The American Sign Language Knowledge Graph: Infusing ASL Models with Linguistic Knowledge

Lee Kezar
Univ. of Southern California

Nidhi Munikote
Univ. of Southern California

Zian Zeng
Univ. of Hawaii

Zed Sehyr
Chapman University

Naomi Caselli
Boston University

Jesse Thomason
Univ. of Southern California

Abstract

Sign language models could make modern language technologies more accessible to those who sign, but the supply of accurately labeled data struggles to meet the demand associated with training large, end-to-end neural models. As an alternative to this approach, we explore how knowledge about the linguistic structure of signs may be used as inductive priors for learning sign recognition and comprehension tasks. We first construct the American Sign Language Knowledge Graph (ASLKG) from 11 sources of linguistic knowledge, with emphasis on features related to signs' phonological and lexical-semantic properties. Then, we use the ASLKG to train neuro-symbolic models on ASL video input tasks, achieving accuracies of 91% for isolated sign recognition, 14% for predicting the semantic features of unseen signs, and 36% for classifying the topic of Youtube-ASL videos.

1 Introduction

Sign language models could play a significant role in making language technologies more accessible to deaf and hard-of-hearing signers. In support of this goal, ACL has called for new technological resources to support this emerging field (Diab and Yifru, 2022). As this work has progressed for American Sign Language (ASL), two major challenges have become clear. First, the number of video examples for training ASL models is orders of magnitude smaller than that of speech (Ardila et al., 2020; Uthus et al., 2024) and likely not enough for large, end-to-end architectures. Second, deaf researchers have shown that pervasive hearing-centric biases in NLP frequently result in sign language models that have limited generalizability to real-world signing, oftentimes the consequence of simplified linguistic frameworks (Desai et al., 2024; Hill, 2020).

As a step towards addressing these challenges, we explored how structured knowledge pertaining to ASL linguistics can improve models' ability to perceive and understand ASL. Structured knowledge, such as a knowledge graph, can help contextualize training examples with relevant information and provide a statistical basis for performing inference (Oltramari et al., 2020), but this capacity has not been directly tested for the case of sign language modeling.

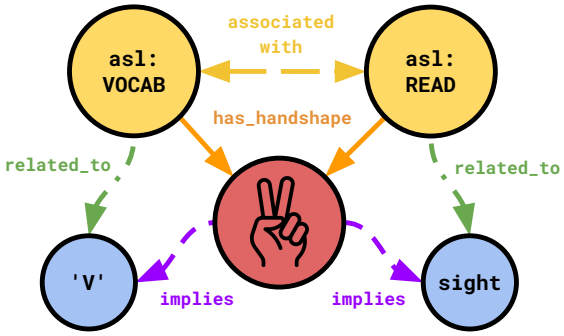

Figure 1: The ASLKG relates the **form** (e.g., *2/V handshape*) and **meaning** (e.g., related to *sight*) of **signs** in the ASL lexicon. We use this knowledge to neuro-symbolically recognize signs (e.g., READ) and infer their meaning.

To empirically test whether structured knowledge can benefit models for ASL, we introduce the **American Sign Language Knowledge Graph** (ASLKG), a collection of 11 knowledge bases containing over 71k linguistic facts related to over 5k ASL signs. The facts in ASLKG primarily relate individual signs to their phonological and semantic properties, for the purpose of building robust sign perception skills. In particular, we show that grounding a video to phonological features in the ASLKG and reasoning about what those features might mean (e.g., signs, semantic features), we achieve 91% accuracy at recognizing isolated signs, 14% accuracy at predicting unseen signs' semantic features, and 36% accuracy at classifying the topic of Youtube-ASL videos (§5).

2 Background

The ASLKG provides a degree of background linguistic knowledge that may be helpful in the process of computationally recognizing and understanding signs. In this initial release, we focus on the phonological and lexical-semantic properties of signs as a symbolic way of reasoning about the identity and meaning of signs.

2.1 Knowledge-Infused Learning

Neuro-symbolic methods combine data-driven pattern recognition with knowledge-driven reasoning over well-defined concepts (Garcez and Lamb, 2023). These methods have been helpful in reasoning over linguistic patterns embedded in high-dimensional data, like video and audio (Hamilton et al., 2022). By accurately grounding these patterns to abstract symbols, models can associate observations with various forms of expert knowledge (Oltramari et al., 2020).

Knowledge infusion describes how expert knowledge informs the parameters of a neural model, and can improve performance and data efficiency (Valiant, 2008). We apply the knowledge-infused learning framework (Gaur et al., 2022) to improve models for sign recognition and understanding. In these approaches, real-world observations are grounded to symbolic knowledge, such as a knowledge graph (KG), which represents facts in the form (*subject*, *predicate*, *object*).[1] For instance, WordNet (Miller, 1995) encodes *hypernym* relationships, such as (*computer*, *is-a*, *technology*).

In this paper, we ground isolated and continuous ASL videos to phonological features in the ASLKG (§3.4.3), then *infer* their corresponding signs and their semantic features (§3.4.4). We additionally develop KG node embeddings, a form of knowledge-infused learning based on fact verification (§3.4.2), to include more holistic linguistic knowledge the inference process.

2.2 ASL Linguistics

Sign languages are complete and natural languages primarily used by deaf and hard-of-hearing people. Sign languages and spoken languages are similar in many ways: there are over one hundred distinct sign languages used by communities around the world (Eberhard et al., 2023); and sign languages demonstrate full phonological, lexical, and syntactic complexity (Padden, 2016; Liddell, 1980; Padden, 2001; Coulter, 2014). Despite these broad similarities, sign languages also differ from spoken languages both in terms of articulation modality (*visual-manual*) and phonology. These differences prevent the extension of standard NLP techniques, like token-based transformers.

2.2.1 Phonology

Understanding the phonological structure of sign languages has direct implications for computational models (Hosain et al., 2021; Albanie et al., 2020; Jiang et al., 2021; Kezar et al., 2023c). In this work, we adopt a phonological description of signs as a means of grounding video data to ASLKG (§3.4.3). Phonemes—the smallest units of language—constitute an inventory of articulatory patterns that can be combined to create words, and are generally not considered to carry meaning. In ASL, signs are distinguished phonologically based on the shape, orientation, movement, and location of the signer's hands, in addition to non-manual markers such as lip shape, eyebrow height, and body shifting (Herrmann and Steinbach, 2011; Michael et al., 2011). These facets are sometimes referred to as *sign language parameters*, and all signs take one or more values for each facet.

2.2.2 Lexical Semantics

Traditionally, linguistic theory has asserted that meaning is conveyed through arbitrary symbols (Locke, 1690; de Saussure, 1916) and their co-occurrence (Firth, 1957). One of the unique affordances of the sign language modality, however, is that signs' phonological forms often physically resemble their meanings (Perniss et al., 2010), and these non-arbitrary associations between phonological form and meaning occur in patterned, systematic ways across the lexicon (Sandler and Lillo-Martin, 2006).

For example, Börstell et al. (2016) found that across ten sign languages, signs representing multiple entities (e.g., SHOES, FAMILY) often use two-handed signs. Similarly, Occhino (2017) showed that in both ASL and Libras,[2] convex and concave shapes (e.g., BALL, BOWL) are frequently represented by signs using a claw handshape. These non-arbitrary forms do not always come about through productive or discrete processes (del Rio et al., 2022)—not all plural signs are two-handed,

[1] There is no widely agreed-upon definition of a knowledge graph. For further discussion, see Ehrlinger and Wöß (2016).

[2] Also known as Brazilian Sign Language

and clawed handshapes do not exclusively denote rounded shapes. Instead, these "systematic" forms are more fluid (i.e., subject to a degree of variation) while expressing similar concepts (Verhoef et al., 2016).

We posit that these flexible and pervasive mappings between form and meaning in sign languages are an important source of information that might improve models' ability to understand out-of-vocabulary signs. As an early step towards computationally modeling the relationships between phonological form and meaning in ASL, we associate signs with their lexico-semantic features like hypernymy and synonymy. We are able to use the ASLKG to confirm that the systematicity of ASL enables *reasoning about the meaning of an unseen sign based on its form.*

3 The ASL Knowledge Graph

In this section, we introduce the American Sign Language Knowledge Graph and describe its intended use. First, we define the elements and structure of the ASLKG (§3.1). Next, we describe how we populated the graph from existing knowledge bases and verified its accuracy (§3.2). Then, we provide graph statistics to characterize the content of the ASLKG (§3.3). Finally, we present three classes of tasks enabled by the ASLKG (§3.4): verifying whether new knowledge can be added to the graph (§3.4.2), grounding video observations v to their corresponding signs, phonemes, or sememes $\mathcal{E}$ (§3.4.3), and inferring implicit relationships from observations (§3.4.4).

3.1 Definition

The ASLKG $\mathcal{G}$ is a set of entities $\mathcal{E}$ interconnected by a set of relations $\mathcal{R}$ according to a set of facts $\mathcal{F}$ pertaining to ASL signs. Formally,

$$\mathcal{G} := \langle \mathcal{E}, \mathcal{R}, \mathcal{F} \rangle \quad (1)$$

$$\mathcal{E} := \{e_1, e_2, \dots e_E\} \quad (2)$$

$$\mathcal{R} := \{r_1, r_2, \dots r_R\} \quad (3)$$

$$\mathcal{F} := \{f_1, f_2, \dots f_F\}, \quad (4)$$

$$f \in \mathcal{E} \times \mathcal{R} \times \mathcal{E} \quad (5)$$

Entities primarily cover ASL signs $s \in \mathcal{E}_{\text{ASL}}$, examples of those signs $\mathcal{E}_v$ ($n = 174547$), English words $w \in \mathcal{E}_{\text{EN}}$, ASL phonemes $\phi \in \mathcal{E}_\Phi$, and ASL semantic features $\sigma \in \mathcal{E}_\sigma$. A number of numeric features, such as the number of morphemes or video duration, are also included in $\mathcal{E}$.

ASL signs ($n = 5802$) and **English words** ($n = 2438$) are associated through expert-annotated labels, constituting a many-to-many relationship between $\mathcal{E}_{\text{ASL}}$ and $\mathcal{E}_{\text{EN}}$. For ASL signs, we additionally include video observations $\mathcal{E}_v$ from Sem-Lex (Kezar et al., 2023b) ($n = 91148$) and ASL Citizen (Desai et al., 2023) ($n = 83399$), which have been manually labeled from a shared vocabulary.

The **phonological features** $\mathcal{E}_\Phi$ ($n = 196$) describe patterns in articulation related to the hands, face, or body. Through the ASL-LEX dataset (Sehyr et al., 2021), $\mathcal{E}_v$ are manually labeled with the phonological features and sign identifiers, enabling machine learning models for phonological feature recognition (Kezar et al., 2023c,a; Ranum et al., 2024).

The **semantic features** $\mathcal{E}_\sigma$ ($n = 319$) describe patterns in meaning, such as semantic associations (Sehyr et al., 2022) and hypernym (is-a) relationships. We supplement the ASL-based features with semantic features of English words ($n = 312$) collected from sources like WordNet, LIWC, and Empath (Miller, 1995; Tausczik and Pennebaker, 2010; Fast et al., 2016). Collectively, the semantic features describe the meaning of lexical items *a priori* at varying levels of abstraction. Given the systematic relationships between phonology and semantics in ASL, these semantic features are potentially helpful in cases where a sign is out-of-vocabulary, such that the *components* of that sign (e.g. a subset of the phonemes) systematically and partially signal the sign's meaning.

3.2 Construction

The ASLKG was constructed in phases, beginning with the identification of candidate knowledge bases, then reformatting the data as RDF triples, and aligning entities across knowledge bases. We then conducted a manual inspection of the ASLKG to verify its facts' accuracy.

3.2.1 Knowledge Base Search

We identified candidate knowledge bases by searching repositories including the Sign Language Dataset Compendium (Kopf et al., 2022), Proceedings of LREC Workshop on the Representation and Processing of Sign Languages (Efthimiou et al., 2024), ACL Anthology, and Semantic Scholar database. A total of 10 knowledge bases pertaining to isolated ASL signs were identified. Two were excluded for inconsistent use of gloss labels (WL-ASL, Li et al. 2020) and WLASL-LEX,

Dataset	Language	|Vocab|	Phon.	Morph.	Syn.	Sem.
ASL-LEX 2.0 (Sehyr et al., 2021)	ASL, En.	2,723	✓	✓	✓	
ASLLVD (Athitsos et al., 2008)	ASL, En.	3,314	✓	✓	✓	
Semantic Associations (Sehyr et al., 2022)	ASL	3,149				✓
IPSL (Kimmelman et al., 2018)	ASL	79	✓	✓		✓
Sem-Lex (Kezar et al., 2023b)	ASL	3,149	✓			
ASL Citizen (Desai et al., 2023)	ASL	2,731				
ASL Phono (de Amorim et al., 2022)	ASL	2,745	✓			
Sensorimotor Norms (Lynott et al., 2020)	En.	39,707				✓
WordNet (Miller, 1995)	En.	155,327		✓	✓	✓
LIWC (Tausczik et al., 2010)	En.	12,000		✓	✓	✓
Empath (Fast et al., 2016)	En.	59,690		✓		✓
ASLKG	ASL, En.	8,240	✓	✓	✓	✓

Table 1: The ASLKG brings together eight sources of knowledge pertaining to the linguistic structure ASL signs, with four English-based sources to supplement morphological, syntactic, and semantic facets via sign-level translation. (*Key: Phon. = phonology, Morph. = morphology, Syn. = syntax, Sem. = semantics*)

Tavella et al. 2022) and one was unavailable for use (ASLNet, Lualdi et al. 2021). Four knowledge bases pertaining to English lexical semantics are included to indirectly supervise the semantic relationships among signs that have an English translation. A snapshot of these knowledge bases' content is shown in Table 1, and we expect to add to this list in future versions.

3.2.2 Conversion to RDF Format

To represent these knowledge bases as a graph structure, we convert their tabular format into Resource Description Framework (RDF) format. For each knowledge base, we identify the column that indexes an ASL sign or English word. In RDF, the values in this column are *subjects* and receive a prefix indicating its vocabulary (`asllrp:`, `asllex:`, `iconicity:`, `en:`).

The remaining columns are labeled according to the kind of linguistic knowledge they convey: `phonological`, `semantic`, etc. In RDF, the headers of these columns are *relations* and the values in that column are *objects*.[3] A Python script performed this conversion for each value in each of the 12 tables.

[3]Some values represent a list of features; these are separated into one fact for each feature.

3.2.3 Entity Alignment

To unify the graph equivalents of these knowledge bases, we identify the labels for signs and phonological features that represent the same thing, and provide one label for all such instances in the graph. Handshape labels in ASL-LEX and ASLLRP were aligned manually and used as supporting evidence for merging the two vocabularies. If two signs have the same gloss (not including word separators `-`, `_` or variant markers `_#`) and the same dominant handshape, then their prefix is replaced with `asl:` to denote a shared vocabulary. 514 pairs of sign labels were merged in this way. English vocabularies are assumed to be identical. For relations/column headers, we identified two pairs that signify the same thing: `has_handshape` and `has_translation`, respectively.

3.2.4 Manual Inspection

The authors who use ASL and have a linguistic background (n=3) contributed to the verification process, wherein 20% of the ASLKGs facts were randomly sampled and verified to be true. We additionally inspected 20 subgraphs generated by depth- or breadth-first search. Minor issues, such as inconsistent number formatting and NaN object values, were addressed on the spot.

3.3 Statistics

Altogether, the $\mathcal{G}$ contains $22,931$ entities and 160 relations distributed across $71,768$ facts. Table 2 shows that ASL is more represented than English in terms of unique entities and number of facts. In general, facts usually take a lexical item as the subject and a descriptor as the object, therefore the average out-degree is greater than the average in-degree. The high standard variation suggests that the knowledge is not evenly distributed across lexical items, aligning with the expectation that more frequent signs are more likely to appear in linguistic data generally. See Appendix A for the distribution of relations and facts by type of linguistic knowledge.

3.4 Tasks

The ASLKG can function as a searchable repository of information related to ASL signs and as a dataset for training NLP models; this remainder of this paper will focus on experimentally testing the latter use case. For interfacing with ASL, we identified the following tasks as most relevant: (1) creating node embeddings that capture localized information in the graph (§3.4.2); (2) grounding an ASL video to its corresponding entities (§3.4.3); and (3) inferring relationships among observed entities (§3.4.4).

3.4.1 Partitioning Facts for ML Tasks

To assist with experimentation, we assign each video example to an *instance fold* ($0 \leq i < 5$) and a *sign fold* ($0 \leq i < 10$). The instance folds $\mathcal{E}_v^{(i)}$ are such that $p(\mathcal{E}_{\text{ASL}})$ is approximately equal across folds, i.e. each sign is equally represented in each instance fold. The sign folds $\mathcal{E}_{\text{ASL}}^{(i)}$ are equally-sized partitions of $\mathcal{E}_{\text{ASL}}$ to facilitate tasks involving *unseen* signs, such as semantic feature recognition. For ISR, we use cross-validation on the instance folds; for SFR, on the sign folds.

3.4.2 Intrinsic Fact Verification

Verification is the task of estimating the existence of some unseen fact $p(f' \mid \mathcal{G})$. Verification is commonly used as a pretraining task for creating **KG embeddings**, where a graph neural network $M_E(f)$ learns to discriminate between true and false facts by means of a scoring function $s(\mathbf{h}, \mathbf{r}, \mathbf{t})$, where $(\mathbf{h}, \mathbf{r}, \mathbf{t})$ are learnable vectors for the head, relation, and tail elements, respectively. Trained in this way, the goal of KG embeddings is to capture the stable, cohesive relationships among entities.

In this work, we apply KG embeddings to subgraphs of $\mathcal{G}$ such that the embeddings capture specific types of relationships (e.g. form, meaning), and experimentally test the extent to which they help with downstream inference tasks.

3.4.3 Grounding ASL Videos to Phonemes

Given video v, let *grounding* model $M_{\mathbf{e}}$ approximate the probability that v is associated one or more KG symbols $\mathbf{e} \subset \mathcal{E}$. Isolated sign recognition (Bragg et al., 2019) may be described as a grounding task, provided that there exists an injective mapping from the ISR model's output classes $\mathcal{Y}$ to signs $\mathcal{E}_{\text{ASL}}$: $M_s(x;\theta) \approx p(\mathcal{E}_{\text{ASL}}|x)$. However, any subset of $\mathcal{E}$ can be the target of grounding. In this work, we explore grounding to phonological features $\mathcal{E}_\Phi \subset \mathcal{E}$ (§4.2) as the first step in ASL-input tasks, like isolated sign recognition.

3.4.4 Inferring Signs and their Meanings

Given a set of grounded symbols $\mathbf{e} \subset \mathcal{E}$, possibly with their associated probabilities $p(\mathbf{e}|x)$, KG *inference* attempts to estimate the presence of some target, for example a novel fact $p(f'|\mathbf{e}) \cdot p(\mathbf{e}|x)$. The benefits of inference in a symbolic medium include: (a) reduced pressure to acquire many training examples; (b) the ability to explain how the model computed its prediction; and (c) deterministic estimations of uncertainty (Oltramari et al., 2020). Each of these benefits is relevant to sign recognition, where neural models are generally less accurate at recognizing signs on the long-tail (Kezar et al., 2023b), and users may want to customize their model or calibrate trust in its output.

Isolated Sign Recognition (ISR). Given a video v_s of a signer demonstrating one sign $s \in \mathcal{E}_{\text{ASL}}$, the model M_s aims to estimate $p(s|v_s)$. M_{S} may be implemented as probabilistic inference over phonological observations $\boldsymbol{\phi} \subset \mathcal{E}_\Phi$ as:

$$p(s|v) \approx p(s|\boldsymbol{\phi}) \cdot p(\boldsymbol{\phi}|v_s)$$

On the note of generalizability to new signs as well as flexibility to user preferences, $p(s|\boldsymbol{\phi})$ can be approximated from relatively few observations $(v_s, \boldsymbol{\phi})$. We compare a number of knowledge-infused methods for this task in Section 4.3.

Semantic Feature Recognition (SFR). For sign fold i, given a video v_s of a signer demonstrating an unseen sign $s \in \mathcal{E}_{\text{ASL}}^{(i)}$, the model M_σ aims to approximate the semantic features $p(\boldsymbol{\sigma}|v_s)$. As with

M_s, in a neuro-symbolic setting, M_σ may be implemented as $p(\boldsymbol{\sigma}|\boldsymbol{\phi}) \cdot p(\boldsymbol{\phi}|v_s)$. We position this task as a first attempt at zero-shot isolated sign understanding, emulating the likely scenario where a sign recognition model encounters an out-of-vocabulary (OOV) sign. In application, such a model could act as a "semantic back-off": when $\max(p(s|v_s))$ is sufficiently low, we might decide that the sign is OOV and, in its place, use semantic features. We explore this inference task, along with its inverse task $p(\boldsymbol{\phi}|\boldsymbol{\sigma})$, in Section 4.4.

Topic Classification. Given a video v_t containing natural, sentence-level signing about genre or topic t, topic classification seeks to approximate $p(t|v_t)$. Compared to isolated sign data, v_t is more realistic and also more information-dense. Using neuro-symbolic methods, we may approach this task as a sequence of independent transformations:

$$v_t \xrightarrow{M_\phi} \boldsymbol{\phi} \xrightarrow{M_s} S \xrightarrow{M_t} t,$$

where $\boldsymbol{\phi}, S$ is a sequence of phonemes and signs, respectively, according to a sliding window over v_t. Such a model could, for example, facilitate searching over repositories of uncaptioned ASL video data on YouTube.

Statistic	ASL	English
# Sources	7	4
# Entities (E)	5802	2438
# Facts (F)	43513	17877
Avg. In-Degree	2.19 (9.4)	1.56 (0.9)
Avg. Out-Degree	33.03 (34.9)	4.23 (2.6)
# Sources per e	1.56 (0.6)	1.99 (1.1)

Table 2: ASLKG statistics by language (std. dev.).

4 Method

We first describe how we collected and formatted the ASLKG data (§**??**). Then, to evaluate the ASLKG's practicality in downstream applications, we apply linguistic knowledge infusion toward three ASL comprehension tasks: isolated sign recognition (§4.3), semantic feature recognition (§4.4)), and topic classification (§4.5). We approach these tasks using the ideas of knowledge-infused learning, in particular by applying linguistic priors to the model architecture, training algorithm, and inference process.

4.1 Intrinsic Fact Verification

As a pretraining task to produce ASLKG embeddings, we train graph neural networks M_E to estimate $p(f)$, where f is either true (sampled from $\mathcal{F}$) or false (randomly constructed such that $f' \notin \mathcal{F}$). We use two implementations of M_E, Trans-E (Bordes et al., 2013) and DistMult (Yang et al., 2015). We train the embedding models (implemented by `kgtk`[4]) for 100 epochs using the subgraph of $\mathcal{G}$ where the tail entities are lexical items $\mathcal{E}_{\text{ASL}} \cup \mathcal{E}_{\text{EN}}$, phonemes $\mathcal{E}_\Phi$ or semantic features $\mathcal{E}_\sigma$.

4.2 Grounding Video Data to ASLKG

To ground video data to $\mathcal{G}$, such as for sign recognition, we use the Sign Language Graph Convolution Network (SLGCN_ϕ) to approximate $p(\boldsymbol{\phi}|v_s)$ following Jiang et al. (2021); Kezar et al. (2023a). This model is 85% accurate at recognizing $n = 240$ phonological features, on average. On the ISR task, for test instance fold i, we train SLGCN_ϕ on instance folds $\neq i$ (and all 10 sign folds). On the SFR task, for test sign fold j, we train SLGCN_ϕ on sign folds $\neq j$. When removing a sign fold, we completely remove all the facts pertaining to those signs in the fold before training.

4.3 Isolated Sign Recognition (ISR)

To estimate $p(s|v_s)$, we compare several models: SLGCN_s, FGM_s, KNN_s, and MLP_s. These models are designed to capture varying degrees of linguistic knowledge. SLGCN_s is a neural baseline trained to predict $v_s \rightarrow s$ directly. Meanwhile, FGM_s and KNN_s are formed from simple heuristics, namely co-occurrence and distance statistics. MLP_s maps embedded representations of $\boldsymbol{\phi}$ to s.

4.3.1 Factor Graph Model

The factor graph model FGM_s approximates $p(\mathcal{E}_{\text{ASL}}|\boldsymbol{\phi})$ according to a partition or *factorization* of $\mathcal{E}_\Phi$, expressed as:

$$\prod_{z_i} p(s|z_i) \text{ s.t. } z_i \subset \mathcal{E}_\Phi \wedge \bigcup_{z_i} = \mathcal{E}_\Phi \wedge \bigcap_{z_i} = \emptyset.$$

Factors are selected based on Brentari's Prosodic Model (Brentari, 1998), grouping phonological features according to articulators (hand configurations), place of articulation (hand location in 3D space), and prosodic features (movements). We employ belief propagation with message passing

[4] https://kgtk.readthedocs.io/en/latest/

(implemented via pgmpy[5]) to infer marginal probabilities across the factors, ensuring efficient computation of $p(s|\phi)$.

4.3.2 k-Nearest Neighbors

The k-nearest neighbors KNN_s model approximates $p(s_a|\phi_b)$ based on the distance between s_a (which is replaced with its ground-truth phonemes ϕ_a) and observations ϕ_b. The distance metric is defined as:

$$d(\phi_a, \phi_b) = 1 - \frac{1}{16}\sum_{i=0}^{16} \delta[\phi_a^i = \phi_b^i] \cdot p(\phi_b^i|x).$$

The final prediction is determined by a majority vote among the nearest k items in $\mathcal{E}_{\text{ASL}}$, using the minimum distance metric to resolve any ties.

4.3.3 Multilayer Perceptron

The multilayer perceptron MLP_s approximates $p(\mathcal{E}_{\text{ASL}}|\phi)$ by learning features for each input. Although less interpretable than other models, MLPs effectively represent many-to-one mappings in training data and could outperform non-parametric and exact inference methods given ϕ. The architecture consists of a randomly initialized embedding layer ($d = 32$) to learn a representation of each phoneme, followed by three hidden layers of sizes $(64/128/256)$ and then a linear projection to the output (size 2723). MLP_s is trained for 100 epochs using cross-entropy loss and the Adam optimizer.

4.4 Semantic Feature Recognition (SFR)

To learn $p(\mathcal{E}_\sigma|\phi)$, we use either MLP_σ or linear regression models REG. Both architectures use a randomly initialized embedding layer ($d = 32$) to learn a coherent representation of each phoneme. Similarly to MLP_s, MLP_σ has three hidden layers of sizes $[64, 128, 256]$ and then a linear projection to the output (size 319).

4.5 Topic Classification

For topic classification, we use Youtube-ASL videos ($n = 11k$), which we divided into 80% train, 10% validation, and 10% test. We first generate topics for each video based on their English captions $c \in C$, then apply a pipeline to each video resulting in (a) a multichannel sequence of phonemes (§4.2), (b) a sequence of signs and their embedding (§4.3), and (c) a single semantic embedding.

[5] https://pgmpy.org

Topic Generation. To generate the topics, we use Latent Dirichlet Allocation (LDA) with $n_t = 10$ on the lemmas in C, weighted by TF-IDF. We use `spaCy`[6] to perform tokenization and lemmatization, and `sk-learn`[7] to perform LDA. We then associate each video with its topic (e.g. `news`, `vlog`) as the topic classifier's final target.

Grounding. To retrofit the phonologizer model $M_\phi(v, \theta)$ to sentence-level data, we use a sliding window approach. We divide each video into a sequence of windows according to a width $W \in \{60, 30, 15\}$ and step $\Delta_f \in \{15, 30\}$ frames.

Inference. We apply an isolated sign recognition model $M_s \in \{\text{FGM}_s, \text{KNN}_s, \text{MLP}_s\}$ to the predicted phonemes to estimate $p(s|\phi)$. We select the most probable sign for each window and form a sequence or *gloss* S. Duplicate signs that are adjacent in the sequence and windows where no $p(s) > 0.1$ are removed.

Embedding. Next, we embed the sequence of signs S using BERT (uncased), implemented by HuggingFace.[8]

$$E(S) = \text{BERT}([E(s_i)\ \forall s_i \in S]) \quad (6)$$

$$E(s) = \sum_{w_i \in t(s)} E_{\text{BERT}}(w_i) * p(w_i|s), \quad (7)$$

where $t(s)$ is the set of translations for s queried from $\mathcal{G}$ and $p(w_i|s)$ is provided by ASL-LEX 2.0, also in $\mathcal{G}$. $E(S)$ is a $d = 768$ vector that we hypothesize will represent the high-level meaning of the ASL sentences in the video.

Topic Classification. Finally, to evaluate the quality of $E(S)$ with respect to topic classification, we train an MLP_t and KNN_t to map $E(S) \rightarrow t$. The MLP model has one hidden layer $d = 100$ and is trained for 50 epochs using a cross-entropy loss. We compare the model performance to random guess and majority class baseline models.

5 Results

We report the results of our experiments on isolated sign recognition, semantic feature recognition, and topic classification. In general, we find that the selected neuro-symbolic methods improve over comparable end-to-end techniques.

[6] https://spacy.io
[7] https://scikit-learn.org
[8] https://huggingface.co

$v_p \to \bullet$	M_s	ACC
s	$\text{SLGCN}_s(v_p, \theta)$	0.64
$(s, \boldsymbol{\phi})$	$\text{SLGCN}_{s,\boldsymbol{\phi}}(v_p, \theta)$	0.66
	$\text{FGM}_s(\hat{\boldsymbol{\phi}}, \mathcal{G}^{\text{TR}})$	0.48
	$\text{KNN}_s(\hat{\boldsymbol{\phi}}, \mathcal{G}^{\text{TR}})$	0.81
$\boldsymbol{\phi} \to s$	$\text{MLP}_s(\hat{\boldsymbol{\phi}}, E_\theta(\mathcal{G}^{\text{TR}}))$	0.85
	$\text{MLP}_s(\hat{\boldsymbol{\phi}}, E_{\text{D}}(\mathcal{G}^{\text{TR}}))$	0.86
	$\text{MLP}_s(\hat{\boldsymbol{\phi}}, E_{\text{T}}(\mathcal{G}^{\text{TR}}))$	**0.92**

Table 3: Top-1 accuracy (ACC) on isolated sign recognition given pose v_p. For embeddings: $E_\theta \sim \mathcal{N}(0, 1)$; E_{D} is DistMult; and E_{T} is Trans-E.

$v_p \to \boldsymbol{\phi}$		$\boldsymbol{\phi} \to s$	ACC$(S \to t)$	
W	Δ_f	M_s	MLP_t	KNN_t
60	15	FGM_s	0.15	0.21
30	15		0.24	0.29
15	15		0.26	0.14
60	30		0.19	0.29
30	30		0.15	0.21
15	30		0.21	0.37
60	15	KNN_s	0.34	0.27
30	15		0.25	0.15
15	15		0.28	0.28
60	30		0.34	0.25
30	30		0.25	0.30
15	30		0.24	0.23
60	15	MLP_s	0.24	0.25
30	15		0.34	0.28
15	15		0.24	0.15
60	30		**0.36**	0.25
30	30		0.31	**0.31**
15	30		0.28	0.23

Table 4: Topic classification top-1 accuracy (ACC) for 10 topics. During $v_p \to \boldsymbol{\phi}$, window width w and step s are varied. Random guess on topic prediction is 0.14 and majority class is 0.21.

5.1 Isolated Sign Recognition

On ISR, we report the top-1 accuracy across models and task configurations in Table 3. These results suggest that shallow knowledge infusion, operationalized as linguistic priors on $p(s|v)$, improves over end-to-end models by 18.9%. We additionally show that intrinsic fact verification, resulting in pretrained embeddings for $\boldsymbol{\phi}$, improves over end-to-end models by 25.2%. The best model is 92% accurate and therefore effective at ISR, a precursor to many ASL-input tasks.

5.2 Semantic Feature Recognition

On the novel task of SFR, we report F_1, precision, and recall in Table 5. As with ISR, intrinsic fact verification as a pretraining task improves over end-to-end models on $\boldsymbol{\phi} \to \boldsymbol{\sigma}$ by 7 points of accuracy. We additionally find several semantic features that are recognized with relatively high F1: signs related to music ($\text{F}_1 = 0.63$), the body ($\text{F}_1 = 0.61$), and family ($\text{F}_1 = 0.50$). The best model is 14% accurate at recognizing semantic features, and in some cases may be useful at recovering from out-of-vocabulary signs. For $\boldsymbol{\sigma} \to \boldsymbol{\phi}$, knowledge infusion improves recognition accuracy by 11% over end-to-end. With further development, this latter task could be helpful in ASL-output tasks, where the intended meaning is already known, and a separate model can translate the phonological features into a coherent sign.

5.3 Topic Classification

We find that the LDA topics thematically align with the those reported in Youtube-ASL, including `vlogs`, `news`, `religion`, and `lessons` (Uthus et al. (2024); ground truth topics were not released to the public). In Table 4, we report the top-1 accuracy with respect to window width W, step Δ_f, ISR model M_s, and topic recognition model M_t. We find that deep knowledge infusion, operationalized as a combination of grounding, inference, and KG embeddings, improves over a majority-class classifier by up to 15%. The best model is 36% accurate at classifying topics, and could assist in searching over large ASL corpora.

6 Discussion

In this work, we introduced the American Sign Language Knowledge Graph containing 71k linguistic facts related to 5.8k signs. We show empirical evi-

M	$\phi \rightarrow \sigma$		$\sigma \rightarrow \phi$	
	F_1	ACC	F_1	ACC
MLP$[E_\theta(\bullet)]$	0.05	0.07	0.18	0.20
REG$[E_\theta(\bullet)]$	0.01	0.02	0.14	0.15
MLP$[E_T(\bullet)]$	0.08	0.08	**0.28**	**0.31**
REG$[E_T(\bullet)]$	0.04	0.05	0.23	0.25
MLP$[E_D(\bullet)]$	**0.12**	**0.14**	0.25	0.27
REG$[E_D(\bullet)]$	0.07	0.09	0.21	0.22

Table 5: F_1 and accuracy (ACC) on semantic feature recognition and the inverse task, semantic-to-phoneme recognition. (*Key:* $E_\theta(\bullet)$ = *randomly-initialized embeddings*; REG = *linear regression model.*)

dence that the ASLKG is an effective resource for modeling American Sign Language input tasks.

On isolated sign recognition, we show that grounding video data to the graph and inferring the sign probabilistically is an accurate, scalable, and interpretable option for large sign vocabularies. Additionally, we show that pretraining on fact verification to produce node embeddings adds an additional 1-7% points of accuracy.

On semantic feature recognition, we show that unseen signs can be partially understood by mapping observed phonological features to semantic labels, such as "related to family", based on form alone. On this task, a simple MLP model is 14% accurate on average, also aided by node embeddings of the input. Future work may explore more sophisticated methods for recognition or apply the recognized features towards understanding tasks.

On topic classification, we sequence grounding and inference models on sentence-level Youtube data, achieving an accuracy of 36% at classifying from ten topics, achieving a 15% improvement over majority class classifier.

As models for ASL attempt to overcome issues with data scarcity and curation quality, our results suggest that including expert-annotated linguistic knowledge through neuro-symbolic mechanisms is an effective path forward. Future versions of the ASLKG will continue to refine the quality of the facts, add additional sources, and ship with more tools for knowledge-infused modeling techniques.

6.1 Limitations

The ASL lexicon is not fixed with respect to the signs in the lexicon, the way those signs are produced, or what those signs mean. Variation exists at all levels of analysis, especially with respect to accent, dialect, and context. These factors are not well-represented in ASLKG, because the primary focus of this work is establishing normative descriptions of ASL structure.

Excluding certain forms of signing disproportionately harms linguistic communities within ASL, such as those who use an underrepresented dialect. We strongly discourage the use of ASLKG towards user-facing applications without the meaningful collaboration of ASL signers, especially those who are deaf and hard-of-hearing.[9]

Phonology Given the tremendous variation in sign language production across signers, despite including many signs in our grounding procedure, our approach does not capture the phonological variation of ASL. Additionally, our approach to discretizing ASL phonology represents only one way to divide an inherently non-discrete system, which is subject from ongoing debate from sign phonologists. For example, ASL-LEX 2.0 (Sehyr et al., 2021) describes eight path movements, while SignWriting describes over 220 (Sutton, 1974). As we continue to refine the ASLKG, we may determine that certain parameterizations of sign phonology are best-suited for different target applications.

Lexical Semantics The use of English data to complement our linguistic knowledge of ASL assumes that there is sufficient overlap in the semantic structure of the two languages. But, these are two independent languages with considerably different structures. Although BERT is based on English grammar and semantics, we here assume that syntax does not play a significant role in capturing the topic of videos, and the meaning of an ASL sign is roughly the meaning of its corresponding English gloss. Both of these assumptions are standard approach in low-resource language modeling, but limit the fidelity of the representation.

[9] To locate potential collaboration with deaf and hard-of-hearing scholars interested in sign language technologies, consider the CREST Network: https://www.crest-network.com.

References

Samuel Albanie, Gül Varol, Liliane Momeni, Triantafyllos Afouras, Joon Son Chung, Neil Fox, and Andrew Zisserman. 2020. BSL-1K: Scaling up co-articulated sign language recognition using mouthing cues. In *European Conference on Computer Vision*.

Rosana Ardila, Megan Branson, Kelly Davis, Michael Henretty, Michael Kohler, Josh Meyer, Reuben Morais, Lindsay Saunders, Francis M. Tyers, and Gregor Weber. 2020. Common voice: A massively-multilingual speech corpus.

Vassilis Athitsos, Carol Neidle, Stan Sclaroff, Joan Nash, Alexandra Stefan, Quan Yuan, and Ashwin Thangali. 2008. The american sign language lexicon video dataset. In *2008 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR) Workshops*.

Antoine Bordes, Nicolas Usunier, Alberto Garcia-Duran, Jason Weston, and Oksana Yakhnenko. 2013. Translating embeddings for modeling multi-relational data. In *Advances in Neural Information Processing Systems (NeurIPS)*.

Danielle Bragg, Oscar Koller, Mary Bellard, Larwan Berke, Patrick Boudreault, Annelies Braffort, Naomi Caselli, Matt Huenerfauth, Hernisa Kacorri, Tessa Verhoef, Christian Vogler, and Meredith Ringel Morris. 2019. Sign language recognition, generation, and translation: An interdisciplinary perspective. In *Proceedings of the 21st International ACM SIGACCESS Conference on Computers and Accessibility*.

Diane Brentari. 1998. *A Prosodic Model of Sign Language Phonology*. The MIT Press.

Carl Börstell, Ryan Lepic, and Gal Belsitzman. 2016. Articulatory plurality is a property of lexical plurals in sign language. *Lingvisticæ Investigationes*, 39(2):391–407.

Geoffrey R. Coulter. 2014. *Current Issues in ASL Phonology: Phonetics and Phonology, Vol. 3*. Academic Press. Google-Books-ID: xyu0BQAAQBAJ.

Cleison Correia de Amorim and Cleber Zanchettin. 2022. Asl-skeleton3d and asl-phono: Two novel datasets for the american sign language. *CoRR*.

Ferdinand de Saussure. 1916. Cours de linguistique générale. *The Modern Language Journal*.

Aurora Martinez del Rio, Casey Ferrara, Sanghee J. Kim, Emre Hakgüder, and Diane Brentari. 2022. Identifying the correlations between the semantics and the phonology of american sign language and british sign language: A vector space approach. *Frontiers in Psychology*, 13.

Aashaka Desai, Lauren Berger, Fyodor O Minakov, Vanessa Milan, Chinmay Singh, Kriston Pumphrey, Richard E Ladner, Hal Daumé III, Alex X Lu, Naomi Caselli, and Danielle Bragg. 2023. Asl citizen: A community-sourced dataset for advancing isolated sign language recognition. In *Proceedings of the 36th International Conference on Neural Information Processing Systems*.

Aashaka Desai, Maartje De Meulder, Julie A. Hochgesang, Annemarie Kocab, and Alex X. Lu. 2024. Systemic biases in sign language AI research: A deaf-led call to reevaluate research agendas. In *Proceedings of the LREC-COLING 2024 11th Workshop on the Representation and Processing of Sign Languages: Evaluation of Sign Language Resources*.

Mona Diab and Martha Yifru. 2022. ACL 2022 D&I Special Initiative: 60-60, Globalization via localization. `https://2022.aclweb.org/dispecialinitiative.html`.

David M. Eberhard, Gary F. Simons, and Charles D. Fennig, editors. 2023. *Ethnologue: Languages of the World*, 26 edition. SIL International, Dallas, Texas, USA.

Eleni Efthimiou, Stavroula-Evita Fotinea, Thomas Hanke, Julie A. Hochgesang, Johanna Mesch, and Marc Schulder, editors. 2024. *Proceedings of the LREC-COLING 2024 11th Workshop on the Representation and Processing of Sign Languages: Evaluation of Sign Language Resources*. ELRA and ICCL, Torino, Italia.

Lisa Ehrlinger and Wolfram Wöß. 2016. Towards a definition of knowledge graphs. In *Joint Proceedings of the Posters and Demos Track of the 12th International Conference on Semantic Systems (SEMANTiCS)*.

Ethan Fast, Binbin Chen, and Michael S. Bernstein. 2016. Empath: Understanding topic signals in large-scale text. *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*.

J. R. Firth. 1957. A synopsis of linguistic theory, 1930-1955.

Artur d'Avila Garcez and Luís C. Lamb. 2023. Neurosymbolic ai: the 3rd wave. *Artificial Intelligence Review*.

Manas Gaur, Kalpa Gunaratna, Shreyansh Bhatt, and Amit Sheth. 2022. Knowledge-infused learning: A sweet spot in neuro-symbolic ai. *IEEE Internet Computing*.

Kyle Hamilton, Aparna Nayak, Bojan Bozic, and Luca Longo. 2022. Is neuro-symbolic ai meeting its promise in natural language processing? a structured review. *Semantic Web*.

Annika Herrmann and Markus Steinbach. 2011. Nonmanuals in sign languages. *Sign Language & Linguistics*, 14(1):3–8. Publisher: John Benjamins.

Joseph Hill. 2020. Do deaf communities actually want sign language gloves? *Nature Electronics*, 3(9):512–513.

Al Amin Hosain, Panneer Selvam Santhalingam, Parth H. Pathak, Huzefa Rangwala, and Jana Kosecka. 2021. Hand pose guided 3d pooling for word-level sign language recognition. *2021 IEEE Winter Conference on Applications of Computer Vision (WACV)*.

Songyao Jiang, Bin Sun, Lichen Wang, Yue Bai, Kunpeng Li, and Yun Raymond Fu. 2021. Skeleton aware multi-modal sign language recognition. *2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*.

Lee Kezar, Tejas Srinivasan, Riley Carlin, Jesse Thomason, Zed Sevcikova Sehyr, and Naomi Caselli. 2023a. Exploring strategies for modeling sign language phonology. In *The European Symposium on Artificial Neural Networks*.

Lee Kezar, Jesse Thomason, Naomi Caselli, Zed Sehyr, and Elana Pontecorvo. 2023b. The sem-lex benchmark: Modeling asl signs and their phonemes. In *Proceedings of the 25th International ACM SIGACCESS Conference on Computers and Accessibility*.

Lee Kezar, Jesse Thomason, and Zed Sevcikova Sehyr. 2023c. Improving sign recognition with phonology. In *European Association for Computational Linguistics*.

Vadim Kimmelman, Anna Klezovich, and George Moroz. 2018. IPSL: A database of iconicity patterns in sign languages. creation and use. In *Proceedings of the Eleventh International Conference on Language Resources and Evaluation (LREC 2018)*.

Maria Kopf, Marc Schulder, and Thomas Hanke. 2022. The sign language dataset compendium: Creating an overview of digital linguistic resources. In *SIGNLANG*.

Dongxu Li, Cristian Rodriguez, Xin Yu, and Hongdong Li. 2020. Word-level deep sign language recognition from video: A new large-scale dataset and methods comparison. In *The IEEE Winter Conference on Applications of Computer Vision (WACV)*.

Scott K. Liddell. 1980. *American Sign Language Syntax*. Walter de Gruyter GmbH & Co KG. Google-Books-ID: 04iFEAAAQBAJ.

John Locke. 1690. An essay concerning human understanding.

Colin Lualdi, Elaine Wright, Jack Hudson, Naomi Caselli, and Christiane Fellbaum. 2021. Implementing ASLNet v1.0: Progress and plans. In *Proceedings of the 11th Global Wordnet Conference*.

Dermot Lynott, Louise Connell, Marc Brysbaert, James Brand, and James Carney. 2020. The lancaster sensorimotor norms: multidimensional measures of perceptual and action strength for 40,000 english words. *Behavior Research Methods*.

Nicholas Michael, Peng Yang, Qingshan Liu, Dimitris Metaxas, and Carol Neidle. 2011. A Framework for the Recognition of Nonmanual Markers in Segmented Sequences of American Sign Language. In *Procedings of the British Machine Vision Conference 2011*, pages 124.1–124.12, Dundee. British Machine Vision Association.

George A. Miller. 1995. Wordnet: A lexical database for english. *Commun. ACM*, 38:39–41.

Corinne. Occhino. 2017. An introduction to embodied cognitive phonology: Claw-5 handshape distribution in asl and libras. *Estudios Ingleses de la Universidad Complutense*, 25:69–104.

Alessandro Oltramari, Jonathan M Francis, Cory Andrew Henson, Kaixin Ma, and Ruwan Wickramarachchi. 2020. Neuro-symbolic architectures for context understanding.

Carol A. Padden. 2016. *Interaction of Morphology and Syntax in American Sign Language*. Routledge. Google-Books-ID: psiVDQAAQBAJ.

Diane Brentari Padden, Carol A. 2001. Native and Foreign Vocabulary in American Sign Language: A Lexicon With Multiple Origins. In *Foreign Vocabulary in Sign Languages*. Psychology Press. Num Pages: 33.

Pamela Perniss, Robin Thompson, and Gabriella Vigliocco. 2010. Iconicity as a general property of language: Evidence from spoken and signed languages. *Frontiers in Psychology*.

Oline Ranum, Gomèr Otterspeer, Jari I. Andersen, Robert G. Belleman, and Floris Roelofsen. 2024. 3D-LEX v1.0 – 3D lexicons for American Sign Language and Sign Language of the Netherlands. In *Proceedings of the LREC-COLING 2024 11th Workshop on the Representation and Processing of Sign Languages: Evaluation of Sign Language Resources*, pages 290–301, Torino, Italia. ELRA and ICCL.

Wendy Sandler and Diane Lillo-Martin. 2006. *Entering the lexicon: lexicalization, backformation, and cross-modal borrowing*, page 94–107. Cambridge University Press.

Sevcikova Zed Sehyr, Naomi Caselli, Ariel Cohen-Goldberg, and Karen Emmorey. 2022. The semantic structure of american sign language: Evidence from free sign associations. In *The 63rd Annual Meeting of the Psychonomic Society*.

Zed Sevcikova Sehyr, Naomi K. Caselli, Ariel 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*.

Valerie Sutton. 1974. SignWriting symbol category 2 and 3: Movement and dynamics. https://www.signwriting.org/lessons/iswa/category2-3.html.

Yla R. Tausczik and James W. Pennebaker. 2010. The psychological meaning of words: Liwc and computerized text analysis methods. *Journal of Language and Social Psychology*, 29:24 – 54.

Federico Tavella, Viktor Schlegel, Marta Romeo, Aphrodite Galata, and Angelo Cangelosi. 2022. WLASL-LEX: a dataset for recognising phonological properties in American Sign Language. In *Proceedings of the 60th Annual Meeting of the Association for Computational Linguistics (Volume 2: Short Papers)*.

David Uthus, Garrett Tanzer, and Manfred Georg. 2024. YouTube-ASL: a large-scale, open-domain american sign language-english parallel corpus. In *Proceedings of the 37th International Conference on Neural Information Processing Systems*.

Leslie G Valiant. 2008. Knowledge Infusion: In Pursuit of Robustness in Artificial Intelligence. In *IARCS Annual Conference on Foundations of Software Technology and Theoretical Computer Science*.

Tessa Verhoef, Carol Padden, and Simon Kirby. 2016. Iconicity, naturalness and systematicity in the emergence of sign language structure. In *The Evolution of Language: Proceedings of the 11th International Conference (EVOLANGX11)*. Online at `http://evolang.org/neworleans/papers/47.html`.

Bishan Yang, Wen-tau Yih, Xiaodong He, Jianfeng Gao, and Li Deng. 2015. Embedding entities and relations for learning and inference in knowledge bases. In *International Conference on Learning Representations (ICLR)*.

A Additional Statistics

See Table 6 for additional details regarding the types of knowledge included in the ASLKG.

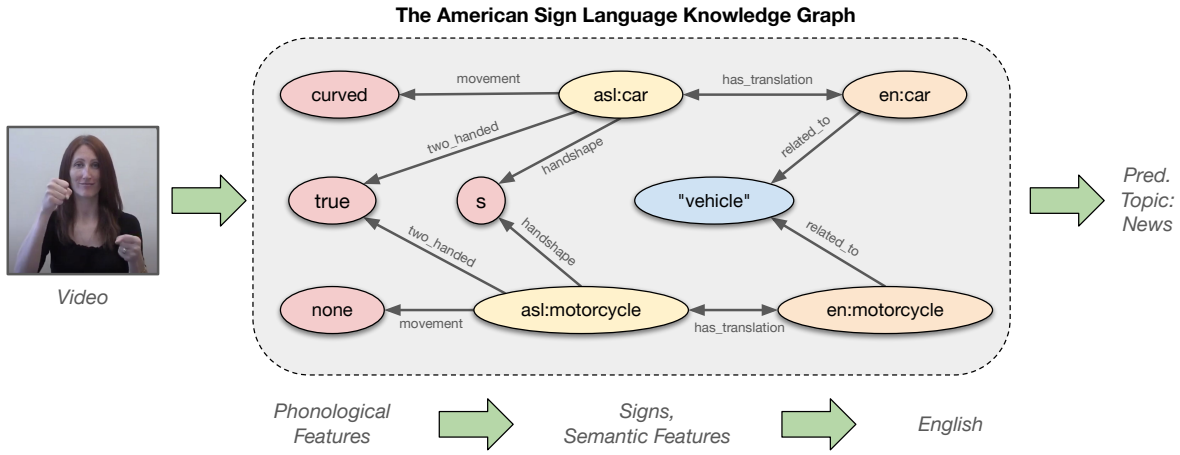

Figure 2: An example of the ASLKG subgraph for two signs (MOTORCYCLE, CAR), illustrating their shared use of two S handshapes and shared association with vehicles. Separately, identified signs and their meaning may be used to predict a video's topic, like news.

Relation Type t	Describes the subject sign's...	$\lvert\mathcal{R}_t\rvert$	$\lvert\mathcal{F}_t\rvert$	Example $r \in \mathcal{R}_t$
phonetic	sub-phonological production	2	2 733	`Sign_Duration`
phonological	phonemes	90	94 218	`Handshape`
morphological	morphemes	1	5 553	`Number_Of_Morphemes`
syntactic	lexical class/part of speech	3	5 657	`Lexical_Class`
semantic	meaning (in isolation)	285	26 581	`Associated_With`
translation	English translation	14	28 925	`Entry_ID`
systematicity	form/meaning interaction	27	29 552	`Initialized_Sign`
statistical	frequency	30	43 763	`Frequency_N`
cognitive	mental representations	2	16 282	`Age_Of_Acquisition`
meta	meta-information	2	9 429	`SignBank_Reference_ID`

Table 6: The types of knowledge in ASLKG. $\lvert\mathcal{R}_t\rvert$ and $\lvert\mathcal{F}_t\rvert$ denote the number of unique relations and facts, respectively, with type t.