Proceedings of the 10th Workshop on the Representation and Processing of Sign Languages (sign-lang@LREC 2022), pages 9–15
Language Resources and Evaluation Conference (LREC 2022), Marseille, 20-25 June 2022

Introducing Sign Languages to a Multilingual Wordnet: Bootstrapping Corpora and Lexical Resources of Greek Sign Language and German Sign Language

**Sam Bigeard[*], Marc Schulder[*], Maria Kopf[*], Thomas Hanke[*],
Kyriaki Vasilaki[‡], Anna Vacalopoulou[‡], Theodoros Goulas[‡],
Athanasia–Lida Dimou[‡], Stavroula–Evita Fotinea[‡], Eleni Efthimiou[‡]**
[*]Institute of German Sign Language and Communication of the Deaf (IDGS), University of Hamburg, Germany
{sam.bigeard, marc.schulder, maria.kopf, thomas.hanke}@uni-hamburg.de
[‡]Institute for Language and Speech Processing (ILSP), Athena Research Center, Greece
{kvasilaki, avacalop, tgoulas, ndimou, evita, eleni_e}@athenarc.gr

Abstract

Wordnets have been a popular lexical resource type for many years. Their sense-based representation of lexical items and numerous relation structures have been used for a variety of computational and linguistic applications. The inclusion of different wordnets into multilingual wordnet networks has further extended their use into the realm of cross-lingual research. Wordnets have been released for many spoken languages. Research has also been carried out into the creation of wordnets for several sign languages, but none have yet resulted in publicly available datasets. This article presents our own efforts towards an inclusion of sign languages in a multilingual wordnet, starting with Greek Sign Language (GSL) and German Sign Language (DGS). Based on differences in available language resources between GSL and DGS, we trial two workflows with different coverage priorities. We also explore how synergies between both workflows can be leveraged and how future work on additional sign languages could profit from building on existing sign language wordnet data. The results of our work are made publicly available.

Keywords: Multilingual Wordnet, Sign Language Wordnet, Semi-automatic resource creation

1. Introduction

Multilingual resources like wordnets are still scarce in the field of sign language research. A multilingual sign language wordnet could open doors for computational linguistics as well as lexicographers working on sign language dictionaries. Sign language resources such as corpora and lexicons are often searchable only through spoken language translation, due to the lack of a common and easy to use sign writing system. Multilingual indexation is hindered by research-specific lemmatisation approaches which are difficult to combine. These conditions are a challenge for both human users and computational applications, the latter lacking machine-readable resources of all kinds for sign languages. To make sign language resources more searchable, more machine-readable, and their sense descriptions more precise and accessible, we aim at creating a multilingual sign language wordnet.

In this paper, we present our approach towards such a sign wordnet. We use a combination of automatic and manual methods to bootstrap the integration of sign languages into a multilingual wordnet. For this we are working on two languages in parallel: Greek Sign Language (GSL) and German Sign Language (DGS). The two languages are very different with regard to available resources, which gives us the opportunity to test different approaches and see which works best for what kind of resource. By describing our method and the issues we have encountered we hope to provide a helpful guide to other researchers working on multilingual sign language resources.

The results of our work are made publicly available and will be updated as our efforts progress.[1]

[1]https://doi.org/10.25592/uhhfdm.10169

2. Background

In this section we outline the relevant background on wordnets, describing the history of spoken language wordnets (Section 2.1) and existing work on sign language wordnets (Section 2.2).

2.1. Wordnets for Spoken Languages

The concept of a wordnet was first introduced by Miller et al. (1990) as the idea of a dictionary based on psycholinguistic principles. The new approach was that words are not organised alphabetically but in so-called synonym sets (synsets), each representing an underlying concept. The synsets are interconnected via directional relations such as hyponymy, antonymy and meronymy. For example, the concept of a dog, the animal, is expressed by a synset consisting of the words ‘dog’, ‘domestic dog’ and ‘Canis familiaris’. This synset is a hyponym (i. e. a more specific form of) the ‘domestic animal’ synset. Other word senses of ‘dog’ are covered by other synsets with their own relations, such as the concept of a reprehensible person expressed by the terms ‘dog’, ‘cad’, ‘blackguard’ and others.

While the original Princeton Wordnet (PWN) was designed for English, wordnets for many other languages have since been created. Several efforts to interconnect these into a multilingual wordnet have been undertaken. The most prominent effort that is still actively supported is the Open Multilingual WordNet (OMW) (Bond and Paik, 2012).

Most wordnet projects use Princeton Wordnet as a basis to expand upon, rather than developing their own wordnet from scratch (Bond et al., 2016). This approach is known as the *expand model*. While this creates a bias toward English, it significantly reduces the amount of work needed to create a new wordnet and connect existing ones.

While the construction of a wordnet for well resourced spoken languages is relatively straightforward, the process has to be revisited for less resourced languages. Commonly used resources like dictionaries, wikis, and others may not be available. Bosch and Griesel (2017) use the *expand model* to create a wordnet of five South African languages. One of their findings is that 'similarities shared on levels such as morphology or grammar and semantics allow the language teams to learn from one another, to share and thus to fast-track the development of the individual wordnets in this way' (Bosch and Griesel, 2017, p. 11). On this basis, we expect that once a wordnet for one sign language is established, subsequent sign language wordnets will be able to build on it, significantly reducing the amount of work needed.

2.2. Wordnets for Sign Languages

Work on creating wordnets for individual signed languages has been reported for Swiss-German Sign Language (DSGS) (Ebling et al., 2012), Italian Sign Language (LIS) (Shoaib et al., 2014) and American Sign Language (ASL) (Lualdi et al., 2021), although no publicly available resource have yet been released. All of these works have in common that they seek to link wordnet structures to existing lexical resources of the respective signed language. This approach allows them to leverage existing video recordings and lexicographic information for individual signs, drastically reducing the cost of creating the wordnet. In the case of ASL, several lexical resources are used to increase the available vocabulary (Lualdi et al., 2021).

Other works do not seek to publish full signed language wordnets, but rather use existing wordnets for a spoken language as an aid to internal work. Troelsgård and Kristoffersen (2018) link entries in their lexical database of Danish Sign Language (DTS) to roughly matching synsets in DanNet. These links are used as an aid to lexicographers and to automatically determine potential synonyms. The authors stress that the wordnet senses do not necessarily correspond exactly to the sign senses. Langer and Schulder (2020) match lexical entries of the DGS Corpus (see Section 3.2) with wordnet lemmas to extract supersense categories for use in coarse semantic clustering for lexicographic work. The matching is done automatically, based on existing German translational equivalents for the signs and does not take into account word sense disambiguation.

3. Resources

Following the approach of other signed language wordnet creation efforts, we build directly on existing resources for Greek Sign Language (GSL), German Sign Language (DGS), Greek, and German. While the resources for GSL and DGS each include a corpus and a lexical resource, their history of creation and resulting available information structures are very different. However, among their main similarities lies the fact that they are both built with their respective sign languages (SLs) as a starting point; in other words, they are SL-based and produced and verified by both deaf and hearing experts of GSL and DGS.

3.1. GSL Lexical Resources

The repository of GSL lexical resources has been collected, built, and annotated for years by the Institute for Language and Speech Processing (ILSP). It mainly consists of the Noema+ bilingual dictionary (GSL and Modern Greek) and the underlying Polytropon parallel corpus, which provides example utterances involving specific signs. These were based on utterances from expert discussions which were then re-recorded in a studio environment and annotated to serve as a 'golden' corpus open to SL technologies research (Efthimiou et al., 2016; Efthimiou et al., 2018). These two resources comprise the most extensive reference pool for GSL to date and include more than 3,600 clauses in GSL.

The lexical database currently consists of approximately 12,000 entries and it has been annotated in its entirety on the basis of the Polytropon corpus. The construction and maintenance of the database is facilitated with the use of a dedicated web-based open environment that supports the creation and interlinking of GSL resources, namely, the SiS-Builder (Goulas et al., 2010).

As the Polytropon corpus consists of isolated utterances chosen to illustrate specific signs, the contribution relating to GSL is more lexicon- than corpus-based. While this has the drawback of not providing the full context and authenticity of natural discourse, the advantage of this more controlled environment is the more explicit correspondence between GSL sign and sense-appropriate Greek translation.

3.2. DGS Corpus Resources

The DGS Corpus is an annotated corpus of 560 hours of natural discourse in DGS (Prillwitz et al., 2008). A subset of the corpus has been released publicly as the Public DGS Corpus (Jahn et al., 2018).

The DGS Corpus implements a type hierarchy, called 'double glossing' (Konrad et al., 2012, p. 88). Each type represents a distinct sign and is further subdivided into subtypes, each of which represents a lexicalised meaning of that sign. Glosses for types and subtypes in the DGS Corpus are available in English and German.

In addition to the gloss name, each subtype can have one or more concept entries associated with it in the lexical database of the DGS Corpus. Concept entries are written with German or English orthography (as opposed to the all-caps glosses) and specify possible meanings. In the DTS corpus, which uses the same lexical database structure as DGS Corpus, these concept entries are in fact used to represent the DanNet synsets (Troelsgård and Kristoffersen, 2018). In the DGS Corpus, however, concepts are only disambiguated in relation to the German and English terms. If sign and word have the same sense ambiguity, only one concept is created. This makes DGS Corpus concepts coarser than wordnet synsets but more fine grained than glosses.

On the basis of the DGS Corpus a digital dictionary for DGS is currently being created, called DW-DGS (Müller et al., 2020). The dictionary provides more nuanced information on signs and their senses. The first pre-release entries are already published[2] and can be used to further feed the sign wordnet for DGS.

[2] http://dw-dgs.de

3.3. Greek WordNet

OMW covers a wide range of spoken languages, created in individual projects. The Greek WordNet included in OMW consists of 18,049 synsets. The Greek synsets were originally developed in the context of BalkaNet, a multilingual wordnet of Balkan languages (Grigoriadou et al., 2004). They were based on a series of Greek lexicons and corpora. In the course of our work we found that the entries of the Greek WordNet that we inspected mainly included glossed explanations of each lexical item with minimal, if any, usage examples.

3.4. GermaNet

The largest wordnet for German is GermaNet (Hamp and Feldweg, 1997). As of version 17 it contains 159,514 synsets. Due to licence restrictions it is not directly integrated into OMW. However, for 28,564 of its synsets a mapping to PWN exists, from which OMW identifiers can be inferred. For our mutlilingual wordnet we decided to use GermaNet and expand the connections to OMW.

4. Wordnet Creation

To create the multilingual sign wordnet both teams — the GSL and DGS team — first work independently on their respective language with frequent exchanges regarding method and implementation.

The GSL team follows a high precision approach of identifying strong synset matches for entries in the GSL lexical database. They prioritise providing at least one sign for many different synsets over specifying every possible synonym. This approach is outlined in Section 4.1.

The DGS team follows a high recall approach of automatically matching its corpus type inventory to wordnet lemmas and then verifying these matches. They prioritise validating many potential synonyms of fewer synsets over partially covering many synsets. This is described in Section 4.2.

The intermediate progress of both teams is compared in Section 4.3.

As work progresses, lists of linked synsets are exchanged between the teams to allow them to prioritise those synsets also covered by the other group. Additional cross-lingual factors are also considered, as described in Section 4.4. This pushes forward the progress towards a large interlingual index.

4.1. Linking the GSL Lexical Resources

The GSL team uses data that is collected by a semi-automatic process of mapping synsets from the Greek part of OMW to the GSL lexical database. As the only common element of both databases are Greek lemmas, this is done by matching the 'Greek equivalent' entries of GSL signs with the Greek lemmas in OMW synsets; at the end of this process, each GSL entry whose Greek equivalent also appears in OMW is mapped to the respective OMW entry.

In the next stage of the process, these automatically generated associations are checked by deaf and hearing GSL experts for validity against the respective videos offered for each sign, resulting in a new 'clean' database of wordnet synsets and their GSL equivalents. Of course, this is hardly a one-to-one connection, as a lot of false equivalents are revealed in the process. For instance, the GSL entry 'ανέκδοτο' has been automatically linked via the Greek equivalent to synset 07220586-n, which matches the English word 'anecdote'. However, this sense of the Greek equivalent does not correspond to the respective entry in the GSL database, where the word is associated with its much more frequent sense of 'joke'. The more fitting synset 06778102-n was not found during automatic matching because it has no Greek entry. All such instances are manually corrected by GSL experts.

The accuracy of the equivalents is tested against GSL examples that are linked to each of the lexical resource's entries to make sure that each corresponds to the correct definition, or rather glossed explanation (Fellbaum, 1998), in Greek WordNet. A secondary way of double-checking whether the correct sense of each entry is selected is reviewing the other available language versions in OMW with which annotators are familiar, namely, English and French. In addition to that, the Greek WordNet proves to be rather limited for the purposes of this experiment, as it comprises 18,049 synsets compared to the English data of PWN, which consists of 117,659 synsets in version 3.0. These numbers limit the linking process even more. To compensate for this, it was decided to extend the mapping of the GSL material to the richer English part of OMW at a second level. At the time of writing, 1819 GSL signs have been linked to 4214 wordnet synsets.

4.2. Linking DGS Corpus

The DGS team uses a three-step method: automatic generation of candidate matches between synsets and subtypes, automatic verification of certain simple cases, and manual verification of all remaining cases.

Automatic matching is done between the lemmas of OMW synsets and the concept entries of DGS Corpus subtypes. Both German and English are used for this, although German is preferred, as concept values are more precise in that languages. Where no concept entry is available for a subtype, its gloss name is used as a fallback.

As mentioned in Section 2.1, we use GermaNet as our German wordnet resource and connect its entries to OMW through its partial mapping to PWN. If a German word is not present in GermaNet or a GermaNet synset has no connection to OMW, the English concept entry or gloss is used instead. For the case that there is no English translation in the DGS Corpus or no corresponding synset in OMW, a fallback solution of automatic translation of the German gloss to English is used.

At the time of writing, automatic candidate matches between 11,856 DGS subtypes and 27,020 synsets were found. Subtypes were associated with a mean of 8.6 synsets and a median of 2. This is a 'long tail' situation, where most subtypes have very few senses, while heavily polysemous terms such as 'have' or 'good' (and their DGS counterparts) have 20 synsets or more associated with them. In many cases, the two synsets associated with the sign represent a basic and a figurative sense.

In a second automatic processing step, candidate matches with a high likelihood of being correct are identified and marked as provisionally validated. This automatic valida-

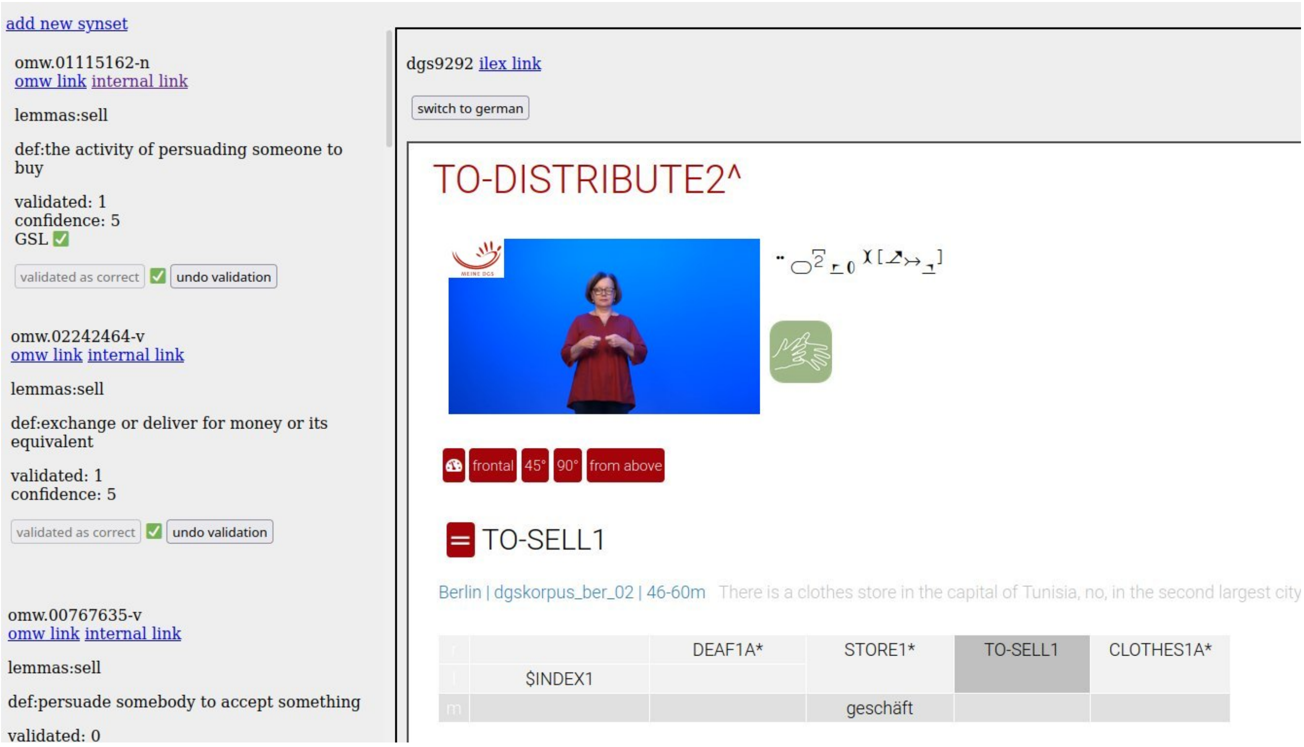

Figure 1: Manual validation interface listing all synsets associated with a specific DGS sign. The left side lists the associated synsets and their validation status for DGS and GSL, while the right side shows the DGS signs' type entry page from the Public DGS Corpus website.

tion step selects subtypes which were matched with only a single synset and using strong match conditions, i. e. not via automatic translation. Such single match-pairs mainly occur among the long tail of homonymous expressions. As they are based on high quality human translations (concept entries or glosses), the chance of such matches introducing incorrect senses for a sign is very low.

In the final step, the remaining automatic matches are validated manually by using corpus evidence and the expert's own acceptability judgements. Ideally such verifications would only be performed by L1 language users. Due to the large number of matches (over 100,000 subtype-synset pairs) this is currently not possible for us. Instead we follow a two-tiered approach in which L2 language users validate cases for which they have high confidence and mark the remaining cases for later review by an L1 user. This method allows us to have more annotators involved, resulting in a quicker workflow.

Figure 1 shows the validation interface for confirming or rejecting all synsets that were automatically matched to a specific sign. At the time of writing, 2230 DGS signs with one or more synsets have been validated.

4.3. Progress

Statistics on the current progress of linking both languages to OMW are outlined in Table 1. A notable difference between the languages can be seen regarding the number of signs and synsets covered relative to the overall number of sign-synset pairs. For GSL the number of validated pairs is close to the number of distinct synsets but greater than the number of signs, meaning that each sign is on average linked to 2.3 synsets, but only few synsets are linked to more than one sign. The number of validated DGS pairs, on the other hand, shows the inverse pattern, with most signs linking to only one synset, but covered synsets containing an average 2.4 signs each.

The difference between the languages is caused by both the difference in workflow between the teams already described in Section 4.1 and Section 4.2 and by the nature of the datasets on which they build. As described in Section 3.1, the GSL lexicon is primarily based on work for the Noema+ bilingual dictionary. Its focus was on providing GSL signs for many concepts. Following in this vein, the GSL team covered a wide variety of different concepts during their validation.

The DGS vocabulary, on the other hand, stems from the sign inventory encountered in the natural conversations of the DGS Corpus (see Section 3.2), reflecting the many variations in participants' vocabulary due to regional differences, age group, register, and other factors, leading to the presence of many synonyms. In addition, the DGS team started out by validating comparatively unambiguous concepts such as the names of months, which are straightforward to validate, but can be expressed by a large number of different signs in DGS. Because of this, comparatively few synsets are covered, but each with a higher number of signs associated with it.

It should be noted that the current ratios are due to the preliminary nature of the work. As the dataset size growths, both languages will start exhibiting the many-to-many ratio of more complete wordnets, with considerably more pairs than synsets or signs. This development is already hinted

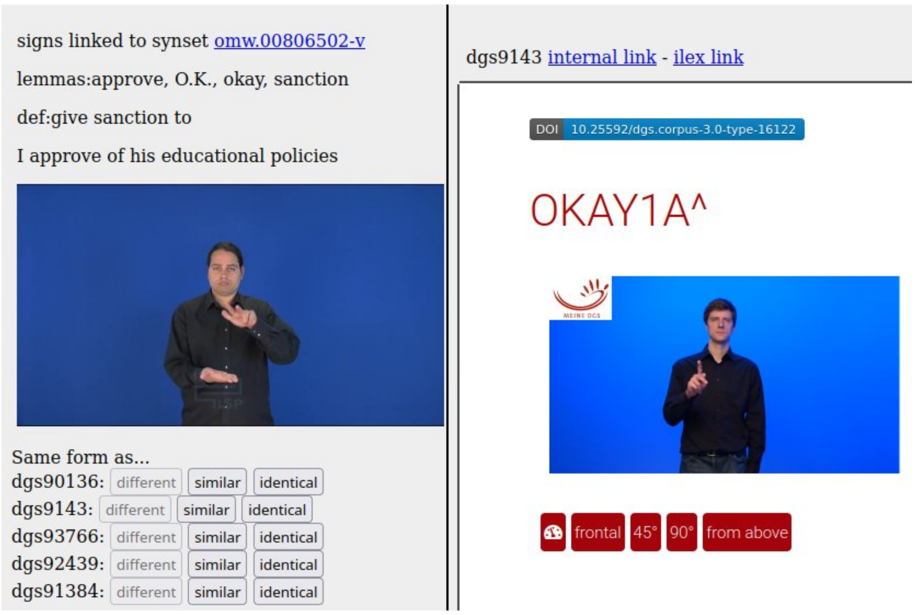

Figure 2: Manual validation interface for comparing GSL and DGS entries associated with the same synset. The interface integrates GSL video from Noema+ (left) and DGS type entry pages from the Public DGS Corpus website (right). The lower left corner lists DGS types that should be compared to the given GSL video to specify whether their sign form is identical, similar or different.

	GSL validated	DGS candidates	DGS validated	GSL/DGS overlap
distinct synsets	4214	27,020	969	278
distinct signs	1819	11,856	2230	*n/a*
sign-synset pairs	4347	138,518	2330	*n/a*

Table 1: Statistics on the current state of linking GSL and DGS vocabulary to OMW.

at by the ratio of automatic DGS candidates, which mirrors the word-synset ratios of the GermaNet entries that they are based on. Some difference, caused by the differences in source data, can however be expected to remain.

4.4. Cross-lingual Connections

Like other wordnet efforts for less-resourced languages, we apply the *expand model* of building on other languages already represented by a wordnet. While spoken language wordnet information is used for this out of necessity, it would be preferable to build on other sign languages where available to be hindered less by modality-specific assumptions.

As we are working on integrating two sign languages in parallel, synergies are used where they present themselves. As the GSL team had already produced a number of synset-sign matches when the DGS team started their manual validation phase, they prioritised synsets which were covered by both automatic DGS matching and GSL.

In addition to validating synset-sign matches, the DGS team also compared the form of the GSL and DGS signs (apart from mouthing) to identify identical and similar signs. The interface for this is shown in Figure 2. Such overlaps between languages can indicate shared iconicity (incidental or otherwise) or other kinds of linguistic relatedness. Annotating these overlaps adds a cross-lingual phonetic relation that is not usually covered by wordnets, but is of great use to research, for example for sign language technologies struggling with data sparsity.

Once signs from both languages are established for a synset, members of either team can inspect which other synsets the sign of the opposing language is connected to. They can then consider whether to expand their own sign to those synsets as well. Synsets with identical/similar forms across languages make particularly good candidates for this step.

5. Conclusion

In this paper we have presented our work on integrating Greek Sign Language and German Sign Language into the Open Multilingual Wordnet. To achieve this, we explore different workflows for working with lexicon-based and corpus-based data and for cross-lingual workflows.

This work has so far resulted in a publicly available dataset of 1819 GSL signs and 2230 DGS signs from existing language resources being linked to 4214 and 969 OMW synsets respectively, including 278 synsets that are covered by both languages. The state of this dataset is preliminary and it will be significantly expanded in size through future updates.[3]

In the long run we intend to add additional languages to this effort. Based on the experience of Bosch and Griesel (2017)

[3]For the latest version of the dataset, see https://doi.org/10.25592/uhhfdm.10168.

with using the *expand model* for less-resourced languages, we expect the required effort for adding new languages will become progressively more manageable as other sign languages can be built upon.

6. Acknowledgements

This work is supported in part by the EASIER (Intelligent Automatic Sign Language Translation) Project. EASIER has received funding from the European Union's Horizon 2020 research and innovation programme, grant agreement n° 101016982.

This publication has been produced in part in the context of the joint research funding of the German Federal Government and Federal States in the Academies' Programme, with funding from the Federal Ministry of Education and Research and the Free and Hanseatic City of Hamburg. The Academies' Programme is coordinated by the Union of the Academies of Sciences and Humanities.

7. Bibliographical References

Bond, F. and Paik, K. (2012). A survey of wordnets and their licenses. In *Proceedings of the 6th Global WordNet Conference*, page 8, Matsue, Japan.

Bond, F., Vossen, P., McCrae, J. P., and Fellbaum, C. (2016). CILI: the collaborative interlingual index. In *Proceedings of the Eighth Global WordNet Conference*, pages 50–57, Bucharest, Romania. University of Iasi.

Bosch, S. E. and Griesel, M. (2017). Strategies for building wordnets for under-resourced languages: The case of African languages. *Literator*, 38(1):1–12. DOI: 10.4102/lit.v38i1.1351.

Ebling, S., Tissi, K., and Volk, M. (2012). Semi-automatic annotation of semantic relations in a Swiss German Sign Language lexicon. In Crasborn, O., et al., editors, *Proceedings of the LREC2012 5th Workshop on the Representation and Processing of Sign Languages: Interactions between Corpus and Lexicon*, pages 31–36, Istanbul, Turkey. European Language Resources Association (ELRA).

Efthimiou, E., Fotinea, S.-E., Dimou, A.-L., Goulas, T., Karioris, P., Vasilaki, K., Vacalopoulou, A., and Pissaris, M. (2016). From a sign lexical database to an SL golden corpus – the POLYTROPON SL resource. In Efthimiou, E., et al., editors, *Proceedings of the LREC2016 7th Workshop on the Representation and Processing of Sign Languages: Corpus Mining*, pages 63–68, Portorož, Slovenia. European Language Resources Association (ELRA).

Efthimiou, E., Vasilaki, K., Fotinea, S.-E., Vacalopoulou, A., Goulas, T., and Dimou, A.-L. (2018). The POLYTROPON parallel corpus. In Bono, M., et al., editors, *Proceedings of the LREC2018 8th Workshop on the Representation and Processing of Sign Languages: Involving the Language Community*, pages 39–44, Miyazaki, Japan. European Language Resources Association (ELRA).

Fellbaum, C., editor. (1998). *WordNet: an electronic lexical database*. Language, Speech, and Communication. MIT Press.

Goulas, T., Fotinea, S.-E., Efthimiou, E., and Pissaris, M. (2010). SiS-Builder: A sign synthesis support tool. In Dreuw, P., et al., editors, *Proceedings of the LREC2010 4th Workshop on the Representation and Processing of Sign Languages: Corpora and Sign Language Technologies*, pages 102–105, Valletta, Malta. European Language Resources Association (ELRA).

Grigoriadou, M., Kornilakis, H., Galiotou, E., Stamou, S., and Papakitsos, E. C. (2004). The software infrastructure for the development and validation of the greek wordnet. *Romanian Journal of Information, Science and Technology*, pages 89–105.

Hamp, B. and Feldweg, H. (1997). GermaNet - a lexical-semantic net for German. In *Proceedings of the ACL Workshop on Automatic Information Extraction and Building of Lexical Semantic Resources for NLP Applications*, pages 9–15, Madrid, Spain. Association for Computational Linguistics.

Jahn, E., Konrad, R., Langer, G., Wagner, S., and Hanke, T. (2018). Publishing DGS Corpus data: Different formats for different needs. In Bono, M., et al., editors, *Proceedings of the LREC2018 8th Workshop on the Representation and Processing of Sign Languages: Involving the Language Community*, pages 83–90, Miyazaki, Japan. European Language Resources Association (ELRA).

Konrad, R., Hanke, T., König, S., Langer, G., Matthes, S., Nishio, R., and Regen, A. (2012). From form to function. a database approach to handle lexicon building and spotting token forms in sign languages. In Crasborn, O., et al., editors, *Proceedings of the LREC2012 5th Workshop on the Representation and Processing of Sign Languages: Interactions between Corpus and Lexicon*, pages 87–94, Istanbul, Turkey. European Language Resources Association (ELRA).

Langer, G. and Schulder, M. (2020). Collocations in sign language lexicography: Towards semantic abstractions for word sense discrimination. In Efthimiou, E., et al., editors, *Proceedings of the LREC2020 9th Workshop on the Representation and Processing of Sign Languages: Sign Language Resources in the Service of the Language Community, Technological Challenges and Application Perspectives*, pages 127–134, Marseille, France. European Language Resources Association (ELRA).

Lualdi, C. P., Wright, E., Hudson, J., Caselli, N. K., and Fellbaum, C. (2021). Implementing ASLNet v1.0: Progress and plans. In *Proceedings of the 11th Global Wordnet Conference*, pages 63–72, Potchefstroom, South Africa. South African Centre for Digital Language Resources (SADiLaR).

Miller, G. A., Beckwith, R., Fellbaum, C., Gross, D., and Miller, K. (1990). Introduction to wordnet: An on-line lexical database. *International Journal of Lexicography*, 3. DOI: 10.1093/ijl/3.4.235.

Müller, A., Hanke, T., Konrad, R., Langer, G., and Wähl, S. (2020). From dictionary to corpus and back again – linking heterogeneous language resources for DGS. In Efthimiou, E., et al., editors, *Proceedings of the LREC2020 9th Workshop on the Representation and Processing of Sign Languages: Sign Language Resources in* *the Service of the Language Community, Technological Challenges and Application Perspectives*, pages 157–164, Marseille, France. European Language Resources Association (ELRA).

Prillwitz, S., Hanke, T., König, S., Konrad, R., Langer, G., and Schwarz, A. (2008). DGS Corpus project – development of a corpus based electronic dictionary German Sign Language / German. In Crasborn, O., et al., editors, *Proceedings of the LREC2008 3rd Workshop on the Representation and Processing of Sign Languages: Construction and Exploitation of Sign Language Corpora*, pages 159–164, Marrakech, Morocco. European Language Resources Association (ELRA).

Shoaib, U., Ahmad, N., Prinetto, P., and Tiotto, G. (2014). Integrating MultiWordNet with Italian Sign Language lexical resources. *Expert Systems with Applications*, 41(5):2300–2308. DOI: 10.1016/j.eswa.2013.09.027.

Troelsgård, T. and Kristoffersen, J. (2018). Improving lemmatisation consistency without a phonological description. the Danish Sign Language corpus and dictionary project. In Bono, M., et al., editors, *Proceedings of the LREC2018 8th Workshop on the Representation and Processing of Sign Languages: Involving the Language Community*, pages 195–198, Miyazaki, Japan. European Language Resources Association (ELRA).