Anna Vacalopoulou/Eleni Efthimiou/
Stavroula-Evita Fotinea/Theodoros Goulas/
Athanasia-Lida Dimou/Kiki Vasilaki

ORGANIZING A BILINGUAL LEXICOGRAPHIC DATABASE WITH THE USE OF WORDNET

Abstract This paper reports on the restructuring of a bilingual (Greek Sign Language, GSL – Modern Greek) lexicographic database with the use of the WordNet semantic and lexical database. The relevant research was carried out by the Institute for Language and Speech Processing (ILSP) / Athena R.C. team within the framework of the European project Easier. The project will produce a framework for intelligent machine translation to bring down language barriers among several spoken/written and sign languages. This paper describes the experience of the ILSP team to contribute to a multilingual repository of signs and their corresponding translations and to organize and enhance a bilingual dictionary (GSL – Modern Greek) as a result of this mapping; this will be the main focus of this paper. The methodology followed relies on the use of WordNet and, more specifically, the Open Multilingual WordNet (OMW) tool to map content in GSL to WordNet synsets.

Keywords WordNet; semantic network; lexicographic database; multimodal database; sign language resources; Greek Sign Language

1. Paper outline

The first section of the paper gives a brief account of all the different parameters of this reorganization process. After a small summary of the Easier project, in the framework of which this research was undertaken (section 1.1), follows a description of the existing lexicographic database, namely, the Noema+ bilingual dictionary, which was based on a multimodal bilingual corpus (section 1.2). The main part of the paper gives an account of the methodology that was used (section 2), first to automatically link the video signs of the dictionary to different synsets in WordNet (section 2.1), as well as to manually review, amend, and validate this mapping (section 2.2.1) through the use of the Greek part of the Open Multilingual WordNet tool. After that, follows an account of the enhancement of the dictionary database based on the semantic and lexical network of relations between words, namely, synonymy, and then between synsets themselves, namely, hyponymy/hypernymy (section 2.2.2). The paper concludes with an evaluation of this semi-automatic enhancement process and some suggestions for future research (section 3).

1.1 The Easier project

Easier (intElligent Automatic SIgn languagE tRanslation)[1] is an ambitious project undertaken by fourteen European institutions that have joined forces to make available, within a period of three years, a unique and innovative service, i. e., an intelligent machine translation framework to bring down language barriers among several spoken/written and sign languages. The main aim of Easier is to develop a service that will facilitate automatic translation between any two pair of European languages, be them sign languages (SL) or spoken/

[1] https://www.project-easier.eu

This paper is part of the publication: Klosa-Kückelhaus, Annette/Engelberg, Stefan/Möhrs, Christine/Storjohann, Petra (eds.) (2022): Dictionaries and Society. Proceedings of the XX EURALEX International Congress. Mannheim: IDS-Verlag.

written languages. To this end, project partners contribute their own tools, technologies, and resources (for different sign languages). The relevant technologies that Easier comprises in this process are SL animation, SL recognition, machine translation, and communication technologies.

As far as the translation component is concerned, one of the main challenges has been to gather enough amounts of data for the required training of the machine translation system. Even though most of the European SLs share a good part of their grammar (including their phonology)[2] and some lexicon (cf. e. g., Pizzuto/Volterra 2000), which is utilised in the project, there is great diversity among various sign languages. To make things harder, the raw data which will feed the machine learning process, however rich, is quite diverse. It, therefore, brings together a collection of different video formats, types of material (e. g., corpora, glossaries, dictionaries), annotation schemes and annotation levels (e. g., in the inclusion of SL-specific aspects such as handedness),[3] transcription focus (i. e., phonology as opposed to meaning and function), etc.

In addition, data is scarce for most SLs, even in the EU context, for a variety of reasons. Apart from the fact that all SLs are minority languages and, as a result, they are more likely to be low on resources compared to more widely used languages, collecting and properly annotating SL material is a time-consuming process, which demands a considerable investment in terms of both financial and human capitals (Vacalopoulou 2020, p. 431). In an attempt to bridge this gap, project partners have been linking different SL resources together and providing them with a common detailed phonological representation system. A central part of this effort is the attempt to harmonize this diverse set of data and make them usable for machine learning purposes. To this end, the project team, led in this task by the Institute for German Sign Language and Communication of the Deaf at the University of Hamburg, have been linking different SL resources together and providing them with a common detailed phonological representation system. The goal is to come up with a transferable system of phonological representation and grammar to be utilized for more under-resourced European SLs.

One of the ways to proceed with this connection is the utilization of the WordNet semantic and lexical database. This selection was largely based on the assumption that the translation of European SLs into their respective spoken/written languages could never accurately grasp the full meaning of the original sign content; inevitably, something would be lost in translation. If one was to project this deviation in meaning across all languages of the project (both signed and spoken), it is not hard to understand that a considerable amount of meaning could be lost in the process. Therefore, the WordNet solution was seen as a way to get round the problem of spoken language interference and find a way to connect signs from different SLs via their meaning (see Bigeard et al. 2022).

[2] According to Brentari/Fenlon/Cormier (2018, Summary), "sign language phonology is the abstract grammatical component where primitive structural units are combined to create an infinite number of meaningful utterances. Although the notion of phonology is traditionally based on sound systems, phonology also includes the equivalent component of the grammar in sign languages, because it is tied to the grammatical *organization*, and not to particular content."

[3] Handedness in sign production refers to the dominant hand a signer uses without altering the intended meaning of the content.

This paper is part of the publication: Klosa-Kückelhaus, Annette/Engelberg, Stefan/Möhrs, Christine/Storjohann, Petra (eds.) (2022): Dictionaries and Society. Proceedings of the XX EURALEX International Congress. Mannheim: IDS-Verlag.

This paper will focus on the experience of the ILSP team with the use of WordNet and, more specifically, the Open Multilingual WordNet (OMW) tool to map content in Greek Sign Language (GSL) to WordNet synsets with a twofold purpose in mind:

1) Contribute to a multilingual repository of signs and their corresponding translations, and
2) Organizing and enhancing the Noema+ bilingual dictionary (GSL – Modern Greek, hereby 'Greek') as a result of this mapping; this will be the main focus of this paper.

1.2 The Noema+ multimodal database

One of the sign resources that have been gathered to contribute to the Easier project is Noema+,[4] a bilingual (Modern Greek – GSL) dictionary currently comprising more than 12,000 entries. This is the most extensive reference work for this language pair to date that has combined various smaller resources and undergone multiple phases of revision and update. The lexical database of the dictionary is a fully annotated multipurpose-multiuse resource.

In terms of its source material, Noema+ was based on the extensive Polytropon bilingual corpus.[5] This resource has been used to build several end products, mainly targeting the bilingual education of deaf children and GSL learning as a second language, as well as services such as the enhancement of the official platform for secondary education in Greece, and an e-class platform as adapted by the Technical Vocational Institute of Athens for accessibility (Efthimiou et al. 2016). Furthermore, it is exploitable for developing a series of SL technologies, including information extraction, Web accessibility tools, incorporation of SL lexical information in natural language processing systems as in the case of machine translation from and into GSL, creation of training material for sign recognition and input to sign synthesis tools enabling signing by virtual signers, i. e., avatars. This extensive corpus was incorporated in the dictionary after an evaluation process of several internal and external stages involving lexicographers, GSL experts, and end-users (Efthimiou et al. 2019). The dictionary was developed in SiS-Builder,[6] a specially designed web-based open environment that enables lexicographers to access other relevant lexical resources and tools in the compilation process (Goulas et al. 2010).

In terms of dictionary microstructure, each video entry consists of one or several translation equivalents in Greek, the use of which is shown in simple, one-sentence examples in both languages. Other microstructure elements contain GSL synonyms, such as *ξυράφι* ('razor') in Figure 1; Greek synonyms, which were added mainly for search purposes to offer a variety of starting points to users who want to look up a certain sign; and multi-word expressions which are cross-referenced to the respective single-word entries.

[4] Noema+ is freely available in ILSP's Sign Language Technologies team website: http://sign.ilsp.gr/signilsp-site/index.php/en/home-3/.

[5] Part of the annotated Polytropon bilingual parallel corpus is freely available via the *clarin:el* repository, the Greek sector of CLARIN, the European infrastructure for language resources and technology: https://inventory.clarin.gr/corpus/835.

[6] SiS-Builder can be accessed in http://sign.ilsp.gr/sisbuilder/.

This paper is part of the publication: Klosa-Kückelhaus, Annette/Engelberg, Stefan/Möhrs, Christine/Storjohann, Petra (eds.) (2022): Dictionaries and Society. Proceedings of the XX EURALEX International Congress. Mannheim: IDS-Verlag.

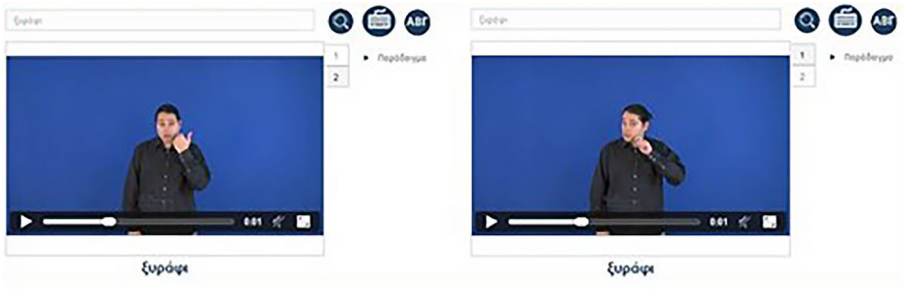

Fig. 1: GSL synonyms (two different signs) for the same meaning (*ξυράφι*, 'razor') in Noema+

2 Linking Noema+ to the WordNet database

The decision to map the entries in Noema+ to the WordNet database was based on the idea to assign concepts directly to signs without having to go through the words of oral language. Compared to more extended sign resources such as corpora, Noema+ is more context-free as it works primarily as a standalone bilingual dictionary; as in any dictionary, its entries consist of items in isolation – in this case, signs – which are then put into context in the examples of use. Starting from work undertaken for the Easier project, this resource's entries were mapped to corresponding WordNet concepts. The goal was, firstly, to enhance the content of the dictionary database with more alternative translations of signs in Greek and, secondly, to attach translations in more languages towards the future end of making the database multilingual.

2.1 Automatic mapping to the Greek WordNet

In the first stage of this process, the dictionary entries were automatically mapped to WordNet synsets through the use of OMW. Synsets are sets of unordered synonyms that correspond to the same concept accompanied by a simple definition that can have the form of an "explanatory gloss" (Fellbaum 1998) and, in some cases, by a domain label (Fellbaum 2006). An example of the two synsets for the Greek word *βιβλίο* ('book') can be seen in Figure 2. Thus, synset 03165211-n corresponds to 'an accounting journal as a physical object' whereas synset 02870092-n covers 'physical objects consisting of a number of pages bound together'. OMW provides access to synsets in a multitude of languages, linking back to the respective WordNets (Bond/Paik 2012). Figure 3 presents an example of the 02870092-n synset in various languages.

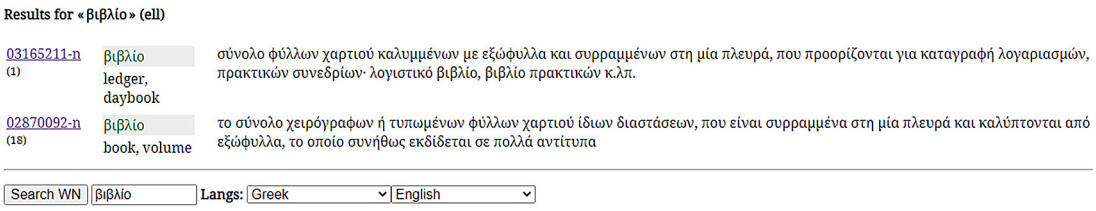

Fig. 2: Search results for the Greek word *βιβλίο* in the Open Multilingual WordNet

This paper is part of the publication: Klosa-Kückelhaus, Annette/Engelberg, Stefan/Möhrs, Christine/Storjohann, Petra (eds.) (2022): Dictionaries and Society. Proceedings of the XX EURALEX International Congress. Mannheim: IDS-Verlag.

02870092-n ✪ 'physical objects consisting of a number of pages bound together'; Search WN

Albanian	*vëllim , libër*
Arabic	*مُجلّد , كِتاب*
Bulgarian	*издание , книга , том*
Catalan	*llibre , volum , exemplar*
Chinese (simplified)	*书籍 , 书*
Danish	*bog*
Greek	*βιβλίο*
English	*book$_{10}$ (▻) , volume$_{8}$*
Basque	*liburu*
Finnish	*volyymi , kirja*
French	*volume , livre*
Galician	*volume , libro , exemplar*
Croatian	*knjiga*
Indonesian	*kitab , jilid , buku*
Icelandic	*bindi , bókarbindi*
Italian	*tomo , volume , libro*
Japanese	*本 , 書 , 冊 , 冊子本 , 冊子 , 図書 , ブック , 篇帙 , 一巻 , -冊 , 書誌 , 書物 , 書籍 , 書史 , 書冊 , 書帙 , 書巻*
Lithuanian	*tomas , knyga*
Dutch	*jaargang*

Fig. 3: Book in different languages in the Open Multilingual WordNet

In this initial phase, the Greek part of WordNet was downloaded from OMW in the form of a text file with tab delimited values, which is available in the rich resources archive of the website. This file comprises approximately 42,200 lines and contains entries that are broken down in several rows. This format was not very practical to use in parallel with the Noema+ database; as a result, the table was converted to a more suitable layout, in which all the contents of the same lexical entry were moved in the same row. Thus, as the goal was to check concepts against GSL video signs without the interference of their Greek equivalents, lexical items sharing a common WordNet ID were grouped together in one entry containing the entire WordNet synset. Figure 4 shows the initial arrangement of the WordNet entry #00006238-v for the Greek *αποχρέμπτομαι* ('expectorate') and the results of the automatic rearrangement in the mapping process. In the first version, the contents of the synset (*αποχρέμπτομαι*, *πτύω*, and *φτύνω*) are aligned horizontally, whereas in the second version, they are arranged vertically to facilitate the automatic mapping process.

Result Grid | Filter Rows: 00006238-v | Edit: | Export/Import: | Wrap Cell Content:

ID	wordnetID	ell:lemma	subID	def
15	00006238-v	αποχρέμπτομαι, πτύω, φτύνω	0	αποβάλλω φλέγματα από τα πνευμόνια με απόπτυση
NULL	NULL	NULL	NULL	NULL

Result Grid | Filter Rows: 00006238-v | Export: | Wrap Cell Content: | Fetch rows:

wordNetID	lemma	subID	Definition
00006238-v	αποχρέμπτομαι		
00006238-v	πτύω		
00006238-v	φτύνω		
00006238-v		0	αποβάλλω φλέγματα από τα πνευμόνια με απόπτυση

Fig. 4: Initial layout (above) and its rearrangement (below) of the OMW entry for *αποχρέμπτομαι* ('expectorate') before mapping it to the Noema+ lexical database

This paper is part of the publication: Klosa-Kückelhaus, Annette/Engelberg, Stefan/Möhrs, Christine/Storjohann, Petra (eds.) (2022): Dictionaries and Society. Proceedings of the XX EURALEX International Congress. Mannheim: IDS-Verlag.

In the next stage, the newly created table (WordNetGR) was joined with the bilingual dictionary database with SQL queries. As the entire synset was now one single entry in the table, it was split by the number of its contents within the SQL query so that it could match dictionary entries and thus create the mappings to the GSL videos. An example can be seen in Figure 5, which shows WordNet ID #00220869-v for *δυναμώνω* ('strengthen'), which has 2 delimiters as the synset contains 3 Greek synonyms: *δυναμώνω*, *ενισχύω*, *ισχυροποιώ*. In order to execute the SQL join query, it was necessary to split the field to its contents. The corresponding dictionary entry to which this synset was mapped was rearranged to a total of three entries, each containing one of the three Greek synonyms. The third entry was included in the SQL-results, because the corresponding equivalent in Greek is not contained in the bilingual dictionary.

The result of this process was the full mapping of dictionary entries to corresponding synsets in WordNet. For quality control purposes, this result was later double-checked by GSL experts, who made corrections if appropriate.

Result Grid | Filter Rows: | Export: | Wrap Cell Content: | Fetch rows:

wordNetID	ell:lemma	Ελληνικό αντίστοιχο (NOHMA)	ell:lemma1	ell:lemma2	ell:lemma3	ell:lemma4
00220869-v	δυναμώνω, ενισχύω, ισχυροποιώ	δυναμώνω	δυναμώνω	ενισχύω	ισχυροποιώ	NULL
00220869-v	δυναμώνω, ενισχύω, ισχυροποιώ	ενισχύω	δυναμώνω	ενισχύω	ισχυροποιώ	NULL
00217499-n	δυστύχημα, καταστροφή, συμφορά	δυστύχημα	δυστύχημα	καταστροφή	συμφορά	NULL
00217499-n	δυστύχημα, καταστροφή, συμφορά	καταστροφή	δυστύχημα	καταστροφή	συμφορά	NULL
00212414-v	διατηρώ, προφυλάσσω, συντηρώ	διατηρώ	διατηρώ	προφυλάσσω	συντηρώ	NULL
00212414-v	διατηρώ, προφυλάσσω, συντηρώ	προφυλάσσω	διατηρώ	προφυλάσσω	συντηρώ	NULL

Fig. 5: WordNet ID #00220869-v for *δυναμώνω* ('*strengthen*') produced two extra entries in the dictionary database after mapping the entries to the corresponding synonyms

2.2 Manual organization of the Noema+ database

The second part of the mapping process was the manual processing of the automatically generated results. The purpose of this stage was to review and amend these results on one hand, and to enhance the content of the dictionary database on the other hand.

2.2.1 Review and amendments

Following the automatic mapping of GSL signs to their corresponding Greek words in OMW, an important goal was to evaluate whether GSL signs and OMW Greek words were semantically equivalent. To this end, each sign was carefully checked by expert GSL signers against its corresponding OMW synset to ensure that they shared the same meaning. In order to decide whether an OMW Greek equivalent indeed corresponded to the meaning of a particular sign, several parameters were taken into consideration including the videos of both isolated signs and their linked examples of use in the Noema+ dictionary, as well as the definitions or explanatory glosses of synsets. An example of this process would be the sign entry for *αστέρι* ('star'), which had been automatically linked to the following two Greek OMW synsets:

1) WordNet ID #09444100-n, i. e., "a celestial body of hot gases that radiates energy derived from thermonuclear reactions in the interior", and
2) WordNet ID #09762509-n "someone who is dazzlingly skilled in any field".

This paper is part of the publication: Klosa-Kückelhaus, Annette/Engelberg, Stefan/Möhrs, Christine/Storjohann, Petra (eds.) (2022): Dictionaries and Society. Proceedings of the XX EURALEX International Congress. Mannheim: IDS-Verlag.

Although the mapping was successful in terms of the equivalents in Greek, it is only the first sense of this polysemous word that corresponds to the video equivalent found in the GSL database. The second sense was, therefore, deleted from the dictionary database. Luckily, almost 70% of the GSL synonyms that had been automatically linked to the same WordNet ID were verified as correct matches at this stage, rendering the automatic mapping process successful.

Following the abovementioned process, all entries in the new combined database were double-checked manually one by one by GSL experts, and items were deleted or verified accordingly, before proceeding to the next stage of dictionary enhancement.

2.2.2 Enhancing the Noema+ database through WordNet

The next, and final, stage in this process was the enhancement of the dictionary database with new entries as a result of their mapping to WordNet synsets. This was pursued mainly by exploiting the richness of each synset, which in some cases contained multiple synonymic items, and by further experimenting with other lexical relations, such as hyponymy and hypernymy.

In regard to the synsets, it was found that they were helpful in the enrichment of the dictionary with multiple Greek equivalents in this relatively quick, semi-automatic process. Thus, more synonymic items from the corresponding synset were added to the database under each GSL entry after being validated by GSL experts. Examples of this type of enrichment include items such as *κατάστημα* ('store') and *μαγαζί* ('shop'), or *αναζητώ* ('search') and *ψάχνω* ('look for').

Apart from this first-level enhancement, the mapped WordNet synsets gave the team of experts the opportunity to further explore polysemy in the context of GSL by adding more senses to specific video signs. For instance, whereas the sign for *άνεση* ('comfort') in GSL had been automatically mapped to WordNet ID #14491889-n, i.e., "freedom from financial difficulty that promotes a comfortable state", careful examination of the Greek WordNet in OMW produced additional senses of the corresponding word in Greek that were found to be translations of the original GSL sign as well. In this case, as shown in Figure 6, the same video was linked to two additional WordNet IDs, namely, #14445379-n ("a state of being relaxed and feeling no pain"), and #07492516-n ("a feeling of freedom from worry or disappointment"), as all three senses are represented by the same sign in GSL. The added value of this procedure was that it allowed for a more detailed documentation of polysemy in GSL.

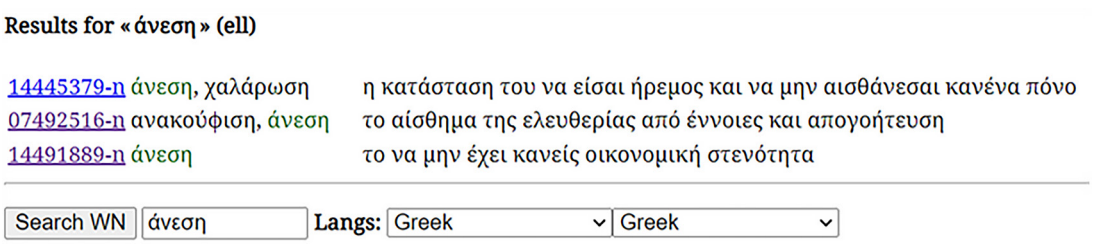

Fig. 6: Multiple WordNet IDs were linked to the same GSL video sign allowing for a more detailed documentation of GSL polysemy

Taking the enhancement process one step further from the synset itself, effort was made to explore lexical relations in which synsets themselves are connected to one another, namely,

This paper is part of the publication: Klosa-Kückelhaus, Annette/Engelberg, Stefan/Möhrs, Christine/Storjohann, Petra (eds.) (2022): Dictionaries and Society. Proceedings of the XX EURALEX International Congress. Mannheim: IDS-Verlag.

hyponymy and hypernymy. Thus, synsets that link to other synsets through such relations were analysed for the possibility of offering ways to further enhance the dictionary database.

Fig. 7: Lexical relations of synset #10794014-n for *συγγραφέας* ('writer') with other synsets

A relevant example is the case of *συγγραφέας* ('writer'), which is connected to multiple hyponyms as shown in Figure 7. As expected, 'writer' is linked to several hyponymic synsets such as 'biographer', 'dramatist', and 'poet'. As observed, this could be a valuable source of creating additional dictionary entries for senses that had not been linked to a dedicated sign by the natural GSL signers, who served as informers in the development of the corpus upon which the first version of the dictionary was based. An example of this would be the hyponym of 'writer' δραματουργός ('playwright') for which there was no entry in the dictionary database as no dedicated sign had been reported to date. In this case, the lexical relations between synsets recorded in WordNet allowed GSL experts to add a compound sign consisting of the simple signs for *θεατρικό* 'play' and *γράφω* 'write' to represent this concept (Figure 8). Indeed, this kind of compounding is a very common technique in sign language morphology (Sandler/Lillo-Martin 2006, pp. 72–75) and a natural way of producing new lexicon within signing communities. Nevertheless, as valuable, and extremely promising as this technique is, it needs to be applied with extra care, as its results are far from self-evident and require multiple levels of validation by native signers and/or against GSL corpora.

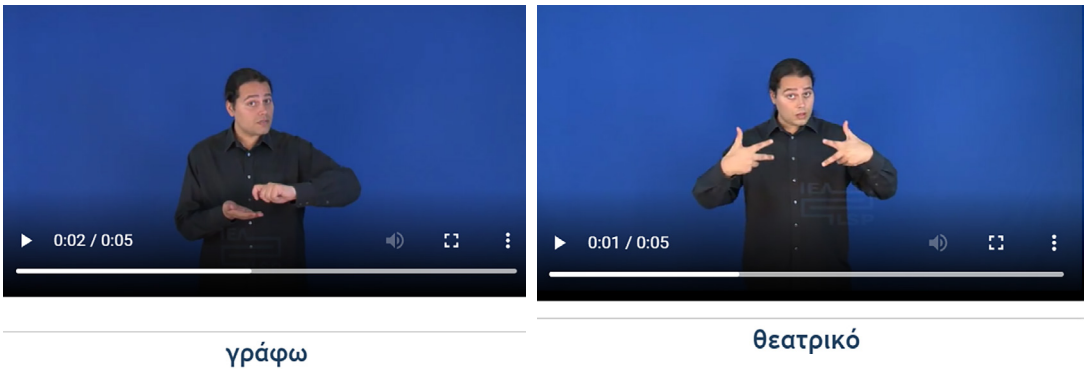

Fig. 8: Combining the signs for *γράφω* (write) and *θεατρικό* (play) to enhance the dictionary database with a new sign for *δραματουργός* (playwright)

3 Conclusion and future steps

This contribution presented an account of the rearrangement and enrichment of the database of the NOEMA+ bilingual GSL – Greek dictionary after mapping its entries to corresponding WordNet synsets. In this ongoing process, more than 1,200 lexical items have already been identified as possible candidates for dictionary enhancement either from analysing the synsets themselves, or by an extended investigation of lexical relations between synsets. These will be evaluated by the team of lexicographers and GSL experts for inclu-

This paper is part of the publication: Klosa-Kückelhaus, Annette/Engelberg, Stefan/Möhrs, Christine/Storjohann, Petra (eds.) (2022): Dictionaries and Society. Proceedings of the XX EURALEX International Congress. Mannheim: IDS-Verlag.

sion in NOEMA+ as this research is still in progress. Based on these promising results, the team will explore further enhancement possibilities towards opening the dictionary to more languages through their respective WordNets with a view of making the resource multilingual.

References

Bigeard, S./Schulder, M./Kopf, M./Hanke, T./Vasilaki, K./Vacalopoulou, A./Goulas, T./Dimou, A.-L./Fotinea, S.-E./Efthimiou, E. (2022): Introducing sign languages to a multilingual WordNet: bootstrapping corpora and lexical resources of Greek Sign Language and German Sign Language. In: Proceedings of 10th Workshop on the Representation and Processing of Sign Languages: Multilingual Sign Language Resources (LREC-2022), Marseille, France, 20–25 June 2022.

Bond, F./Paik, K. (2012): A survey of wordnets and their licenses. In: Proceedings of the 6th Global WordNet Conference (GWC 2012). Matsue, pp. 64–71.

Brentari, D./Fenlon, J./Cormier, K. (2018): Sign language phonology: Oxford research encyclopedia of linguistics. https://oxfordre.com/linguistics/view/10.1093/acrefore/9780199384655.001.0001/acrefore-9780199384655-e-117#acrefore-9780199384655-e-117-div1-1 (last access: 14-05-2022).

Efthimiou, E./Fotinea, S-E./Goulas, T./Vacalopoulou, A./Vasilaki, K./Dimou, A.-L. (2019): Sign language technologies and the critical role of SL resources in view of future internet accessibility services. In: Technologies 7 (1), 18, pp. 1–21.

Efthimiou, E./Fotinea, S.-E./Dimou, A.-L./Goulas, T./Karioris, P./Vasilaki, K./Vacalopoulou, A./Pissaris, M./Korakakis. D. (2016): From a sign lexical database to an SL golden corpus – the POLYTROPON SL resource. In Proceedings of 7th Workshop on the Representation and Processing of Sign Languages: Corpus Mining (LREC-2016), Portorož, Slovenia, 23–28 May 2016, pp. 63–68.

Fellbaum, C. (ed.) (1998): Wordnet: an electronic lexical database. Cambridge, MA.

Fellbaum, C. (2006): WordNet(s). In: Brown, Keith et al. (eds.): Encyclopedia of language and linguistics. Second Edition. Oxford, pp. 665–670.

Goulas, T./Fotinea, S.-E./Efthimiou, E./Pissaris, M. (2010): SiS-Builder: a sign synthesis support tool. In: Proceedings of the 4th Workshop on the Representation and Processing of Sign Languages: Corpora and Sign Language Technologies (CSLT 2010), Satellite Workshop of the LREC-2010 Conference, Valetta, Malta, 17–23 May 2010, pp. 102–105.

Grigoriadou, M./Kornilakis, H./Galiotou, E./Stamou, S./Papakitsos, E. (2004): The software infrastructure for the development and validation of the Greek Wordnet. In: Romanian Journal of Information Science and Technology 7 (1–2), pp. 89–105.

Noema+ (2022): http://sign.ilsp.gr/signilsp-site/index.php/en/home-3/ (last access: 09-03-2022).

Open Multilingual Wordnet (OMW): http://compling.hss.ntu.edu.sg/omw/ (last access: 14-03-2022).

Pizzuto, E./Volterra, V. (2000): Iconicity and transparency in sign languages: a cross-linguistic cross-cultural view. In: Emmorey, K./Lane, H. (eds.): The signs of language revisited: an anthology to honor Ursula Bellugi and Edward Klima. Mahwah, pp. 229–250.

Polytropon Parallel Corpus, Version: 1.0.0 (2022): https://inventory.clarin.gr/corpus/835 (last access: 21-03-2022).

Princeton University (2010): About WordNet. https://wordnet.princeton.edu/ (last access: 07-03-2022).

Sandler, W./Lillo-Martin, D. (2006): Sign language and linguistic universals. Cambridge.

SiS-Builder (Sign Synthesis Builder) (2022): http://sign.ilsp.gr/sisbuilder/ (last access: 21-03-2022).

This paper is part of the publication: Klosa-Kückelhaus, Annette/Engelberg, Stefan/Möhrs, Christine/Storjohann, Petra (eds.) (2022): Dictionaries and Society. Proceedings of the XX EURALEX International Congress. Mannheim: IDS-Verlag.

Vacalopoulou, A. (2020): Sign language corpora and dictionaries: a multidimensional challenge. In: Proceedings of the EURALEX XIX Congress of the European Association for Lexicography – Lexicography for inclusion, Alexandroupolis, 7–11 September 2021. Proceedings Book Volume 1, pp. 427–434.

Contact information

Anna Vacalopoulou
Institute for Language and Speech Processing, ATHENA RC
avacalop@athenarc.gr

Eleni Efthimiou
Institute for Language and Speech Processing, ATHENA RC
eleni_e@athenarc.gr

Stavroula-Evita Fotinea
Institute for Language and Speech Processing, ATHENA RC
evita@athenarc.gr

Theodoros Goulas
Institute for Language and Speech Processing, ATHENA RC
tgoulas@athenarc.gr

Athanasia-Lida Dimou
Institute for Language and Speech Processing, ATHENA RC
ndimou@athenarc.gr

Kiki Vasilaki
Institute for Language and Speech Processing, ATHENA RC
kvasilaki@athenarc.gr

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 paper is part of the publication: Klosa-Kückelhaus, Annette/Engelberg, Stefan/Möhrs, Christine/Storjohann, Petra (eds.) (2022): Dictionaries and Society. Proceedings of the XX EURALEX International Congress. Mannheim: IDS-Verlag.