FP7 - Information and Communication Technologies

Collaborative Project

Dicta-Sign

Sign Language Recognition, Generation and Modelling with application in Deaf Communication

Grant Agreement Number 231135

Deliverable D4.5: Report on the linguistic structures modelled for the Sign Wiki

Leading Partner: CNRS-LIMSI

V1.2 (2012-02-10)

February 2012

INTRODUCTION

This deliverable describes the progress performed during the last year of the project on two supporting grammars, dedicated to recognition and generation. These two supporting grammars cover different aspects of SL that are correlated to the progress respectively in recognition and generation systems. Because sentence processing is more mature for generation than for recognition, the generation supporting grammar cover properties than can be found at various level, from sentence to lexical unit.

Recognition supporting grammar focuses only on lexical unit, and is aimed at enhancing recognition results in case of contextual signs, which production depends on the sentence context.

Generation supporting grammar focuses on various levels, from lexical unit to discourse structure, and is aimed at enhancing generation during the creation of a content, or after as a post-processing step.

In both cases, the point is to provide with schemes to formalise Sign Language aspects with a multi-SL perspective. Of course, it is irrational to claim full grammar specification for four languages, and indeed the first year review rightly advised the consortium to prioritise and focus on a limited set of elements before any modelling effort was carried out.

Parts of the formalisations presented here have been implemented in the recognition and generation software to be used in the final demonstration, the Sign Wiki (WP8).

The first section of this deliverable describes the progress in recognition supporting grammar, and the second one, in generation supporting grammar.

1. RECOGNITION SUPPORTING GRAMMAR

This section described the recognition supporting grammar (WP4 task4.4) that has been defined in the revised DoW as follows:

"This task will provide unified linguistic models usable for recognition. WP1 and WP2 teams work on extracting HamNoSys parameters. This extraction may contain errors, or be incomplete, causing the sign recognition to fail. A grammar will be designed in order to reduce the number of choices for the HamNoSys values. The additional information coming from the knowledge base of this grammar then will be fed back to the recognition system further to reduce or verify the validity of the choices. Task 4.4 consists in designing such a grammar. It will be implemented and used in Task 7.3 (Sign Wiki), as a help function to analyse input in Dictation mode."

The HamNoSys parameters from the WP1 and WP2 teams work on extracting HamNoSys serve as the input for a grammar supporting recognition, based on prototypes reported in previous periods. The objective of the grammar will be to identify errors in the recognition and to infer missing information where recognition is incomplete.

Given that partner NTUA is working on recognition GSL this work will be designed for limited linguistic structures and for a limited vocabulary starting with GSL. The structures will nevertheless contain signs that vary according to context, and hence an extensible and flexible framework will be created, in contrast to traditional recognition techniques where only fixed signs are included. The additional information coming from processing the input through this grammar will be fed back to the recognition system further to refine and improve the choices.

The grammar work is based on the output of NTUA's phonetics-based subunits framework (see D2.3 for further information). This is an extension of the work in phonetics-based PDTS computational subunits. For more information see the paper by V. Pitsikalis et al 2011 and the chapter by Vogler 2011 in Deliverable D4.3. "PDTS" stands for "Postures, Detentions, Transitions, Steady Shifts".

After an agreement on the linguistic structures and vocabulary on GSL, the first trials were made for 'I go to Athens' (see Figure 1). This structure was chosen as a typical and in the annotated corpus frequently occurring example. It is planned that the lexicon will contain some variations of this phrase, where the subject pronoun and the goal changes (e.g.: he/she and airport/Paris). Partner ILSP provided us with information on frequency of signs and phrases in the annotated GSL corpus.

Figure1: Frames ME GO ATHENS, extracted from the GSL corpus.

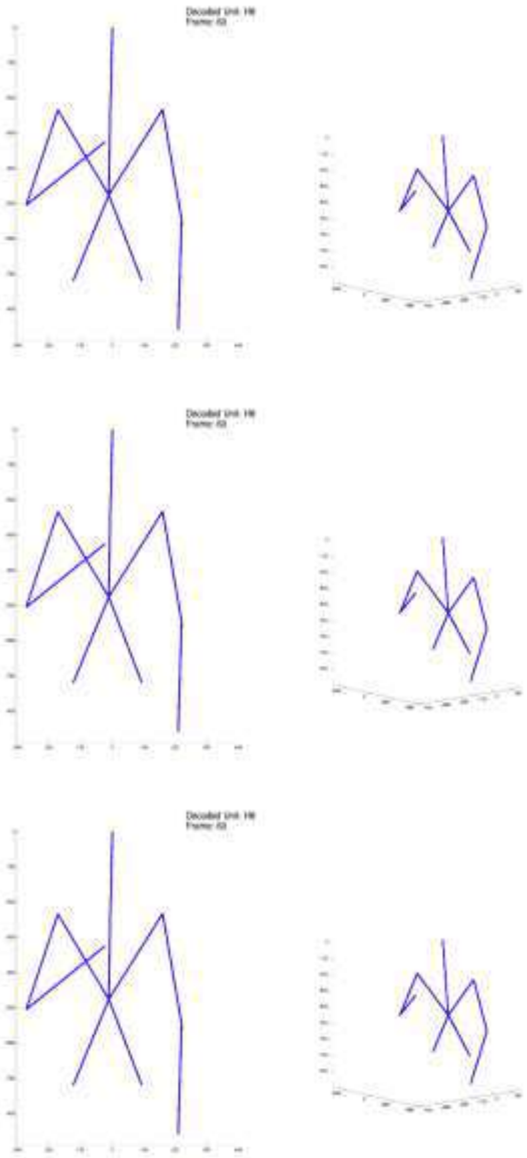

Figure2: Snapshots of recognised SU for 'I go to Athens'.

The processing begins with parsing the subunit list of NTUA (see Figure 2), which looks like the following:

```
(1) [h6,m26,ml18,mb_m17,m3,hc1],
```

where `h` stands for postures, `m` for transitions of the dominant hand, `ml` transitions of the non-dominant hand, `mb` for movements of both hands and `hc` hand constellation of both hands. Hand configuration is not known. In this list we try to find further segmentation on sign level by matching SUs against the lexicon. We also intend to recover HamNoSys from the SUs therefore a lexical entry consists of a list of SigML HamNoSys mnemonics, PDTS structure and the SUs:

```
(2)
[[go],
 [hamflathand,hamthumboutmod, Efd, hampalml, Heightsubj, Distsubj,
  hamparbegin,Direction, hamsmall,hamreplace, Efd,Heightobj,
  Distobj,hamswinging, hamparend],
 [non_dom:e_list,
  dom:(p1:[location:[Heightsubj], distance:[Distsubj]],
       t1:[direction:Direction, size:hamsmall,
           motion:hamswinging],
       p2:[location:[Heightobj],distance:[Distobj],
           movement:[Direction,hamsmall]])],
 SUverb] ---> Phon,Syn,Sem.
```

The HamNoSys contains uninstantiated elements (indicated by capital letters), because this is not a fixed verb but a directional one, therefore the location and direction of the movement can change depending on the subject and the object that it takes. SUverb stands for the SUs that come from the NTUA list and which has to be checked in the SYN feature of the verb whether a match is possible. The value e_list means that the feature non_dom (non-dominant hand) is empty being a one-handed sign. On the right hand side of the arrow the phonetic (Phon), syntactic (Syn)and semantic (Sem) feature structures describing the behaviour and characteristics of the lexical item on each level, similarly to the HPSG feature structure in the ViSiCAST project but extended with the Liddell&Johnson (1989) PDTS structures and NTUA subunits. The PDTS structure for the example verb 'GO' starts with a posture (p1) that is an associated position with the position of the subject. It is followed by a movement (transition t1) that is a small swinging movement towards a certain position that is defined by a second posture (p2), which is the object position (see further details below). The second posture also contains a movement feature. This is a result of the HamNoSys Sigml PDTS segmentation, when the position is not defined by the HamNoSys, it can be inferred from the movement in t1. The example lexical entry 'GO' can be found in Appendix Example (1) below.

In this subset we know that the lexical elements' segmentation will be either a posture (P) or a sequence of P, T (transition), P. However, in the visual recognition there might be missing elements, where P for example is not occurring. When designing the matching algorithm against the lexicon, this possibility had to be taken into account. The lexical item's SYN feature structure contains a feature called CONTEXT_IN. This feature holds the possible subunits for that entry but segmented according to the PDTS structure that characterises that lexeme. The incoming SUs are matched against this list. The list contains the most probable positions and movements for that particular entry, e.g. in case of 'GO' an outward, left or right curved or straight movement is possible, but upward and downward movements are less likely and therefore not put in the list.

```
(3) subunits:(p1:[h6,h7],t1:[m26,m40,m13,m14,m25],p2:[h8,h7,h10])
```

Here we have to note though that the SUs of movement are imprecise, in case of curved movement the combination of directions is not possible, which is a source for error in recognised movement. Also, because of the size of the training set and the fact that the SUs of positions are not directly coupled with HamNoSys, many in HamNoSys terms distinctive positions are in the same cluster. The syntactic parsing of the lexical items will have to try to recover these imperfections, whenever it is possible. In cases where the SUs are consistently wrong such a recovery might not be possible by using a grammar.

The algorithm tries to find the longest possible sequence first to match it against the lexicon. If this choice proves not to be allowed at a certain stage, it backtracks to find next possible but shorter sequence iteratively. After the match against the lexicon, we get a list of SUs segmented into signs:

```
(4)
[[[me],[HNS],[PDTSme],(h6)],
[[go],[HNS2],[PDTSgo],(m26)],
[[athens],[HNS3],[PDTSAthens],(m3,hc1)]]
```

The aim of parsing this list with the grammar is to identify the missing elements in the recognition, like not recognised positions, and associate the SUs with HamNoSys. The syntactic parsing also rules out matched signs, which are syntactically not permitted in the sequence. The latter will have an important role when the lexicon is expanded and several matches might be possible. The grammar therefore has conditions for valid sequences, which is a unification based dependency grammar (Covington, 2001), written in Prolog:

```
(5)
syn:(precomps:[(@ np2(Ph,Gloss,Plm,Efd,Index2,Precomp1,
                      Sg,Pldistr2,Heightobj,Distobj)),
               (@ np2(W,Glosssubj,Plm2,EfdT,Index1,Precomp2,
                      Num,PLdistr,Heightsubj,Distsubj)) ],
     postcomps:[],
```

The syntax feature structure contains two precomplements (precomps) for this verb.

In this case they have to be nouns/pronouns (np2 macro indicated by @ with its relevant features; Heightobj and Heightsubj. These agree with the end and start positions of the verb). Within the syntax feature structure there is also a head feature. The head feature contains the relevant type of the part of speech in the lexical entry:

(6)

`head:(dirverb_lxm,...)` for a directional verb lexeme or

`head:(loc_or_noun,...)` or `head:(pr_noun,...)` for a locatable noun or pronoun

Currently any sequence of three signs can be parsed and accepted, so as not to restrict the order and account for some flexibility for other SLs. However, the precomps (precomplements) feature is an option; it determines the complements required to precede the given sign. We also define that certain features of the complements have to match the signs features. For the directional verb `GO', the beginning/source of the sign has to match the subject; in this case the pronoun location. Now the location of the pronoun and the start of the verb sign are not necessarily the same. We can associate a position with the complement in the following feature in SYN:HEAD:

```
(7)
syn:(precomps:[],
      head:(pr_noun,
      context:(context_in : (height:[hamchest,hamshoulder],
                             dist:hamclose),
                 context_out: Cin),
```

The sign (here 'GO') therefore can start either at chest or shoulder level but still being the same sign. The same is true for the goal position, where ATHENS, the object complement also contains different possibilities for position. If the SU code for position contains these HamNoSys positions then parsing can be successful. However, as mentioned earlier SUs are often vague in HamNoSys terms or even missing, therefore the movement information on both hands is being exploited to refine the likely hand positions.

In order to do this, we changed the original matching algorithm slightly: in the SU sequence (see example 1.) we are looking more closely at the ml SUs, which reflect the movement of the non-dominant hand towards a certain position. This helps us to recover the missing position information of the non-dominant hand that in this case is the final position of the directional verb. For example in the case of Athens we only have position information on the dominant hand that starts at head or shoulder level and after a downward movement the two hands have a constellation marked by an SU hc1. Therefore there is no indication of the side where this sign was actually signed. However, if the possible sequence for Athens is preceded by a non-dominant hand movement SU, like ml14, which is an outward to the right movement, we can infer that the sign was signed to the right of the signer. This work is currently ongoing.

The output is a recognised sequence of signs with their syntactic (predicate and its dependents) and semantic structure (semantic roles) and their recovered HamNoSys. The current implementation still needs refinements and and an extension of the vocabulary based on further data. It is planned that the vocabulary has 10-15 entries. Also it is envisaged that NTUA provides a refinement to the available SUs list that probably would allow a more direct mapping of the SUs to HamNoSys in the grammar.

References

- Liddell SK, Johnson RE (1989) American sign language: The phonological base. Sign Language Studies 64:195-277
- Michael A. Covington. 2001. A fundamental algorithm for depencency parsing. Proceedings of the 39th Annual ACM Southeast Conference.
- C. Vogler, D4.3 Deliverable on Linguistic Models for Recognition, February 2011
- V. Pitsikalis, S. Theodorakis, C. Vogler and P. Maragos, "Advances in Phonetics-based Sub-Unit Modeling for Transcription Alignment and Sign Language Recognition",CVPR Workshop Gesture Recognition, Colorado Springs, June 2011.

Appendix

```
(1)
%%% lexical entry for 'GO'
[
 [go],
 [hamflathand,hamthumboutmod, Efd, hampalml, Heightsubj,Distsubj,
  hamparbegin,Direction, hamsmall,hamreplace, Efd,Heightobj,
  Distobj,hamswinging, hamparend],
 [non_dom:e_list,
  dom:(
       p1:[location:[Heightsubj], distance: [Distsubj]],
      t1:[direction:Direction, size:hamsmall, motion:hamswinging],
      p2:[location:[Heightobj],distance: [Distobj],
          movement:[Direction,hamsmall]])],
 SUverb
]

--->

word,SUverb, gloss:go,

 phon:
  (hns:(allow_weak_drop:minus,
        face:brow:Brow,
        mouth:pict:'go',
        man:(ndh:hns_string,
             hsh:hamflathand,
```

```
                ori:(plm:hampalml, efd:Efd),
                mov:[src:(por:Index1, rep:R2, gl:Glosssubj,
                          height:Heightsubj, dist:Distsubj),
                          gol:(por:Index2, gl:Gloss, rep:R1,
                              height:Heightobj, dist:Distobj),
                              fob:minus],
                const:hns_string)),
      pdts:(non_dom:e_list,
            dom:(
                p1:[location:Heightsubj, distance:Distsubj],
                t1:[direction:Direction, size:hamsmall,
                    motion:hamswinging],
                p2:[location:Heightobj,distance:Distobj,
                    movement:[Direction, hamsmall]]))),

      syn:(precomps:[(@ np2(SU,Ph,Gloss,Plm,Efd,Index2,Precomp1,
                           Sg,Pldistr2,Heightobj,Distobj)),
                     (@ np2(SU2,W,Glosssubj,Plm2,EfdT,Index1,Precomp2,
                           Num,PLdistr,Heightsubj,Distsubj))],
           postcomps:[],
           allow_pl_repeat:yes_loc_indiv_finite,
           allow_pl_sweep:no,
           head:(dirverb_lxm, agr:(num:(number:SSg,collordist:Coll),
                per:per,gref:Index2),
                aux:minus,
                context:
                  (context_in :
                     (subunits:(p1:[h6,h7], t1:[m26,m40,m13,m14,m25],
                      p2:[h8,h7,h10]),
                      add_list :
                       [(glossref:[(ref:Index2,glossr:Gloss)],
                                  locat:locatefd:Efd,
                                  distance:Distobj,
                                  heights:Heightobj),
                        (glossref:[(ref:Index1,glossr:Glosssubj)],
                                  locat:locatefd:Efdsubj,
                                  distance:Distsubj,
                                  heights:Heightsubj)],
                      delete_list:[]),
```

```
                    context_out: list_context),
              prodrp_subj:(first:can, second:cant, third:cant),
              prodrp_obj:(first:can, second:cant, third:cant)),
              arg_st:[(sem:(index:Index2,Precomp1)),
                     [(sem:(index:Index1,Precomp2))]),

   sem:(index:(ref,Ind),
        restr:[(sit:Ind3, reln:go, arg:nil,
                location:Index2, addressee:nil,
                figure:nil, ground:nil,
                instance:nil, quant:nil,
                act:Index1, thm:nil, sense:1,
                args:[(index:Index1,Precomp2),
                      (index:Index2,Precomp1)])]).
```

2. GENERATION SUPPORTING GRAMMAR

This deliverable completes the deliverable D4.2, related to Linguistic modelling. It first explains the choices we made as a consortium in terms of linguistic structures to be studied (section 2.1). A second section (2.2) gives the methodology we used to study the DictaSign corpus and derive grammatical rules. Then we give our corpus observations with respect to the structures previously chosen (section 2.3). The next sections provide a set of rule formalisations (2.4) and illustrations (2.5). Lastly (section 2.6), we talk about the implementation progress.

2.1. Choice of structures

The choice of the linguistic structures to be studied in the scope of DictaSign was decided between the project partners. The initial point was to build a set which would serve both recognition and synthesis software, and be useful for the project's applications, in particular the Sign Wiki. After the second year review, it was decided to extend the set of studied linguistic structures, but to be used only for generation. Their use for recognition is not impossible (see the “Triggering from recognition output” sub-section for each structures in D4.3) but were considered not achievable within the duration of the current project.

The first list built intuitively, and then adjusted according to findings along the way was the following:

- enumeration;
- alternative constructions;
- rhetorical questions;
- weak hand persistence.

The deliverable D4.2 report on each of those structures, describing what was observed and described about them.

After the second review, this list was adjusted and extended, with the aim to cover various levels from lexical unit to discourse structure, and useful structures to be used in the Sign Wiki. Thus the list of structures we worked on during the third year is the following:

- enumerations;
- alternative constructions;
- qualification/naming;
- neutral questions;
- quantifiers (small and big);
- announcing titles.

The first three cover part of utterances, the fourth one whole sentence, the fifth one lexical unit and the last one discourse structure.

2.2. Methodology

A “linguistic structure” here is not tied to a specific layer of language such as lexical agreement or syntactic clause order. The notion is more general, and allows generalising a production rule from the consistent matches found in our video data between a semantic operation and its surface implementation in the language. A rule can therefore represent the way to build a two-sign association, how to turn an assertion into a question or how to structure a whole discourse from the contents of its outlined titles and sections (see §2.3.6).

Here is a recall of the methodology used in D4.2, slightly adjusted in this deliverable for a less LSF-biased study, used for each structure listed in §2.1:

1. define what will be called an occurrence of the structure;
2. make a list of hypothetical parameters to be annotated in the corpora for eventual generalisation, both phonetic (like “gaze direction”) and semantic for the structure (e.g. for a list: “list is exhaustive”);
3. ensure “eye blinks” in this list, to monitor rapid eye closure systematically as they do not pop up if not tracked on purpose;
4. also add whether the occurrence is mixed with a relevant use of space, as this semantic operation often modifies surface productions which may then need to be addressed separately;
5. using this ; it is important to find various occurrences and observe multiple informants and tasks;
6. derive a rule by generalising the logical links between the criteria values, and finding the invariants over all occurrences.

2.3. Rule descriptions

For each linguistic structure given in section 2.1, a sub-section here gives:

- a **definition** of what is considered an occurrence;
- a **general observation** on what invariants were found following the methodology above, what optional features co-occurred, and which most frequent properties were observed;
- the **variation across SLs**, where applicable, reporting on other Sign Language observations from the partners when provided.

A formalisation using AZalee is given for selected structures in subsequent section 2.4. For each sub-section 2.3.x, example occurrences in a corresponding section 2.5.x illustrate the observed phenomenon with snapshots from the project corpora. It is to be noted that BSL is excluded from this study, due to the delay in the BSL corpus capture and annotation.

2.3.1. Enumerations

Definition

A set of unordered elements, signed in sequence but with no relevance to chronology.

General observation

The following invariant, initially found with the LSF corpus, was confirmed for all observed languages.

Invariant: For each item of the enumeration, the head (consider chin) moves from a forward position reached near the end of the signed item and back to its original position.

Variation across SLs

In the DGS data, enumerations equally do or do not use lexical numbering signs such as [FIRST], [SECOND], etc. Also, while head movement on each item was confirmed by our DGS team, they were reported as “more or less visible”.

2.3.2. Alternative constructions

Definition

An alternative construction is when different options of a choice are listed sequentially, and the semantics is that either one (only) of the (usually few) options is to be chosen or considered possible.

General observation

The following invariant, initially found with the LSF corpus, was confirmed for all observed languages.

Invariant: The base of the chin reaches a new position towards the end of each option, and moves downwards into a short rest before being released before the next item.

Optional: An optional lexical sign like [OR] can also be used but is not always observed either.

Optional: As often in SL, items can be assigned locations in the signing space, allowing subsequent (pointing) references.

Variation across SLs

In the LSF observed, we found many examples of this structure, and it appeared clearly that chin alone was a safe indicator of alternatives.

In the DGS data, alternative constructions are mostly expressed with an additional use of different hand locations for each options in front of the signer, together with a head tilt or turn. In some cases the upper body moves as well.

2.3.3. Qualification/naming

Definition

This section groups all structures serialising signs where one sign in the sequence (called S_0) is an entity specified by the others, with either of the following semantics:

qualifying S_0 with an adjectival construction, e.g. "a red box";

identifying S_0 with a name-sign or a finger-spelt word, e.g. "Oxford Street" or "Greece, the country".

We are only considering those where the hands are "geometrically" unrelated, thereby excluding things like pointing to the weak hand, classifier predicates, etc. Those are clearly more iconic and may need to be dealt with very specifically.

For the observations below, we only considered the occurrences where S_0 was signed first in the sequence, even though this is not part of our definition for a qualifying or naming structure. The reason for this is that our corpus study for this rule initially focused on a phonetic form (weak-hand persistence, see D4.2), rather than a semantic function.

General observation

There are different possible combinations for the occurrences of this structure, in terms of the handedness of the signs involved in the sequence. We observe the following property quite consistently:

Frequent property: When S_0 is a 2-handed sign followed by one or more qualifying/naming 1-handed signs, the weak hand tends to be held strongly in its last S_0 posture while the other signs are performed with the strong hand, as illustrated below:

```
Strong hand:   |__ S0 __| |__ 1-handed signs __|
Weak hand:     |__ S0 _______ held from S0 __________
```

We have not yet found any general invariant, but human observation of the candidate occurrences do give the feel that they group the signs of the sequence together. This is to do with the dynamics, and cannot be formalised without numeric measurements. For the time being, we propose the following hypothesis:

Invariant hypothesis: The shoulder line does not move during the sequence, and the time between S_0 and the following signs is shorter than average.

Variations across SLs

The frequent property above was confirmed for all Sign Languages observed.

2.3.4. Neutral questions

Definition

Many questions are in fact mixed with surprise or doubt. These inflict dramatically on the facial articulators and were excluded from this study. For this section, we only considered the "neutral" questions, i.e. the ones where the speaker is genuinely asking for an answer which he is not able to predict.

Different types of questions were considered:

- Yes/no questions: self-explanatory
- Closed questions: leave only a list of possible answers
- Open questions: all other questions

General observation

The point of this categorisation was to give a refined expression of the invariant for each category, as it was expected that types of questions would behave differently. In fact, regardless of the sign language, observation has proven this wrong. Questions do behave consistently, but the categories listed above seemed not to be relevant.

Every neutral question is marked by one or more of the following features:

A. a head tilt, mostly to the side;
B. lower eyelids raised, giving a sort of “interrogative squint”;
C. eyebrows raised;
D. eyebrows lowered.

Invariant: The invariant observed for neutral face-to-face questions, using the named propositions above, can be expressed as the following logical combination:

```
A or ((B or D) xor C)
```

where “`A`” is very likely, and “`(B or D) xor C`” is quite likely, which implies that the first “or” is often rather an “and” operation.

In other words, in addition to a head tilt which is a very good question marker, a visual cue for questions is the change in distance between the cheekbones and the eyebrow line, whether it is increased by raising the eyebrows or reduced by lowering them and raising the lower eyelid.

Contrarily to what is generally accepted as the most prominent question marker across most SLs, we see that eyebrows are not the only reliable marker for questions in a dialogue situation.

Optional: A rapid and small head shake.

Variation across SLs

The invariant holds for all three observed SLs: LSF, DGS and GSL. However, the use of features (A) through (D) is not consistent and here are the main differences found:

LSF seems to use (A) more than GSL, which uses eyebrows in most of the examples reported to us.

In DGS data, the features (C) and (D) are used frequently, but without any definable systematic consistency either. Instead of (A) or in combination with a head tilt sometimes the whole body moves forwards and backwards.

We are aware that these hypotheses can be the result of personal styles of signing, as only a few informants for each language were filmed in this corpus. They were consistent enough, however, for us to report them here, following our objective to compare SLs, as they can initiate further studies on the topic.

2.3.5. Small quantifier

Definition

Anything that quantifies a small size or quantity.

Observation

Various non manual events have been observed for the studied SLs, but nothing could be declared properly invariant. These are:

1. eye squint, with the lower eyelids going up;
2. eyebrows raised or lowered;
3. tip of tongue in contact with upper teeth;
4. tongue protruding;
5. mouth corners moving apart with teeth made visible;
6. elbows touching the body with shoulders slightly up;
7. cheeks sucked in.

However, item 1 was observed quite often across all languages, especially in LSF.

Frequent property: lower eyelids going up, eyes semi-closed.

The greater number of facial events above did not allow us to conclude on any invariants, but let us hypothesise that they each really fulfil different semantic purposes, which can be accumulated. One example that emerges from our observations is the tongue protrusion, which seems to be associated with a pejorative point of view of the signer more than with the quantifier itself.

Variation across SLs

In DGS data, item 3 was concluded a typical mouth gesture, whereas LSF makes a weaker use of it.

2.3.6. Big quantifier

Definition

Anything that quantifies a big size or quantity.

Observation

Here again, no invariant can be proposed, but the following various events occur more or less frequently to express this quantifier:

1. cheek or lip puff;
2. body or head leaning forward/backwards;
3. eyebrows lowered;
4. tongue made visible but inside the mouth;
5. mouth patterns like “bam!” or “pah!”;
6. eyes closing;
7. jaw lowered.

This list is obviously long enough to think it is not exhaustive, which leads us to think once again that they each take on different semantic meanings.

This category, and indeed the previous one, definitely needs some refinement and further corpus analysis to put forward invariant form-meaning associations.

Variation across SLs

Regarding item 2 in the list above, it was observed that when the body does lean on a “big” quantifier, it does so forwards in LSF. In DGS data, the body is used as a stressing moment or as a support for the expressed content. This body usage could be a movement forwards or backwards (in dependency on what the signer is going to express).

Again, this does not happen every time, but we have not seen it happen the other way round.

2.3.7. Announcing titles

Definition

A titling structure is one that would represent the signed utterance announcing the topic of a discourse section, i.e. its title in the outline of the discourse.

Observations

In GSL, several examples of such topic announcement structure were found, which all conform to the following invariant. In the other two languages observed (LSF and DGS), we have not managed to find any occurrence for this structure.

Invariant: A head nod after the end of the topic or title announced, accompanied with a closure of the eyes, not necessarily synchronised.

Variation across SLs

The only corpus containing examples of this structure was the GSL corpus. Nevertheless, this structure is not unfamiliar to LSF and many examples of it were encountered in other LSF corpora. We note that these corpora are mostly of the narration genre, which is different from the DictaSign dialogue set-up.

This difference in use of titles between GSL and LSF may raise an interesting comparative sociolinguistic study between the two SLs.

2.4. AZalee specification of selected structures

In this section, we give a formalisation, using AZalee (see D4.4, §B.1), of three linguistic structures, namely alternative lists, enumerations and section titles. These are those for which formal invariants were found, and can easily serve as tests for integration in the Sign Wiki (see D7.3).

Enumerations

Let *E1*, …, *En* be the *n* items of the enumeration. Each *Ei* is wrapped in a structure *ENUMi* defined as follows:

```
ENUMi = [[
  |hd = sgn<|,
  |hd| = short

|| sgn: USE Ei
|| hd:
   Start with chin forward
   Move it backwards
   Finish at its original position
]]
```

where `hdretract` is a short-duration TI starting where the head has reached the forward position and ending with the head back in its original position, some short while after the end of TI `sgn`, which contains the sign utterance for item *Ei*.

ENUM1, …, *ENUMn* are then signed sequentially.

Alternative constructions

Let *E1*, …, *En* be the *n* alternate options. Each *Ei* is wrapped in a structure *OPTi* defined below:

```
OPTi = [[
  |hd = sgn<|,
  |hd| = short

|| sgn: USE Ei
|| hd:
   Start with chin at location Ki different from K(i-1)
   Give it a small downward movement
   Hold it for the rest of the TI
```

```
]]
```

where `hdset` is a short-duration TI establishing the head in a new location *HDi* for option *Ei*, starting immediately before the end of `sgn`. Like the enumeration list, the *OPTi* structures are signed in a sequence for *i*=1,...,*n*.

Announcing titles

This is the azalisting representing the topic announcement *title*.

```
[[
  |nod = tit<|,
  |nod| = short,
  |blink > tit|,
  blink| < nod|

|| tit: USE title
|| blink: Close eyelids
|| nod:
   Move chin down a little
   Move chin up back to initial position
]]
```

2.5. Illustrations

For each structure and each language, we indicate the video from which the pictures are extracted and the corresponding glosses. After the pictures, we specify the observed non-manual events.

Alternative constructions

(a) LSF and DGS examples in D4.2 p.15-16

(c) GSL Pair 05 Task 2aI9

MORNING - EVENING


Left+forward followed by right+forward movement

Enumeration

(a) LSF and DGS examples in D4.2 p.13-14

(c) GSL Pair 03 task 4 I6

CREAM - RAZOR


Forward/backward body movement, head movement

Qualification/Naming

(a) LSF examples in D4.2 p.18-19

(b) DGS naming1

TICKET – rectangular object

(c) GSL Pair 01 Task6a 2b

SALAD - DELICIOUS

Neutral questions

(a) LSF T4S2 B15 @ 00'00

WANT – LEAVE – PARIS – OR - ABROAD

Eyebrow lowered and head tilt

(b) DGS Dicta08_B_1260447

small bag – PUT – INTO – BAG - YOU

Eyebrow raised

(c) GSL Pair 03 task 6b I5

LEAVE - WHERE

Eyebrow lowered

Small quantifier

(a) LSF T6S4 B @ 4'58

Small portion of food

Eye squint,, eyebrow raised, mouth corners moving apart and tongue protruding

(b) DGS quant06

Small measure

Tip of the tongue in contact with upper teeth

(c) GSL P06 Task 8 I12

Small dog

Eye squint,, eyebrow lowered and tongue protruding

Big quantifier

(a) LSF T4S9B @ 3'51

Big suitcase

Cheek puffed, body leaning forward, eyebrows lowered

(b) DGS quant5

Big bottle

Jaw puffed, body leaning forward

(c) GSL Pair 05 Task 3 I9

Big building

Eyes closing

Announcing titles.

(a) GSL P08 Task 6b I16

"PLANE" - */pause head nod/*

Pause, head nod, closure of the eyes

(b) LSF in LS-COLIN corpus Chev1Lau3vu

TITLE – STORY - HORSE

Pause, closure of the eyes

CONCLUSION

This deliverable describes the progress performed during the last year of the project on two supporting grammars, respectively dedicated to recognition and generation.

Recognition supporting grammar is aimed at enhancing recognition results in the case of contextual signs, whose production depends on the sentence context. The study has focused only on lexical units, especially directional verbs. After an agreement on limited GSL vocabulary and linguistic structures, the first tests were made for ‘I go to Athens’. This structure was chosen as it is typical and occurs frequently in the annotated corpus. The input is composed of NTUA’s subunit (SU) output and ILPS's HamNoSys+PTDS work. The output is a recognised sequence of signs with their syntactic and semantic structure and their recovered HamNoSys.

Generation supporting grammar is aimed at enhancing generation during the creation of content for the Sign Wiki, or as a post-processing step. We have focused our studies on various levels: lexical units, phrase, sentence and discourse structure. We have selected six linguistic structures and their occurrences in the DictaSign corpus, varying tasks, informants and languages: quantifiers on the lexical level; enumeration, alternative lists and naming/qualifying on a phrase level; neutral questions on a larger utterance level; and announcing titles on the discourse level. For some of the structures (enumeration, alternative, questions) all observations made so far seem to confirm not only the existence of invariants for the rules in a language, but also, to a significant extent, a certain generalisation across languages. For others (quantifiers, naming/qualifying), only frequent properties were found, but no common and systematic features. From a systematic criterion observation, we managed to derive and formalise rules for enumeration, alternatives, and title announcement, with invariant specification for synthesis. A system was implemented, capable of applying rules on sequences of signs, turning SiWikiML into LimSiWikiML, as stated in D7.3.

These two grammars have been implemented, at least on typical examples or structures, for foreseen integration to the project demonstrator and interaction with the recognition and generation modules.