

Sign language generation in an ALE HPSG

Ian Marshall

University of East Anglia

Eva Safar

University of East Anglia

Proceedings of the 11th International Conference on
Head-Driven Phrase Structure Grammar

Center for Computational Linguistics, Katholieke Universiteit Leuven

Stefan Müller (Editor)

2004

Stanford, CA: CSLI Publications

pages 189–201

Marshall, Ian & Eva Safar. 2004. Sign language generation in an ALE HPSG.
In Stefan Müller (ed.), *Proceedings of the 11th International Conference on Head-Driven Phrase Structure Grammar*, Center for Computational Linguistics, Katholieke Universiteit Leuven, 189–201. Stanford, CA: CSLI Publications. DOI: 10.21248/hpsg.2004.11.



Abstract

During the past fifty years sign languages have been recognised as genuine languages with their own syntax and distinctive phonology. In the case of sign languages, phonetic description characterises the manual and non-manual aspects of signing. The latter relate to facial expression and upper torso position. In the case of manual components these characterise hand shape, orientation and position, and hand/arm movement in three dimensional space around the signer's body. These phonetic characteristics can be notated as HamNoSys descriptions of signs which has an executable interpretation to drive an avatar.

The HPSG sign language generation component of a text to sign language system prototype is described. The assimilation of SL morphological features to generate signs which respect positional agreement in signing space are emphasised.

1 Introduction

A prototype English text to sign language (SL) translation system has been developed¹. English text is analysed into a Discourse Representation Structure Representation (DRS). The DRS is transformed into HPSG sem components as input to an ALE based SL generation component for British Sign Language (BSL) The Hamburg Notation System (HamNoSys)[Prillwitz et al. (1989); Hanke and Schmalting (2001); Hanke (2002)] provides the SL phonetic description which is subsequently input to a 'virtual human' avatar (Figure 1)²[Kennaway (2001)].

A constraint based lexicalist framework of HPSG [Pollard and Sag (1994); Ginzburg and Sag (2000)] is appropriate as sign order in SLs is largely lexically determined: verbs and adjectives typically determine whether their complements precede or succeed them. In addition the uniform representation of phonetic, syntactic and semantic information facilitates constraining of a complex though determinate information flow between lexical items and sign space representation for determining morphological constituents.

2 A Brief Characterisation of Sign Language Phenomena

This discussion concentrates upon the use of HPSG for generation of the manual components of signing in British Sign Language (BSL). In particular, we note the following phenomena [Brien (1992); Sutton-Spence and Woll (1999)]

1. some nominals can be signed at specific positions in signing space and these locations then have syntactic significance

¹This work was initiated within *ViSiCAST*, an EU Framework V supported project which builds on work supported by the UK Independent Television Commission and Post Office. The project develops virtual signing technology in order to provide information access and services to Deaf people.

²The avatar illustrated was developed by Televirtual, Norwich within *ViSiCAST*.

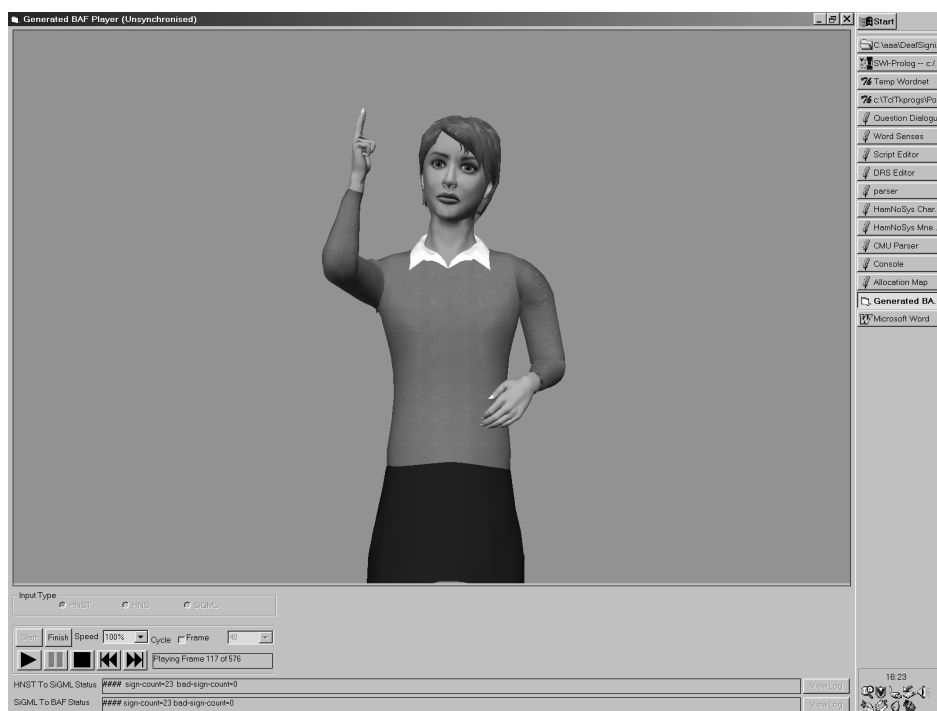


Figure 1: Avatar sign realisation

2. nominals which cannot be located in this way can be positioned in signing space by indexing a particular location after the sign
3. nominals can be referred to anaphorically by inclusion of classifier handshapes within manipulator verbs
4. directional verbs must be syntactically consistent with the locations of their subject and object
5. verbs exhibits syntactic agreement for number with their arguments.

In addition, particular positions in signing space (see Figure 2) can be populated by more than one object or person though typically these can be distinguished by different classifier handshapes. The sentence *I take the mug.* is glossed as

MUG (px) TAKE (px, p1, manip_handshape (MUG)) I (p1)

where the original position of the 'mug' and 'I' must agree with the start and end positions of the sign for 'TAKE'. The fully instantiated generated HamNoSys phonetic form for this sentence is:

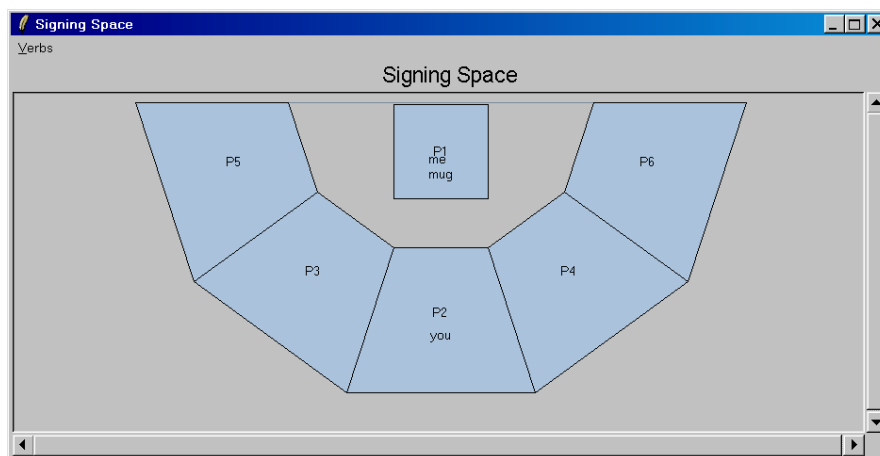


Figure 2: Signing Space Positions

```
[ [ mug ],
  [ non_raised ],
  [ hamfist, hamthumbacrossmod, hamextfingerol,
    hampalml, hamshoulders, hamclose, hamparbegin,
    hammoveu, hamarcu, hamsmallmod, hamreplace,
    hamextfingerul, hampalmdl, hamparend ] ],
[ [ take ],
  [ non_raised ],
  [ hamceall, hamextfingeror, hambetween,
    hamextfingerr, hampalml, hamshoulders, hamlrat,
    hamarmextended, hamreplace, hamextfingeror,
    hambetween, hamextfingerr, hampalml, hamchest,
    hamclose ] ],
[ [ me ],
  [ non_raised ],
  [ hamfinger2, hamthumbacrossmod, hamextfingeril,
    hampalmr, hamchest, hamtouch ] ]
```

3 The Feature Structure

The HPSG feature structure was designed for parallel development of lexicons for a number of different national sign languages³. To account for the SL phenomena above SL constituents are subclassified as *sentence*, *sent*, *phrase*, *word* and *leer* (German 'empty' denoting dropped items). *Word* is the feature structure for an in-

³Colleagues at the University of Hamburg designed the initial structure, subsequently revised to accommodate a number of refinements as work progressed.

dividual sign, and is subclassified as *verb*, *noun* and *adjective*. *Verb* is subclassified to distinguish *fixed*, *directional* (parameterised by start/end positions), *manipulative* (parameterised by a proform classifier handshake). Combinations of these types are permitted, for example 'take' is a *dir(ectional)manip(ulative)* verb. *Adj* is subclassified to indicate whether an adjective requires pre or post complements. *Noun* is subclassified as *person*, *object* or *location*, each of the former two may be either invariant or locatable in signing space. Each constituent has a relatively standard division of SEM, SYN and PHON features, and in addition a (English textual) GLOSS.

The most significant deviations from more standard HPSG grammars are the complexity of the phonetic component and the use of a CONTEXT feature within SYN.

4 The Lexicon and Grammar Rules

[[take], [Brow], [teIk, Nhd, Hsh, Efd, Plm, Const, Heightobj,
Distobj, R1, hamreplace, Efd, Plm, Heightsubj, Distsubj, R2]] \implies

word	
gloss	take
phon	see Fig 4
syn	see Fig 5
sem	see Fig 6

Figure 3: take's LHS

Figures 3, 4, 5, 6, 7, and 8 illustrate a typical lexical entry for 'take'. The lexicon exhibits the most significant adaptation of ALE to accommodate the BSL grammar. The standard ALE implementation calls the predicate 'gen' which, for a successful derivation, returns a sequence of words in one of its arguments. The ALE implementation has been modified so that this result is a list of sign phonetic descriptions, each element of which is a 3 tuple of sign gloss, non-manual and manual descriptions determined by the left hand side of a lexical items (Figure 3). The non-manual and manual descriptions are each lists of HamNoSys sign 'phonemes'. Thus, the structuring of the PHON feature duplicates information collated in the LHS of lexical items. The flattened phonetical non-manual and manual lists are determined locally by the lexical item (rather than by a post generation tree walk). However, as variables in the LHS of lexical items cannot be referenced non-locally in grammar rules, the PHON feature functions as a 'scratch area' via which non-local bindings and constraints are established.

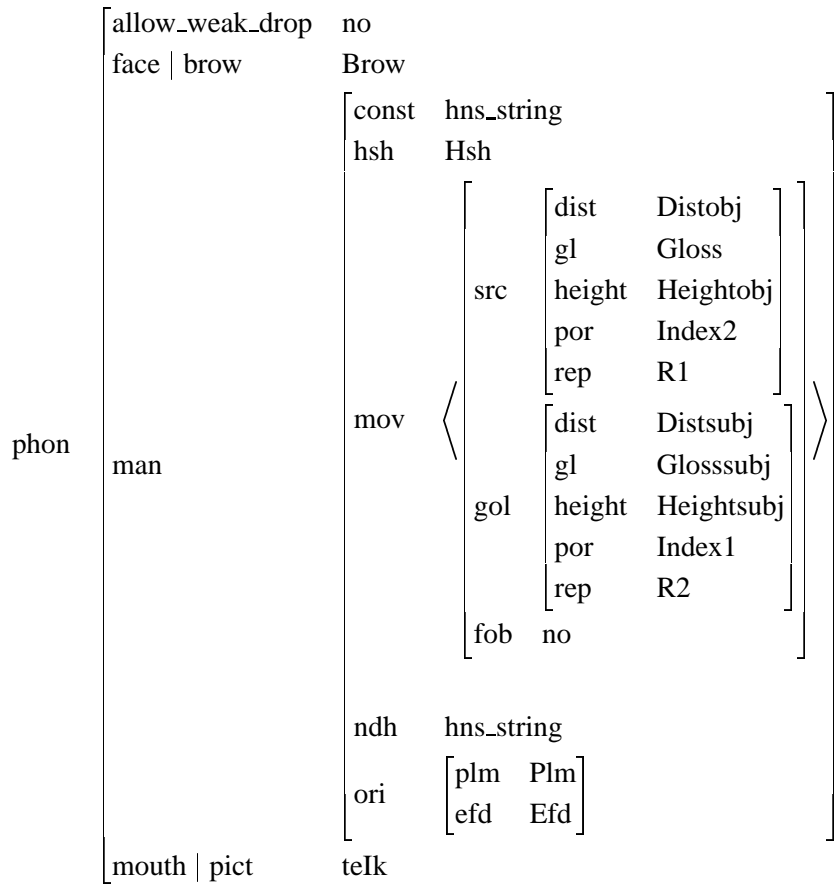


Figure 4: take's Phon

The SYN component contains features which determine sentence mode, pluralisation, (pro)noun drop and placement (anaphoric reference in signing space). In addition the usual interpretations of pre- and post- complements governs grammar rule selection and thus determines aspects of sign order.

Sentence MODE is propagated throughout a HPSG structure for a sentence by associating sentence type (declarative, yes-no question, wh-question) with eye BROW position (normal, raised, furrowed) in PHON and propagating this throughout all mother, head daughter (sem_head) and daughter (cat) nodes in both chain and non-chain rules. Currently this is overly simplistic as eye brow position is significant at the ends of questions (rather than throughout the entire proposition).

Anaphoric relationships are achieved by reference to positions within signing space. Nominals can be located at specific positions (either by being signed at that location if they are not fixed signs or by pointing). Subsequently, these positions can be used to refer to the nominal. In the case of directional verbs, such positions

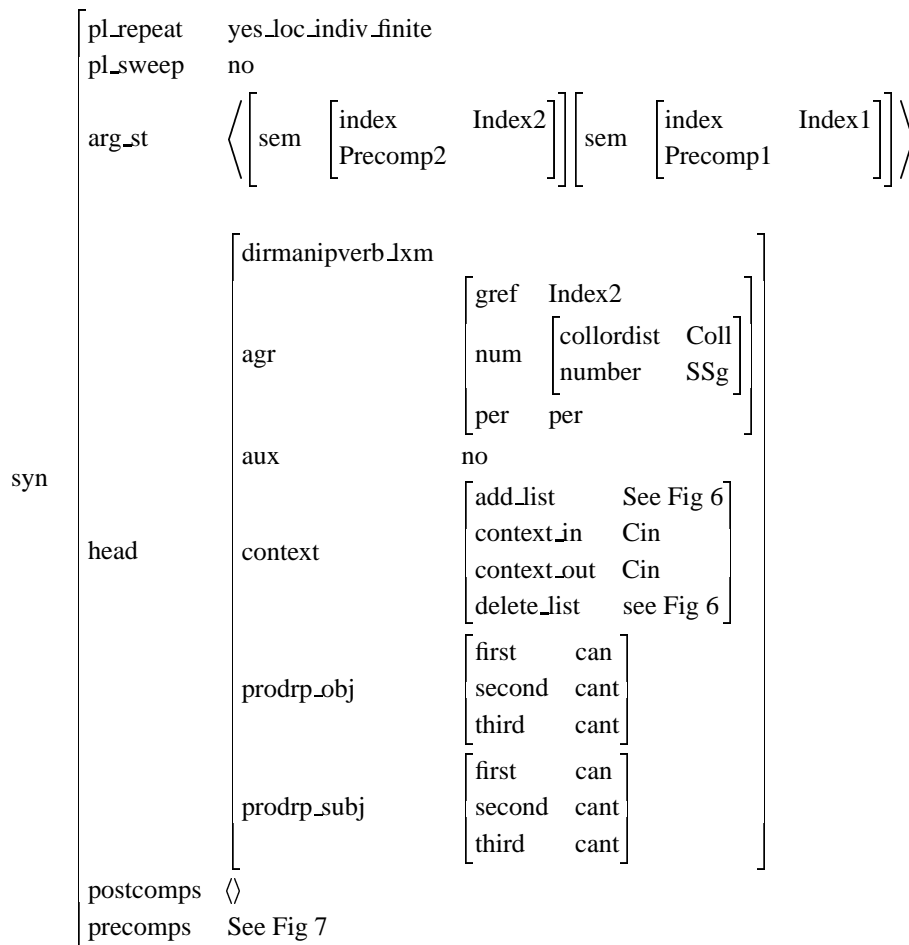


Figure 5: take's Syn

are obligatory morphemes of the sign and must agree with the appropriate position of the sign. Such agreement is achieved by propagating a map of sign space positions (phonemes for pointing, moving towards and moving to each location) through a derivation in the SYN:HEAD:CONTEXT feature. CONTEXT_IN is the current state of signing space which at the start of the derivation indicates positions of 'I' and 'you' and other non-allocated available positions. Nominals which are referred to anaphorically and those which are arguments of directional verbs of movement have to be registered within signing space. This is achieved by associating planning system style ADD_LIST and DELETE_LIST features with verbs and propositions. The ADD_LIST is an ordered sequence of registrations which are instantiated to indexical information associated with its nominal arguments and are used to register the start and end positions. The DELETE_LIST is used with directional verbs of movement in order to model movement of an object or person to the

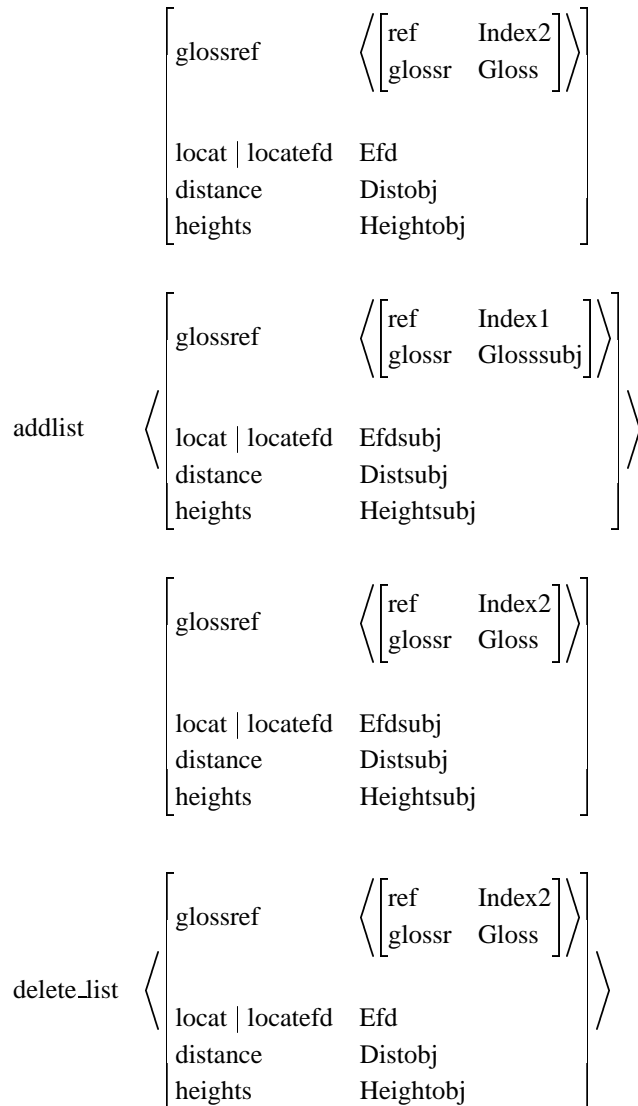


Figure 6: take's add and delete lists

destination location. The CONTEXT_OUT feature records the result of such registrations in order that an up-to-date model of signing space is propagated through a derivation.

The values Hsh, Const, Plm and Efd in the LHS of the rule (Figure 3) which determine the handshape are propagated from the classifier features associated with PreComp2 in SYN:PRECOMPS (Figure 7). Precomp2 is instantiated to the feature structure of the nominal, and hence its classifier proforms.

ALE's head-driven generation algorithm [Carpenter and Penn (1999)] is appropriate as the modelled BSL constructions are analysed as consisting of an identifi-

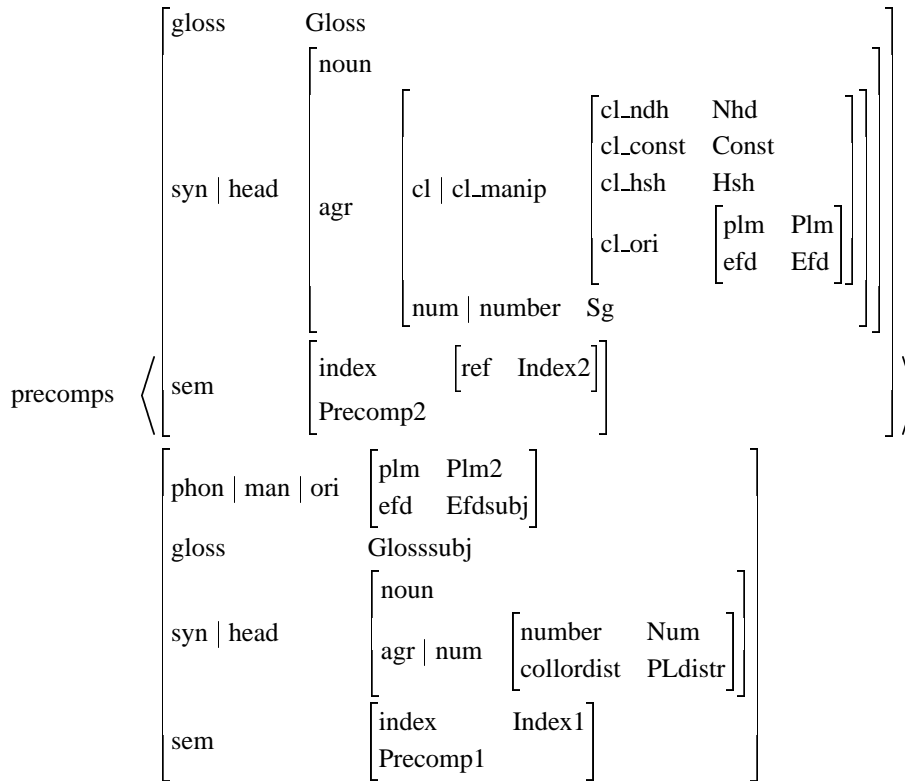


Figure 7: take's pre complements

able head-daughter. A nested SEM structure is derived from the DRS representation and an initial allocation map for SYN:HEAD:CONTEXT:CONTEXT_IN are used to initiate generation. The *semantics* predicate propagates both the SEM and CONTEXT_IN component of SYN when non-chain rules are used. Bottom-up chain rules have a conventional form as illustrated in Figure 9.

The CONTEXT_IN feature contains the available allocation map positions and reflects the state of occupied and free positions at a given stage of generation. The following illustrates a single position using HamNoSys mnemonics for the second person singular location (glossed as YOU).

```
syn:head:context:context_in:
  [(glossref: [(glossr:you, ref:Ind1)],
    halfway_mov: [hammoveo],
    locat: (locatplm: [hampalm1],
      locatefd: [hamextfingero]),
    distance: [hamarmextended],
    heights: [hamchest]),
    ...]
```

The movement of the sign towards this position is an outward movement from

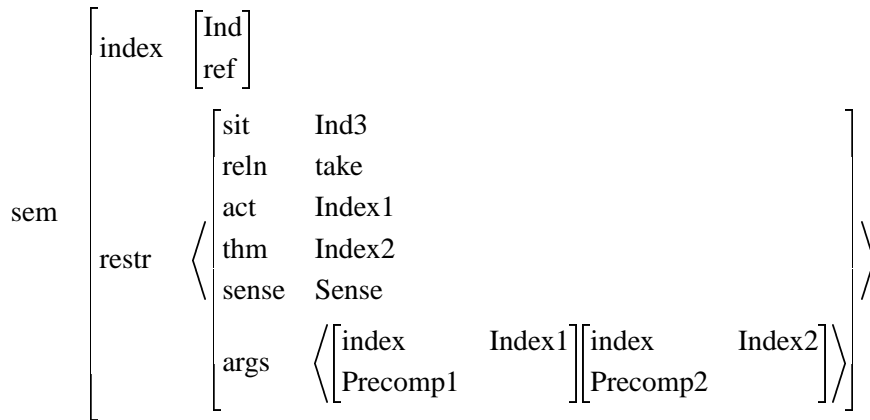


Figure 8: take's Sem

the body (hammoveo), the orientation of the palm is to the left (hampalm), extended finger direction is outward (hamextfingero). The position of the hand relative to the body is chest level (hamchest) at an extended arm distance (hamarmextended). Currently the allocation map consists of five such positions all at chest level. This needs to be extended further to include locations at differing heights to allow for locating of objects at naturally occurring locations.

The propagation of CONTEXT_IN and CONTEXT_OUT are governed by allocation map propagation principles (Figures 10, 11 and 12). Principle 1 requires that verbs and propositions propagate (unmodified) the CONTEXT_IN value to CONTEXT_OUT. As head daughters, verbs and prepositions are responsible for registering the objects/persons in their complements in signing space. Principle 3 applies for their final complements, however addition (\oplus) is addition without replication (if the object is already located in signing space then the allocated position is used). Many verbs have empty delete lists, but directional verbs of movement specify removal (\ominus) of the starting location and the add list determines the new destination location. The rule for non-final complements (Principle 2) requires the add and delete lists to be inherited by the mother node.

The ordering of head-daughter followed by daughter supports proposition-final subject pronouns. Procedural attachment of the pro_drop principle allows such a pronoun to be optionally deleted due to 'take's PRODRP.SUBJ:FIRST feature. Such deletion requires that the non-head daughter is derived as an explicit pronoun (recorded in Syn) modified (as Syni) to generate its pronominal features even though its LHS lexical realisation is empty (see Figure 9). If two ALE cat> daughters are generated to achieve this then both lexical LHSs would appear in the resulting sign sequence. Hence derivation of the pronominal form is achieved by a recursive call to ALE ('gen').

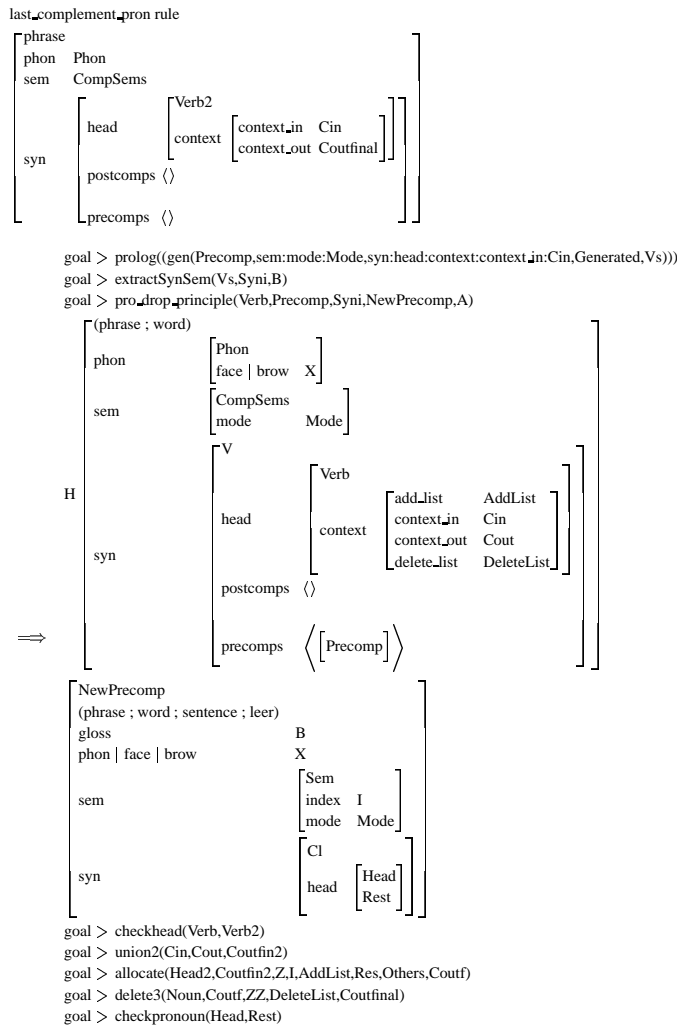


Figure 9: Last Complement Pronoun Rule

5 Conclusions

In passing we propose a comparison with the following English sentences:

1. I held the foot of the ladder on which John stood as he passed the book down/?up to me.
2. Leslie went to London where he had a gender change operation but now ?she/?he has returned to Newcastle.

The former involves semantic relationships which must be invoked to explain the anomalous reading which is sign language would be inherent in the placement of the ladder, book and individuals. In English it is difficult to manufacture examples where syntactic agreement dynamically alters mid sentence, but as we illustrate above this phenomena is not uncommon in sign language.

$$\left[\begin{array}{l} \text{word} \\ \text{syn:head:} \left[\begin{array}{l} \text{verb | preposition} \\ \text{context:} \left[\begin{array}{l} \text{context_in: } A \\ \text{context_out: } A \end{array} \right] \end{array} \right] \end{array} \right]$$

Figure 10: Principle 1 : Allocation Map - Default Lexical Propagation

For precomp: P1, postcomp: P2, where $P1 \oplus P2 \neq [C]$

$$\left[\begin{array}{l} \text{phrase} \\ \text{syn:head:context:} \left[\begin{array}{l} \text{context_in: } A \\ \text{context_out: } Z \\ \text{add_list: } \text{ADD} \\ \text{delete_list: } \text{DELETE} \end{array} \right] \end{array} \right]$$

$$\rightarrow H \left[\begin{array}{l} \text{phrase} \\ \text{syn:head:context:} \left[\begin{array}{l} \text{context_in: } A \\ \text{context_out: } Z \\ \text{add_list: } \text{ADD} \\ \text{delete_list: } \text{DELETE} \end{array} \right] \right] \left[\begin{array}{l} \text{phrase} \\ \text{syn:head:context:} \left[\begin{array}{l} \text{context_in: } A \\ \text{context_out: } Z \end{array} \right] \right]$$

$$\left[\begin{array}{l} \text{phrase} \\ \text{syn:head:context:} \left[\begin{array}{l} \text{context_in: } A \\ \text{context_out: } Z \\ \text{add_list: } \text{ADD} \\ \text{delete_list: } \text{DELETE} \end{array} \right] \right]$$

$$\rightarrow \left[\begin{array}{l} \text{phrase} \\ \text{syn:head:context:} \left[\begin{array}{l} \text{context_in: } A \\ \text{context_out: } Z \end{array} \right] \right] H \left[\begin{array}{l} \text{phrase} \\ \text{syn:head:context:} \left[\begin{array}{l} \text{context_in: } A \\ \text{context_out: } Z \\ \text{add_list: } \text{ADD} \\ \text{delete_list: } \text{DELETE} \end{array} \right] \right]$$

Figure 11: Principle 2 : Allocation Map - Default Phrasal Propagation Principles

References

- Brien, D. (Ed.). 1992. *Dictionary of British Sign Language/English*. London, Boston: Faber and Faber.
- Carpenter, B. and Penn, G. 1999. *The Attribute Logic Engine. User's Guide. Version 3.2 Beta*. Bell Labs.
- Ginzburg, J. and Sag, I.A. 2000. *Interrogative Investigations - The Form, Meaning, and Use of English Interrogatives..* Stanford, California: CSLI Publications.
- Hanke, T. 2002. HamNoSys in a sign language generation context. In R. Schulmeister and H. Reinitzer (eds.), *Progress in sign language research. (In honor of Siegmund Prillwitz)*, International Studies on Sign Language and Communication of the Deaf; 40, pages pp. 249–264, Hamburg.

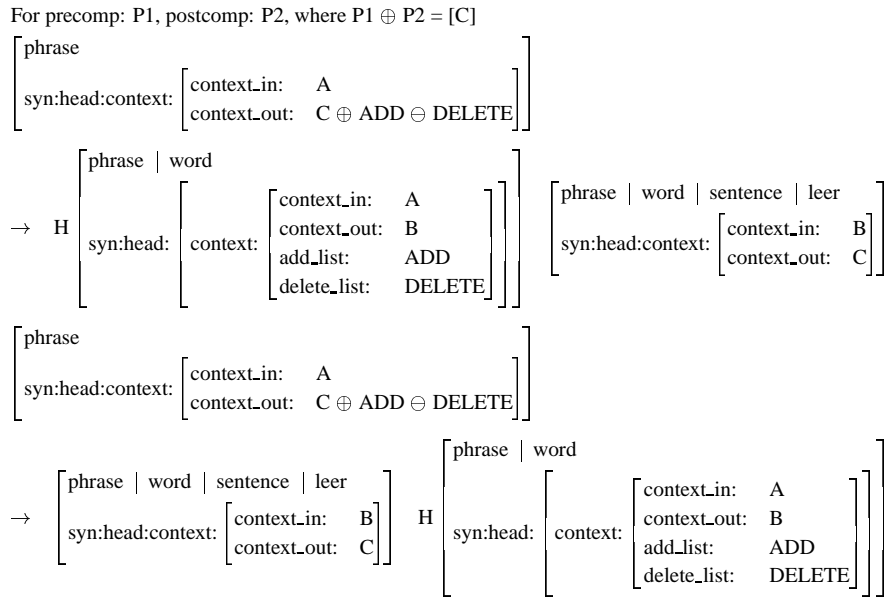


Figure 12: Principle 3: Allocation Map - Last Complement Propagation Principles

Hanke, T. and Schmalig, C. 2001. A HamNoSys-based phonetic transcription system as a basis for sign language generation. In *Gesture Workshop 2001*, London.

Kennaway, J.R. 2001. Synthetic Animation of Deaf Signing Gestures. In *The Fourth International Workshop on Gesture and Sign Language Interaction, Gesture Workshop 2001 (GW2001)*, City University, London, UK.

Pollard, C. and Sag, I.A. 1994. *Head-Driven Phrase Structure Grammar*. Chicago: The University of Chicago Press.

Prillwitz, S., Leven, R., Zienert, H., Hanke, T., Henning, J. et al. 1989. *Hamburg Notation System for Sign Languages - An Introductory Guide*. Institute of German Sign Language and Communication of the Deaf, University of Hamburg: International Studies on Sign Language and the Communication of the Deaf, Volume 5.

Sutton-Spence, R. and Woll, B. 1999. *The Linguistics of British Sign Language. An Introduction*. Cambridge: University Press.