

**ViSiCAST Deliverable D5-1: Interface Definitions  
Peer Review**

<b>Project Number:</b>	IST-1999-10500
<b>Project Title:</b>	ViSiCAST Virtual Signing: Capture, Animation, Storage and Transmission
<b>Document Type:</b>	Peer Review

<b>Deliverable Number:</b>	D5-1
<b>Contractual Date of Delivery:</b>	December 2000
<b>Actual Date of Delivery:</b>	March 2001
<b>Title of Deliverable:</b>	Interface Definitions
<b>Work-Package contributing to the Deliverable:</b>	WP 5 (Language and Notation)
<b>Nature of the Deliverable:</b>	RE (Report)
<b>Author(s):</b>	Thomas Hanke (ed.)

<b>Peer Review Author(s):</b>	Dr Trevor Johnston
<b>Date of Review:</b>	12/04/01

**ViSiCAST Deliverable D5-1: Interface Definitions  
Peer Review**

**Project Number:** IST-1999-10500  
**Project Title:** ViSiCAST  
 Virtual Signing: Capture, Animation, Storage and Transmission  
**Document Type:** Peer Review  
**Deliverable Number:** D5-1  
**Title of Deliverable:** Interface Definitions  
**Work-Package contributing to the Deliverable:**  
 WP 5 (Language and Notation)  
**Peer Review Author(s):** Dr. Trevor Johnston  
 Senior Research Associate, University of Newcastle, Australia

**Qualifications and Expertise as Peer Reviewer:**

I am a sign language linguist and have researched into and published materials on the lexicon and grammar of Auslan (Australian Sign Language). Auslan is very closely related to BSL, one of the three output sign languages of the ViSiCAST project. Indeed, I maintain that Auslan is just one of several dialects or varieties of BSL. I have had extensive experience with using and developing HamNoSys, especially in extending its range to deal with non-manual features of signing and developing conventions for dealing with stems and their related inflected or modified forms (e.g., locatable and directional signs, and 'classifier' signs). I am particularly familiar with functional theories and approaches to grammatical description and analysis, but have only superficial knowledge of Discourse Representation Theory & HPSG and their related formalisms so my ability to comment on some aspects of this deliverable/project is limited.

<b>X</b>	<b>I have read the report of the deliverable</b>
<b>X</b>	<b>I have read the Work-package description relating to the deliverable</b>

**Rating of the report as a whole:**

1	2	3	4	5	Poor =1, Excellent = 5	Explanation
			4		<b>Overall quality</b>	Needs more examples and more illustrations to ensure all participants are clear of each other's meaning. There is a lack of reference to or citation of BSL, NGT & DGS reference grammars. The style (including lack of references) may be appropriate for an internal document (assuming shared knowledge), but best practice would counsel greater clarity at all stages to facilitate monitoring or identify areas requiring follow up.
			4		<b>Reflecting the state-of-the-art</b>	Clearly participants are experts in their fields and are at the forefront of sign language transcription and virtual sign animation but could have used more of their own resources (e.g., HamNoSys) self-referentially when citing signs.
			4		<b>Meeting the objectives of the work-package</b>	With respect to the overall project aim of a 'semi-automatic translation system operating on a restricted domain' this deliverable indicates adequate, timely and appropriate progress is being made.

		3		<b>Meeting a real need</b>	Would score 4 or even 5 if the need was stated as “semi-automated sign language transcription with generated virtual signing representation”, but only scores 3 because I am not convinced automated machine translation of English text (especially on a restricted domain) into virtual signing meets a real deaf community need in itself or that cannot be better addressed in other ways.
			4	<b>Contributing to this field</b>	Significant progress towards providing the transcription and coding tools required for generating virtual signing, significant progress towards providing interface and program capability for semi-automation of sign transcription, and progress towards project aim of semi-automatic text translation in a very restricted domain with signing avatars output. Unsure about progress towards more meaningful and open-ended automatic translation of this type.
		3		<b>Style and clarity of the report</b>	Style too condensed (for outside reviewer and for any final publication) but may be acceptable to insiders. Chapters vary. Overall 3. (Chapter 1 scores 4, Chapter 2 scores 3, Chapter 3 scores 4, Chapter 4 scores 4, Chapter 5 scores 3).

**General Comments on Deliverable: as outlined in the following pages.**

#### *General comments*

In my overall opinion, this deliverable gives a clear indication of the state of the project. I believe that the content and quality of this first deliverable suggests that the three-year project is, at the end of its first year, making satisfactory progress towards its primary stated aim of a semi-automatic translation system operating on a restricted domain of English written text giving signing avatar output representation in each of three signed languages (BSL, NGT and DGS). However, I do question the feasibility of the general long term aim, the realisation of which this project positions itself only as the very first step, of automatic machine translation from English-text into a virtual signing avatar version of a specified number of signed languages. The reason for this doubt is the lack of available published descriptive grammars of the signed languages in question, especially NGT and DGS, but also including BSL. Regardless of whether or not the long term aim be judged desirable, useful or meeting a real need (on which I reserve my judgement), the more limited goal of this project is certainly achievable but even here the lack of descriptive grammars will pose some problem. I do not wish to underestimate the significance and relevance of a possible application of this technology insofar as it facilitates automation of transcription and the generation of a parallel virtual signing output. I think *this* is very important and very useful.

#### *Chapter 1 Introduction*

This chapter represents a good clear summary of the intended form of the final product, though the style is somewhat dense. I suspect that the contents of this chapter probably formed the basis of any initial submission to undertake this project. Consequently, I would have expected greater clarity and thorough exemplification of all general points in this first chapter of the first deliverable. For example, the project team should have taken one English sentence and illustrated its decomposition and recomposition at each stage in the five modelled steps of the translation and virtual signing generation process. This would have made the explanation of the processes clearer. However, this was not done. Furthermore, the illustrative freeze-frames of virtual signing in the example screen dumps for the morphology and animation screens all use the same illustration. I would have expected the report to include at least four distinct and appropriate freeze-frames to illustrate each sign in the morphology and animation screens of the example sentence(s) used. One would have more confidence that real progress was being made if this relatively simple step (even if only a mock-up) had been taken.

The author of the chapter makes it quite clear that the end product will need to have its language domain severely restricted (out of necessity) simply to be able to run from the English module to the final animation module at all. Furthermore, the author states that it appears inevitable that human manual

intervention will be needed to make any non-trivial automatic output correct, especially if co-reference or spatial and or directional modifications are required (e.g., if the text is of any substantial length, which may mean as little as two or more sentences/clauses if they involve anaphora in any way).

At several points in this text, too much assumption is made about the reader knowing what the project participants are talking about. For example, ‘signing assistants’ are mentioned on page 3. A footnote mentions that they will be discussed in work-package 2, but I can find no explanation of what ‘signing assistants’ may actually be. More explanation is needed. (If it has been provided as part of other documentation linked with the project, then the reviewer needs to see it in order to form a proper judgement.)

### *Chapter 2 Discourse Representation Structure theory*

I am less qualified to comment on this chapter than any other in the deliverable, having only a non-specialist familiarity with Discourse Representation Theory, Discourse Representation Structures and HPSG.

This chapter focuses on the representation of the content of English sentences into a DRS representation (to be further represented using HPSG then represented into the target sign languages in subsequent phases of the project). Consequently, some account of sign language grammars must be made at this stage so that the DRS captures information, which would otherwise be lost, in an appropriate form so that it can be later realised when coding into BSL, NGT or DGS, further on in the project.

In this chapter, some of features of sign language grammars that the DRS team feel are important (Safar & Marshall, 2001) are discussed. Namely:

- (i) pronominal reference, co-reference and placement
- (ii) classifiers
- (iii) temporal marking
- (iv) distributive aspect
- (v) topic comment structure (sign order)

In the chapter itself, the source of the language judgements is not made clear (i.e., reference to published documentation on the relevant signed languages is not made). From Safar & Marshall (2001) one can assume the sources for BSL are Magill & Hodgson (2000) and Sutton-Spence & Woll (1999). (I question the appropriateness of Magill & Hodgson (2000) as a source text on grammatical structures. Its discussion of grammar is perfunctory and oversimplified in the extreme. Even as a source of BSL lexicon, great caution should be exercised due to dialect and regional variation in Britain. The lexicon is closer to some varieties of Auslan than it is to some other varieties of BSL!) There is also no indication, here or elsewhere in the deliverable, what the equivalent grammatical overviews may be for NGT and DGS. (I doubt that they exist.)

I would have expected that a project of this nature, given its aims, to clearly reference key descriptive grammars of each of these sign languages. The paper does provide clear reference to the literature on DRS and should similarly provide references to accepted descriptive grammars of each of the target signed languages. This criticism may be unjustified insofar as this is an internal project document and this information may already be known to other participants (and known to be known by the authors).

*[It appears from D5-1, Chapter 5 “Lexicon structure” (footnote 27) and from the accompanying documentation regarding the work-package description and statement of milestones that a description of the grammar module is scheduled for delivery in D5-3, July 2001, so my comments here may be addressed by that deliverable. However, it seems that it is the project itself will provide a statement of the ‘grammar for rules of sign formation’ in each of the three target signed languages. This seems a somewhat ambitious task given the state of the literature.]*

The Safar & Marshall (2001) paper provides some background information concerning what the consortium members who are working on the DRS representation consider to be key features of ‘sign language grammars’. Most are discussed in this chapter (directional verbs are discussed in other chapters of this deliverable.) One assumes the information on location and placement in the signing space will be used to control deixis (e.g., pronouns) *and* verb directionality.

I find the discussion of the major characteristics of sign languages rather superficial and somewhat misleading. Let me give an example of the problem with respect to ‘time lines’ in BSL. I would argue that time lines, as such, are far less important than temporal adverbials for signalling temporal information. That is, the important and essential observation, with respect to BSL and Auslan at least, is that the signalling of temporal information is made lexically or ‘periphrastically’, using ‘time signs’ or temporal adverbials rather than morphologically, inflecting verb forms for tense. These temporal adverbials themselves may or may not be articulated with reference to a number of time lines found in BSL and Auslan. (In other words, there are some temporal adverbials that appear to have no ‘time line’

component or actually contradict the supposed temporal directionality of the time lines). Reference to descriptive grammars would be of assistance here in clarifying the source and status of the information the authors present. I consider this to be essential with these target output languages because they are “less intensively researched [...] languages” (D5-1, Chapter 1, page 3) than the (written) languages that automatic machine translation programs and routines usually deal with.

For example, the observation (Brennan 1983) that BSL has four time lines has been long contested (Johnston 1989 unpublished doctoral dissertation, University of Sydney) and more recently questioned (Sutton-Spence & Woll 1999). Regardless of how many time lines there may or may not be in BSL, I believe this fact is actually immaterial to the project-at-hand. Tense needs to be coded separately so that it can be later realised in signed output with a lexical temporal adverbial (fully specified). It is true that some signs representing temporal units, such as MONTH, may be projected along a ‘time line’ to signify ‘monthly’ or ‘every month’ but they are few in number and not related to tense as such. Moreover, they can also be expressed perfectly correctly lexically as EACH MONTH or REGULAR MONTH. Consequently, one wonders how the supposed articulation of some signs with respect to any number of time lines is relevant at all to the coding exercise. I use this example to show that it is not clear from the documentation and its referencing what participants think the consequence of identifying this and some other aspects of the grammar of BSL etc. is, and what they actually understand by the feature.

As for the DRS itself, the subsequent documentation is beyond my expertise to evaluate. In so far as I can follow the formalisation, it appears to me that some progress has been made in allowing for some of the five ‘sign language characteristics’ in the DRS representation discussed in the introduction to the chapter. What appears to be presented in this first deliverable is a standard ‘formalisation’ of DRS, drawn from the literature, for a number of illustrative cases. I am unable to comment further.

### *Chapter 3 Encoding manual aspects of sign language: HamNoSys 4.0*

For the reader who is familiar with HamNoSys, this chapter succinctly states additions, updates and alterations to HamNoSys that address most of the previously identified short-comings of earlier versions of the transcription system. I have no major problems with any of the proposed changes with the system. It is clear, to the point and brief. Indeed, it is somewhat *too* brief because it lacks illustrations and adequate exemplification. Only someone very familiar with the system could possibly follow this chapter as it now stands.

It seems to me that the usefulness of the final documentation for HamNoSys 4.0, which I assume is in production, to the ViSiCAST partners—those who will be responsible for coding sign language data or lexicon writers—will be minimal until fully illustrated and clearly exemplified HamNoSys 4.0 document is published. Finally, with respect to HamNoSys, there is no indication in the deliverable documentation if the morphology module of the ViSiCAST translation process will include some automatic system for checking the well-formedness of transcription strings in HamNoSys, either by reference to certain ‘syntax rules’ or to a generated dictionary of previously accepted transcribed forms. I think this should be addressed as soon as possible if it has not already.

### *Chapter 4 Encoding non-manual aspects of sign language*

This is another clear and straightforward chapter. The team has achieved its aim of providing the ViSiCAST members with a simple, easy-to-key-in coding system for all the major non-manual behaviours thought to be significant by sign language linguists. I cannot think of any non-manual behaviours that have not been identified. There will, however, be little problem in expanding the non-manual codings to accommodate any newly identified behaviours.

It was decided not to use HamNoSys symbols to encode non-manual behaviours because of an unwillingness to add new semantics to established symbols to facilitate this. It is argued that it goes against the original idea of HamNoSys (a phonetic script for the writing of a manual sign form in one continuous string?). It is possible that computer generated signing output has always been linked with HamNoSys, but I suspect that the *writing* function for the system has been to the forefront until relatively recently. Now that HamNoSys (in version 4.0, see chapter 3 of the deliverable) has been extended to be increasingly phonetically accurate *in order to provide a base for automatic routines to generate a virtual signed form through an avatar* it is actually no longer relevant what the coding system actually *looks like*—it is simply made available to a program to generate an image or animation, not for anyone to *read* (except when keying in and correcting on the morphology screen). So, rather than ‘pollute’ the original idea of HamNoSys, short strings of Roman letters, as defined in this chapter of the deliverable, are to be used on separate tiers in the sign representation to code for each non-manual behaviour. However, it is by no means obvious to the reviewer that doing so is any easier for a transcriber than staying within the repertoire of HamNoSys symbols but giving them a different semantics (or value) when entered on

different tiers of the sign representation. It may actually be easier for a transcriber to stay within the script when referring to sign form—only experience will tell.

Another, better reason for opting for the solution presented in this chapter has been communicated to me—using modified HamNoSys symbols may be very unstable in this environment in terms of dealing with hitherto unidentified facial expressions. It would be difficult to modify the symbols' semantics or value yet again, further down the track, if new types of facial expressions are identified..

These considerations aside, the solution offered in this chapter is entirely acceptable and practical, though it may not be the best.

I find the distinction between states and transitions overstated because non-manual behaviours are clearly very dynamic, which the writers readily admit. (One reason for making this 'overstated' distinction appears to be to justify not using HamNoSys movement operators and thus HamNoSys symbols in the coding for non-manual aspects—they deal with movements, not states. However, I maintain that it is relatively an easy matter to change the semantics of individual symbols and combinations of symbols. But this is a separate and, now, irrelevant question.) There is little doubt that some non-manual behaviours are often 'held' for considerable periods of time relative to the manual sign stream. Therefore, it is simply a practical consideration to code the state and define its domain rather than describe a transition. (One, of course, assumes that the generation routine will automatically build in transition states from neutral to non-neutral facial expressions.)

No indication is given of how the multiple non-manual tiers will be incorporated into the morphology screen (or screens) of the ViSiCAST integrated editing environments. This may be part of some future deliverable. The earlier this is addressed, the better.

### *Chapter 5 Lexical structure*

The style of this chapter is particularly condensed, to the point of sometimes being cryptic. It could be much clearer. It appears incomplete. I would like some indication of how aspect and plurality are to be dealt with here or, at least, when they will be dealt with. (Footnote 30 suggests plurals will be described in deliverable D5-4.)

Questions of grammar are necessarily involved in the discussion of the structure of the lexicon, so once again I expect to be able to see a point by point summary of the grammatical features of BSL, NGT and DGS or, at least, references to documentation that does this. Footnote 27 in this chapter suggests that there is a grammar module in a later deliverable (D5-3). However, it is not clear if that module simply formalises aspects of the grammar (with or without systematic reference to published accounts of the three signed languages) or actually presents a summary of the signed language grammars in question.

Before I continue, I am struck by the fact so little attention is paid to exemplification and illustration of sign language examples. True, this chapter does include numerous examples from DGS, and in many instances I either happen to know the DGS sign or can make an intelligent guess of sign form based on iconicity. However, I cannot really understand why the authors do not use *both* HamNoSys *and* illustrations to make their meaning as clear as possible. It does not inspire confidence that in the very environment where HamNoSys (let alone digitised graphics and freeze-frames) have their undisputed proper application—in technical discussions regarding sign language—it is not being used. After all, this is precisely the context in which HamNoSys should be used. Instead we only get glosses, as if HamNoSys did not actually exist.

To return to grammar and the structure of the lexicon, it is not clear to me what is being proposed in this chapter. The author does not come out and simply and unambiguously state it. I assume that non-inflecting (or invariant) signs will be written in traditional HamNoSys strings (using version 4.0). I then assume that inflectional signs will be dealt with in one of two ways: (i) have some kind of fully specified stem notated in HamNoSys to which some kind of HPSG feature structure, specifying the modification, will be attached or linked; or (ii) there may be no fully specified 'stem' at all, and the HPSG-like feature structure will take its place according to the three different models described. The feature structure would then be reduced to a HamNoSys string to feed the animation machine. If any stem existed at all for these types of signs it is a skeleton of the feature structure that contains only the invariant (HamNoSys) values for that sign. (I have been informed in a personal communication that this second description is closer to the real situation.) The model examples we are given are 'heavily abbreviated', so I am unsure what a full form would look like.

I have two queries. First, if there really is no traditional stem written in HamNoSys at all for inflecting signs then surely the majority of lexical entries would be like this. (It is my belief that only a small minority of lexical signs are actually 'invariant' in these signed languages—if not all signed languages.) In this case, why not enter the information for all signs in the same way? That is, generate all HamNoSys strings from a limited set of feature structures. Second, do these inflectional signs really have no stem?

Surely most relate to a citation sign of which they can be seen to be a modification. Is it not just as simple to substitute for a citation value in a stem, as it is to fill an empty slot with some 'calculated' value? There may be good technical reasons for this of which I am unaware.

With reference to the types of inflectional signs in the three signed languages and the three major subtypes that are suggested, I believe more discussion and justification may be required of the 'typology'. Classifier signs are a particular problem area in sign linguistics. I wonder if the definition given really is compatible or consistent with most approaches (as suggested in footnote 29): "a sign that inherits some form aspects from the semantics of one of its arguments [is] a classifier sign, the information inherited is called a classifier".

I have little problem with the recognition of two of the types of inflectional signs: locatable signs and directional signs. The three types of location change (relocating, relocating and rotating, body locating) are appropriately identified but the explication is so condensed I have no real idea of how this will be coded into HamNoSys strings, or how the parameter framework will be used. (In this context, 'handconstellation' need to be defined clearly.) Similarly, with respect to directional signs the SRC and GOL morphemes and the POR, HEIGHT and DIST features pose no problems. I just do not understand how it is to be implemented or coded into the lexicon.

Before the project progresses much further I would expect to be able to see a sizeable sample lexicon of several hundred signs from each of the target signed languages tagged according to inflectional type. Each sign should be categorised as inflectional or non-inflectional, and locatable and directional (according to sub-types). This would test the categories and indicate the number of signs of each type in the lexicon. The work-page descriptions and milestones indicate that modelling will use test sign lexicons of 500 and 1000 items at various stages, so this point may well be addressed as a matter of course as the project unfolds.

**Comments on Contribution to objectives of Work-package and ViSiCAST project:**

This deliverable is part of WP-5 "Language and notation" and its aim was to define "the computational and communication interfaces between various system components of the project". I have given a rating score of 4 for "meeting the objectives of the Work-package" because from reading the documentation it is as yet unclear how all the various components will, in fact, be integrated, except in the broadest possible terms. I have little doubt the project team knows this very well, it is just not conveyed very well, from my perspective. This is apparently a task of future work-packages. This deliverable does nonetheless identify and define the system components adequately. With respect to the overall project aim of a 'semi-automatic translation system operating on a restricted domain' this deliverable indicates adequate and appropriate progress is being made within the timetable.

**Signed**

**Date:**