Sign Language Translation Mobile Application and Open Communications Framework

Deliverable 4.1: First symbolic intermediate representation - InterL-S

This project has received funding from the European Union's Horizon 2020 Research and Innovation Programme under Grant Agreement No. 101017255

Project Information
Project Number: 101017255
Project Title: SignON: Sign Language Translation Mobile Application and Open Communications Framework
Funding Scheme: H2020 ICT-57-2020
Project Start Date: January 1st 2021

Deliverable Information
Title: First Routines for transformation of text from and to InterL
Work Package: WP 4 – Transfer and InterLingual Representations
Lead beneficiary: Katholieke Universiteit Leuven
Due Date: 30/06/2022
Revision Number: V0.3
Authors: Tim Van de Cruys, Maud Goddefroy, Ineke Schuurman, Euan McGill
Dissemination Level: Public
Deliverable Type: Other

Overview: This document presents the first steps in creating a symbolic language-independent intermediate representation (interL-S), including (1) the creation of a rule-based system for data augmentation with synthetic glosses, (2) the implementation of a word sense disambiguation (WSD) scheme to link to WordNet and the creation of SignNets and (3) the investigation of a language-independent representation of the meaning of input sentences (abstract meaning representation, AMR).

Revision History

Version #	Implemented by	Revision Date	Description of changes
v0.1	Tim Van de Cruys, Maud Goddefroy, Ineke Schuurman	10/06/2022	Initial text and structure
v0.2	Euan McGill	24/06/2022	Added LSE subsection
v0.3	Tim Van de Cruys	28/06/2022	Integration of comments from consortium partners

The SignON project has received funding from the European Union's Horizon 2020 Programme under Grant Agreement No. 101017255. The views and conclusions contained here are those of the authors and should not be interpreted as necessarily representing the official policies or endorsements, either expressed or implied, of the SignON project or the European Commission. The European Commission is not liable for any use that may be made of the information contained therein.

Approval Procedure

Version #	Deliverable Name	Approved by	Institution	Approval Date
V0.1	D4.1	Aoife Brady	DCU	23/06/2022
V0.1	D4.1	Marcello Scipioni, Marco Giovanelli	FINCONS	27/06/2022
V0.1	D4.1	Vincent Vandeghinste	INT	22/06/2022
V0.1	D4.1	Olatz Perez de Viñaspre	UPV/EHU	22/06/2022
V0.1	D4.1	John O'Flaherty	MAC	22/06/2022
V0.1	D4.1	Horacio Saggion	UPF	22/06/2022
V0.1	D4.1	Irene Murtagh	TU Dublin	22/06/2022
V0.1	D4.1	Mathieu De Coster	UGent	22/06/2022
V0.1	D4.1	Jorn Rijckaert	VGTC	23/06/2022
V0.1	D4.1	Ellen Rushe	NUID UCD	27/06/22
V0.1	D4.1	Henk van den Heuvel	RU	23/06/2022
V0.3	D4.1	Tim Van de Cruys	KU Leuven	28/06/2022
V0.1	D4.1	Rehana Omardeen	EUD	22/06/2022
V0.1	D4.1	Mirella De Sisto	TiU	25/06/2022
V0.1	D4.1	Lorraine Leeson	TCD	22/06/2022

Acronyms

The following table provides definitions for acronyms and terms relevant to this document.

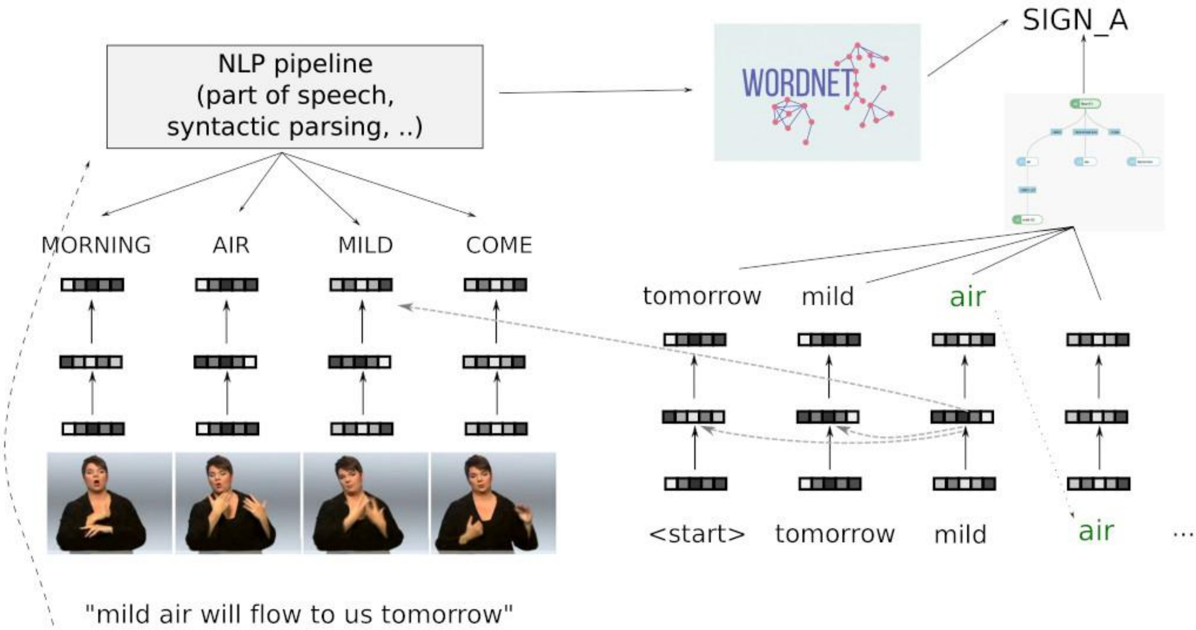
Acronym	Definition
AMR	abstract meaning representation
BLEU	Bilingual evaluation understudy metric
WSD	word sense disambiguation
VGT	Vlaamse Gebarentaal (Flemish Sign Language)
NGT	Nederlandse Gebarentaal (Sign Language of the Netherlands)
LSE	Lengua de Signos Española (Spanish Sign Language)
NER	Named Entity Recognition
NLP	Natural Language Processing
NLU	Natural Language Understanding
OVS	Object-Subject-Verb
PoS	Part of Speech
RDF	Resource Description Framework
SL	Sign Language
SLR	Sign Language Recognition
SVO	Subject-Verb-Object
WP	Work Package
InterL(-E, -S)	Interlingua (embedding, symbolic)

Table of contents

1. Introduction

This document describes the work carried out in the context of task 4.1 of work package (WP) 4 "First symbolic intermediate representation - InterL-S" in the scope of the project "SignON: Sign Language Translation Mobile Application And Open Communication Framework". In WP4, we investigate a language independent intermediate representation, InterL. In a previous deliverable (D4.3), we introduced InterL-E, an intermediate representation based on embeddings. In parallel, we investigate a symbolic approach (InterL-S), based on traditional NLP annotations. The goal within the project is to explore both approaches, investigate their respective strengths and shortcomings, and eventually combine them in order to achieve better results. A general overview of the different subtasks of 4.1 is given in Figure 1.

Figure 1. An overview of the different subtasks of 4.1, viz. data augmentation (section 2), the induction of WordNet senses (section 3), and the prediction of abstract meaning representation (AMR) structures (section 4).



Firstly, we designed a rule-based system that generates synthetic glosses based on spoken language input (section 2). At the moment, we designed a system for Flemish Sign Language (VGT), and we are working on systems for Sign Language of the Netherlands (NGT) and Spanish Sign Language (LSE). Each system is based on linguistic sign language knowledge. The result can improve the performance of the InterL-E system by providing additional training data (i.e., data augmentation).

We also focused on WordNets (section 3). WordNet is a lexical database for English that groups words with the same meaning into synsets[1] (Fellbaum, 1998), and subsequently different versions have been developed for many languages. Synsets in different WordNets have been linked by means of Interlingual Indices (ILI), making it possible to look up the translation of a synset. Initially, we considered the development of a WordNet for VGT (since one does not exist yet) which we refer to as a SignNet (3.1). In order to make use of WordNet as an interlingual resource, we need to link incoming text to the appropriate synsets. This task is called Word Sense disambiguation (WSD, 3.2). The resulting synsets are useful as part of an InterL-S, since they link meanings over different languages in a symbolic way.

Finally, we investigated abstract meaning representation (AMR), which is used to capture the semantic structure of an input (section 4).

2. Rule-based spoken language to gloss system

Sign languages do not have a written form (Baker et al., 2016). For linguistic research purposes, there are several transcription/annotation approaches that are taken. For example, in linguistic research, glosses can be used to annotate sign videos. A gloss represents the meaning of a sign as closely as possible using a word from a spoken language. Glosses are often close to (one of) the possible translations of the sign, but should not be seen as a translation. Often, more than one word is required to gloss a single sign or, conversely, one word may capture the intent of several signs in a sequence. Often, only the manual component of the established lexicon is glossed, making glosses an incomplete representation of any sign language. Despite these shortcomings, researchers have successfully used glosses within sign2text MT systems, called sign2gloss2text (Yin & Read, 2020) and sign2(gloss+text) (Camgöz et al., 2020; De Coster et al., 2021).

Although previous research shows the importance of gloss-annotations, when focusing on the sign languages included in the SignON project, we do not have enough gloss-annotated data to successfully train such a system available for most of them. This is because annotating with glosses is a time-consuming, manual task that has not yet been accurately automated. For this reason, we employ a rule-based system to augment existing parallel data containing sign videos and spoken language translations with synthetic glosses. This data augmentation technique is a continuation of the technique proposed by Moryossef et al. (2021) and can, as proposed by them, also be applied to monolingual data in order to create a corpus for training gloss2text models. Compared to their

[1] A group of words (*synonym set)* that can all refer to the same concept.

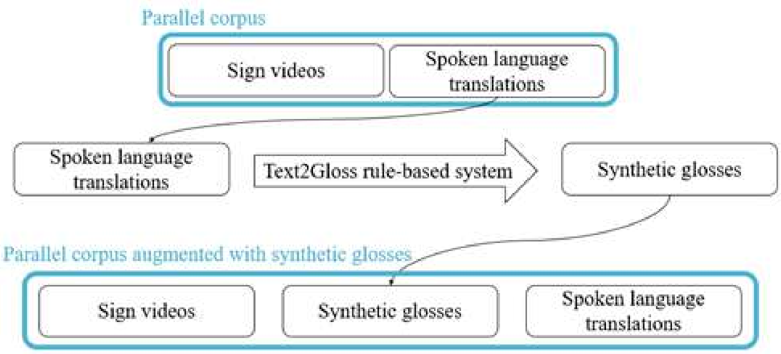
system, we propose a larger set of rules that are all linguistically informed. With the augmented data, it will be possible to train sign2gloss2text or sign2(gloss+text) systems.

2.1 Data augmentation

To alleviate the lack of parallel data for training neural MT models in low-resource languages, multiple techniques have been developed, such as data augmentation and back-translation. In back-translation, new pseudo-parallel data is created from monolingual corpora in the target language in an automatic way (Sennrich et al., 2016). A reasonable amount of data is still needed to train the model performing the back-translation, which is lacking for low-resource languages, such as sign languages (Xia et al., 2019).

This is why Moryossef et al. (2021) propose the use of data augmentation instead, creating synthetic glosses from monolingual corpora in the target language by means of a rule-based system. In addition to this application, we propose to use the system to augment existing parallel data that lack gloss-annotations with synthetic glosses (cfr. Figure 2). With the augmented data, sign2gloss2text or sign2(gloss+text) systems can be trained for sign languages that do not have a (large enough) corpus available with gloss annotations.

Figure 2. Illustration of how the synthetic glosses can be used for data augmentation. We use the translation from a parallel corpus without gloss-annotations to generate synthetic glosses that can be used to augment the corpus.



There are two methodological differences between Moryossef et al.'s (2021) study and our current work. Firstly, Moryossef et al. (2021) studied German Sign Language (DGS) and American Sign Language (ASL), two languages not included in the SignON project. Secondly, while they focused on general and some language-specific rules, our system is composed of a large set of rules that are linguistically informed and operate on different levels.

2.2 Dutch to synthetic VGT or NGT glosses

The goal of our system is to transform Dutch sentences into a sequence of glosses that represent a possible translation in NGT or VGT of the input. To this end, the NLP pipeline (presented in D3.5) is used to obtain the short and detailed part-of-speech tag (PoS-tag), the lemma, the dependency label and the head in the dependency parse. In the first step, we process each word separately. In the second step, the sentence is split up into clauses and each clause is largely processed separately. The system was first developed for VGT synthetic glosses; we are currently transforming it into a system that generates for NGT synthetic glosses. Table 1 contains 2 examples for VGT.

Table 1. Two examples of text (A) to synthetic VGT gloss (F) translations. The following processes are applied: (B) lemmatization, (C) rules activated based on PoS-tag and (D) head and (E) label in dependency parse.

(A) Je	hebt	het	mij	zelf	gegeven	. (You gave it to me yourself.)
(B) JE	HEBBEN	HET	MIJ	ZELF	GEVEN	.
(C) WG-2	*delete*	WG-3	WG-1	ZELF	GEVEN	*delete*
(E) WG-2		*delete*	*delete*	ZELF	2-GEVEN-1	
(F) ZELF 2-GEVEN-1 WG-2						
(A) Hij	weet	niet	wat	erin	zit	. (He does not know what is in it.)
(B) HIJ	WETEN	NIET	WAT	ERIN	ZITTEN	.
(C) WG-3	WETEN	NIET	WAT	IN	*delete*	*delete*
(D) WG-3	WETEN-NIET		WAT	IN		
(F) WETEN-NIET WG-3 WAT IN						

2.2.1 Word Processing

Since glosses are often the unconjugated form of the word with corresponding meaning and are capitalised by convention, we initialise the gloss of each word with its capitalised lemma. Then, we change this representation where needed.

ID-glosses

The NGT Signbank, a lexical database of NGT (Crasborn et al., 2016), and the dictionary of VGT (Van Herreweghe et al., 2004; Vlaams GebarentaalCentrum, 2019) are used to detect the correct ID-glosses (Johnston & Schembri, 2010). An ID-gloss is an agreed-upon label for one specific sign. On the one hand, there might be multiple signs with approximately the same meaning, often regional variants. To distinguish them, a hyphen and letter are added in the lexical resources. Since we do not want to choose one specific variant, we remove this letter (e.g., 'AF-A' and 'AF-B' are changed to one entry 'AF'). On the other hand, the same sign can have multiple possible translations. We look up each lemma in the dictionary and replace the gloss-representation with the ID-gloss when possible.

There are also signs that translate to multiple words. We take this into account by also looking for combinations of words in the dictionary.

Some word-specific changes or gloss conventions were not included in the dictionaries. Numerals, for example, are glossed with their numerical representation in the corpus NGT, but written out in full in the VGT corpus. Another example are separable verbs, i.e. verbs that have a particle that appears separate from the core of the verb in some conjugations. Since the infinitive incorporates the particle, we make sure the particle is added to the verb before we look up the correct ID-gloss in the dictionary.

Deletions

There are certain (types of) words that do not appear in NGT and/or VGT, such as determiners, expletives (e.g., 'it') and linking verbs (e.g., 'to be'). In NGT, negation is usually expressed via non-manual features only (headshakes), which is why we delete negation glosses as well (Klomp, 2021). In VGT, the gloss for negation is randomly deleted in 50% of cases because it is not obligatory (Van Herreweghe & Vermeerbergen, 2006).

Other glosses that are deleted include those for relative pronouns introducing a complement clause, prepositions that are part of a fixed preposition-verb combination and interjections for both languages. This is either because the sign does not exist or because it is almost never used. For some other glosses, it might depend on the context and/or signer whether the sign is included in the sentence. A conditional clause, for example, can be introduced by a sign meaning 'if', but this is not obligatory in NGT (Klomp, 2021) nor in VGT (Vermeerbergen, 2010).

2.2.2 Clause processing

Once we have determined the glosses, we need to order them, which is done per clause. Firstly, a sentence is split up into clauses. In some cases, the ordering of the clauses itself needs to be adapted. For example, when there is a conditional conjunction, the condition has to appear before the main clause in NGT (Baker et al., 2016) and VGT (Vermeerbergen, 2010). Another example is conjunctions with a temporal relation in VGT: when one action occurs before the other one, VGT signers prefer to use the real-life order.

At the moment, there is no evidence that passive constructions as we know them in Dutch (and English) occur in either sign language (Baker et al., 2016; Klomp, 2021), although there exists little research. For this reason, we transform passive constructions into the active equivalent.

There is also more research needed into the word order in both sign languages. Current studies show that there are multiple possible orderings, depending on the type of the verb or its arguments (Baker et al., 2016; Klomp, 2021; Vermeerbergen, 2010). A distinction can be made between locative and non-locative sentences and based on whether the arguments are reversible (e.g., 'The boy pushes the girl' is reversible). We were not able to create a system that can detect these characteristics and apply a randomised system including all the options for each type at this moment. Instead, we used the subject-object-verb (SOV) word order as the default word order for NGT (Klomp, 2021) and SVO as the default word order for VGT (Vermeerbergen, 2010). Only when the subject is a pointing sign (except the first person singular), the order changes to OVS for VGT. This procedure constitutes a first step, and it will be finetuned in a later phase of the project.

Both in NGT and VGT, the temporal phrase appears typically at the start of the sentence (Baker et al., 2016; Klomp, 2021; Vermeerbergen, 2010). We combine the named-entity recogniser and some hand-made rules to detect this phrase and put it at the front of the clause. In VGT, the question phrase appears at the front, at the end or both at the front and the end of the sentence (Van Herreweghe & Vermeerbergen, 2006), while the ordering in NGT does not change (i.e. if the question phrase has the function of object in the sentence, it appears between the subject and the verb) (Klomp, 2021). The ordering rules take this into account. All of the code for our synthetic gloss-system can be found on GitLab[2].

2.2.3 Spanish Sign Language

For Spanish Sign Language (LSE), we are also investigating data augmentation methods to boost the performance of neural machine translation models to transform Spanish text into LSE glosses. LSE is an *extremely* low resource language for language processing applications, and has very few resources even by the standards of SLs in Europe. In Chiruzzo and colleagues (2022), we leverage an available parallel corpus within the restricted domain of renewing government documents (ID/DL)[3] which contains 466 utterances in Spanish and LSE respectively.

One experiment augments the monolingual Spanish corpus AnCora (Taulé et al., 2008) by transforming the original text with rules crafted based on the rules of LSE (e.g. Rodríguez González, 1992) and what we observe in the ID/DL corpus. Table 2 describes these ten rules. We use this large body of synthetic data to pretrain an OpenNMT neural machine translation model, before fine-tuning

[2] SignON / WP4 / Rule_based_gloss_generation / Rule_based_dutch_to_gloss · GitLab

[3] National Identity Card and Driving Licence (ID/DL) corpus (San-Segundo et al., 2008)

with the real-world parallel data from the ID/DL corpus. Compared to a baseline of the test partition of an OpenNMT model trained on the ID/DL corpus parallel utterances, we observe a best word-level BLEU score of 58.98[4] - an improvement of 28.11 over the baseline.

Inclusion of explicit morphological markers	Example
1) Add the "PLURAL" token before any plural word.	perros → PLURAL PERRO
2) Add the "FUTURO" token before any verb in future tense.	comerá → FUTURO COMER
3) Add the "TÚ" token before any verb in second person.	vienes → TÚ VENIR
4) Change a possessive determinant to "PROPIO" + the pronoun.	mi madre → PROPIO YO MADRE
Removal of words not used in LSE	**Example**
5) Remove determinants (except the possessive, which are changed by rule 4).	el perro → PERRO
6) Remove prepositions "de" and "en".	de tarde →TARDE
Particular lexical transformations	**Example**
7) Copula words are changed to the token "SE-LLAMA".	esto es importante → ESTO SE-LLAMA IMPORTANTE
8) Sequences whose lemmas correspond to the sequence "TENER QUE", are changed to "NECESITAR".	tiene que llevar → NECESITAR LLEVAR
9) Instances of "denei" are changed to "DNI".	llevar denei → LLEVAR DNI
10) All other words are represented as their uppercase lemmas.	perros → PERRO

Table 2: Rules used to create synthetic LSE data from Spanish

We must, however, take care when interpreting these results. Firstly, the domain used for translation and testing is restricted and moreover the vocabulary size (unique tokens) in the ID/DL corpus is only 290 glosses. Next, we found that a large number of synthetic glosses generated in the rule-based process had no grounding in real-world signs. In other words, they do not exist within the LSE-Sign (Gutierrez-Sigut, 2016) database. This poses a problem in the real-world translation paradigm, and an overarching aim of SignON - end-to-end translation between a spoken language and outputting sign language. One research avenue we are following is to design a heuristic which would detect an out-of-vocabulary synthetic gloss before transforming it into a suitable real-world gloss, availing of techniques in WSD described later in this document.

2.3 Conclusion

A goal of the SignON project is to investigate two pathways, a symbolic approach and an approach using embeddings to achieve one optimal translation system. We investigated possible symbolic systems that can improve the InterL-E by automatically generating extra, synthetic data.

For the symbolic system, data augmentation is preferred over back-translation for sign2text translation, since sign languages are *extremely* low-resource. We set up a rule-based data

[4] Considered as "Very high quality, adequate and fluent translations" by the Google Cloud AutML documentation

augmentation system for VGT and are working on systems for NGT and LSE. In the future, we should investigate whether we can design a system for Irish Sign Language as well as further extend the systems for NGT and LSE. We will also investigate the value of the resulting synthetic glosses in sign2gloss2text models.

3. WordNets and SignNets

A symbolic representation that has been developed and investigated for many languages is WordNet. We will first discuss what WordNets are and which WordNets can be used in the SignON project (3.1) and then introduce the idea of SignNets (Schuurman et al., 2022), an equivalent to WordNets for sign languages that we are developing, starting with VGT (3.2). Lastly we investigate how we can use WordNets and SignNets in our translation system by means of a word sense disambiguation (WSD) tool (3.3).

3.1 WordNets

A WordNet is a lexical database that contains senses (nouns, verbs, adjectives and adverbs). Synonyms are grouped together in synsets, which also give a definition and example sentences. The synsets are also linked based on different types of relations, such as hyponymy or meronymy. Homonyms such as 'bank' have multiple meanings that will be part of different synsets, one for each possible meaning. Researchers have created WordNets for many different languages and linked them in a multilingual WordNet, which made it possible to automatically generate the translation of a certain synset. Table 3 gives an overview of the available languages in the Open Multilingual WordNet (Bond and Foster, 2013) relevant for the SignON project.

Table 3. Number of synsets, words and senses per WordNet relevant in the SignON project.

Language	Synsets	Words	Senses
Spanish	47,737	47,762	74,848
English	117,659	151,688	213,480
Irish	5,043	5,179	6,267
Dutch	33,925	46,552	69,088

A more recent development is BabelNet. BabelNet is a multilingual dictionary that contains information from the Open Multilingual WordNet as well as many other sources, including named entities. BabelNet extends the synsets from WordNet to contain multilingual lexicalisations (Navigli et al., 2021).

3.2 SignNets

SignNets are lexical databases for sign languages equivalent to WordNets. Since sign languages are not composed of words but of signs, this name is more appropriate. Building a WordNet for sign languages, i.e. a SignNet, entails more than adding videos of signs to an existing WordNet as sign languages are independent natural languages.

None of the sign languages that are part of the SignON project have a SignNet. As a matter of fact, not a single sign language has a fully developed SignNet to our knowledge. There has been some work on SignNets for ASL (Lualdi et al. 2020, 2021; Wright 2021). ASLNet is for the time being concentrating on 'easy' nouns. And very recently, building WordNets for Greek sign language (GSL) and German sign language (DGS) has been started (Bigeard et al. 2022). Within the SignON project, we are developing a model that works for sign languages in general, while also taking into account other PoS categories (like verb, preposition, adjective) as far as these are contained in the dictionary of that sign language. But our model should also be able to deal both with single signs translating into multiple words, and single words translating into multiple signs. And finally also with idiomatic expressions. The SignNets will be in a new RDF-format, constructed in collaboration with OntoLex-Lemon.

Note that when building a SignNet, the starting point is the gloss. This one gloss may represent a series of signs all expressing the same concept. *In se* the gloss constitutes the SL synset. Often, but not always, such signs originate from different parts of the country.

In the SL dictionaries we are aware of, the gloss is accompanied by a few 'guiding' translations in the spoken language. And these we use to find the proper links with the corresponding WordNet and Princeton WordNet (or Open English WordNet, which is more up-to-date). These guiding translations are contained in the SignNets, as are the semantic categories to which the signs belong (Nature, Law, Technology, etc.), the gender of the Deaf people using a specific sign (especially relevant for ISL), and so on. The links will be included using interlingual indices from the Collaborative Interlingual Index (CILI).[5] Note that this way the SLs are also linked to one another. We often discover that the same gloss comes with translations a) as both nouns and verbs, but even when that is not the case, b) with translations belonging to different synsets in spoken language. This needs to be 'solved' in our SignNets.

Some examples:

1) AFBREKEN-A, AFBREKEN-B, AFBREKEN-C, AFBREKEN-H, AFBREKEN-I, AFBREKEN-J[6] (~demolish)
    a) possible translations (in Dutch): afbreken (demolish), afbraak (demolition), slopen (knock down)

    'afbraak' is a noun, 'afbreken' and 'slopen' are verbs

2) HANGEN-B (~hang up)
    a) possible translations (in Dutch): hangen (hang), aanhangen (hitch on), ophangen (hang (up)), aanhaken (hook up, couple)
    b) InterLingual Indexes: i29124, i30105 (PWN 3.1)

In SignNet we split such combined noun/verb signs into two entries, one for each part of speech. One of the reasons for doing so is that, when not just manual movements are taken into account, there may be some differences: (non) iteration of a gesture, mouthing, … When merging or splitting synsets is necessary, we will do so adhering to international standards.[7]

[5] https://github.com/globalwordnet/cili
[6] AFBREKEN-D/E/F/G are contained in the VGT SignBank, but not (yet) in the dictionary
[7] For example using Global WordNet Grid

Our SignNet will also explicitly mention homophone signs, cf homonyms in WordNet. For example, in VGT one of the signs used to express GEEL (yellow) is identical to one of those used to express DONDERDAG (Thursday). Also hyponym/hypernym (more general vs. more specific words) and holonym/meronym (part-whole relations) will be expressed (as far as the relevant glosses are contained in the dictionary).

So, whereas ASLNet focuses on nouns, we intend to build a model covering all types of signs covered by dictionaries, for the time being focusing on the vocabulary typically used in the SignON use cases (such as 'hospitality').

Of course, the SignNets will contain the videos linked to the signs represented by the gloss (not the gloss itself). Whenever a notation is available in Sign_A, SignWriting, HamNoSys or SiGML, these will also be added in SignNet. In WordNet often examples are included, clarifying the meaning of a word. In SignNet we intend to do the same, both in sign language and in spoken language, both in the national languages and in signed English (ASL) and written English. Until the signed examples are available, adding a relevant photograph (Imagenet, the Noun project, ...) may be a good idea (to be discussed with the community), cf dictionary of Nederlands Gebarencentrum (www.gebarencentrum.nl).

3.3 Word sense disambiguation

Word sense disambiguation (WSD) is the task of identifying the correct meaning of a word in a certain context. Homonyms are a clear example: depending on the context, the same word form may mean something entirely different. The task of WSD is often defined as finding the correct synset in a WordNet or, more recently, in BabelNet for each verb, noun, adjective and adverb. Table 4 shows two examples of WSD, generated with the online version of the AMuSE-WSD tool[8] (Orlando et al., 2021).

AMuSE-WSD is, as stated on the website, an All-in-one Multilingual System for Easy Word Sense Disambiguation. The tool can be used via a website or can be downloaded as a Docker image. This Docker image includes several pre-processing units as well as the main module including the trained model for performing WSD. We extracted the main module and linked it to our NLP Pipeline (cfr. D3.5) in order to integrate it into our system.

[8] AMuSE - WSD | All-in-one Multilingual System for Easy Word Sense Disambiguation (uniroma1.it)

The WordNet or BabelNet synsets resulting from WSD can be (part of) the InterL-S, because nouns, verbs, adjectives and adverbs in different languages referring to the same concept are linked. In the future, we will investigate how to use these synsets for better sign language generation. In order to achieve this, SignNets will play a vital role.

Table 4. Two example sentences with BabelNet senses generated by the AMuSE-WSD tool.

I **light**	a **candle**	with a **match.**
bn:00088230v Definition: Begin to smoke	bn:00015211n Definition: Stick of wax with a wick in the middle	bn:00036522n Definition: Lighter consisting of a thin piece of wood or cardboard tipped with combustible chemical; ignites with friction
They **deserved**	to **win**	the **match**.
bn:00086670v Definition: Be worthy or deserving	bn:00095777v Definition: Be the winner in a contest or competition; be victorious	bn:00053772n Definition: A formal contest in which two or more persons or teams compete

3.4 Conclusion

(Part of) the Open Multilingual WordNet as well as BabelNet are candidates for becoming our InterL-S. For sign languages, there aren't any equivalent lexical resources available at the moment. We are working on a sign language lexical resource (a SignNet) for VGT. It is possible to automatically link incoming text to synsets in a WordNet or BabelNet with a WSD tool. We investigated the AMuSE-WSD tool, which covers English, Dutch, Spanish and Irish, and integrated it with our NLP pipeline. In the future, we will proceed to work on the VGT SignNet and investigate how the synsets resulting from the WSD can improve the translation results, for example for sign language generation.

4. Abstract meaning representation

4.1 Introduction

In this section, we present the experiments we carried out with regard to the automatic prediction of semantic representations using the Abstract Meaning Representation (AMR) formalism within the context of a sign language. The AMR formalism was introduced by Banarescu et al. (2013); an AMR structure consists of a directed, acyclic graph that represents the meaning of a sentence. An example of an AMR graph for the sentence 'The boy slapped me on the face' is given below:

```
(z0 / slap-01
   :ARG0 (z1 / boy)
   :ARG1 (z2 / face
          :part-of (z3 / i)))
```

An AMR consists of concepts (e.g. slap-01) that are linked to variable names (e.g. z0). Concepts can be related to each other, which is indicated by a colon (:) followed by the relationship between the concepts (e.g. ARG0, indicating the agent).

As a formal semantic representation, AMR abstracts away from the specific syntactic structure, in the sense that sentences that have the same basic meaning will also have the same AMR. Moreover, AMR can be used as an interlingual representation; even though it is heavily biased towards English, sentences in different languages but with the same basic semantic content can equally be assigned the same AMR.

Within the SignON project, formal semantic representations are specifically used in order to drive the sign synthesis avatar by means of the Sign_A framework (D5.4). While Sign_A in its current form relies on the use of the logical-semantic structures as defined by Role & Reference Grammar (Van Valin, 2005), we have focused within our experiments on AMR with one important objective in mind: The automatic prediction of semantic representations using a machine learning approach. A large corpus of AMR annotations is readily available, which means machine learning systems can be automatically trained to predict the correct semantic representation for a given input sentence.

4.2 Predicting AMR representations

A number of neural approaches exist that are able to perform Text-to-AMR parsing (Konstas et al., 2017; Damonte & Cohen, 2019; Ge et al., 2019). Here, we focus on the approach by Bevilacqua et al. (2021). This architecture leverages the pretrained BART model for Text-to-AMR parsing, which makes it an instance of transfer learning. A sentence is encoded using the model's encoder, and the decoder then auto-regressively predicts a linearized version of the AMR graph. The model is trained (fine-tuned) on the AMR 3.0 (LDC2020T02) corpora release, which consists of 59,255 manually-annotated sentence-AMR pairs.

Quite suitably, the BART encoder-decoder architecture is similar in structure to the mBART architecture that lies at the core of the SignON pipeline, and the model can straightforwardly be extended into a multilingual model (cfr 4.3), such that AMR graphs can be induced for any of the languages that are included within the SignON project.

Below, we present the automatically induced AMR graphs for a number of example sentences, in parallel with the logical-semantic structure of the Sign_A framework. The examples indicate that many of the elements that are necessary for adequate logical-semantic sign language structures can be induced automatically by predicting AMR graphs.

Figure 3(a). Three example sentences, represented as AMR graphs and the accompanying logical-semantic structure of the Sign_A framework.

English Sentence	AMR	Logical Structure
Excuse me	(z0 / excuse-01 :ARG0 (z1 / you) :ARG1 (z2 / i))	[**do’**(1sg, **say’**(1sg, **excuse’**(2sg, 1sg)))]
Thank you	(z0 / thank-01 :ARG1 (z1 / you))	[**do’**(1sg, **say’**(1sg, **thank’**(1sg, 2sg)))]
Do you use SignON?	(z0 / use-01 :ARG0 (z1 / you) :ARG1 (z2 / product :wiki - :name (z3 / name :op1 "Signon")) :polarity (z4 / amr-unknown))	[**be’**(Q-do, **use’**(2sg, SignON)]

Figure 3(b). Two example sentences, represented as AMR graphs and the accompanying logical-semantic structure of the Sign_A framework.

English Sentence	AMR	Logical Structure
Hello, nice to meet you	(z0 / multi-sentence :snt1 (z1 / hello) :snt2 (z2 / nice-01 :ARG1 (z3 / meet-03 :ARG1 (z4 / you))))	[[**do’**(1sg, **say’**(1sg, hello))] & [**be’**(3sg.neuter, **nice’**(**do’**(1sg, **meet’**(1sg, 2sg))))]
Sorry, I do not know sign	(z0 / sorry-01 :ARG1 (z1 / i) :ARG2 (z2 / know-01 :polarity - :ARG0 z1 :ARG1 (z3 / sign)))	[[**do’**(1sg, **say’**(1sg, sorry)] & [<neg<[**know’**(1sg, sign)]>>]]]

Figure 3(c). An example sentence, represented as an AMR graph and the accompanying logical-semantic structure of the Sign_A framework.

English Sentence	AMR	Logical Structure
Excuse me, could you clarify that?	(z0 / and :op1 (z1 / excuse-01 :mode imperative :ARG0 (z2 / you) :ARG1 (z3 / i)) :op2 (z4 / clarify-10 :mode imperative :polite + :ARG0 z2 :ARG1 (z5 / that)))	[**do**'(1sg, **say**'(1sg, **excuse**'(2sg, 1sg)))] & [<MODAL.ABILITY.can<(**clarify**'(2sg, that))>>] **OR** [**do**'(1sg, **say**'(1sg, **excuse**'(2sg, 1sg)))] & **[request**'(1sg, [**do**'(2sg, **clarify**'(2sg, that))])]

4.3 A multilingual set-up

As the core of the architecture for sign language translation consists of a multilingual encoder-decoder architecture, the approach can straightforwardly be extended to a multilingual setup. The source message (in textual form), can be encoded using any of the source languages included in the project, and the decoder then predicts the same AMR for the different languages, effectively using the AMR graph as a semantic interlingua for all of the spoken languages included within the SignON project. Multilingual generalisation is achieved in two ways. First, according to the principle of transfer learning that lies at the heart of pre-trained neural architectures, knowledge that has been acquired by training on the English dataset is automatically transferred to the other spoken languages included within the multilingual model. Secondly, we are strengthening the intrinsic transfer capabilities of the model by automatically translating the English dataset into the other spoken languages included within the project, using designated neural architectures for (spoken language) neural machine translation (Tiedemann & Thottingal, 2020); for each of the (English) sentence/AMR pairs, the sentence is translated into Spanish, Dutch, and Irish, while the AMR remains the same. The mBART encoder-decoder is then further fine-tuned using the augmented AMR dataset in order to predict the correct AMR structure for each of the source sentences in the different languages.

4.4 Conclusion & future work

By extending the multilingual mBART encoder-decoder model that lies at the core of the SignON translation architecture, we are able to automatically induce symbolic interlingual representations from embedded language representations. The prediction of these interlingual AMR structures constitutes a first step in the induction of appropriate representations for sign language synthesis. In a second phase, we will focus on the induction of appropriate surface representations, that take into account the syntactic and lexical characteristics of the specific sign languages (as detailed in deliverable D5.4). This task will be carried out in close collaboration with work package 5.

5. Summary & Future Work

In this deliverable, we have presented the first steps that have been taken in the creation of a symbolic language-independent intermediate representation (InterL-S). First, we presented a rule-based system for data augmentation with synthetic glosses, which will be used to improve the training process of our sign language translation architecture. Secondly, we presented the implementation of a WSD scheme to provide a linkage to WordNet, as well as the creation of appropriate SignNets, which function as a pivotal element within the symbolic representation. Finally, we presented an approach for the induction of language-independent meaning representations in the form of abstract meaning representation (AMR), which will be used to interface with the representation of sign language representations that will drive the sign language synthesis component.

The elements summarised above constitute the first version of the language-independent symbolic representation (InterL-S). In the second part of the project, our aim is to extend this representation, and improve upon it in several ways. We aim to extend our data augmentation approach to the different languages of the project, and exploit the results in order to improve the training procedure of our recognition and translation model. Secondly, we aim to improve the coverage of the SignNet resources for the different languages within the project. And thirdly, we aim to integrate the AMR structures within the sign synthesis pipeline, making sure the semantic representations are able to drive the avatar. To this end, we will investigate the development of sign language-specific representations that take into account the linguistic characteristics of the specific sign languages used within the project—in close collaboration with the research carried out in WP5.

6. References

Baker, Anne, Beppie van den Bogaerde, Roland Pfau, and Trude Schermer, eds. *The linguistics of sign languages: An introduction*. John Benjamins Publishing Company, 2016.

Laura Banarescu, Claire Bonial, Shu Cai, Madalina Georgescu, Kira Griffitt, Ulf Hermjakob, Kevin Knight, Philipp Koehn, Martha Palmer, and Nathan Schneider. 2013. Abstract Meaning

Representation for Sembanking. In Proceedings of the 7th Linguistic Annotation Workshop and Interoperability with Discourse, pages 178–186, Sofia, Bulgaria. Association for Computational Linguistics.

Bevilacqua, M., Blloshmi, R., & Navigli, R. (2021). One SPRING to Rule Them Both: Symmetric AMR Semantic Parsing and Generation without a Complex Pipeline. Proceedings of the AAAI Conference on Artificial Intelligence, 35(14), 12564-12573.

Sam Bigeard, Marc Schulder, Maria Kopf, Thomas Hanke, Kyriaki Vasilaki, Anna Vacalopoulou, Theodore Goulas, Athanasia-Lida Dimou, Stavroula-Evita Fotinea and Eleni Efthimiou (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 (June 25th, 2022), LREC 2022, Marseille, France.

Bond, F and R. Foster, 2013, Linking and Extending an Open Multilingual WordNet. Proceedings of the 51st Annual Meeting of the Association for Computational Linguistics, pages 1352–1362, Sofia, Bulgaria, August 4-9 2013.

Chiruzzo, L., McGill, E., Egea Gómez, S. and Saggion, H. (2022) *Translating Spanish into Spanish Sign Language: Combining Rules and Data-driven Approaches.* [Manuscript submitted for publication]. Large Scale Text Understanding Lab, Universitat Pompeu Fabra.

Crasborn, Onno, R. Bank, I. Zwitserlood, E. Van der Kooij, E. Ormel, J. Ros, A. Schüller et al. "Ngt signbank." *Nijmegen: Radboud University, Centre for Language Studies* (2016).

Damonte, Marco, and Shay B. Cohen. "Structural Neural Encoders for AMR-to-text Generation." In *Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (Long and Short Papers)*, pp. 3649-3658. 2019.

Fellbaum, Christiane. (1998, ed.) WordNet: An Electronic Lexical Database. Cambridge, MA: MIT Press.

Ge, DongLai, Junhui Li, Muhua Zhu, and Shoushan Li. "Modeling Source Syntax and Semantics for Neural AMR Parsing." In *IJCAI*, pp. 4975-4981. 2019.

Gutierrez, S., Costello, B., Baus, C. and Carreiras, M. (2016) LSE-Sign: A lexical database for Spanish Sign Language. *Behaviour Research Methods, 48*, 123-137.

Johnston, Trevor, and Adam Schembri. "Variation, lexicalization and grammaticalization in signed languages." *Langage et société* 131, no. 1 (2010): 19-35.

Konstas, Ioannis, Srinivasan Iyer, Mark Yatskar, Yejin Choi, and Luke Zettlemoyer. "Neural AMR: Sequence-to-Sequence Models for Parsing and Generation." In *Proceedings of the 55th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pp. 146-157. 2017.

Klomp, Ulrika. *A descriptive grammar of Sign Language of the Netherlands*. LOT, 2021.

Lualdi, Colin, 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, pages 63–72, University of South Africa (UNISA). Global WordNet Association.

Lualdi, Colin, Jack Hudson, Christiane Fellbaum, and Noah Buchholz. 2019. Building ASLNet, a WordNet for American Sign Language. In Proceedings of the 10th Global WordNet Conference, pages 315–322, Wroclaw, Poland. Global WordNet Association

Navigli, R., M. Bevilacqua, S. Conia, D. Montagnini and F. Cecconi. Ten Years of BabelNet: A Survey. Proc. of IJCAI 2021, pp. 4559-4567.

Orlando, Riccardo, Simone Conia, Fabrizio Brignone, Francesco Cecconi, and Roberto Navigli. "AMuSE-WSD: An all-in-one multilingual system for easy Word Sense Disambiguation." In *Proceedings of the 2021 Conference on Empirical Methods in Natural Language Processing: System Demonstrations*, pp. 298-307. 2021.

Peng, Xiaochang, Chuan Wang, Daniel Gildea, and Nianwen Xue. 2017. Addressing the Data Sparsity Issue in Neural AMR Parsing. In Proceedings of the 15th Conference of the European Chapter of the Association for Computational Linguistics: Volume 1, Long Papers, pages 366–375, Valencia, Spain. Association for Computational Linguistics.

Rodríguez-González, M. A. (1992) *Lenguaje de signos.* Barcelona, Confederación Nacional de Sordos de España and Fundación ONCE.

San-Segundo, R., Barra, R., Córdoba, R., D'Haro, L. F., Fernández, F., Ferreiros, J., Lucas, J. M., Macías-Guarasa, J., Montero, J. M., and Pardo, J. M. (2008) Speech to sign language translation system for Spanish. *Speech Communication, 50*(11), 1009-1020

Schuurman, Ineke, Vincent Vandeghinste, Caro Brosens, Margot Janssens, and Thierry Declerck (2022). WordNets? SignNets! (poster). Computational Linguistics at the Netherlands, June 17th 2022, Tilburg, The Netherlands.

Taulé, M., Martí, M. A., and Recasens, M. (2008) AnCora: Multilevel Annotated Corpora for Catalan and Spanish. In *Proceedings of the Sixth International Conference on Language Resources and Evaluation (LREC'08)*. Marrakech, Morocco.

Tiedemann, Jörg, and Santhosh Thottingal. "OPUS-MT--Building open translation services for the World." In *Proceedings of the 22nd Annual Conference of the European Association for Machine Translation*. European Association for Machine Translation, 2020.

Van Herreweghe, Mieke. "Negatie in de Vlaamse Gebarentaal." *Handelingen-Koninklijke Zuid-Nederlandse maatschappij voor taal-en letterkunde en geschiedenis 54* (2000).

Van Herreweghe, M., Vermeerbergen, M., Van Mulders, K., De Weerdt, K. (2004). *Woordenboek Vlaamse Gebarentaal : Nederlands-Vlaamse Gebarentaal/Vlaamse Gebarentaal-Nederlands (online)*. Gent: Universiteit Gent.

Van Herreweghe, M., Vermeerbergen, M. (2006). Interrogatives and negatives in Flemish Sign Language. In: U. Zeshan (Eds.), *Interrogative and negative constructions in sign languages*, (225-256). (Sign Language Typology Series, 1). Nijmegen: Ishara Press. ISBN: 90-8656-001-6.

Van Valin, R. (2005). Exploring the Syntax-Semantics Interface. Cambridge: Cambridge University Press.

Vermeerbergen, M. (2010). 'Woordvolgorde' in de Vlaamse Gebarentaal. *Onderzoeksrapport VGTC*, Affligem: Vlaams GebarentaalCentrum.

Vlaams GebarentaalCentrum. 2019. Woordenboek Vlaamse Gebarentaal. Accessed 28 June 2022. https://woordenboek.vlaamsegebarentaal.be/.

Wright, Elaine (2021), Signs in the Mind: Constructing ASLNet. Princeton University Senior Theses

Xia, Mengzhou, Xiang Kong, Antonios Anastasopoulos, and Graham Neubig. 2019. Generalized data augmentation for low-resource translation. In Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics, pages 5786–5796, Florence, Italy. Association for Computational Linguistics.

