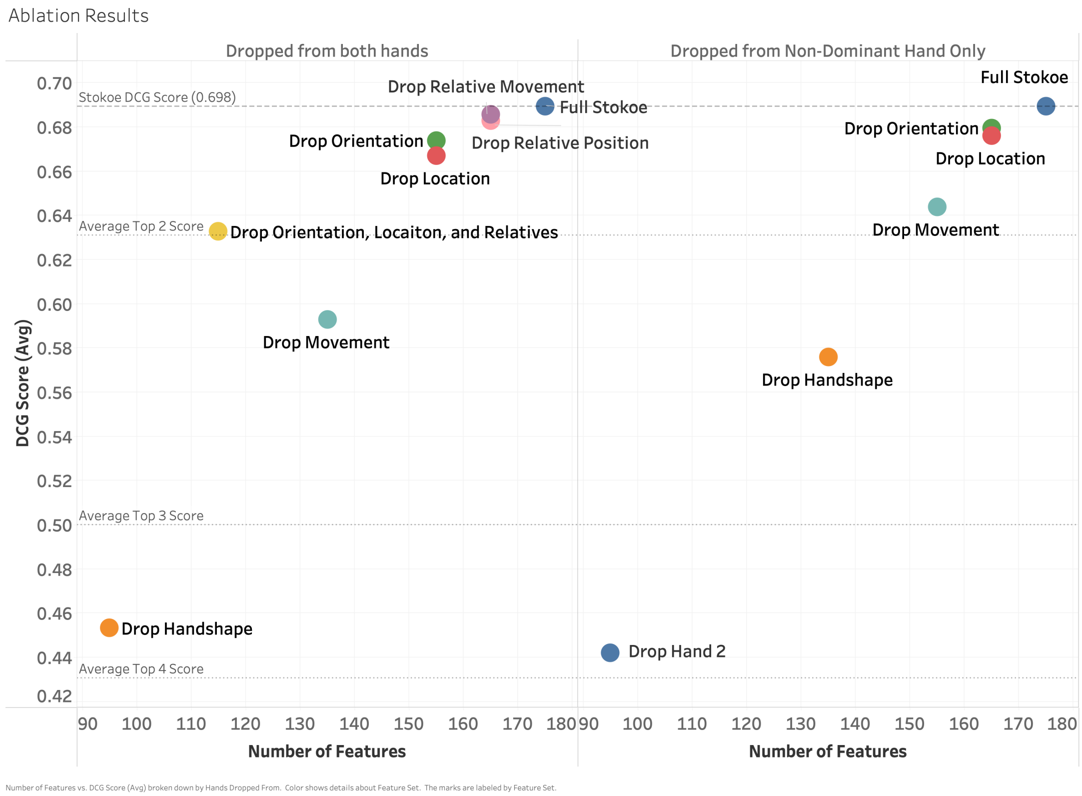


Figure 1: The average Discounted Cumulative Gain (DCG) performance for feature sets when each feature category was removed, when hand 2 and all of its encoded features were removed, and when the orientation, location, relative position, and relative movement features were removed.

Combinatorial Features. We also experimented with reducing handshape features using a smaller set of features that could combinatorially describe any handshape, as opposed to all of the previous sets of features which are disjoint. This smaller set of combinatorial handshape features were based on the theoretically guided phonological coding system used in ASL-LEX 2.0 [8, 29], designed to capture a large amount of information while minimizing coding variability and effort. This reduction strategy reduced 41 handshapes to 14. To describe a handshape using these features, the user must identify "selected" fingers, and the flexion, spread, and thumb contact of those fingers, in addition to thumb position. Throughout this paper, we refer to this set of reduced handshape features as *Selected Fingers Handshapes* . For more information about the rationale and development behind these features, refer to [29].

Figure 2 shows the effect of replacing the baseline handshape category with each of our proposed reduced feature sets on test set retrieval accuracy. We observe that all reduced feature sets are effective at maintaining accuracy relative to the old *Stokoe Handshapes* features, The *No Variation* feature set was most effective at maintaining performance, dropping by only 0.01 DCG, followed by the *Clustered Handshapes* (dropping by 0.05 DCG), *Selected Fingers Handshapes* (dropping by 0.08 DCG), and *Number of Fingers Handshapes* (dropping by 0.09 DCG) feature sets. The *Number of Fingers Handshapes* feature set was most effective at reducing the number of features, followed by *Clustered Handshapes* (dropping by 34%), *Selected Fingers Handshapes* (dropping by 29%), and *No Variation* (dropping by 17%).

3.5 Reducing Movement Features

Next, we started with the 17 movements in the *Stokoe features* – *Stokoe Movements* –and designed reduced movement feature sets like we did with the handshapes.

Movements Clustered by Visual Similarity. Similarly to how we clustered handshapes by visual similarity, an author who is a native ASL signer proposed initial clusters movements by similarity. After several iterations of feedback from several other fluent signers and ASL learners, we finalized the following clusters: vertical movements, sideways movements, movements towards or away from the signer movements, and circular movements. With an additional "other" feature, this reduced the original 17 movements down to 5. Throughout this paper, we refer to this set of reduced movement features as *Clustered Movements* .

Replacing Non-Dominant Hand with Binary "Is Moving". We additionally experiment with a variant of the *Clustered Movements* feature set by replacing all movement features in the non-dominant hand with a single binary feature indicating whether that hand is moving or not. This was done based on the Dominance and

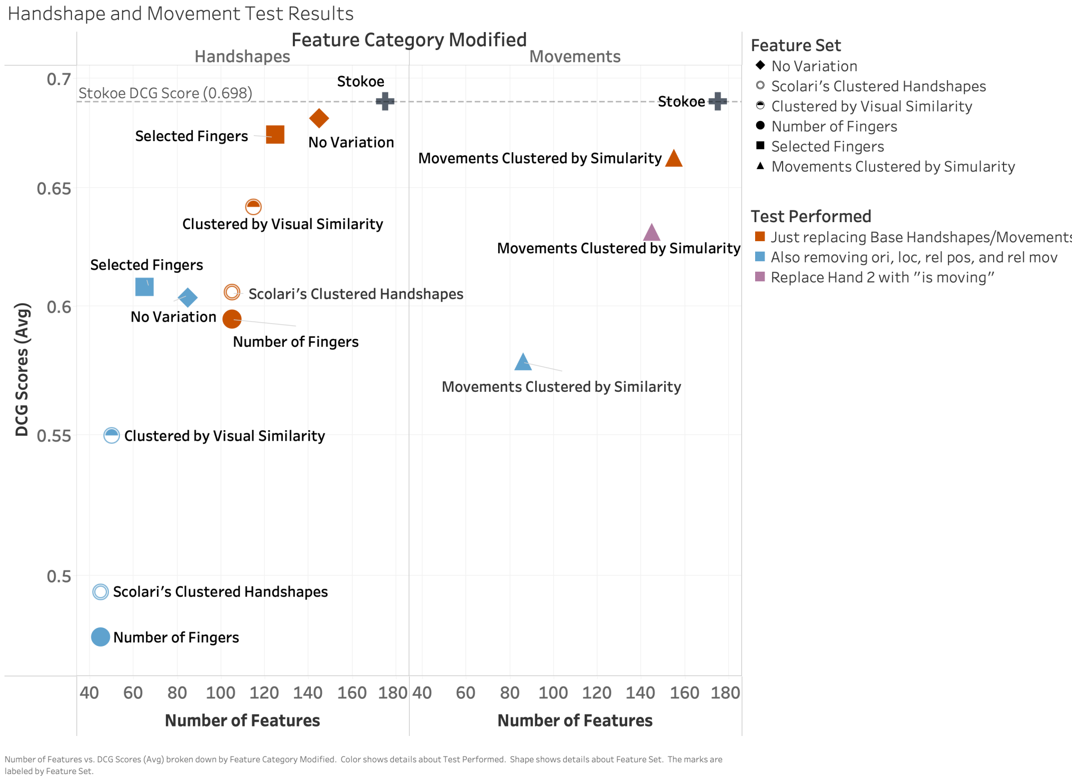


Figure 2: A summary of the average Discounted Cumulative Gain (DCG) performance for our reduced handshape feature sets. On the left shows the average DCG performance of our reduced handshape feature sets (*No Variation* , *Clustered Handshapes* , *Number of Fingers Handshapes* , and *Selected Fingers Handshapes*), while the right shows the average DCG of our reduced movement feature set (*Clustered Movements*). DCG performance of the original Stokoe handshapes and movements are also shown. Blue dots show the average DCG performance of each reduced feature set when orientation, location, and the relative features are also dropped.

Symmetry Conditions, which implies that for all sign types in ASL (one-handed, two-handed symmetrical, two-handed asymmetrical with same handshapes, etc), the movement of the non-dominant hand can be inferred from the movement of the dominant hand.

Figure 2 and Table 2 shows the effect of replacing the baseline movement category with our proposed reduced feature sets on test set retrieval accuracy. We observe that clustering movements by visual similarity (*Clustered Movements*) was also effective to maintaining accuracy, dropping accuracy only by 0.03 DCG while reducing the number of features by 11%. When replacing the non-dominant hand movement features with a single binary feature—*is moving*—we found that the search accuracy only drops by 0.06 DCG from baseline with a 17% reduction in the number of features.

3.6 Removing Location, Orientation, and Relative Features

In our previous ablation study, we observed that removing location, orientation, and relative features from the baseline features only negligibly impacted retrieval performance. We next sought to confirm that removing these feature sets also negligibly impact retrieval performance when using our reduced feature sets for handshape and movement. These results are shown in Figure 2 and Table 3.

After removing the location, orientation, relative position, and relative movement features, we observe that *Selected Fingers Handshapes* was the most effective at maintaining performance, dropping by only 0.08 DCG, followed by the *No Variation* (dropping by 0.09 DCG), *Clustered Handshapes* (dropping by 0.14 DCG), and *Number of Fingers Handshapes* (dropping by 0.29 DCG). Dropping the location, orientation, relative position, and relative movement features from both hands reduced the number of features by 35%. *Clustered Movements* did very well maintaining accuracy, only dropping accuracy by 0.11 DCG with a 51% reduction in features.

3.7 Combining Reduced Feature Sets

Since our previous ablation study showed that both handshapes and movements are needed for effective sign retrieval, we sought to explore how retrieval performance was impacted if we were to combine all of our feature reduction strategies for both the handshape and movement features while still dropping the location, orientation, relative position, and relative movement features from both hands. The results are shown in Table 4.

Features Dropped	Number of Features	DCG Scores (Avg)	Percent Drop in Features From Stokoe	Difference in DCG From Stokoe
Stokoe	177	0.689		
No Variation	145	0.682	17%	0.01
Scolari Handshapes	105	0.606	4%	0.08
Handshapes Clustered by Visual Similarity	115	0.642	34%	0.05
Number of Fingers	105	0.595	4%	0.09
Selected Fingers	125	0.674	29%	0.02
Movements Clustered by Similarity	155	0.663	11%	0.03
Movements Clustered by Similarity (with binary ('is moving' for NDH)	145	0.631	17%	0.06

Table 2: A tabular summary of the average DCG performance of simulated ASL-to-English look up using the Stokoe features but with the handshapes or movements replaced by our reduced handshape and movement feature sets.

Features Dropped	Number of Features	DCG Scores (Avg)	Percent Drop in Features From Stokoe	Difference in DCG From Stokoe	Difference in DCG with Ori, Loc, and Rel
Stokoe	177	0.689	-	-	-
No Variation	85	0.604	51%	0.085	0.078
Scolari Handshapes	45	0.495	74%	0.194	0.111
Clustered by Visual Similarity	50	0.55	71%	0.139	0.092
Number of Fingers	45	0.48	74%	0.209	0.115
Selected Fingers	65	0.608	63%	0.081	0.066
Movements Clustered by Similarity	86	0.578	51%	0.111	0.085

Table 3: A tabular summary of the average DCG performance of simulated ASL-to-English look up using the Stokoe features but with the handshapes or movements replaced by our reduced handshape and movement feature sets AND the orientation, location, and relative features dropped.

We observe that after removing the location, orientation, and relative features, only using the *Selected Fingers Handshapes* and *Clustered Movements* features for the dominant hand and the *Selected Fingers Handshapes* and binary "is moving" feature for the non-dominant hand is the most effective at maintaining retrevial accuracy, dropping only by 0.14 DCG. When using the *Clustered Handshapes* instead, DCG dropped by 0.21 DCG. When using *Number of Fingers Handshapes* , DCG dropped by 0.27 DCG. In terms of reducing the number of features the most, applying all of these feature reduction strategies while using the *Number of Fingers Handshapes* had the largest reduction in features (89%), followed by *Clustered Handshapes* (86%), then followed by *Selected Fingers Handshapes* (79%).

Table 4: A summary of the average DCG performance when combining our different feature reduction strategies.

Handshapes Used	Movements Used for DH	Number of Features	DCG Scores (Avg)	Percent Drop in Features from Stokoe	Difference in DCG from Stokoe
Stokoe	Stokoe	177	0.689		
Stokoe	Clustered by Similarity	86	0.663	0.11	0.03
Clustered by Similarity	Clustered by Similarity	25	0.482	0.86	0.21
Number of Fingers	Clustered by Similarity	19	0.424	0.89	0.27
Selected Fingers	Clustered by Similarity	37	0.557	0.79	0.14

4 User Study of Human Preferences

In the previous section, we found that accurate ASL-to-English dictionary search only requires reduced sets of handshape and movement features. To explore the usability of our reduced sets of handshapes and movements with ASL learners, we ran an IRB approved user study where participants experienced inputting dictionary queries with different feature sets and provided feedback on the experience.

4.1 Participants

We recruited 75 participants in total through relevant email lists, social media, and snowball sampling. Out of the 75 participants, one participant preferred to provide no demographic information. A majority of our participants (58, 77%) were ASL learners who identified as hearing. The remaining 17 participants (23%) identified either as DHH (12, 14%), a child, sibling, or grandchild of d/Deaf adults (3, 4%), a hearing parent of a Deaf child (1, 1%), or interpreter in training (1, 1%). A majority of participants self-identified as female (54, 72%). The majority of participants were in their early 20 to mid 20s, with the average age being 25 years and the median age being 21 years (min = 18, max = 50). A majority identified as White (48), while 17 identified as Asian, 1 identified as Black or African American, 8 identified as multi-racial, and 3 preferred not to answer.

All participants reported having experience with ASL, with varied levels. 28 participants had one year or less of ASL experience (37%), 19 participants had one to four years of ASL experience (25%), and 28 participants had more than 4 years of experience

(37%). When asked, only six of the seven participants who self-identified as d/Deaf reported that they used ASL as their primary form of communication (8%), with the remaining participants considering it a secondary form of communication (69, 92%).

Participants were also asked about their prior experience with ASL-to-English dictionaries. A large majority of participants noted that they had previous experience with using a digital or physical tool to search for the meaning of a sign in ASL. 59 participants (79%) mentioned that they had previously used a physical or digital tool to look up the meaning of a sign. 35 of those participants (47%) had previously only used a digital dictionary to do so, three (4%) only had used a physical dictionary, and 17 (22%) had used both digital and physical dictionaries. Participants were given an option to specify which resources they had prior experience with searching for signs, but only 11 out of 59 responded to this. Six of those 11 participants mentioned that they did not use online dictionaries, but instead inputted sign descriptions into Google and browsed video results until they saw one that matched their intended query (which often required trial and error and didn't always help them find the sign they were looking for). The three of the 11 participants who mentioned a reverse dictionary resource only mentioned Handspeak's reverse dictionary, which only allows participants to search for signs by handshape (requiring them to scroll through a large list to find the desired sign). Three of the 11 participants cited resources that only support English-to-ASL, suggesting that some participants may have misunderstood the question.

4.2 Procedure

We conducted a Qualtrics survey with novice and experienced ASL users to evaluate the usability and intuitiveness of our reduced handshape and movement sets. Participants were screened for basic ASL proficiency prior to receiving the survey link. The survey took about 30 minutes to complete and participants were monetarily compensated for their time.

We selected three sets of reduced features from our simulation study for our user study: *Number of Fingers Handshapes* , *Selected Fingers Handshapes* , and *Clustered Movements* . We chose to evaluate only some of the feature sets from our simulation study to ensure the user study did not exceed 30 minutes. We excluded *Clustered Handshapes* features because these were a variant of the *Number of Fingers Handshapes* features with similar performance in simulated retrieval, but more features. We also excluded the *No Variation* features because they had the lowest performance in our simulation study.

For each reduced feature set, participants were shown videos of three ASL signs and asked to describe either their handshapes or movements (3 signs × 5 feature sets). To compare against the original *Stokoe Handshapes* and *Stokoe Movements* features, participants were asked to do the same task using the *Stokoe Handshapes* and *Stokoe Movements* features. An example of the interface participants used to input handshape and movement features is shown in Figure 3

To provide a consistent experience across feature sets, we used the same three signs for all the handshape feature sets (LOOK-FOR, ART, and WAKE-UP) and a separate set of three signs for all the movement features (LIKE, DISCUSS, and STEPBROTHER). For the three signs shown to participants to describe handshapes, we chose a one-handed sign, a two-handed sign, and a sign that had multiple handshapes. The sign we picked for each of the three types was chosen arbitrary from introductory ASL vocabulary, and spanned varied handshapes, movements, and complexity. None of the signs had the same handshape, which was done on purpose to give participants the opportunity to test the systems with as many different handshapes as possible. We did the same for the three movement signs, and none of the signs had the same movement for the same reason. The order in which users experienced each feature set and each of the three signs was counterbalanced. For one-handed signs, participants were only asked to input features for the dominant hand alone, as we assume they wouldn't normally bother specifying what the second hand is doing for a one handed sign when inputting the sign in real life. For two-handed signs and signs with multiple handshapes, participants were asked to input features for both the dominant and non-dominant hands.

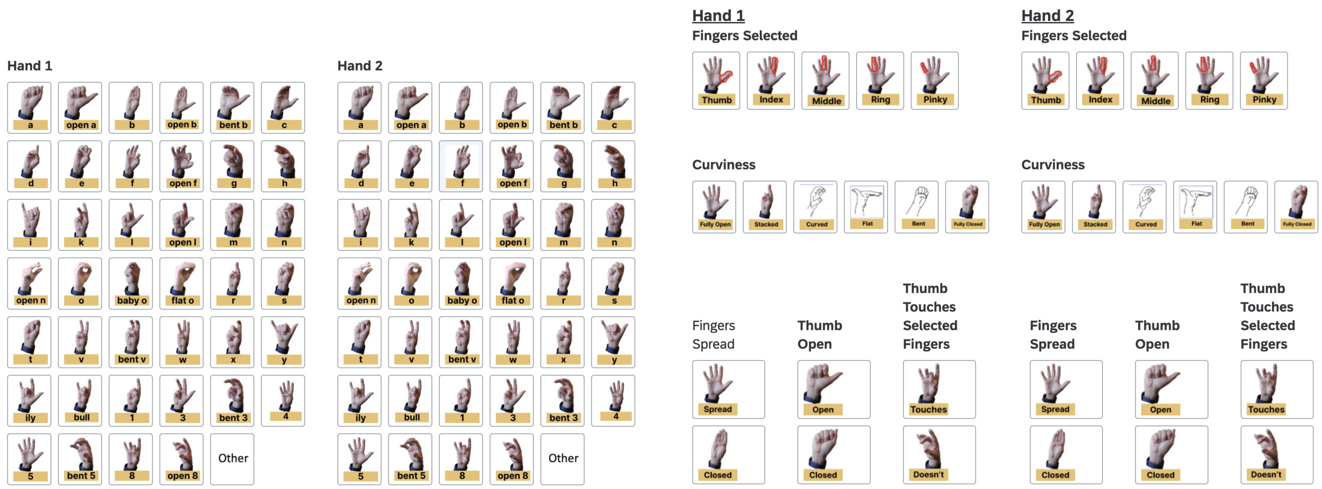
Usability was primarily measured using a subset of System Usability Scale (SUS) questions, a standardized questionnaire that assesses user satisfaction and perceived ease of use through ten statements rated on a five-point Likert scale [9] After experiencing each handshape or movement feature set, participants were asked to rate its usability through our SUS questions and had the opportunity to provide open feedback. We also measured how long it took each participant to input the handshapes or movements.

4.3 Results

In this section, we present the results of our user study: the perceived usability of each feature set, their timed efficiently, and common themes found in the open feedback. We find that in terms of describing handshapes, *Stokoe Handshapes* and our new *Clustered Handshapes* features were generally preferred to the *Selected Fingers Handshapes* features. In terms of describing movements, the *Clustered Movements* features were generally preferred over the *Stokoe Movements* features and found to be more efficient.

4.3.1 SUS Responses. When prompted for agreement with the statement "I think that I would like to use this system frequently" for handshapes, participants reported they were most likely to use the *Stokoe Handshapes* features [$\mu = 3.893, \sigma = 1.021$], slightly less likely to use our *Clustered Handshapes* features [$\mu = 3.48, \sigma = 0.935$], and least likely to use *Selected Fingers* features [$\mu = 2.667, \sigma = 1.212$], with significant difference between all three sets of features [$X^2(2) = 38.85$, $p < 0.001$, Kruskal-Wallis]. For the movement feature sets, participants reported a preference for using *Clustered Movements* [$\mu = 3.32, \sigma = 1.187$] over *Stokoe Movements* [$\mu = 3.053, \sigma = 1.229$], with significant difference [$p < 0.05$, Mann Whitney].

In response to the prompt "I found this system unnecessarily complex" (SUS Q2), participants generally rated *Stokoe Handshapes* and *Clustered Handshapes* highest for handshape features, and *Clustered Movements* highest for movement features (see Fig4b). Participants generally disagreed (77% responded 2 or 1 on the likert scale) that *Stokoe Handshapes* [$\mu = 2.013, \sigma = 0.966$] and *Clustered Handshapes* [$\mu = 2.12, \sigma = 0.972$] were unnecessarily complex, while



(a) A screenshot of the user interface participants used to input *Stokoe Handshapes* features in the user study.

(b) A screenshot of the user interface participants used to input *Selected Fingers Handshapes* features in the user study.

Figure 3: An example of the user interface that each participant was asked to input features using. While this only shows the user interface for inputting handshapes using the *Stokoe Handshapes* and *Selected Fingers Handshapes* features, the *Number of Fingers Handshapes* , *Stokoe Movements* , and *Clustered Movements* feature sets used similar-looking interfaces.

participants were split on *Selected Fingers Handshapes* being unnecessarily complex [$\mu = 3.267, \sigma = 1.189$]. The difference in perceived complexity between the handshape feature sets was significant [$X^2 = 41.16, p < 0.001$, Kruskal-Wallis], as was the difference for movement feature sets [$p < 0.001$, Mann Whitney].

When asked how much they agree with "I thought this system would be easy to use" (SUS Q3), participants again generally rated *Stokoe Handshapes* and our *Clustered Handshapes* highest for handshape features, and *Clustered Movements* highest for movement features (see Fig4b). Most participants agreed (4 or 5 on the likert scale) that *Stokoe Handshapes* [$\mu = 3.987, \sigma = 0.937$] and *Clustered Handshapes* [$\mu = 3.907, \sigma = 0.903$] were easy to use, while disagreeing that *Selected Fingers Handshapes* features were easy to use [$\mu = 2.52, \sigma = 1.044$]. For movements, participants generally found our *Clustered Movements* easier to use [$\mu = 3.773, \sigma = 1.06$] compared to the *Stokoe Movements* features [$\mu = 3.027, \sigma = 1.127$]. This difference in perceived easiness between the handshape feature sets was significant [$X^2 = 70.79, p < 0.001$, Kruskal-Wallis], as was the difference between in perceived ease between movement feature sets [$p < 0.001$, Mann Whitney].

In response to the SUS prompt "I would imagine that most people would learn to use this system very quickly," participants equally agreed that the *Stokoe Handshapes* and our *Clustered Handshapes* were quick to learn [$\mu = 4.067, \sigma = 0.963$ and $\mu = 3.827, \sigma = 0.95$ respectively]. However, participants felt that our *Selected Fingers Handshapes* features were not quick to learn [$\mu = 2.72, \sigma = 1.047$] with significant difference [$p < 0.001$, Kruskal-Wallis]. The difference between perceived learnability between the handshapes was significant [$X^2 = 61.39, p < 0.001$, Kruskal-Wallis]. For movements, most participants found *Clustered Movements* quick to learn (rated 4 or 5) [$\mu = 3.733, \sigma = 1.06$] with significant difference [$p < 0.001$, Mann Whitney] compared to the *Stokoe Movements* features [$\mu = 3, \sigma = 1.078$].

4.3.2 Efficiency of Each Feature Set. These trends in perceived complexity are also consistent with how long it took participants to use each feature set. On average, participants took longer to input signs using *Selected Fingers Handshapes* [$\mu = 36.33$ seconds, $\sigma = 19.44$ seconds] than *Stokoe Handshapes* and *Number of Fingers Handshapes* , which took around same amount of time to input dominant hand features [$\mu = 22.56$ seconds, $\sigma = 16.95$ seconds and $\mu = 20.06$ seconds, $\sigma = 14.17$ seconds respectively]. The difference in time to use the three different handshape systems was significant for both the dominant hand [$X^2(2) = 123.66, p < 0.001$, Kruskal-Wallis] and non-dominant hand [$X^2(2) = 111.81, p < 0.001$, Kruskal-Wallis]. Interestingly, *Stokoe Handshapes* and *Number of Fingers Handshapes* did not appear to be faster for inputting WAKE-UP, a sign with multiple handshapes. *Non-dominant hand* features also took slightly longer to input using *Stokoe Handshapes* than *Clustered Handshapes* [$\mu = 9.89$ seconds, $\sigma = 5.71$ seconds and $\mu = 7.19$ seconds, $\sigma = 3.24$ seconds respectively].

On average, it took longer to input both dominant and non-dominant hand features using *Stokoe Movements* [$\mu = 25.95$ seconds, $\sigma = 14.41$ seconds and $\mu = 8.82$ seconds, $\sigma = 5.66$ seconds respectively] than *Clustered Movements* [$\mu = 7.91$ seconds, $\sigma = 4.29$ seconds and $\mu = 5.78$ seconds, $\sigma = 3.06$ seconds respectively], with statistical significance [$p < 0.001$, Mann Whitney].

4.3.3 Common Themes in Open Feedback. Some participants noted in their open feedback that the *Stokoe Handshapes* features had "too many options" (P23) and that they liked the simplicity of *Clustered Handshapes* . Some participants mentioned that because they had never seen the *Clustered Handshapes* features before, they required a learning curve, but were easier to use once learned. Similarly, some participants liked the simplicity of *Clustered Movements* , mentioning that it was "faster," "less repetitive" (P43), and "less confusing" (P47, P51). This trend of wanting simplicity can also be seen in the open feedback given for *Selected Fingers Handshapes* . While participants felt that *Selected Fingers Handshapes* allowed for more detailed description, they often noted that it required too much prior linguistic knowledge (e.g. understanding what it means for fingers to be "selected") to be efficiently used (P6, P40, P20, P50, P71).

On the other hand, some participants felt that *Clustered Handshapes* may be "a little too general" (P27), citing specifically that it wasn't always clear which clustered feature aligned most with certain handshapes (P8, P65, P42, P64, P7). Some participants also preferred *Stokoe Movements* over *Clustered Movements* for the same reason: "It was easier to understand the movement labels [using the *Clustered Handshapes* features], but hard to classify the signs into different movements" (P40). Some participants suggested a middle ground between *Stokoe Handshapes* and *Clustered Handshapes* : "I think this one has too many choices. Somewhere between this and the previous that had five options would feel best to me" (P68).

For inputting signs that have multiple handshapes or movements, several participants suggested that it would "make things clearer" (P33) if the interface allowed you to specify which handshapes or movements were at the beginning of the sign and which were at the end, which existing feature-based dictionaries do not currently support. This was most apparent with the *Selected Fingers Handshapes* features, which were commonly noted as being not suitable for describing signs with multiple handshapes: "I especially think this system is difficult when the handshape changes during the sign [...]" (P55). With *Selected Fingers Handshapes* , a user selects multiple features (fingers) to describe a single handshape, whereas with the other feature sets a user select a single feature (a handshape or set of handshapes) to describe a single handshape. Because *Selected Fingers Handshapes* already requires multiple selections to describe a single handshape, the representation of multiple handshapes became particularly muddied. This is apparent in P9's open feedback: "I wasn't sure what to do for the last sign, that had the hands going from one shape to another - which shape do I try to describe? Or do I just try to do both?" (P9).

Some participants noted that using the *Selected Fingers Handshapes* features to describe handshapes "requires too much cognitive thinking," (P71) and "felt complicated and a bit confusing when having to determine which finger was used vs not" (P40). The general consensus amongst participants was that *Selected Fingers Handshapes* features were not necessarily "complex[, but just] a little unclear" (P64). Ambiguities are apparent in P6's open feedback: "Say the hand is in a 'C' shape, would I select every finger because every finger is 'in use' per say? Or would I just select the 'C' hand shape and leave it at that?" (P6). P6 touches on two common confusions that participants had with *Selected Fingers Handshapes* : (1) the system depends on the user understanding which fingers are "selected fingers" in a sign (a concept that requires some background or instruction) and (2) since the system allows for multiple possible combinations to describe the same sign, will they all return my desired result if I were to use this system for dictionary search?

5 Discussion

In this work, we designed reduced feature sets for sign language dictionary retrieval that maintain retrieval accuracy while being rated as usable in a user study. In our simulation study, we found that only reduced sets of handshape and movement features are needed for effective ASL-to-English retrevial. In our user study, we found that these reduced sets of handshapes and movements smaller can improve the user experience in some cases. In this section, we discuss the higher-level value of our findings in both the simulation and user studies, as well as limitations and opportunities for future work.

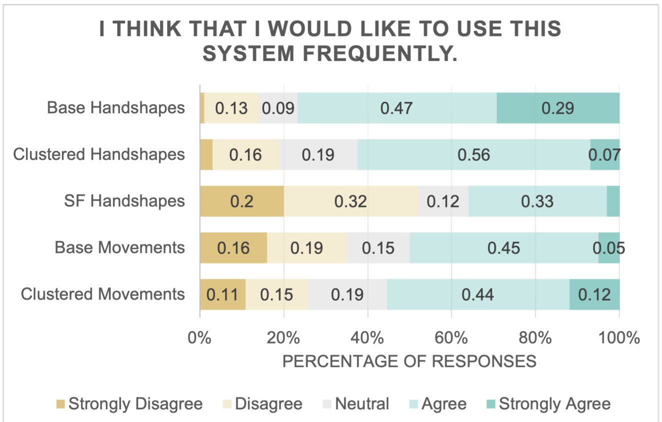
5.1 Simplifying Accurate Feature-Based Sign Language Interfaces

Our simulation study suggests that it is possible to do effective sign language search with a smaller amount of intuitive features using a simple and efficient machine learning algorithm (indexing via LSA). We propose multiple methods for feature reduction that can be used individually or in combination, with the removal of the location, orientation, relative position, and relative movement being arguably the most beneficial change. For a 34% reduction in the number of features, you significantly simplify sign language search by reducing the number of categories that the user needs to think about from six to two while still maintaining top two accuracy on average for returned search results. Additionally incorporating the reduced sets of features we propose for these two remaining categories–handshapes and movements–reduces the number of features by 89% (177 to 19 features) and introduces an effective digital ASL-to-English dictionary that maintains top-4 accuracy on average while using the handshapes (*Number of Fingers Handshapes*) and movements (*Clustered Movements*) found most usable in our user study (or above top-2 accuracy on average with features if using equally preferred *Stokoe Handshapes* instead). There is also more opportunity for both researchers or builders of a digital ASL-to-English dictionary interface to play around with these feature reduction strategies and explore different tradeoffs (e.g. if more accuracy is needed, movements can be added back into the non-dominant hand).

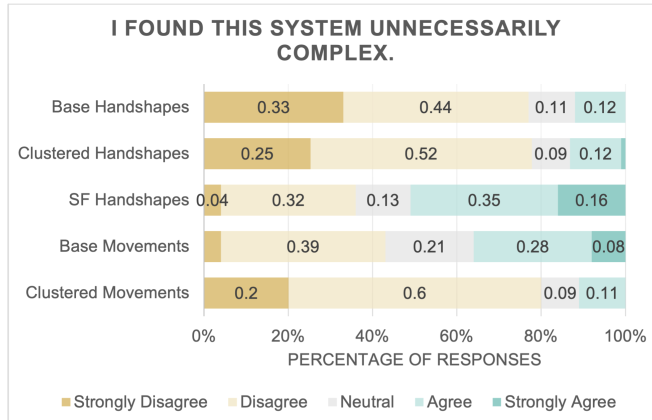
While combining this change with our other feature reduction methods, like our reduced sets of handshape and movement features, does reduce the accuracy of the search results between being in the top two to three results on average (especially after additionally removing location, orientation, and the relative features), there is previous work to show that users are still satisfied with sign language dictionary systems as long as the result they are looking for still appears above the fold on the first page of results [1]. While it depends on the interface how many search results can fit above the fold in a sign language dictionary search system, it is reasonable to assume that the first three results will always appear above the fold.

5.2 User Study

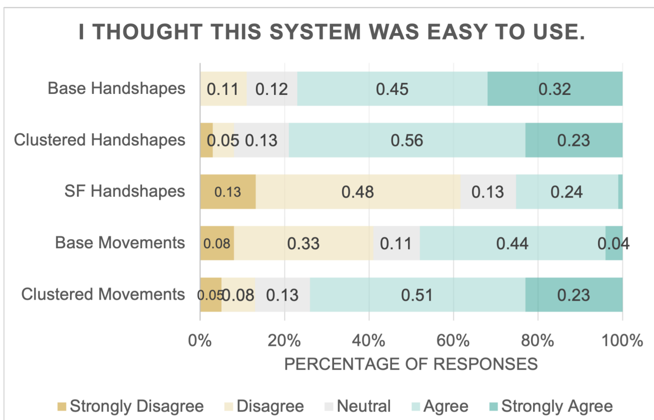
Throughout the user study, participants responded very similarly for our *Clustered Handshapes* and the old *Stokoe Handshapes* features, while clearly having less preference for the *Selected Fingers*



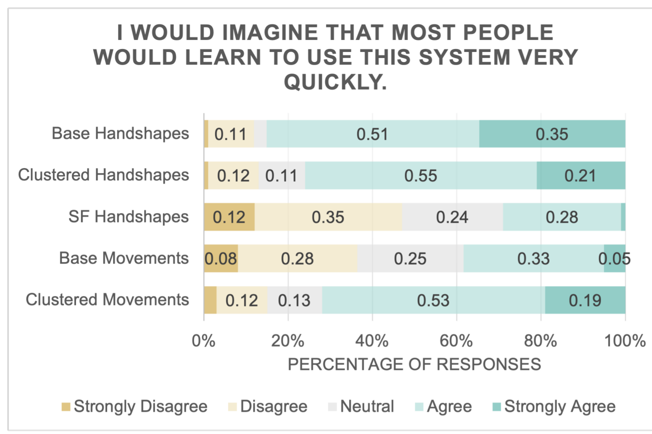
(a) A summary of how frequently participants would like to use each feature set.



(b) A summary of perceived complexity for each feature set.



(c) A summary of how easy each feature set felt to participants.

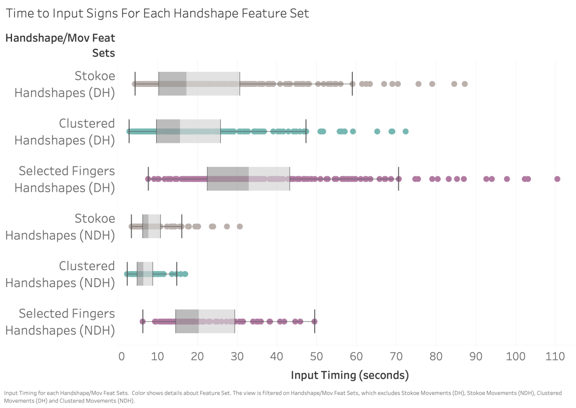


(d) A summary of the how learnable each feature set felt to participants.

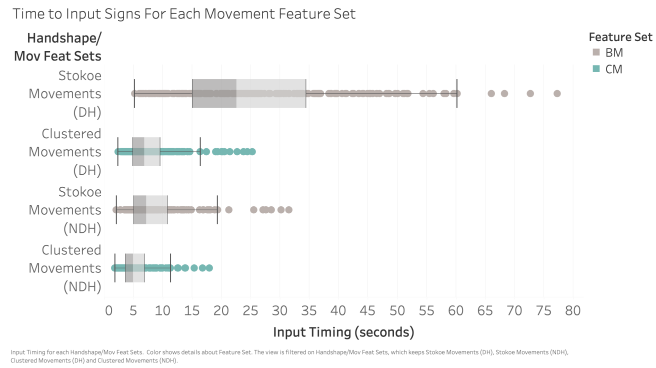
Figure 4: A summary of likert responses from participants when asked about the usability across the three handshape feature sets and two movement feature sets.

Handshapes . It is possible that similarity in preference for participants between the *Clustered Handshapes* and *Stokoe Handshapes* features comes from familiarity bias towards *Stokoe Handshapes* . ASL learners and fluent signers have often already been exposed to the Stokoe handshapes due to their common use in many ASL-to-English dictionaries [22, 37] and the fact that many of the Stokoe handshapes are made up of the A-Z letter and classifier handshapes, both of which are commonly taught in ASL curricula [31]. Several participants directly mentioned this bias was present their open feedback. Because participants weren't as familiar with the *Clustered Handshapes* features, they often were unsure about which handshapes in a sign belonged in which clusters, or didn't realize that the dictionary search system was robust enough to allow for uncertainty. It may be possible that if given adequate time to become familiar with the *Clustered Handshapes* features, signers may ultimately prefer them over the *Stokoe Handshapes* features due to simplicity. Amongst the handshape feature sets evaluated, participants commonly noted that the *Selected Fingers Handshapes* features had the least preference due to its complexity, highlighting that less features might not always mean simpler and users of an ASL-to-English dictionary prefer simplicity over descriptive power.

The apparent similarity in overall preferences may not indicate equal favorability for both the *Stokoe Handshapes* and *Clustered Handshapes* among all participants, but rather a division in preference. Approximately half of the participants preferred the greater specificity of the *Stokoe Handshapes* , while the other half favored the simplicity of the *Clustered Handshapes* . This divergence, along with some participants mentioning that they would prefer some kind of balance between the two feature sets (i.e. not as simple as the *Clustered Handshapes* but not as detailed as the *Stokoe Handshapes*), suggests that exploring a hybrid handshape system—one that still clusters by visual similarity but incorporates a greater number of clusters—may be a worthwhile direction for future research. One possible direction to explore was suggested by P40: "I didn't like how every choice was grouped into a huge clump that you had to look through. I'd rather a combination between the



(a) Timing results (in seconds) for the original *Stokoe Handshapes* and our proposed *Clustered Handshapes* and *Selected Fingers Handshapes* feature sets.



(b) Timing results (in seconds) for the original *Stokoe Movements* and our proposed *Clustered Movements* feature sets.

Figure 5: A summary of how long it took each participant to input features (in seconds) for each feature set evaluated in the user study. The timing results for features inputted for the dominant hand (DH) and non-dominant hand (NDH) are separated.

first and second systems. (Broken up into how many fingers, but naming each specific handshape within those options)." The alternative, smaller handshape and movement feature sets introduced in this paper, along with the prototyped interfaces designed for sign language dictionary search in the user study, are intended not as final solutions but as foundational steps for further refinement and adjustment. Because participants were new to the *Clustered Handshapes* , *Selected Fingers Handshapes* , and *Clustered Movements* features, feedback from participants indicate that it may be beneficial to incorporate tutorials and examples for how to use each feature set if implemented in a digital ASL-to-English dictionary.

5.3 Limitations and Future Work

There were some limitations with our simulation study. Since all of the 6,180 annotations we collected were created only using the 177 *Stokoe features* features, all of the smaller sets of features that we developed in part one were created by linearly mapping each of the 6,180 annotations x to every new feature y in the alternate set of features. This meant that the feature sets that we created were constrained to features that could only be derived from the 177 *Stokoe features* features. There is potential for future work to explore creating smaller sets of intuitive handshape and movement features for ASL-to-English dictionary search, or digital sign language input generally, that are not constrained by the *Stokoe features* features.

In our ASL-to-English dictionary retrieval simulations, we only used a vocabulary of 1145 signs, which is much smaller than the full vocabulary of a practical ASL-to-English dictionary. Future work is needed to explore if these trends continue with a more real-world vocabulary set. Additionally, even though we use annotations from real ASL learners, we predict the average search accuracy of each set of features via simulation and not from real users.

There have been no previous feature-based sign language dictionary systems that have explored the use of auto-complete–giving suggestions for the sign you are re looking for while you input the signs. This is another potential direction of research in making digital sign language dictionaries easier to use.

Additionally, while not the focus of this study, reducing the number of features may have implications for improving the usability of ASL-to-English dictionaries on mobile devices, as smaller feature sets could facilitate faster, more accessible look-ups. However, further research is needed to explore the specific impact on mobile interface design.

It is also important to note that the participants' preferences and efficiency with some of the handshape and movement feature sets evaluated may have also been impacted by limitations in our design of the input interfaces used for them in the user study. When designing the input interface for the *Clustered Handshapes* features in the User Study, we gave examples of what commonly seen handshapes might belong to each cluster, but chose not to include an example for every possible handshape as to not introduce too much clutter. This led to some participants expressing confusion about what to do if a handshape wasn't listed in any of the examples: "I didn't see an L handshape so it caused a bit of confusion/hesitation when selecting the HS" (P41). This highlights potential future work exploring how to best design an interface for handshape features that are clustered visually and the importance of some kind of tutorial when designing search interfaces using visually clustered handshape features.

Additionally, a common point of confusion across all five feature sets was that participants were not interacting with the full sign search dictionary interface, but rather with isolated handshape or movement features. As one participant noted, "I would use this system frequently if it was combined with a system that described handshape" (P46), suggesting that the lack of an integrated search interface may have impacted their experience. This confusion may have impacted participants' responses for certain feature sets during the user study, although ideally evenly across all feature sets.

Along with the user interface, responses in the user study may have also been affected by the limited number of signs that we evaluated for each feature set. Further work needs to be done to see if all three handshape feature sets are generally more difficult to use with more complex signs.

We also acknowledge that for the two main groups who find sign language dictionary search systems useful—sign language learners and sign language experts/natives—the sample in the user study was predominately ASL learners, with only 15 participants (20%) either DHH or in the Deaf community and 28 participants (37%) having four or more years of ASL experience. Dictionaries are important to both sign language learners as well as experts, including both perspectives is important– however, ASL novices and experts may use dictionaries for different reasons and may recall signs differently based on their expertise. Therefore, we believe expert's experiences might not generalize to novices and vice versa. In our work, we aimed to collect perspectives from both groups but our sample was not large enough to analyze the differences between the two systematically – we suggest this as a promising direction for future work.

Our findings in this work can extend beyond just American Sign Language. It is often the case that different sign languages have different sets of commonly used handshapes and movements–or phonemic inventories. Sign language character systems have been meticulously designed over decades to generalize across many sign languages. For example, SignWriting includes characters that represent handshapes, movements, non-manual markers, dynamics and timing, and more from across over forty different sign languages [36]. However, their significant learning curve limits their practicality for use in sign language dictionary search. Our proposed reduced handshape and movement feature sets may actually be abstract enough to describe all possible handshapes, regardless of what sign language and phonemic inventory is being used. We recommend future work explore how well our proposed reduced handshape and movement feature sets generalize to other sign languages. These contributions can also be extended beyond just ASL-to-English dictionaries to any interface that requires feature-based sign language input, including any sign language to spoken language dictionary and sign-language to other sign language dictionaries.

5.4 Ethics

Our work was motivated by the need to create feature sets that were easy use in the context of sign language dictionary search. This required we explore how to reduce feature sets, i.e., drop features. However, we caution against doing so in other sign language tasks such as translation, where the context offered by each of the features is even more important. For a dictionary context involving isolated signs, using the full set of features (handshapes, movements, locations, and orientations) for both hands doesn't fully uniquely identify our 1,145 signs. In translation context with continuous signing, recognition further complicated by coarticulation of signs (i.e., signs are influenced by their neighbors) and sentence structure. Dropping features may then adversely impact the accuracy in translation contexts.

Additionally, while our user study found that both sign language learners and experts alike had preference towards the specific reduced feature sets that we developed in this paper, we also found that there was variability in what features people found important. Different people may find different features important for search, and it is important to consider this in the development of not just ASL dictionary tools, but sign language dictionary tools generally. If Deaf community members want to customize the features used in dictionary search or even create their own feature sets, it is important that they are given the power to do so. We envision future ASL search interfaces, and sign language search interfaces generally, being adaptive or allowing for customization based on user preference, and we see our work as preliminary work in this space through the methodology we explored in creating our reduced ASL feature sets.

6 Conclusion

Sign language dictionaries are important for language learning and documentation, but there is currently no easy way to look up signs for sign languages. Feature-based dictionaries help overcome this challenge, but have previously been cumbersome due to their use of large numbers of complex features.

In this work, we explored different methods of reducing the set of features commonly used for feature-based sign language search and not only found that only handshapes and movements were needed for effective search, but effective search could be maintained with smaller sets of handshape and movement features. We then explored user preference with these smaller handshape and movement features and found that some of them improved the user experience in some cases. As a next step, we envision the development of a well-supported digital ASL-to-English dictionary that incorporates the findings in this work, along with previous works to create an ASL-to-English that is used widely by both the ASL learning community and the native ASL community.

Acknowledgments

This research, in part, was supported by the Center for Research and Education on Accessible Technologies and Experiences (CREATE) at the University of Washington, the Undergraduate Research Internship program at Microsoft Research, New York, and the Mary Gates Research Scholarship from the University of Washington.

References

[1] Oliver Alonzo, Abraham Glasser, and Matt Huenerfauth. 2019. Effect of automatic sign recognition performance on the usability of video-based search interfaces for sign language dictionaries. In *Proceedings of the 21st International ACM SIGACCESS Conference on Computers and Accessibility*. 56–67.

[2] Willaim Woods ASL. 2016. Sign language around the world.

[3] Robert A Augustus, Elsie Ritchie, and Suzanne Stecker. 2013. The official american sign language writing textbook. *Los Angeles, CA: ASLized* (2013).

[4] Robbin Battison. 1974. Phonological deletion in American sign language. *Sign language studies* 5, 1 (1974), 1–19.

[5] Robbin Battison. 1978. Lexical borrowing in American sign language. (1978).

[6] Danielle Bragg, Oscar Koller, Mary Bellard, Larwan Berke, Patrick Boudreault, Annelies Braffort, Naomi Caselli, Matt Huenerfauth, Hernisa Kacorri, Tessa Verhoef, et al. 2019. Sign language recognition, generation, and translation: An interdisciplinary perspective. In *Proceedings of the 21st International ACM SIGACCESS Conference on Computers and Accessibility*. 16–31.

[7] Danielle Bragg, Kyle Rector, and Richard E Ladner. 2015. A User-Powered American Sign Language dictionary. In *Proceedings of the 18th ACM Conference on Computer Supported Cooperative Work & Social Computing*. 1837–1848.

[8] Diane Brentari. 1998. *A prosodic model of sign language phonology*. Mit Press.

[9] J Brooke. 1996. SUS: A quick and dirty usability scale. *Usability Evaluation in Industry* (1996).

[10] Fabio Buttussi, Luca Chittaro, and Marco Coppo. 2007. Using Web3D technologies for visualization and search of signs in an international sign language dictionary. In *Proceedings of the twelfth international conference on 3D Web technology*. 61–70.
[11] Helen Cooper, Nicolas Pugeault, and Richard Bowden. 2011. Reading the signs: A video based sign dictionary. In *2011 IEEE international conference on computer vision workshops (ICCV Workshops)*. IEEE, 914–919.
[12] DeafTEC. [n. d.]. https://deaftec.org/stem-dictionary/
[13] Aashaka Desai, Lauren Berger, Fyodor Minakov, Nessa Milano, Chinmay Singh, Kriston Pumphrey, Richard Ladner, Hal Daumé III, Alex X Lu, Naomi Caselli, et al. 2024. ASL Citizen: A Community-Sourced Dataset for Advancing Isolated Sign Language Recognition. *Advances in Neural Information Processing Systems* 36 (2024).
[14] Ethnologue.com. [n. d.]. Ethnologue: Languages of the World. http://ethnologue.com/
[15] Institute for Disabilities Research and Inc Training. 2013. The Ultimate American Sign Language Dictionary.
[16] Lynn A Friedman. 1977. Formational properties of American sign language. *On the other hand, New perspectives on American sign language* (1977).
[17] Nancy Frishberg. 1975. Arbitrariness and iconicity: historical change in American Sign Language. *Language* (1975), 696–719.
[18] José L Fuertes, Ángel L González, Gonzalo Mariscal, and Carlos Ruiz. 2006. Bilingual sign language dictionary. In *International conference on computers for handicapped persons*. Springer, 599–606.
[19] Thomas Hanke. 2004. HamNoSys–representing sign language data in language resources and language processing contexts. In *sign-lang@ LREC 2004*. European Language Resources Association (ELRA), 1–6.
[20] Saad Hassan, Akhter Al Amin, Alexis Gordon, Sooyeon Lee, and Matt Huenerfauth. 2022. Design and evaluation of hybrid search for american sign language to english dictionaries: Making the most of imperfect sign recognition. In *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems*. 1–13.
[21] Kalervo Järvelin and Jaana Kekäläinen. 2002. Cumulated gain-based evaluation of IR techniques. *ACM Transactions on Information Systems (TOIS)* 20, 4 (2002), 422–446.
[22] J Lapiak. [n. d.]. Handspeak. https://www.handspeak.com
[23] Dennis Looney and Natalia Lusin. 2018. Enrollments in Languages Other than English in United States Institutions of Higher Education, Summer 2016 and Fall 2016: Preliminary Report.. In *Modern Language Association*. ERIC.
[24] Ross E Mitchell, Travas A Young, Bellamie Bachelda, and Michael A Karchmer. 2006. How many people use ASL in the United States? Why estimates need updating. *Sign Language Studies* 6, 3 (2006), 306–335.
[25] J Murray. 2020. World Federation of the deaf.
[26] Don Newkirk. 1987. *SignFont Handbook*. San Diego: Emerson and Associates.
[27] David Peterson. 2013. Sign Language International Phonetic Alphabet (SLIPA). http://dedalvs.conlang.org/slipa.html
[28] Sergio Scolari, Onno Crasborn, and Gilson Braviano. 2022. Searching for Signs: Developing a handshape taxonomy based on visual similarity. *International Journal of Lexicography* 35, 3 (2022), 320–341.
[29] Zed Sevcikova Sehyr, Naomi Caselli, Ariel M Cohen-Goldberg, and Karen Emmorey. 2021. The ASL-LEX 2.0 Project: A database of lexical and phonological properties for 2,723 signs in American Sign Language. *The Journal of Deaf Studies and Deaf Education* 26, 2 (2021), 263–277.
[30] ShuR. [n. d.]. SLintoDictionary. http://slinto.com/us
[31] Cheri Smith, Ella Mae Lentz, and Ken Mikos. 2008. *Signing naturally*. DawnSignPress San Diego, CA.
[32] William C Stokoe, Dorothy C Casterline, and Carl G Croneberg. 1976. A dictionary of American Sign Language on linguistic principles. *(No Title)* (1976).
[33] William C Stokoe Jr. 2005. Sign language structure: An outline of the visual communication systems of the American deaf. *Journal of deaf studies and deaf education* 10, 1 (2005), 3–37.
[34] Thomas Stone. 2009. ASLSJ. https://sites.google.com/site/aslsignjotting/
[35] Samuel Supalla. 2012. ASL-phabet. http://www.aslphabet.com/home.html
[36] Valerie Sutton. 2010. The SignWriting Alphabet. *Read and Write any Sign Language in the World. ISWA Manual* (2010).
[37] Richard A Tennant and Marianne Gluszak Brown. 1998. *The American sign language handshape dictionary*. Gallaudet University Press.
[38] Haijing Wang, Alexandra Stefan, Sajjad Moradi, Vassilis Athitsos, Carol Neidle, and Farhad Kamangar. 2012. A system for large vocabulary sign search. In *Trends and Topics in Computer Vision: ECCV 2010 Workshops, Heraklion, Crete, Greece, September 10-11, 2010, Revised Selected Papers, Part I 11*. Springer, 342–353.
[39] Sherman Wilcox, Joanne Scheibman, Doug Wood, Dennis Cokely, and William C Stokoe. 1994. Multimedia dictionary of american sign language. In *Proceedings of the first annual ACM conference on Assistive technologies*. 9–16.
[40] Bencie Woll. 2017. The Linguistics of Sign Languages: An Introduction.
[41] Kayo Yin, Chinmay Singh, Fyodor O Minakov, Vanessa Milan, Hal Daumé III, Cyril Zhang, Alex X Lu, and Danielle Bragg. 2024. ASL STEM Wiki: Dataset and Benchmark for Interpreting STEM Articles. *arXiv preprint arXiv:2411.05783* (2024).

Received 12 September 2024; revised 10 December 2024; revised 10 December 2024; accepted 16 January 2025