Sign Language Machine Translation and the Sign Language Lexicon: A Linguistically Informed Approach

Irene Murtagh irene.murtagh@adaptcentre.ie
Department of Informatics, TU Dublin, Ireland
Víctor Ubieto Nogales victoremilio.ubieto@upf.edu
Josep Blat josep.blat@upf.edu
Departament de Tecnologies de la Informació i les Comunicacions Human-Computer Interaction, graphics and Educational Technologies.

Abstract

Natural language processing and the machine translation of spoken language (speech/text) has benefitted from significant scientific research and development in recent times, rapidly advancing the field. On the other hand, computational processing and modelling of signed language has unfortunately not garnered nearly as much interest, with sign languages generally being excluded from modern language technologies. Many deaf and hard-of-hearing individuals use sign language on a daily basis as their first language. For the estimated 72 million deaf people in the world, the exclusion of sign languages from modern natural language processing and machine translation technology, aggravates further the communication barrier that already exists for deaf and hard-of-hearing individuals. This research leverages a linguistically informed approach to the processing and modelling of signed language. We outline current challenges for sign language machine translation from both a linguistic and a technical prespective. We provide an account of our work in progress in the development of sign language lexicon entries and sign language lexeme repository entries for SLMT. We leverage Role and Reference Grammar together with the Sign_A computational framework within this development. We provide an XML description for Sign_A, which is utilised to document SL lexicon entries together with SL lexeme repository entries. This XML description is also leveraged in the development of an extension to Bahavioural Markup Language, which will be used within this development to link the divide between the sign language lexicon and the avatar animation interface.

1. Introduction

Sign Languages (SLs) are visual gestural languages articulated within a three-dimensional signing space and have no written form (Murtagh, 2019a). Many deaf and hard-of-hearing individuals use SL on a daily basis as their first language. For the estimated 72 million deaf people in the world, the exclusion of sign languages from modern natural language processing and machine translation technology, further
aggravates the communication barrier that already exists for deaf and hard-of-hearing individuals (Allen, 2013). We outline our research work in progress in the development of a SL lexicon architecture, including SL lexicon entries and SL lexeme repository entries for a sign language machine translation (SLMT) system. We provide some background information on the Role and Reference Grammar (RRG) and the Sign_A framework, which are leveraged within this development. We discuss the XML specification for the Sign_A computational framework,

which we leverage to define SL lexicon entires and SL lexeme repository entries. We also discuss the extension to the specification for Behavioural Markup Language in the development of a planner for SL translation.

2. SignON Project

We draw here on work we are engaged in for the Horizon 2020 funded SignON project, which seeks to create a service that translates between sign and verbal languages, facilitating new resource generation over time, which in turn will further improve the service[1] (Shterionov et al., 202; Saggion et al., 2021). SignON – Sign Language Translation Mobile Application and Open Communications Framework – seeks to reduce the communication gap that exists between deaf sign language users, hard-of-hearing and hearing people. SignON targets Irish, British, Dutch, Flemish and Spanish Sign Language, together with English, Irish, Dutch and Spanish spoken language. The overarching project goal is to increase inclusiveness through accessible translation services powered by state-of-the-art artificial intelligence (AI). The co-creation process lies at the core of this project, with tight collaboration from European deaf and hard-of-hearing communities. This collaboration informs the co-design and co-development of the SignON service and application, while also enabling continuous assessment of quality.

3. Sign Language

Sign languages are linguistically complete, very rich and complex languages (Murtagh 2019). Communication across sign languages encompasses manual features (MFs) and non-manual features (NMFs). MFs include hand shapes, hand locations, hand movements and orientation of the palm of the hands. NMFs include the use of eye gaze, facial expression, mouthing, head and upper body movements. The visual gestural realisation of a word in SL involves the simultaneous and parallel expression of a varied number of MFs and NMFs, each with their own duration, orientation and relative configuration and movement.
The SignOn project targets Irish Sign Language (ISL) Flemish Sign Language (VGT), British Sign Language (BSL), Spanish Sign Language (LSE) and Dutch Sign Language (NGT). We take Irish Sign Language (ISL) as our sign language of focus within this research paper, as this is our initial language under linguistic investigation within the SignON project.

3.1. Sign language machine translation challenges

Challenges for sign language machine translation (SLMT) exist within two separate realms. On one hand, we must consider the linguistic challenges and on the other hand, the technical challenges. While spoken language communication occurs within auditory-oral modality, sign language communication occurs within visual gestural language that is articulated within three-dimensional (3D) space (Leeson and Saeed, 2012). The modality difference for human-to-human communication together with the fact that there is no written or aural form for sign language introduces many interesting challenges for SLMT. With regard to challenges facing SLMT, (Murtagh et al., 2021), outline linguistic and technical challenges and report on the critical importance of: close engagement and co-construction of MT agendas with Deaf communities; the inclusion of deaf experts on MT project teams; the need for interdisciplinary approaches to MT work on sign languages; the need for robust data sets; and the need to manage expectations around what can be achieved to a high level as we progress with work in this domain.

[1] https://signon-project.eu

Linguistic challenges (Murtagh et al., 2021) report on the linguistic phenomena that must be addressed, but which have not been documented sufficiently to date as a result of under-resourcing. Irish Sign Language (ISL) was used as the SL of focus, but the point regarding the under-documentation of ISL 'holds equally for most sign languages of the world'. For ISL, these under-described areas include: description of the non-discrete lexicalised elements in Irish Sign Language including simultaneous constructions, body partitioning, motivations underpinning use of signing space; the absence of an ISL SignBank; the need for more research on the syntax, semantics and pragmatics of ISL; and the need for a broader base of data from which to generate linguistic rules and train MT ISL receptive models.

Technical challenges There are also many technical challenges involved in machine translation (MT) between spoken and signed language and vice-versa. Research shows that when SLs and spoken languages are compared, it is speech plus co-speech gesture rather than speech alone that should be considered as an equivalent to signing (Leeson and Vermeerbergen, 2022).

The reliance of SL on the use of space for linguistic purposes together with the (more) simultaneous organisation of SL compared to the (more) sequential organisation of spoken language are two important linguistic phenomena that pose a challenge for SLMT from a technical perspective (Leeson and Vermeerbergen, 2022). Further challenges are posed by the sign language lexicon. The SL lexicon refers to both an established lexicon and a productive lexicon. The established lexicon accounts for established signs, which are highly conventionalised in both form and meaning, whereas signs encompassed within the productive lexicon are constructed using conventional strategies to fit contextual needs (Leeson and Saeed, 2012). These strategies form the productive lexicon. The productive lexicon is composed of sets of language-specific handshapes that can combine with a wide range of movements, orientations of the palm of the hand, and locations of articulation within in the signing space/gestural space to articulate meaning. We refer to these as manual features (MFs) in SLs.These may also be accompanied by non manual features (NMFs) (e.g mouth gestures, eye-gaze, brow-raises/brow-furrows, ...) to represent clauses or sentences encoding a particular character perspective.This is particularly challenging for verbal language to SLMT and vice versa.

4. The sign language lexicon

We implement our lexicon leveraging RRG (Van Valin and La Polla, 1997; Van Valin, 2005), together with the Sign_A framework (Murtagh, 2019) to create lexicon entries that will sufficiently accommodate SL. RRG views language as a system of communicative social action. RRG defines grammatical structures in relation to both semantic and communicative functions. Syntax is viewed as being relatively motivated by semantic and pragmatic factors. RRG is sufficiently flexible and robust to accommodate SL at a semantic, syntactic and pragmatic level. It allows us to address certain characteristics that have proven problematic for head driven phrase structure grammar (HPSG), which was utilised in the development of a computational lexicon for British Sign Language (BSL) (Sáfár and Glauert, 2012). Many of the rules found in the HPSG literature do not apply to SLs, and therefore, to adequately represent SLs, we leverage the use of RRG and extend its capability using the Sign_A framework, allowing us to develop a lexicon architecture that is sufficiently robust in nature to cater for the linguistic phenomena pertinent to SLs.

4.1. Role and Reference Grammar

Role and Reference Grammar, henceforth termed RRG, is a model of grammar, which incorporates many of the points of view of current functional theories of grammar (Van Valin, 2005). In RRG, the description of a sentence in a particular language is formulated in terms of its logical structure and communicative functions, and the grammatical procedures that are available in the language for the expression of these meanings. Semantic decomposition of predicates and their semantic argument structures are represented as logical structures. The lexicon in RRG takes the position that lexical entries for verbs should contain unique information only, with as much information as possible derived from general lexical rules.

Figure 1 from Van Valin (2005) provides an illustration of the organisation of the RRG architecture including constructional schemata. Van Valin (2005) takes the position that constructions within RRG are utilised to capture language specific idiosyncratic linguistic behavior.

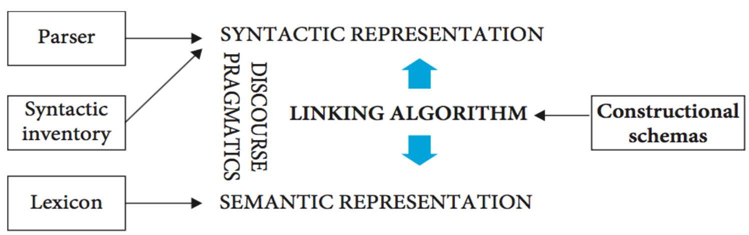

Figure 1. The organisation of the RRG architecture, Van Valin (2005)

4.2. Sign_A computational framework

The Sign_A framework, was developed by Murtagh (2019) with the "A" within this term representing 'Articulatory Structure Level'. As there is no current agreed standard with regard to the documentation of SLs, the Sign_A framework was developed with a view to accommodating the representation of sign languages within the SL lexicon. 'Articulatory Structure Level' extends the theory of the generative lexicon (GL) (Pustejovsky 1991), introducing a fifth level of lexical representation, which accounts for the essential (computational) phonological parameters of an object as defined by the lexical item.

4.3. SignON sign language lexicon architecture

With regard to our SL lexicon architecture, Figure 2 provides a high level view of the RRG + Sign_A framework architecture. It is important to note that each SL added into the architecture will have a separate lexicon, lexeme repository etc. for each respective SL. We include a lexeme repository, which maintains the NMF and MF lexemes for each SL. We also include a morpheme store, which maintains those grammatical units that demonstrate no conceptual meaning. We propose a morpheme store and a lexeme repository to cater for SL morphemes and SL lexemes respectively. We use the context of an utterance to decipher whether an item should be placed within the morpheme store or within the lexeme repository of the SL lexicon architecture. An item may exist within the morpheme store and also exist within the lexeme repository depending on its context within any given sentence. SL morphemes, which demonstrate grammatical function, but lack any conceptual meaning will be placed within a morpheme store, while SL lexemes or those morphemes that function in grammatical terms, while also exhibiting conceptual meaning will reside within a lexeme repository. The lexicon within this figure maintains the RRG + Sign_A rich logical structures for each SL. The grammar component is responsible for maintaining and assembling the clause, ensuring that word order, agreement features, tense etc. are aligned and assembled correctly.

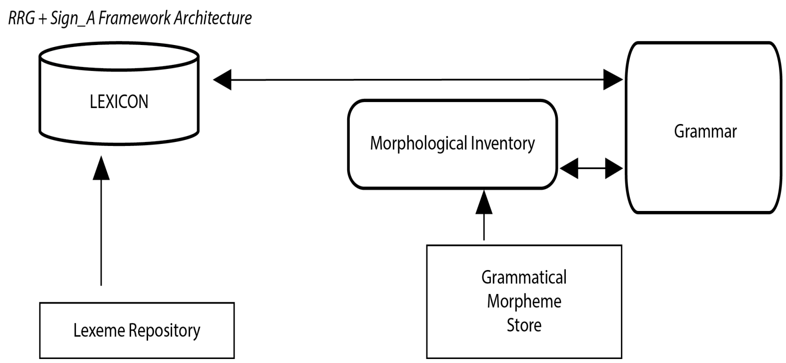

Figure 2. RRG + Sign_A Lexicon Architecture (Murtagh, 2019: 248)

4.4. Sign language lexicon and lexeme repository entries

Categories currently included in the SL lexicon are nouns, classifiers and verbs. As an example, we refer to SL verbs with regard to lexicon and lexeme repository entries. In order to provide some context, we provide a brief discussion of RRG, the theoretical model of grammar that we use in the development of this lexicon architecture. RRG semantic representation is based on a system of lexical representation and semantic roles. RRG employs the system of lexical decomposition proposed by Vendler (1967). Saeed (2016) defines the task of a semanticist as showing "how the inherent semantic distinctions carried by verbs, and verb phrases, map into a system of situation types". Saeed (ibid.: 119) identifies Vendler's influential approach to doing this (Vendler, 1967: 97-121).

Within RRG, verbs are represented in the lexicon according to their Aktionsart classification. Verbs can be divided into four distinct classes: states, activities, achievements and accomplishments. These four classes can be further defined by three features: [±static], [±punctual], and [±telic] (Binns-Dray, 2004). Static indicates if a verb represents something happening. If one can answer the question, "What happened?" or "What is happening?" then the verb is seen to be static. Telic represents whether a verb describes a state of affairs that has a terminal end point. Achievements and accomplishments are telic, or bounded, as in "The clothes are drying on the line", while states and activities are atelic, or unbounded, as in "John is running in the park". Punctual represents whether a telic verb (achievements and accomplishments) has internal duration or not (Binns-Dray, 2004). There are two additional classes; active accomplishments, which describe telic uses of activity verbs (e.g. devour) and also semelfactives (punctual events; Smith, 2009).

SL verbs will be represented in the SL lexicon according to their Aktionsart classification (Vendler, 1967). A single verb can have more than one Aktionsart interpretation. For example the verb 'march' would be listed in the lexicon as an activity verb, and lexical rules would derive the other uses from the basic activity use. The lexical representation of a verb or other predicate is termed its LOGICAL STRUCTURE [LS]. State predicates are represented simply as predicate´, while all activity predicates contain do´. Accomplishments, which are durative, are distinguished from achievements, which are punctual. Accomplishment LSs contain BECOME, while achievement LSs contain INGR, which is short for ‘ingressive’. Semelfactives contain SEML. In addition, causation is treated as an independent parameter that crosscuts the six Aktionsart classes. It is represented by CAUSE in LSs. The lexical representations for each type of spoken language verb shown above are provided in Table 1.

Aktionsart Class	**Logical Structure**
State	**predicate'** (x) or (x, y)
Activity	**do'** (x, [**predicate'** (x) or (x, y)]}
Achievement	INGR **predicate'** (x) or (x, y), or INGR do' (x, [**predicate'** (x) or (x, y)]}
Accomplishment	BECOME **predicate'** (x) or (x, y), or BECOME **do'** (x, [**predicate'** (x) or (x, y)]}
Active accomplishment	**do'** (x, [**predicate**$_1$,' (x, (y))]) & BECOME **predicate**$_2$; (z, x) or (y)
Causative	α CAUSE β where α, β are representations of any type

Table 1. Lexical representation for Aktionsart classes, Van Valin and La Polla (1997: 109)

Table 2 provides a sample sentence in ISL from Murtagh (2019: 142), where the Aktionsart class or event type is provided, together with the tripartite verb class (Padden, 1988).

Gloss and English Translation	**ISL Verb**	**ISL Verb Class**	**Transitivity**	**Event Type**	**Reference**
REAL LOVE MY JOB ‘I really love my job’	LOVE	plain	transitive	State	SOI Corpus Noeleen (03) Personal Stories (Dublin)

Table 2. ISL sentence with event type and tripartite verb type, Murtagh (2019b)

Murtagh (2019) provides a broad analysis of ISL verbs covering all event types, however, in this case, for purpose of illustration, we will focus on an ISL plain verb, according to the traditional tripartite verb class. We use the information in Table 2 to produce an RRG + Sign_A logical structure (LS) lexicon entry, capable of representing SL within our SL lexicon. An illustration of the sentence in Table 2, taken from the SOI corpus (Leeson et al., 2006), is provided in Example 1 below.

Example 1

REAL LOVE MY JOB
‘I really love my job’
SOI Corpus Noeleen (03) Personal Stories (Dublin)

Plain verbs are typically not marked for person or location (McDonnell 1996: 116). The participant is referring to the fact that ‘she loves her job’, with job being introduced and established earlier in the discourse. The situation type for the ISL plain verb ‘LOVE’ within this sentence is state. Table 3 provides the Sign_A + RRG logical structure, which will be used as the lexicon entry for the sentence “I really love my job”. This table also provides the lexeme repository XML description, based on the Sign_A computational framework. Section 5 provides an overview of the Sign_A framework XML description.

Gloss	REAL LOVE MY JOB
English Translation	‘I really love my job’
RRG+Sign_A Logical Structure	**LOVE´** <TEMPORAL><MF><NMF> (1sg, JOB)
ISL Lexicon XML SL Verb Entry	
<ISLGlossTranslate="LOVE" IPA=”/lʌv/" LogicalStructure= "LOVE´<TEMPORAL> <LOCATION><MF><NMF> (1sg, JOB);" NumberVerb="sg" P.O.S="PlainVerb" personVerb="3rd" tenseVerb="PRES" love/>	
Lexeme Repository Sign_A XML description for Manual Features <MF> of SL verb LOVE	
<HAND><dh>"right"</dh><ndh>"left"</ndh></HAND>	
<HS><HSMode>unique</HSMode><HSID><value>24</value></HSID>	
<AM><Spatial><SOURCE>“[a]locus”</SOURCE><GOAL>“ [b]locus”</GOAL>EDti></EDti><EDtn></EDtn> <TLti></TLti><TLtn></TLtn> </SPATIAL><AM>	
<PO><p2><p2_i><EDti></EDti><EDtn></EDtn></p2_i><p1_n><EDti></EDti><EDtn></EDtn></p2_n><TLti></TLti><TLtn></TLtn></p1></PO>	
Lexeme Repository Sign_A XML description for Non Manual Features <NMF> of SL verb LOVE	
<MOUTHING><VERB_ONE_TO_ONE><VERBIPA>”/lʌv/”</VERBIPA></VERB_ONE_TO_ONE></MOUTHING	

Table 3. ISL plain verb lexicon and lexeme repository XML description

5. Sign_A framework XML description

The Sign_A framework XML specification was developed with a view to documenting and accommodating SL lexicon entries in computational terms[2]. We report on MF specifications, NMF specifications, and finally TEMPORAL specifications.

5.1. Manual feature specifications

With regard to SL MFs, William Stokoe (1960) originally identified the various parameters, which are relevant for the analysis of SL. He suggested that the articulation of a sign encompassed three different parameters. A designator*,* which was used to refer to the specific combination of hand configuration, abbreviated to *dez*. A tabulation, used to refer to the location of the hands and abbreviated to *tab*, and a *signation* used to refer to the movement of the hands and abbreviated to *sig*. Dez, tab and sig were examples of what he called *cheremes*, the signed equivalent of phonemes (Murtagh. 2019b).

Later research refers to these parameters of SL as *handshape*, *location* and *movement.* (Sutton-Spence & Woll (1999) : Valli & Lucas (1995)). Battison (1978) claimed that a fourth parameter is necessary in order to be able to fully transcribe signs. This fourth parameter is called *orientation*, and denotes the orientation of the hands and fingers during the articulation of the sign. The abbreviation of orientation is *ori.*

The Sign_A MF XML specification includes a specification for <HAND>, handshape <HS>, hand movement <HM>, palm orientation <PO>, arm movement <AM>, forearm <FA> and upperarm <UA>. For illustrative purposes we will include the <HAND> MF here. Example 2 illustrates n XML computational description for the hands, where the ‘dominant hand’ is defined as <dh> and the non dominant hand as <ndh>. This example provides an illustration of initialising the right hand as the dominant hand.

Example 2
```
<MF>
        <HAND>
                <dh>"right"</dh>
                <ndh>"left"</ndh>
        </HAND>
...
</MF>
```

5.2. Non-Manual feature definitions

(Murtagh, 2019b) reports that the existence of NMFs within signed languages has been well documented by researchers, including Liddell (1980), Nolan (1993), Coerts (1990), Bellugi and Klima (1990), Baker and Padden (1978b). NMFs consist of various facial expressions such as eyebrow movement, movement of the eyes, mouth patterns, blowing of the cheeks head tilting and shoulder movement. NMFs areused to convey additional information to the meaning being expressed by manual handshapes. While NMFs are normally accompanied by a signed lexical item, they can be used to communicate meaning independent to manual accompaniment (Leeson and Saeed, 2012).

[2] https://signon-project.eu/wp-content/uploads/2022/01/SignON_D5.4_First-Sign-Language-Specific-Lexicon-and-Structure_v1.0.pdf

Sign_A NMF XML specifications include specifications for describing articulations relating to the head <HEAD>, eyebrow <EB>, Eyelid <EL>, eye gaze <EG>, cheek <CHEEK>, mouth <MOUTH>, tongue <TNG>, nose <NOSE>, Shoulder <SHOULDER>, mouthing <MOUTHING> and mouth gesture <MOUTHGESTURE>. Example 3 provides an XML computational description of <MOUTHING>. We include an International Phonetic Alphabet (IPA) description of the respective nouns and verbs within the lexicon to cater for the one-to-one mapping between the sign and the respective noun or verb being mouthed.

Example 3

```
<MOUTHING>
  <NOUN_ONE_TO_ONE><NOUNIPA> </NOUNIPA></NOUN_ONE_TO_ONE>
  <VERB_ONE_TO_ONE><VERBIPA></VERBIPA></VERB_ONE_TO_ONE>
  <EDti></EDti><EDtn></EDtn>
  <TLti></TLti><TLtn></TLtn>
</MOUTHING>
```

5.3. Temporal feature specifications

Temporal feature specifications refer to timing information associated with both the MFs and NMFs. The *event duration* parameter <ED> is used as an attribute together with each distinct phonological parameter, for both MF and NMF. It functions linguistically at the morphological-phonological interface, defining the duration or time taken for any given MF or NMF phonological parameter to be realised. The visual gestural realisation of an ISL MF and NMF phonological parameter is considered to be an *event* within the Sign_A computational framework. The realisation of each event has a specific duration bound to it. This can be referred to as an *event duration* <EDtn>. The event duration parameter is used to allow us to synchronise the timing information relating to when each distinct MF or NMF phonological parameter, providing information on when an event may execute along a larger timeline parameter. Due to the visual gestural nature of sign language and the fact that parameters for MFs and NMFs may be articulated simultaneously along a timeline to articulate an utterance, the event duration parameter plays an essential role within the Sign_A framework. The eventDuration <EDtn> parameter of each MF and NMF phonological parameter will be executed in relation to the timing information of the entire utterance or the timeline parameter <TL>.

The following example provides an XML description for the event duration timeline parameter, where the initial eventDuration <EDti> element is responsible for storing the event start time in relation to the timeline parameter <TLtn> and end event duration element <EDtn> is responsible for storing the actual duration that a phonological parameter will play out for.

Example 4

```
<EDti></EDti> <!--initial time relative to the timeline -->
<EDtn></EDtn> <!--end time relative to the timeline -->
```

The *timeline* parameter <TL> refers to a linear timeline representing the overall time taken from the moment an ISL utterance begins until the moment an entire utterance or articulation is completed or terminates. An utterance refers in this case to an ISL lexeme, phrase or sentence that communicates something meaningful. The timeline parameter will play a central role within our computational framework as it is responsible for synchronisation and keeping track of the sequence in which each phonological parameter event will be realised.

The example below provides an XML description for the timeline parameter <TLtn>, where the initial timeline <TLti> element is responsible for providing the event duration start time <EDti>. This value is used as input to the initial event duration <EDti> and is used to allow for synchronisation. The end timeline element <TLtn> is responsible for storing the overall duration that an entire utterance will take.

Example 5
<TLti></TLti> <!—initial time relative to the sign language utterance -->
<TLtn></TLtn> <!—end time relative to the sign language utterance -->

6. Linking the divide between the lexicon and animation interface

We extend the specification for Behavioural Markup Language in the development of a planner for SL translation. This planner will be responsible for the translation from the Sign_A XML specification within the lexicon architecture, to a BML-based script for driving a SL embodied conversational agent. We extend the BML specification with a view to accommodating the Sign_A XML definitions. Table 4 below provides an example of the specification for BML, which has been extended to cater for Sign_A XML hand MF <HAND>. We refer to example 2, illustrated previously. Other Sign_A XML definitions which have been extended with regard to the BML specification include handshape <HS>, hand movement <HM>, palm orientation <PO>, arm movement <AM>, upper arm <UA>, head <HEAD>, eyebrow <EB>, eyelid <EL>, eyegaze <EG>, cheek <> CHEEK, mouth <MOUTH>, mouthing <MOUTHING>, tongue <TNG>, shoulder <SHOULDER>, body anchored locations <BA>, signing space locations <SPATIAL>, event duration <ED> and timeline <TL>.
Table 4

Feature	Defined in Sign_A	Defined in BML extension
Hand	XML element inside <MF>. <HAND> <dh>"right"</dh> <ndh>"left"</ndh> </HAND>	DomHand attribute of the BML block. <bml id=”bml1” characterID=”Eva” domHand=”RIGHT” end=”5”> [behavior blocks should go here] </bml>

7. Conclusion

We have outlined of work in progress in the development of sign language lexicon entries and sign language lexeme repository entries for SLMT. We have also outlined an XML description for Sign_A, which is leveraged within the SL lexicon entries together with SL lexeme repository entries of this development. We provide an overview of the SL lexicon architecture used within this development. We also outline current work in progress in the development of a planner for translation of our XML description with a view to synthesizing SL. Future work will focus on further developing the lexicon architecture and indeed the SL lexicon, to take into account linguistic phenomena associated with Flemish Sign Language (VGT), British Sign Language (BSL), Spanish Sign Language (LSE) and Dutch Sign Language (NGT). Future work also includes further development of routines to automatically compute these SL LSs, while also working on further developing the the BML planner and realiser in this cutting edge development.

Acknowledgements

This work is supported by the SignON project, which is funded by the European Union's Horizon 2020 Research and Innovation program, under grant number 10107255.

References

Allen, C. (2013): Equality for Deaf People: How Do We Get There? Paper presented at the Centre for Deaf Studies Occasional Lecture Series, Trinity Long Room Hub, Trinity College Dublin.

Baker C. and Padden, C. (1978). Focusing on the Non-Manual Components of American Sign Language. In: P. Siple, ed., *Understanding Language Through Sign Language Research.* London: Academic Press.

Battison, R. (1978). "Lexical Borrowing in American Sign Language", Silver Spring, MD: Linstok Press.

Bellugi, U. and Klima, E. (1990). Properties of Visual Spatial Languages. In: S. Prillwitz and T. Vollhaber eds., *Sign Language Research and Application, International studies in sign language and the communication of the deaf,* Hamburg: Signum-Press, pp. 115-143.

Binns-Dray, K. R. (2004). *Content Questions in American Sign Language: An RRG Analysis*. Unpublished PhD dissertation. State University of New York at Buffalo.

Coerts, J. (1990). The Analysis of Interrogatives and Negations in Sign Language of the Netherlands. In S. Prillwitz and T. Vollhaber eds., *Current Trends in European Sign Language Research*, *International studies in sign language and the communication of the deaf*, Hamburg: Signum-Press. 9, pp. 265-277.

Leeson, L., Saeed, J., Byrne-Dunne, D., Macduff, A. and Leonard C. (2006). *Developing a Digital Corpus of Irish Sign Language. The 'Signs of Ireland' Corpus Development Project. Information Technology and Telecommunications Conference.* Carlow: Carlow Institute of Technology, Ireland.

Leeson, L. and Saeed J. (2012). *Irish Sign Language* . Edinburgh, UK: Edinburgh University Press.

Leeson, L. and Vermeerbergen M. (2022). Introducing Sign Languages. Simultaneity, Multimodality and other Characteristics. Talk presented at the SignON Internal Seminar 4.

Liddell, S. (1978). Non-manual signals in ASL: A many layered system. In: W. C. Stokoe ed., *Proceedings of the First National Symposium on Sign Language Research and Training*, 1977, Chicago: National Association of the Deaf, pp. 193-228.

Nolan, E. (1993). *Non-manual features in Irish Sign Language*. Unpublished essay. Horizon deaf studies project, Centre for Language and Communication Studies (CLCS), University of Dublin, Trinity College, Dublin, Ireland.

Murtagh, I. (2019a). Motivating the Computational Phonological Parameters of an Irish Sign Language Avatar. In: B. Nolan and E. Diedrichsen (eds.), *Linguistic Perspectives on the Construction of Meaning and Knowledge Representation.* Cambridge Scholars Publishing, pp.323-339.

Murtagh, I. (2019b) *A Linguistically Motivated Computational Framework for Irish Sign Language*. PhD Dissertation. Dublin: Trinity College Dublin.

Murtagh, I., Leeson, L., and Moiselle, R. (2021). Sign Languages and Language Technology: Linguistic and Technical Challenges. In *Proceedings of the the 2021 Annual Irish Association of Applied Linguistics Conference: Applied Linguistics in the 21st Century*, Online, University College Dublin, Ireland.

Pustejovsky, J. (1991a). The Generative Lexicon. *Computational Linguistics*, 17(4), pp. 209–441.

Saeed, J. (2016). *Semantics*. Fourth ed. Malden: Wiley-Blackwell.

Sáfár, E.and Glauert J. (2012). Computer Modelling. In: R. Pfau, M. Steinbach, B.Woll eds., *Sign Language: An International Handbook*, Berlin: Mouton de Gruyter.

Saggion, H., Shterionov, D., Labaka, G., Van de Cruys, T., Vandeghinste V. and Blat J. (2021). SignON: Bridging the gap between Sign and Spoken Languages. In *Proceedings of the XXXVII Spanish Society for Natural Language Processing conference.*

Shterionov D., Vandeghinste V., Saggion H., Blat J., De Coster M., Dambre J., van den Heuvel H., Murtagh I., Leeson L., and Schuurman I. *(2021)* The SignON project: A Sign Language Translation Framework. Meeting of Computational Linguistics in the Netherlands #CLIN31

Van Valin, R. and La Polla, R. (1997). Syntax: Structure, Meaning and Function. Cambridge: Cambridge University Press.

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

Vendler , Z. ([1957] 1967). *Linguistics in Philosophy*. Ithaca: Cornell University Press.

Smith, C. (2009). *Linguistics of American Sign Language, 2nd Edition*. Washington: Gallaudet University Press.

Stokoe, W. (1960). Sign Language Structure: An Outline of the Visual Communication System of the American Deaf. In: Studies in Linguistics Occasional Papers, number 8. Buffalo, New York: University at Buffalo.

Sutton-Spence R and Woll B. (1999). The Linguistics of British Sign Language; an introduction. Cambridge:Cambridge University Press.

Valli, C. and Lucas, C. (1995). Linguistics of American Sign Language, 2nd Edition. I Washington.