The Challenges of Cross-Modal Translation:

English to Sign Language Translation in the Zardoz System

Tony Veale[1], Bróna Collins[2], Alan Conway[3]

[1]*School of Computer Applications, Dublin City University, Glasnevin, Ireland.*

[2]*Department of Computer Science, Trinity College, Dublin, Ireland.*

[3]*Iona Technologies, Dublin, Ireland.*

Abstract

The sign languages used by deaf communities around the world represent a linguistic challenge that natural language researchers in A.I. have only recently begun to take up. This challenge is particularly relevant to research in Machine Translation, as natural sign languages—such as ISL (Ireland), BSL (Britain) and ASL (U.S.A.)—have evolved in deaf communities into efficient modes of gestural communication, which differ from English not only in modality but in grammatical structure, exploiting as they do a higher dimensionality of spatial expression. In this paper we describe Zardoz, an on-going AI research system that tackles the cross-modal machine-translation problem, translating English text into fluid sign language. The paper presents an architectural overview of Zardoz, describing its central blackboard organization, the nature of its Interlingual representation, and the major components which interact through this blackboard to both analyze the verbal input and generate the corresponding gestural output in one of a number of sign variants.

Keywords: Sign Language, Space, Gesture, Cross-Modal Translation, Metaphor

The Challenges of Cross-Modal Translation:

English to Sign Language Translation in the Zardoz System

Tony Veale[1], Bróna Collins[2], Alan Conway[3]

[1]*School of Computer Applications, Dublin City University, Glasnevin, Ireland.*[*]

[2]*Department of Computer Science, Trinity College, Dublin, Ireland.*

[3]*Iona Technologies, Dublin, Ireland.*[*]

1. Introduction

Recent years have seen the acceptance of sign-language by the linguistics community as a fully-featured, *first-class* natural language, one that exhibits the full range of traditional linguistic phenomena, as well as a host of expressive powers unique to gestural communication (e.g., see Klima and Bellugi 1979; Liddell 1980). In parallel, the sign languages used by deaf communities around the world represent a linguistic challenge that natural language researchers in A.I. have only recently begun to take up. This paper describes the architecture and methodology of Zardoz, a multilingual sign translation system designed to translate textual/spoken language (ostensibly English) into a variety of graphically animated sign-language variants, in particular ASL (American), ISL (Irish) and JSL (Japanese). This goal of fluid articulation of sign language gestures from English language input embodies the unique linguistic challenge of *cross-modal* translation, one which possesses significant social, commercial and theoretical implications.

Sign translation raises a variety of interesting issues for the way MT is used. For instance, there exists a sizable body of sign language users world-wide, for which such technology will provide valuable educational tools: the technology will not *replace*, but *empower* and educate new sign interpreters. Indeed, contrary to the perceived A.I. goal of humanizing machinery, sign translation systems are tailor-made for social situations where an obviously *non-human* translator is required, for a machine will not violate the doctor/patient and lawyer/client confidentiality expected by a signer. From a linguistic and cognitive perspective, the pursuit of cross-modal translation further challenges our preconceptions about what constitute language universals. And from a pragmatic A.I. perspective, sign-language MT is a unifying goal which provides an ideal opportunity for the synthesis of existing A.I. theories and techniques into a workable and socially-relevant application in the short term. As there exists an ever-growing body of research concerning the structural properties of sign-languages, for instance the treatment of ASL due to Liddell (1980), this paper complements this work in discussing the purely A.I. considerations of sign communication.

Zardoz thus represents a compromise between pure theory and practical utility. In many respects the design choices made in Zardoz are not final but convenient, employing A.I. techniques chosen for their pragmatic applicability to the project rather than their cognitive or theoretical import. Given these pragmatic and theoretical goals, the rest of this paper assumes the following structure: section two introduces the sign language medium, which serves to place the contents of the paper in some focus. Section three then presents an overview of the system architecture of Zardoz, which is conceived and implemented around the blackboard control metaphor. Section four discusses the motivations and mechanics of conceptual interlingual representation as employed in Zardoz; central to this discussion is the notion of conceptual *schematization* in sign-languages, that is, the mechanism by which concepts are *chunked* into manageable units, and the manner in which these chunks are manifest at the lexical/gestural level.

---

* The research reported in this paper was initially conducted by the first two and third authors as part of an ongoing development effort by Hitachi Dublin Laboratory from 1993 – 1995. The first author may be contacted by email at *tonyv@compapp.dcu.ie*.

Consideration of sign-language at a conceptual level offers an abundance of evidence for the proposition that different languages conceptualize the world in different ways (i.e., the weak Sapir-Whorf hypothesis). In particular, because sign-languages tend to emphasize concepts and metaphors that make explicit the spatial dimensions of the actions involved, a sign MT system requires a deep understanding of conceptual schemata and their spatial underpinnings to produce natural sign generation. Likewise, an appreciation of the cultural boundaries of such metaphors and schemata must also be a major consideration in any interlingua-based translation system. Section five then turns to grammatical issues, providing a description of the syntactic formalism—based upon *Spatial-Dependency* graphs—which Zardoz employs to specify the output syntax of sign languages in a flexible and robust manner. Having considered the broad syntactic/semantic issues of sign language, section six then turns to matters lexical, discussing the mechanisms employed by Zardoz for sign lookup and dynamic sign invention, a topic which exploits principles of spatial metaphor in sign. Since much of spoken language, such as English, is *localist* in nature (see Lyons 1977), structured as it is around a host of core spatial metaphors (such as the orientation metaphors of Lakoff & Johnson 1980), it should not be surprising then that sign language, which uses space not only as a conceptual medium but as an expressive canvass, should be rich in exploitable spatial metaphor. Space, and its cognitive role in sign thought are thus stressed throughout as the unifying theme of this paper. Section seven delves deeper into this vein, discussing more pressing issues of spatial awareness in sign, such as *gestural anaphora* and the systematic allocation of *spatial indices* in multi-entity situations. The paper concludes with a summary and some closing remarks in section eight.

2. Sign Language as a Communication Medium

There is a strong tendency among the speaking community to trivialize the capacity of sign as a full communication medium. It is not an uncommon assumption that sign language, being iconic in nature, is a universal language shared by the deaf communities of the world. It therefore comes as no small shock to holders of this view that variants of sign language differ widely from country to country, and that nations which ostensibly share the same spoken language (e.g., English in the cases of Britain, Ireland and America) do not necessarily employ the same form of sign (e.g., BSL, ISL and ASL respectively). These assumptions derive from two common misconceptions: firstly, that sign language is primarily iconic in nature, and secondly, that sign language is a gesturally-coded form of spoken language. Certainly iconicity plays a stronger role in sign language than sound symbolism does in spoken language, but as with any full language there exists a strong tendency to move from iconicity to arbitrariness (see Klima & Bellugi 1979). And while sign language can often be employed as mere gestural *coding* of a spoken language, *native* sign language possesses a syntax which is independent of any spoken language.

However, the difficulty in specifying and storing sign gestures, as opposed to lexemes, severely limits the range of lexical resources available within the medium. It follows that the consensus core of a sign language (the body of signs known to most users) is considerably smaller than that of a language such as English (as defined by the Oxford English Dictionary, say), and thus sign language is often seen to be lexically (though not expressively) *impoverished*. It is thus necessary for a sign generator to exhibit some degree of creativity in assigning concept-to-sign correspondences. Metaphor-based measures for assigning such correspondences *on-the-fly* are discussed in section five.

2.1. Notational Conventions

At this juncture it is perhaps useful to introduce the notation employed throughout this paper to distinguish words, concepts and signs. A lexeme is denoted in roman face within quotation marks, while the underlying, language-independent concept is capitalized in courier. Thus, “Headache” denotes the English word for same, while `Headache` denotes the interlingual token corresponding to that lexeme. Signs, being symbolic frame names for the purposes of the system, are also capitalized, but are additionally qualified by a particular sign variant identifier. The token `ASL-Headache` thus denotes the sign token for "Headache" in ASL. In addition, the notation `Part::Sign` indicates that Sign is made at, or with (if Part is a hand) a particular body Part. Thus `Left-Hand::ASL-Man` directs that the sign for "Man" in ASL is to be made *with* the left hand, while `Elbow::ASL-Hurt` directs that the ASL sign for "Pain" (as "Hurt" and "Pain" are synonymous in ASL) be made *at* the elbow. Specific gestural signals, such as *Tilt-Head-Backwards*, which have a specific meaning in a sign language but are not performed with the hands (rather the face, posture, etc.), are denoted using a capitalized italics face. And finally, because the letters of the English alphabet "A" ... "Z" are represented in a sign language like ASL by distinct hand-position signs, they are denoted here as `ASL-A` ... `ASL-Z` for the purposes of finger-spelling.

2.2. Non-Manual Features

The task of text to sign language translation is intuitively more akin to the task of *speech to speech* MT—as employed in the ITVox system of Werhli (1996; 1992) say—than to the task of traditional *text to text* MT. This is because the target language must be considered at the phonological, rather than simply morphological, level, if fluid articulation of sign language is to be achieved (see Padden and Perlmutter 1987). Thus, as described in Conway and Veale (1994), Zardoz employs the sign-phonology model of Sandler (1989) as the representational basis of its graphical animation component. Nonetheless, to maintain a coherent focus, the current paper will address only the morphological level of sign, that is, the level at which gestural features can be directly associated with semantic features.

Of primary concern then is the treatment of non-manual features in sign, i.e., gestural features that are not signed by the hands, but by the face, the shoulders, the head, and so on. These features are generally articulated in parallel with some other sequence of manual signs, causing those manual features to be construed in a certain, inflected sense. For instance, an opening of the lips and a protrusion of the tongue during the articulation of an adjective in ASL directs the receiver to construe that adjective as being used in an adverbial sense. This feature is thus the equivalent of the suffix morpheme *+ly* in English. Liddell (1980) also isolates the following non-manual features, amongst others, in ASL:

- *Nod Head Slowly* : *Assertion / Existence Marker*
- *Raise-Brow ,Tilt Head back slightly* : *Topic Marker*
- *Raise Brow, Cheek and Upper Lip, Tilt Head back* : *Relative Clause Marker*
- *Shake Head side to side, purse Lips and frown* : *Negation Marker*
- *Eyebrows downward* : *Wh-Question*

These features have scope over the clauses with which they are co-articulated, allowing Zardoz to view them as binary switches that may be turned on and off at the boundaries of well-formed constituents. For instance, in section five we shall see how Zardoz inserts symbolic markers for these features (such as *Eyebrows-Downward*) into the target surface syntax to signal the beginning of the relevant non-manual articulation, while using the generic marker *Resume-Previous-Face* (as also employed in Liddell 1980) to signal the end of this co-articulation.

3. System Architecture: An Overview of the Zardoz system

Zardoz is constructed as a modular system organized around a central blackboard control structure (see Cunningham and Veale 1991, Veale & Cunningham 1992). This blackboard is in turn built upon the frame-based Knowledge Representation language Krell (see Veale & Smyth 1992), whose generic frame format and rich demonology is also suited to the representation of, and seamless communication between, the concept network, the Interlingua level, and language-specific lexicons.

A process-oriented view of the system is illustrated in Figure 1, which presents the blackboard as compartmentalized into distinct *panels*. Task-specific knowledge agencies (composed of autonomous, write-activated demons) communicate by both reading from and writing to these panels.

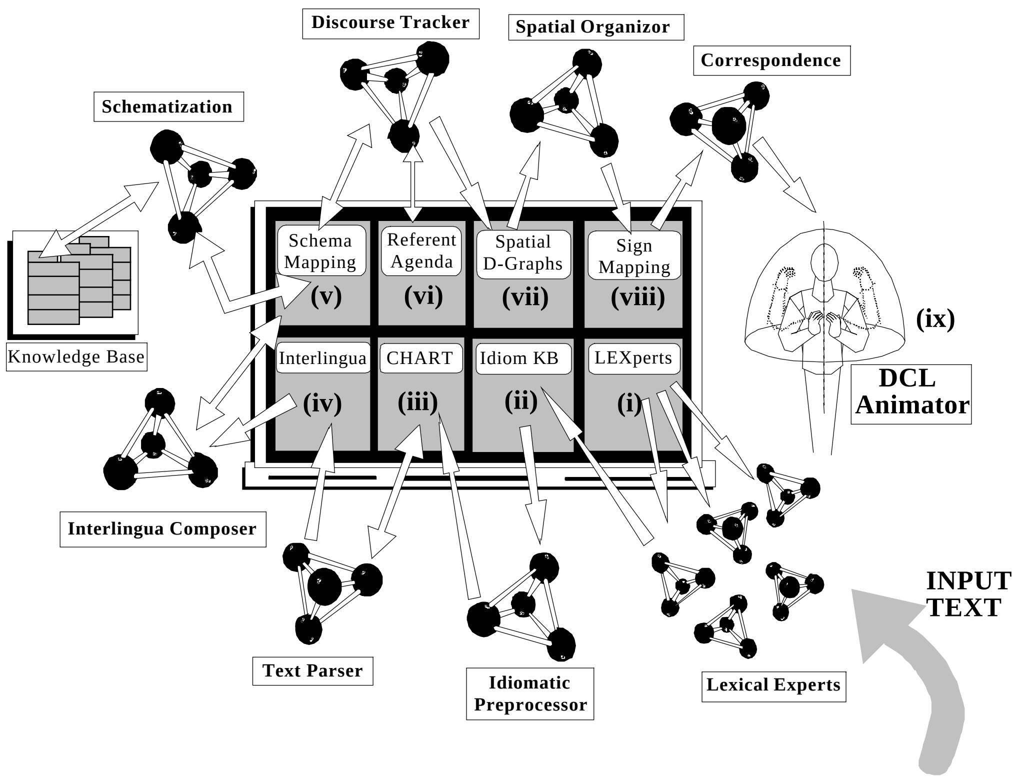


Figure 1: The ZARDOZ Blackboard Architecture: A communication medium for diverse knowledge agencies.

Taking a clockwise tour around Figure 1, system operation proceeds as follows: (i) the incoming text stream is processed by a swarm of *Lexperts*—lexical experts in the sense of Adriaens & Small (1988) specified as autonomous demons—which individually implement both morphological rules and heuristic measures for recognizing and representing compound word constructs in terms of known concepts (e.g., “Bellyache” can be decomposed into the concepts `Stomach + Pain` and, at a later stage, signed accordingly, by analogy with `Headache`, which is known to the system as an expression of pain that is articulated in the head region; see Liddell and Johnson 1986). The digested text then undergoes (ii) *idiomatic reduction*, where known idioms are replaced with equivalent phrases more amenable to compositional analysis, before it is subsequently (iii) *parsed* (the parse agency employs a PATR-based unification grammar; see Shieber 1986) to produce a deep syntactic / semantic representation. From the resultant unification structure a first-cut Interlingua representation is then (iv) composed into an *interlingual frame* format (in a fashion described in Cunningham & Veale 1991); however, before this representation can be considered truly language-independent, metaphoric and metonymic structures specific to the source language are removed by a process of (v) *schematization* (a chunking process described in the next section). The interlingua representation proper provides grist for the (vi) *discourse tracking* agency (anaphoric resolution is a sensitive issue even in sign MT, as will be discussed in section six), before being passed to (vii) the *sign syntax* agency, which employs a robust scheme of Spatial Dependency (SD) graphs (described in section five; see also Veale & Conway 1994) to generate the linear order of the gestural translation, and (viii) the *sign mapping* agency, which employs direct lookup or a variety of heuristic measures to assign concept-to-sign correspondences to the tokens that comprise the interlingua structure (described in section six; see also Veale & Collins 1996).

The syntax and mapping agencies are responsible for transducing the interlingua structure into a flat output stream of sign tokens, which eventually forms the compilation basis for a *Doll Control Language* (DCL) program. A DCL program, when executed, manipulates an on-screen animated doll, causing the correct gesture sequence to be articulated graphically to the user by (ix) a *DCL animator* (see Holden and Roy (1992) for a discussion of the main issues in sign language animation; the particular mechanics of sign animation in Zardoz are described in Conway & Veale 1994).

4. Interlingual Representation

To ensure maximal decoupling of the input languages (e.g., English, Japanese) from the output sign variants (ISL, ASL, or JSL), Zardoz eschews the *Transfer* approach (originated in Yngve (1957) and more recently advocated in a sign translation context by Lee & Kunii 1992) in favor of the *Interlingua* methodology (originated in Weaver (1955), and more recently employed for sign-language MT purposes by Mitamura *et al*. 1991), which places a language-independent interface between source and target. Broadly speaking, an interlingua may capture the generic fact-stating capacity of language using two quite different strategies: the first attempts to construct a *universal grammar* that generalizes over the semantic nuances of many languages, while the second attempts to model the world directly. This second strategy is knowledge intensive, but allows for the incorporation of heterogeneous common-sense inference processes into the translation process.

The English-to-ASL translation system of Patten and Hartigan (1993) employs an interlingua closer in spirit to the first strategy above. However, as the Zardoz architecture is built upon the foundations of the TWIG knowledge-based text-understanding system (described in Veale and Cunningham 1992), we opt for the knowledge-based approach as our methodology of choice on practical design grounds. The present approach therefore emphasizes the representation of content over form, albeit with some concessions to surface style. In theory, an ideal representation of meaning will capture the meaning of any nuance of surface style that affect the semantics of interpretation at the receiving end of a communicative act. In practice however, it is virtually impossible to separate form from content, for the expressive style of an utterance often contributes nuances of meaning which are not easily captured by a strictly compositional representation. We prefer instead to mark the semantic/conceptual representations created by Zardoz with indications of the original surface form from which these representations were derived, and attempt to replicate these surface features in the target whenever possible. Should Zardoz fail to capture all the meaning of an utterance then, this compromise may nevertheless help to communicate those nuances that would otherwise be lost.

4.1. Schematization and Conceptual Representation

The first-cut Zardoz representation of an utterance is derived compositionally from stored lexeme-to-concept correspondences. However, as different languages employ a multitude of conventional metonymies and metaphors, these cultural conventions must subsequently be *spirited away* to achieve a truly neutral interlingual representation. This situation is illustrated in Figure 2, which demonstrates the use of the core English metaphor `Possession-as-Abstract-State` (see also Veale & Keane 1992 for a computational treatment of this kind of metaphor). The logical form of the utterance “I have a terrible headache” suggests that the interlingua frame instantiation `Have-0` be created, with the concepts `*Speaker*` and `Headache-0` filling the `Possessor` and `Possession` slots respectively. With a first-cut representation in hand, the system can then proceed to locate the most suitable schema set that describes the current situation. Thus, upon finding the schema `Suffer-from-Ailment`, the concepts `*Speaker*` and `Headache-0` are subsequently remapped into the more appropriate slots `Sufferer` and `Ailment`.

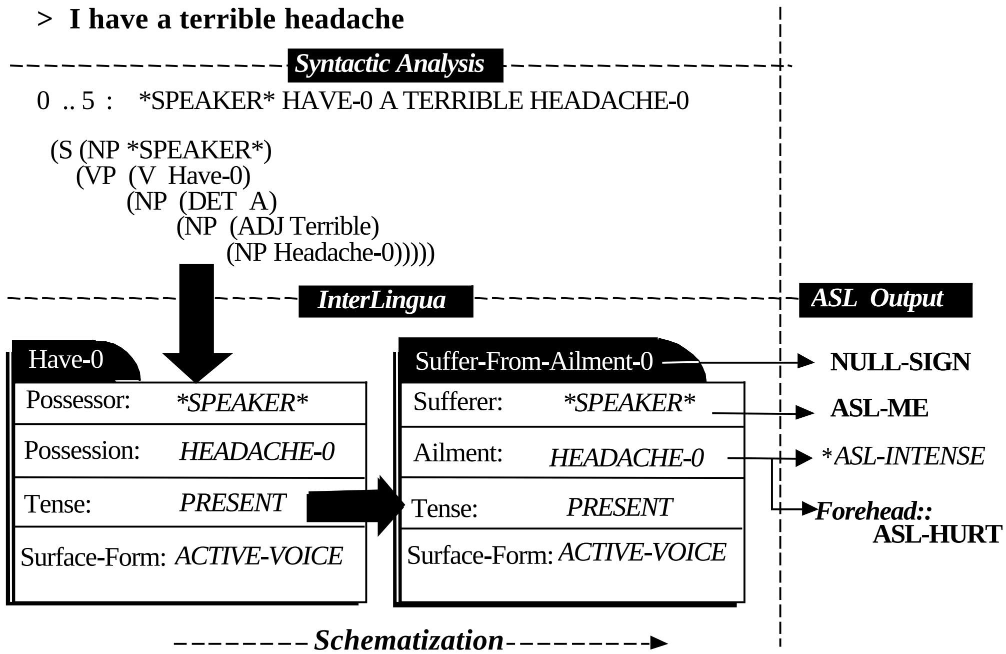


*Figure 2: Sample Syntactic and Interlingual Analysis with ASL output. ASL tokens prefixed with * are sign modifiers, rather than first-class sign gestures.*

Schematization is a *search-and-match* task which employs spreading activation to locate the most apt schema (in this case, activation is spread from the matriarch nodes `Have`, `*Speaker*` and `Headache`). A *preference-based* case representation of each schema is then used (in the fashion of Wilks 1975) as the basis of a *frame subsumption* test to determine which marked schema most suits the situation concerned. The importance of the schematization phase is recognized when one considers that ASL supports a sign for `Have` (possession), but dictates that the sign for `Suffer-From` be elided (thus Figure 2 shows `NULL-SIGN` as a translation): the metaphor simply does not travel to ASL (or indeed to spoken languages like Irish), and must be side-stepped to produce a natural translation. Other common scenarios requiring schematization concern *polymorphic verbs* such as “Paint”; this verb is articulated in ISL as a backward and forward sweeping motion, the wrist swiveling to indicate a brushing motion but remaining fixed to suggest the use of a roller-pad, while the articulation plane of the sign reflects whether it is a wall, ceiling, floor or canvas being painted.

4.2. Surface Annotations

Note the importance of preserving the surface form characteristics of the original utterance when moving to the target language. Although not particularly vital in the headache example of Figure 2, surface form often carries nuances of pragmatic/semantic meaning that would otherwise require a scrupulously detailed and rich interlingua to capture; Zardoz takes the less scrupulous (but more practical) position of annotating the interlingua form with surface details that will later be used when attempting to reexpress the conceptual structure in a target language. Consider for instance an act of marriage: in a system such as Zardoz which employs an uncomplicated frame:slot:filler meaning representation, an identical conceptual representation is given to both “Why did Mary marry Bill?” and “Why did Bill marry Mary?”, two distinct questions which may pragmatically demand quite different responses. In such a compositional representation it is thus necessary to do more than mark the corresponding proposition (e.g., “Bill married Mary”) as a question. In lieu of a complete model of meaning, Zardoz makes an effective compromise in augmenting the representation with a record of syntactic-to-semantic case mappings, e.g., `Groom` → Subject, and to preserve these mappings in generation of the target language. By ensuring that `Bill` (say) occupies the subject position in the target language, Zardoz is able to carry over the semantics of the original utterance cleanly.

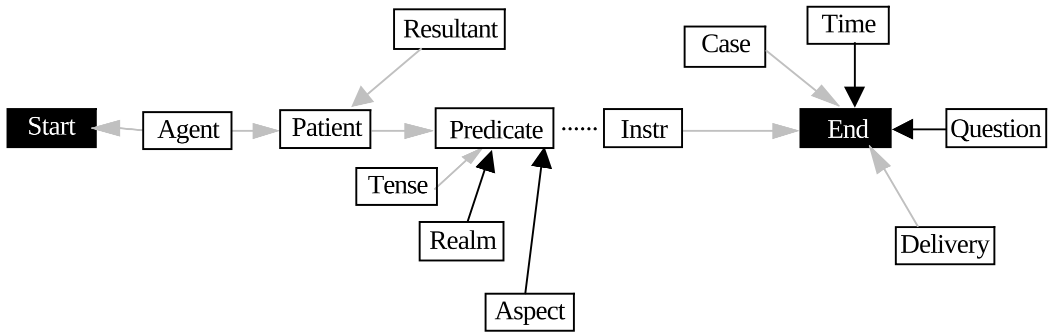
Our discussion of interlingual issues does not end here. We shall return to the issue of metaphor in sign language, and how it must be represented and handled at the conceptual level before a true interlingual state is achieved, in section six. Additionally, the frequent requirement for explicit spatial depiction of events at the interlingual level will also be discussed in section seven.

5. Sign Language Generation: Syntactic Issues

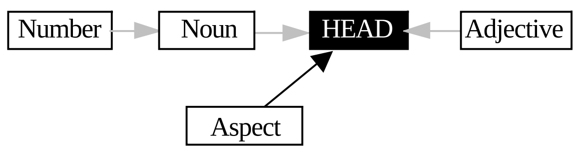
If parsing is a process in which structure is imposed upon a flat input stream, then generation can be viewed as involving a complementary process whereby structure is removed from a meaning representation to produce a flat output stream. The heart of a generation system is essentially a *linearizer* which both selects and orders elements of the meaning representation. The robustness of this linearizer is of particular importance as the decoupling of source and target languages may mean that the interlingua is capable of specifying features not lexically expressible in a given target language.

5.1. Spatial Dependency Graphs

The syntactic framework introduced in this section, that of a *Spatial Dependency Graph,* is designed to score well in the areas of expressiveness, effectiveness and robustness. A spatial dependency graph (or SD-graph) is a partial ordering of case types drawn from a syntactic/semantic case ontology, indicating which elements are to be selected from the interlingua structure, and the relative ordering those elements are to assume in the output stream. Tapping into the case hierarchy affords an SD-graph with greater expressive scope to describe potential syntactic arrangements; for example, the node `Instr`ument in an SD-graph will bind with any member of the INSTRUMENT class, such as `Weapon` (e.g., in verbs like "Hit", "Kill"), `Tool` (e.g., "mend", "build") and `Utensil` (in "Eat", "Cook"). Similarly, since an SD-graph represents a collection of linear-precedence *preferences* rather than hard-constraints, it exhibits considerable robustness in generation (e.g., it always produces *some* total ordering of the input).



(a) Basic Sentence Syntax (SOV, etc.)



(b) Basic Noun Phrase Syntax

Figure 3: Spatial Dependency Graph Representation of core ASL Sign Syntax. (a) depicts the default ASL sentence syntax, while (b) depicts the default ASL noun-phrase syntax. (Key: left to right arrows indicate Before; right to left arrows indicate After; vertical arrows indicate Same Position As; Black arrows indicate Closer Proximity than gray arrows; Gray nodes indicate Sign Literals as opposed to constituent types, while black nodes represent the fixed points of the graph).

An SD-graph represents a syntactic context, or structural preference (such as SVO versus SOV), rather than a strict rule of grammar; in effect, an SD-graph comprises a collection of soft constraints folded in together. Figure 3 depicts the SD-graph representation of the default, or core, syntax of ASL, using a pictorial form for expository purposes: 3(a) depicts the core ASL sentential context, which assumes a topicalized agent in the indicative mood, while 3(b) depicts the core ASL noun-phrase context. These graphs represent complete, or *stand-alone*, syntactic contexts, inasmuch as they are capable of transforming (i.e., *linearizing*) an interlingual frame without recourse to additional syntactic information. An SD-graph is thus a collection of spatial constraints for ordering the elements of an interlingua frame structure. In Figure 3, gray arrows illustrate weak constraints of simple before and after ordering, while black lines illustrate strong constraints that also demand close proximity of constituents. Following the constraints to Figure 3(a) then, the linearizer will place the occupants of the `Agent` and `Aspect` cases *before* the verb/predicate in the output, but will also ensure that the `Aspect` *follows* the `Agent` and directly precedes the verb.

While the SD-graphs of Figure 3 represent *stand-alone* syntactic contexts, those of Figure 4 represent partial *augmentations* which can only be applied relative to the core syntax of Figure 3.

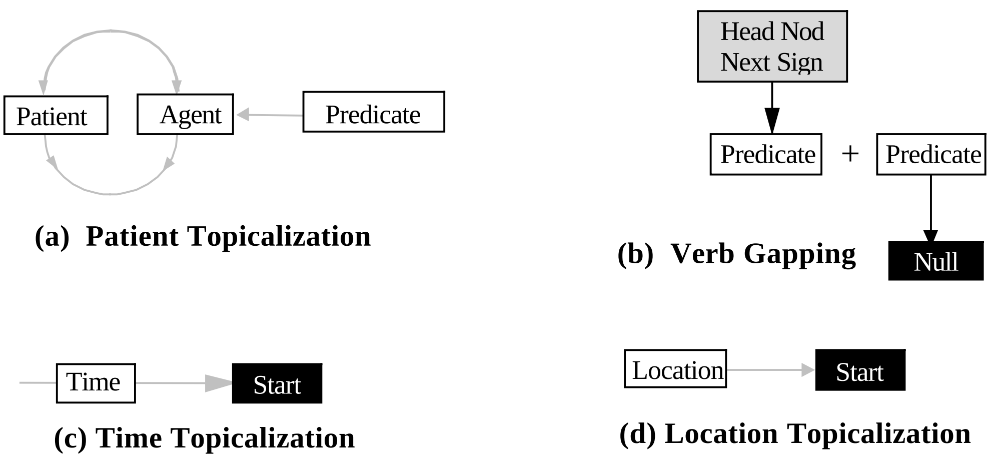


Figure 4: SD-Graphs representing augmentations to the core syntax of Figure 3(a).

There should exist such a partial graph for each syntactic variation of surface form the system wishes to capture in the target language; for instance, Figure 4 illustrates those graphs responsible for the patient topicalization, time and location topicalization, and verb gapping in ASL (the exact nature of these transformations is described next in section 5.2). These augmentations to the core syntax are triggered by *style markers* carried in the `Surface-Form` slot of each interlingual frame (as shown earlier in Figure 2). When linearizing the contents of an interlingual structure, the style-markers stored in each frame are thus used to suggest further SD augmentations to the core syntax graph, which combine to generate a new SD-graph tailored to the frame at hand.

Augmentation graphs are merged with a core graph by pooling the spatial constraints of each into a new graph, in such a way that the constraints of the augmentations take precedence over the core. The constraints of the new graph (core + augmentations) are then instantiated with the contents of the current frame, and resolved relative to the fixed nodes `Start, End, Head` (for noun phrases) and `Null` to produce the final linearized ordering.

5.2 Word Order in Sign

Whereas spoken languages like English must suffice with temporal, left-to-right ordering to impose structure, sign language has the capacity to employ not only temporality but all three spatial dimensions. Thus word order, while being the dominant syntactic constraint in English, assumes a significantly reduced role in sign (see Liddell 1980).

In ASL, for example, major case bindings (e.g., `Agent` and `Patient`) are often established not by sign order, but by an indicative sweep of the signing hand(s) as the matrix verb is articulated. If, for example, `Bill` is signed on the left (with the left hand, perhaps), and `Mary` on the right, then a left to right sweep while signing the verb, `ASL-Chase` say, indicates `Bill` as the pursuer and `Mary` as the pursued, while a right to left sweep indicates the reverse. Thus patient topicalization is realized in ASL by simply reversing the order of the agent and patient constituents. Of course, the verb/predicate will now have to be signed *after* both agent and patient have been articulated, as the sign ordering may now go against the sweeping motion of the verb (e.g., the patient is signed first on the *right*, the agent next on the *left*, then the verb with a *left* to *right* sweep). The SD-graph for this transformation is presented in Figure 4(a).

Case topicalization is similarly handled in ASL. SD-graphs 4(c) and 4(d) ensure the movement of the topicalized case to the left (or first) of the target translation by constraining the topicalized case to appear before the `Start` node of the output stream (for the frame being linearized). This constraint then overrides those of the core syntax presented in Figure 3(a).

Finally, verb gapping is achieved in ASL by dropping the verb sign (naturally), and nodding while the object of the gapped verb is signed. Thus in translating “Bill prefers pizza, Tom hamburgers and Harry hot-dogs”, the signer nods while articulating the signs for `Pizza`, `Hamburger` and `Hot-Dog`. SD-graph 4(b) performs this transformation by mapping the verb onto the NULL position in the output stream (this node acts rather like */dev/null* in UNIX), thus deleting it, while inserting a sign literal *Head-Nod-on-Next-Sign* to take the verb’s place.

5.3 Content-Dependent Syntactic Contexts

The graphs of Figures 3 and 4 represent *content-independent* syntactic contexts, inasmuch as they are applicable to an interlingua frame regardless of its conceptual content. To produce a natural translation into the target language, however, it is often necessary to employ *content-dependent* syntactic contexts, that is to say, SD-graphs which are triggered by particular content-bearing tokens of the interlingua structure. Such a context is depicted in Figure 5.

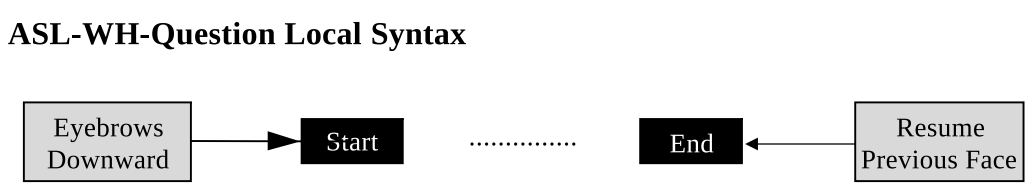


Figure 5: The Local Syntax SD-Graph for WH-Question words in ASL.

For example, WH-question words in ASL require that the signer maintain an eyebrows-downward facial pose for the duration of the interrogative context (see Liddell 1980). Thus whenever the target translation contains the sign `ASL-Why`, `ASL-Where`, or any member of the sign class `ASL-WH-Question`, the non-manual feature indicators *Eyebrows-Downward* and *Resume-Previous-Face* must also be present. The content-dependent context of Figure 5 is therefore associated with the sign token `ASL-WH-QUESTION` in the sign hierarchy, where it can be inherited by all ASL wh-question words. These dependent contexts are simply merged into the SD-graph for a particular frame, where they can augment the output with the sign-specific caveats necessary to produce a natural translation.

5.4. Robust Parsing and Mode-Interleaved Sign Generation

Perhaps the most significant distinction one can make regarding the manner in which sign is exploited by different communities lies between what is termed *native* sign and what is termed *borrowed* sign. Native sign language usually evolves within a deaf community, over a number of years, in much the same way that spoken languages evolve in hearing communities. Native sign is thus sign at its most natural and unfettered, and is the obvious manifestation employed by deaf signers. In contrast, another form of sign usage, termed *coded* or *borrowed* sign language, is not a natural outgrowth of deaf sign usage, rather a gestural coding of an existing spoken language. For example, Signed Exact English (SEE) is a variant of sign language which employs standard English syntax, expressed not in standard English morphemes, but in gestures (either individual signs or finger-spellings of English words and suffixes). While native sign is the dominant form for communication between deaf signers, coded sign is most often used for educational purposes (where hearing signers are involved), and for such ends as signed news summaries on television. As a result, most native signers are comfortable with both manifestations of sign, and encounter little difficulty in segueing between each.

This ability affords an MT system such as Zardoz an increased level of robustness, inasmuch as it provides a base-level performance that can be guaranteed by the system. Should the source parser be unable to generate a full syntactic structure that spans the entire input utterance, the system is still in a position to produce a full output representation, interleaving both native and coded signing strategies. This situation is somewhat analogous to the use of multi-engine MT systems in which a poor quality yet robust engine is used as a fallback when a higher-quality yet brittle engine fails. In Zardoz, the native sign component is thus invoked for those fragments of the input which produce case-frame interlingual representations, while the coded sign component is invoked for the troublesome link words which cause the parser to fail. This interleaving strategy is essentially a form of *artificial code-switching*, allowing the system to segue between native and coded sign in a manner which follows the constituent ordering dictated by the input utterance. This fallback position obviates the need for Zardoz to complicate its parsing process with word-skipping heuristics (e.g., see Yamada 1996) or other, chart-based techniques for parsing ill-formed input (e.g., see Mellish 1989).

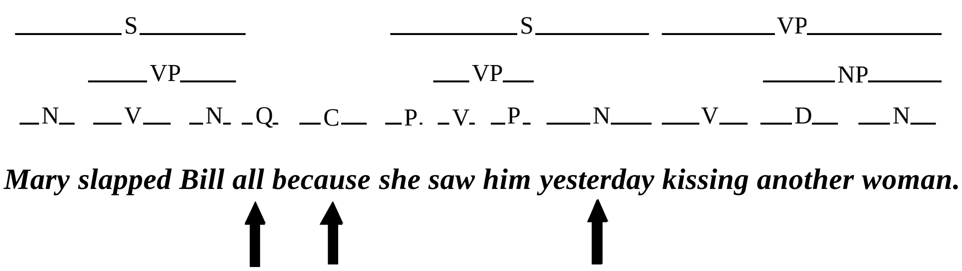


Figure 6: A fragmented parse of an input utterance; Labeled horizontal lines represent the extents of inactive chart edges, while vertical arrows indicate the troublesome constituents which blocked a global parse.

The example sentence of Figure 6, "Mary slapped Bill all because she saw him yesterday kissing another woman", provides just such a case where mode interleaving proves useful. The parser is thrown by the speaker's use of the colloquial form "all because" which is a hole in the system's idiomatic knowledge. The parser also runs aground when it encounters the temporal adjunct "yesterday" leading into a verb phrase, a construct not anticipated by a grammar designer who naïvely expects such adjuncts to always precede or follow sentential clauses. Nevertheless, an interlingual frame representation is still created for each of the verb phrase/sentence parse fragments, and a target native-sign translation is produced for each. These fragmentary outputs are then *glued together* by the mode interleaver, preserving the constituent order observed in the original utterance, in combination with individual sign translations for the trouble zones "all", "because" and "yesterday", which then act as *glue signs* for the overall translation.

6. Sign Lookup and Creation

After linearization, the next stage in the generation process is the assignment of language-specific signs to the concept tokens of the interlingua. When available, Zardoz exploits pre-stored sign→word correspondences which are retrieved via direct lookup, but in many cases the system is required to demonstrate some measure of creativity in mapping concepts to gestures.

A system like Zardoz has the ultimate fallback position of finger-spelling those words it cannot translate. However, not only does this presuppose a familiarity with the source language on the part of the addressee, it becomes unduly cumbersome when overused. Nonetheless, though this unwieldiness is necessary when first introducing proper names into a narrative, it can subsequently be dropped in favor of a spatial designation, a finger-spelling of the first letter, or both. Thus, if Japan is first signed `Left-hand::[ASL-J, ASL-A, ASL-P, ASL-A, ASL-N]`, it can later be referred to simply as `Left-hand::ASL-J`, or perhaps even left implicit in the signing of the governing verb (as described in section 7).

Zardoz also exploits the concept hierarchy to mimic the basic inventiveness of native signers and derive new sign correspondences as they are needed. For instance, if the system lacks an ASL mapping for `Aspirin`, it is an easy matter to create, *on-the-fly*, the gestural concatenation `ASL-A + ASL-MEDICINE`, as distinct to that for `Tylenol`, which becomes `ASL-T + ASL-MEDICINE`, and so on. While perhaps not the preferred native sign, this is nevertheless a preferable solution to finger spelling or simplistic code-switching, and one that exercises the idiom of the target language.

More specific spatial knowledge is employed whenever it is necessary to inflect the base concept with something more than a discriminatory first letter, as when the base concept specifies a body location at which the derived sign is to be articulated. This arises when signing concepts such as Headache and Backpain, which are iconically associated with the specific body locales Forehead and Lower-Spine. The ASL sign sequence for Headache is thus Forehead::ASL-Hurt (where ASL-Hurt is inherited from Pain), while that for Backpain is Lower-Spine::ASL-Hurt.

6.1. Exploiting Structural Metonymy for Sign Creation

However, the most effective heuristic for sign creation employs an *interlingual gloss* that captures the broad meaning of a concept. Such glosses can be pre-specified by the lexicon designer, when a sign informant is unavailable to provide the native form, or inferred dynamically by the system at run-time. For example, the designer may specify a gloss for Pill as [MOUTH MEDICINE], or for Sandwich as [Food Inside Bread]. Thus the ASL mapping for Pill is heuristically defined to be Mouth::ASL-Medicine. If no gloss is available at run-time, the underlying Krell frame manager is called upon to provide one. Krell does this by examining the frame:slot:filler structure of the concept involved to determine a sequence of appropriate metonyms which have a known sign articulation. The gloss for Ham, for example, is [Pig Meat], as Ham appears in the Meat slot of Pig. The ASL mapping of Ham-Sandwich is thus ASL-Pig + ASL-Meat (which subsumes ASL-Food) + ASL-Inside + ASL-Bread.

6.2. Exploiting Spatial Metaphor for Sign Creation

Spatial metaphor is frequently argued to provide a descriptive framework expressive enough to describe many of the conceptual structures underlying everyday language (see Lakoff & Johnson 1980; Veale & Keane 1992a, 1992b). Naturally, this argument is applicable not only to spoken/written language, but also to other modalities of expression, such as sign language. It should not be surprising then to find that sign languages, which employ space not only as a conceptualization framework but also as an expressive medium, is steeped in highly productive and coherent spatial metaphor.

The Conceptual Scaffolding model of metaphor, proposed by Veale & Keane (1992), is a skeletal meaning representation itself built upon spatial metaphor; the rationale for such an approach is provided by the work of Lakoff & Johnson, who argue that conceptual structures must be experientially grounded in physical reality. This in turn follows the empiricist tradition which claims that our linguistic/conceptual map of the world is acquired via sensory experience, and is thus structured in those terms (see Lyons 1977). This would suggest that spatial metaphor, combining a physical origin with an abstract descriptive power, provides both the physical experience, and the conceptual framework, upon which to base a general model of meaning.

The Scaffolding model specifies a spatial calculus which is defined upon the metaphor constructors `Up`, `Down`, `Connect` and `Disconnect`. These constructors are posited as cognitively-real *building blocks* of meaning, from which the semantics of many everyday concepts—both concrete and abstract—may be composed. The `Up` and `Down` constructors model the fundamental orientation metaphor, as exhibited in such conventional metaphors as “Food prices *soared*”, “IBM *fell* into a *slump*” and “The market *rose* out of a *depression*”, while the `Connect` and `Disconnect` constructors similarly model the fundamental connection metaphor. The association of two ideas/concepts is seen as conceptual *connection*, while the disassociation of ideas is seen as conceptual *disconnection*. Our previous work has demonstrated the connection/disconnection metaphor at work in social concepts such as FRIENDSHIP, MARRIAGE, DIVORCE, as well as at a more abstract level, in corporate relations such as company mergers and rivalries; we now argue that this metaphor, in conjunction with the orientation metaphor, can be highly productive in the creation of new signs.

Empirical evidence for the cognitive reality of these spatial constructors is found in Japanese Sign Language (JSL), which seems to exploit the scaffolding philosophy in a regular and coherent manner. Consider for example Figure 7, which presents a representative montage of spatial metaphor in JSL. Employing the *classifier handshapes* of 7(a), which are essentially a type of *semantic anaphor* or *class restriction*, the scaffolding constructors `Up, Down, Connect` & `Disconnect` are used to construct the meanings of signs 7(b) ... 7(n). Given a sign language which supports a rich class of classifier handshapes (and most sign languages do, such as ASL and ISL), and a knowledge-base specified around the spatial semantics of the scaffolding model (such as that of Cunningham & Veale 1991), then signs (b) ... (k) represent simple applications of this spatial knowledge. Note how the classifier handshapes of 7(a) are used in signs (b) ... (k) almost as slot fillers in a conceptual schema (or arguments to a polymorphic function). The ability to insert different classifiers into the same spatial articulations thus allows Zardoz to dynamically create new, metaphor-grounded, signs on the fly if a gestural correspondence does not already exist in the target sign-language.

The scaffolding philosophy holds, as a defining claim, that because these metaphor constructors are derived from shared experience of the world, they are cognitively realized (in some form) in most cultures, and thus any reasonably uncomplicated composition of constructors should be readily interpretable by the end-user.

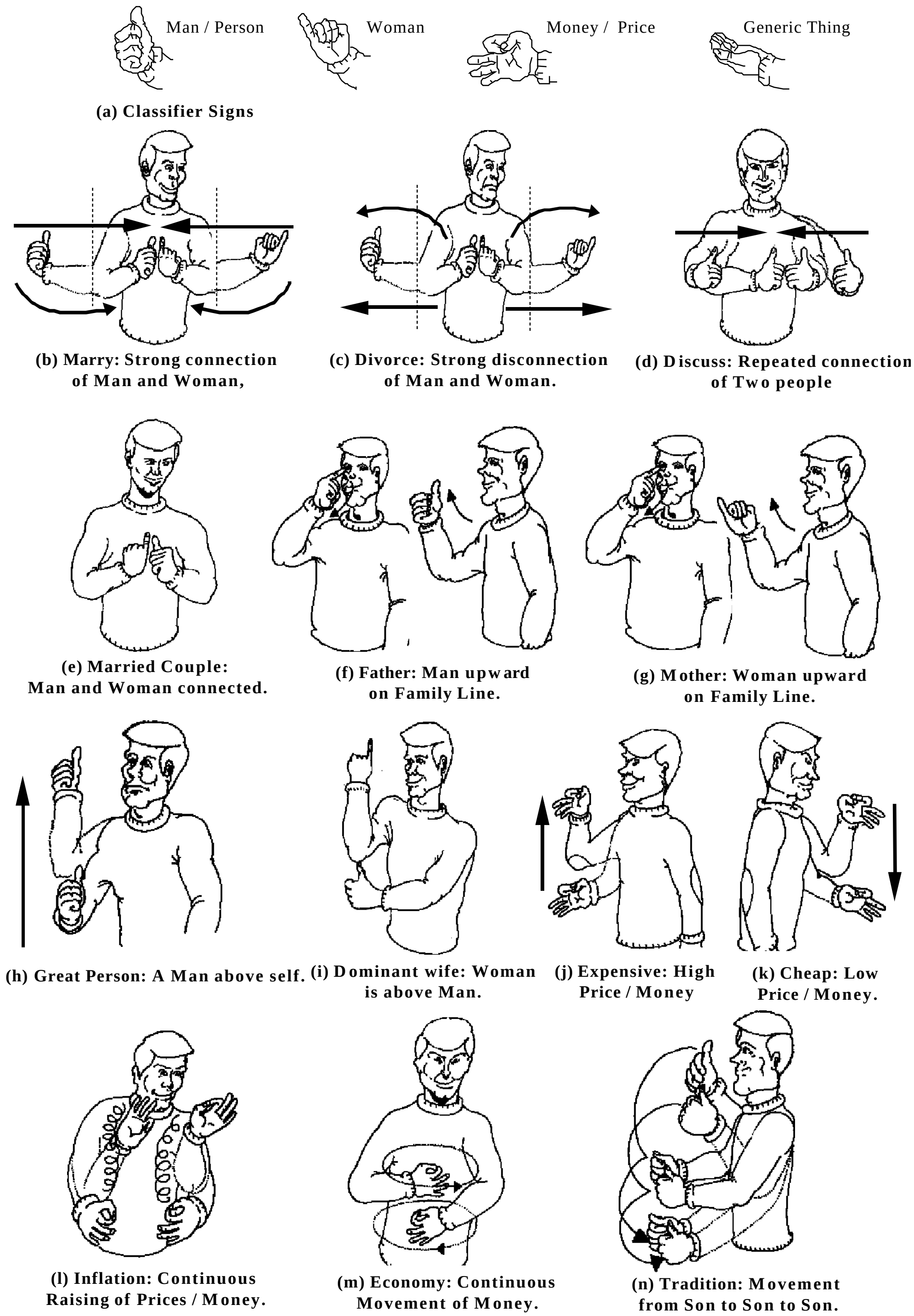


Figure 7: A montage of spatial metaphors in JSL. Part (a) presents the classifier handshapes that are coherently exploited in signs (b) ... (n).

This view is pragmatically attractive despite the suggestion that there exist subtle cultural differences between hearing and non-hearing language users, which is reflected in their use of different metaphor models of the world. For instance, Grushkin (1995) argues that non-hearing ASL users conceptualize anger differently than hearing speakers of English; namely, ASL users conceptualize anger as a mental condition, and thus rely heavily on the sign `ASL-Mind`, while native English speakers have a more visceral model of anger (e.g., *"You turn my stomach*!") However, because Zardoz limits its assumptions of near universality to spatial domains, we expect that dynamic signs based on spatial metaphor should nevertheless be relatively portable between both signed and verbal languages, though these assumptions have yet to be given strong empirical validation in our work.

For example, a sign language that provides a classifier for `Company` or `Institution` will thus support a metaphoric definition of `Employee` as the connection of `Company` and `Person`, effectively one who is married to the company, in the fashion of 7(b). Likewise, a corporate merger might be metaphorically articulated as the connection of two companies, whilst corporate rivalry can be signed as a disconnection of companies—a corporate divorce in effect (one is reminded of IBM and Microsoft)—in the fashion of 7(c).

More complex examples of localist meaning are depicted in 7(l) ... (n), which illustrate how composite, or *aspectually inflected*, forms of the scaffolding constructors are used to represent more abstract concepts. As further evidence of the claim argued in Veale and Keane (1992a,b)—that the `Abstract State-Change as Movement` metaphor schema can be exploited to structure a wealth of diverse verbs—Figures 7(l)...(n) demonstrate that when organized around spatial underpinnings, abstract concepts such as `Tradition` can be communicated as the successive concatenation of other, less abstract concepts, here the `Father/Daughter` metaphors of 7(f)/(g).

7. Dimensionality and Spatial Depiction in Sign Language

Space is exploited in sign language in two distinct fashions. The first, and most obvious usage of space is as an expressive medium in which to articulate different concepts — just as sound is exploited by spoken language to shape and combine phonetic structures, space lies at the phonological heart of sign language. The conventional metaphor of the time-line, for example, is often employed to convey temporal concepts such as past and future tense in spatial terms (see Klima and Bellugi 1979). Spatial nuances are also applied during articulation to express different sign inflections and aspectual modifications (such as *Continuous*, *Resultative*, etc.), and to conflate adjectival descriptors into their associated noun gestures (for instance, a "wide road" is not articulated as two successive signs in ISL, but as one sign, "road", where the interactions of the hands are *widened* to convey broadness). The second usage of space follows from the visual qualities of sign language, in which descriptions of spatial scenarios are mirrored in a *re-constructive* fashion by the signer. Language is frequently used to describe spatial relations between entities (such as "The car park is to the left of the department store"), but verbal languages such as English often leave much of the spatial reasoning inherent in a statement implicitly coded, placing the onus on the hearer to mentally reconstruct the given spatial situation. Sign language, however, in its capacity to exploit all three spatial dimensions, is used to convey such spatial reasoning explicitly (for example, the signer will literally articulate `Car-Park` to the left of `Dept-Store`).

From an MT perspective then, natural generation of sign can be a much more complex task than that for verbal output, as a translation system must actually apply some spatial common-sense to *understand* the situation being conveyed. This need for extra spatial processing in sign-MT systems mirrors the experimentally-determined belief (see Emmorey, 1995) that because of the increased spatial demands of their language, native sign users have heightened powers of spatial reasoning.

7.1. Sign Space

Zardoz partitions the sign-space of the virtual signer/doll into several discrete areas; spatial areas of particular importance include those labeled Left-Space, Right-Space, and Middle-Space, where each denote a different area along the *standard signing line* (just below chest height) of the signer. Other areas of importance include Upper-Left-Space, Upper-Middle-Space and Upper-Right-Space, which lie along the *upper signing line* above shoulder height, and Lower-Left-Space, Lower-Middle-Space and Lower-Right-Space which lie along the *lower signing line* below waist height.

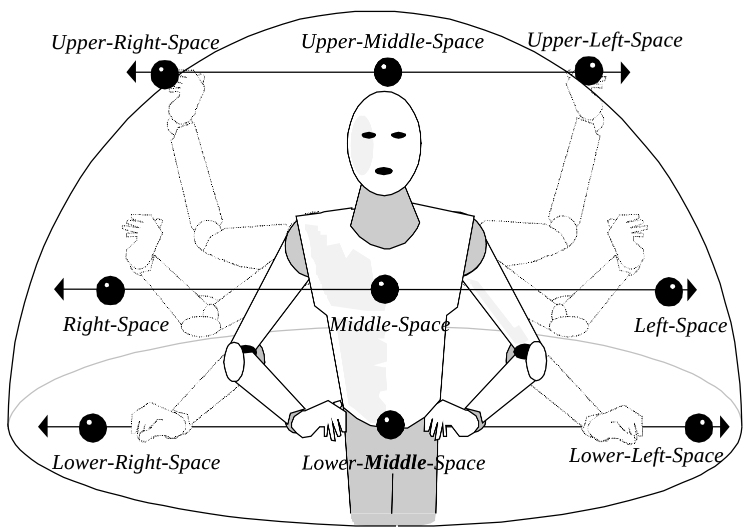


Figure 8: The Signing Space in which a sign is articulated by Zardoz.

The standard line positions are generally allocated by Zardoz as spatial indices to the principle thematic roles of `Agent, Patient` and `Locus` (the focal-point of an action that involves movement (fictive, or otherwise) along a path), while the upper and lower line positions are reserved for entities with explicit orientational associations (e.g., Mountain Summit/Base, concepts which are Abstract/High-Status/Low-Status) and for additional narrative entities which the standard line is full. The organization of these positions in the signer's space (from the signer's perspective) is illustrated in Figure 8.

7.2. Sign Hierarchies

Both *sign space* and *world space* are modeled in Zardoz using the same representational strategy, wherein sign concepts are organized around an object-oriented inheritance hierarchy which supports method attachment at different levels of sign specification. Zardoz also employs a representational isomorphism between frames, objects and blackboard-panels, and between demons, methods and knowledge-sources, where each is simply a different perspective upon the same underlying representation. The knowledge-base thus becomes its own control architecture, as the blackboard and concept hierarchy are cut from the same cloth. This allows for maximal integration of knowledge in the system, and allows for a uniform treatment of space in sign generation.

Associated with each frame in this sign-concept hierarchy is one or more DCL code-segments, which when collectively assembled under inheritance, provide the articulatory basis for each sign gesture. Sign inflections are in turn modeled as method-activating messages which are passed to the sign concept under inflection, with the expectation that a local or inherited method can adapt either its DCL segments, or the DCL variables over which these are defined, to induce the correct articulatory behavior (e.g., the stressing or repetition of the sign).

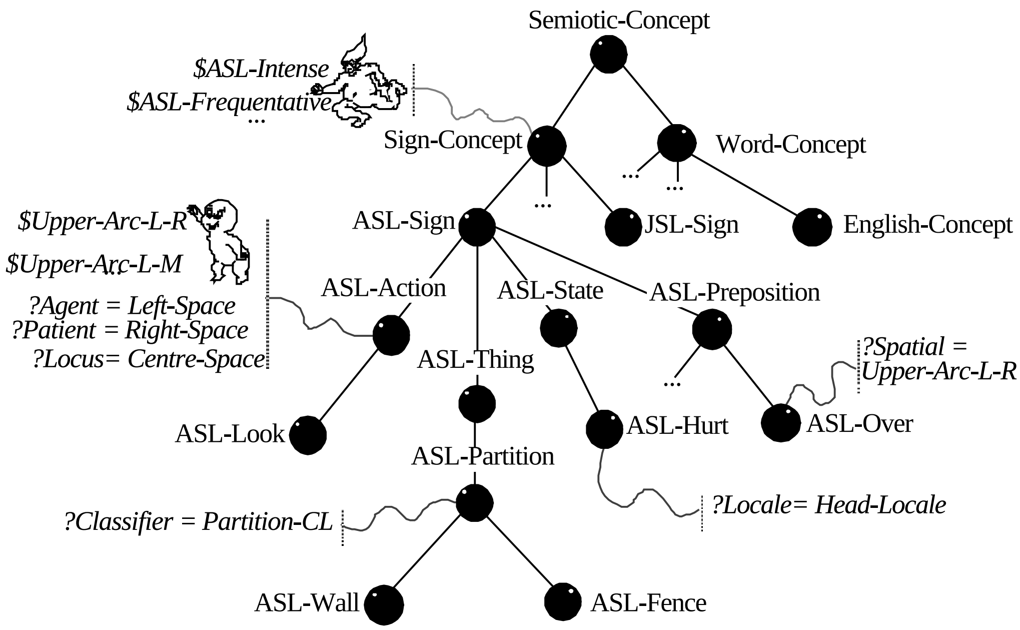


Figure 9: The ZARDOZ Concept Network and Object Hierarchy. Statements prefixed with "?" indicate local DCL variable assignments, while tokens prefixed "$" indicate demon attachments for given message types.

As illustrated in Figure 9 for example, the message `ASL-Intense` is mirrored at the highest point of the sign hierarchy, `Sign-Concept`, by a demon which magnifies the local values of the DCL variables that control the spatial extent of a gesture, `?in`, `?out`, `?left` and `?right`. Provided then that a hyponymic sign, such as `ASL-Hurt`, makes reference to these variables in its DCL specification, this message will induce the correct gestural behavior i.e., the sign will be articulated with broader, more urgent, motions. Similarly, body locales may be employed in sign language as inflectional messages—when the message `Head-Locale` is passed to `ASL-Hurt`, a corresponding demon (again inherited from `Sign-Concept`) modifies the local value of the DCL variable `?locale`, thus ensuring that the sign is articulated at the forehead rather than its default stomach location. Such a sign hierarchy is illustrated in Figure 9.

7.4. Spatial Reasoning and Depiction

The second form of spatial usage in sign language—the explicit modeling of spatial assumptions — is also supported by this sign hierarchy organization. It particular, it is in the allocation of anaphoric spatial reference points that the burden of spatial reasoning is most pressing. Zardoz currently exploits inheritance in the sign hierarchy to consistently allocate such reference points: as illustrated in Figure 9, inherited DCL variables, under the auspices of dedicated spatial demons, provide a default assignment of spatial indices to the participants of a communicated scenario, say the act of looking over a wall. However, the system must also concern itself with speaker viewpoint, for although this is a largely pragmatic issue in verbal language, changes in viewpoint can have considerable effects on the articulation of a sentence in sign. For instance, in ASL, the verb "Enter" can be expressed from either of two viewpoints, via the signs `ASL::I-Enter` and `ASL::You-Enter`. In ASL then, the sentence *"There is a chair on the left as one enters"* can be signed either as `ASL::I-Enter Left-Hand:ASL::Chair Left-Hand:ASL::There`, or alternatively as `ASL::You-Enter Right-Hand:ASL::Chair Right-Hand:ASL::There`. Depending on the chosen viewpoint, the system may have to perform a spatial rotation of 180^{o} upon the given spatial indices to maintain consistency of reference throughout (see Emmorey 1995).

We examine a relatively straightforward scenario here, one in which an agent 'Bill' uses binoculars to look over a wall and spy upon a patient, 'Mary', on the other side. This scenario is the basis of the Zardoz trace illustrated in Figure 10.

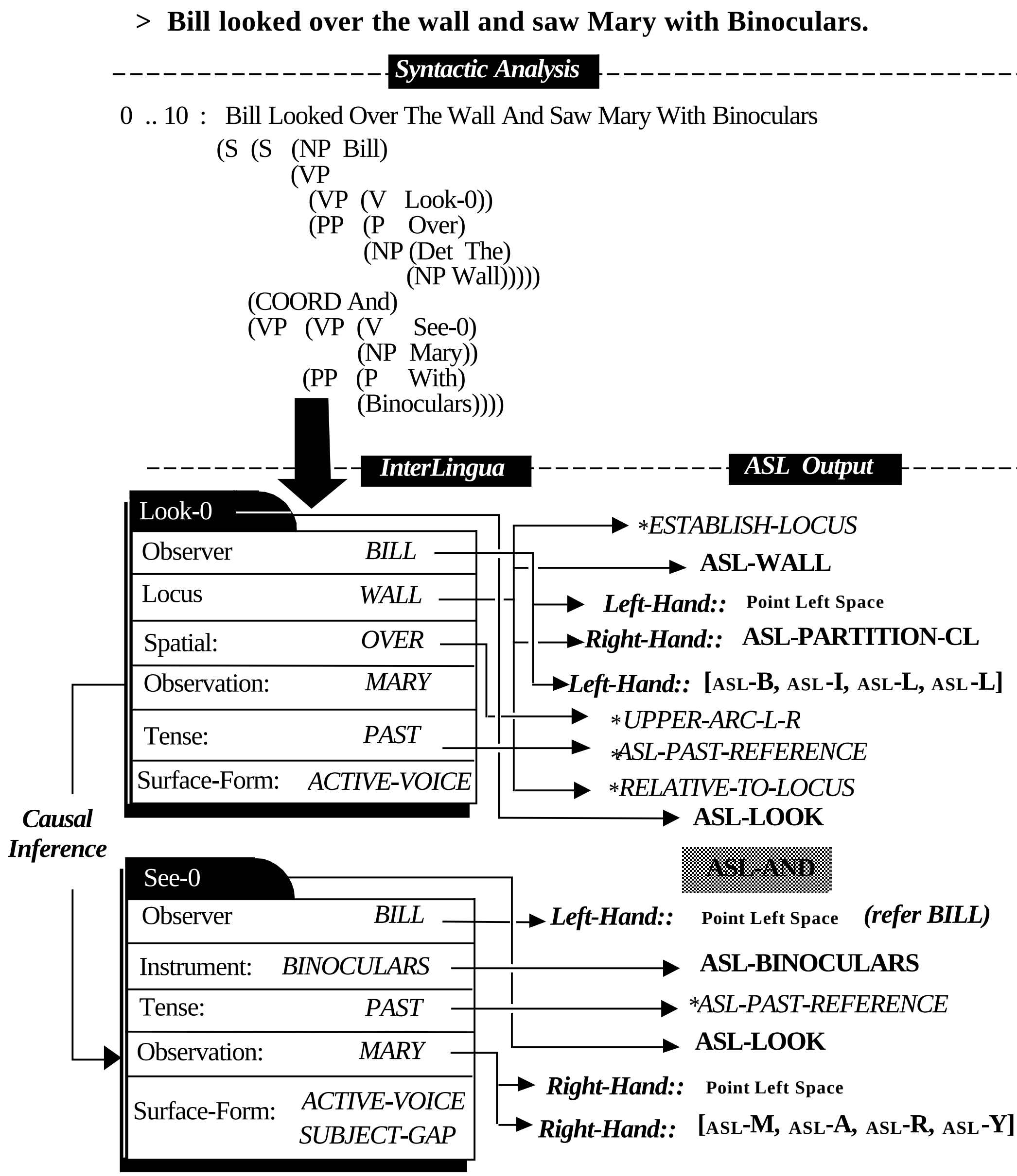


Figure 10: Analysis of a sentence requiring spatial reasoning and causal inference. (CL is here used as a notational shorthand for Classifier).

The first form of pragmatic reasoning demanded in this context requires Zardoz to recognize that the act of looking frequently leads to the act of seeing, and thus, if Mary is the patient/observation of the latter (`See-0`), she is most likely the observation of the former (`Look-0`) also. With both participants thus bound within the same frame, they are allocated, by inheritance (as in Figure 9), the coherent spatial bindings Left-Space (for Bill, the agent) and Right-Space (for Mary, the patient) respectively. Because the concept Wall is recognized to serve the role of locus point for the action Look-0 (and by the previous pragmatic inference, See-0 also), it receives the default spatial index Middle-Space. Thus the entities Bill, Mary and Wall are assigned spatial indices which explicitly convey the implicit spatial organization of the original English input—that Bill and Mary are on opposite sides of the wall.

A reference to the central preposition of the action, “Over”, must also be made in deriving this assignment of indices. In this case, however, the actions of the demon $Upper-Arc-L-R, inherited by the concept ASL-Look and mirrored as a message specifier in the concept ASL-Over, do nothing to alter the default assignment of indices inherited from ASL-Action. In contrast, however, were the sentence under analysis “Bill looked *onto* the wall and saw a squirrel with binoculars”, the demon $Upper-Arc-L-M would be invoked accordingly, causing the local value of ?Patient at ASL-Look to become temporally set to the value Upper-Middle-Space (or more precisely, a spatial index suited to the local assignment of ?Locus). Thus, the sweeping arc of the Look gesture would terminate above the locus position—the hand classifier for Wall, ASL-Partition-CL, held in Middle-Space—while the hand classifier for Squirrel, ASL-Animate, is signed accordingly above this place-holder for Wall.

8. Summary and Conclusions

We conclude by reiterating our commitment to an interlingua methodology, though the grounds for this commitment are perhaps more pragmatic than theoretical. The knowledge-intensive interlingua approach, where Zardoz attempts to model the world directly, has been pursued for two main reasons: (a) to ensure maximal decoupling of source and target languages, as Zardoz is intended to possess competence is several different sign languages, and possibly even multiple input languages (Japanese is currently being investigated); and (b), an interlingual stage of representation allows the system to bring common-sense inference to bear upon the translation process. This latter benefit is in many ways a necessary one, as sign generation requires a level of understanding and spatial reasoning that is clearly outside the realm of traditional linguistic analysis, demanding instead an A.I. knowledge-based comprehension system. As the Zardoz blackboard shell, object hierarchy and demonology are all woven from the Krell frame system, message-passing between heterogeneous agencies is supported, ensuring the optimum integration of linguistic, conceptual and pragmatic knowledge.

Finally, a commercial rationale for using an interlingua springs from the difficulties of scalability a system such as Zardoz must inevitably experience, since sign lexicography presents a more significant challenge than verbal lexicography. It is expected that systems such as Zardoz will operate best in restricted domains for which the lexical/gestural acquisition bottleneck is considerably less vexing, and where its common-sense rules of inference and spatial awareness can best be utilized. Such domains as weather reports, and other well-circumscribed news topics (financial updates perhaps), seem most accommodating of interlingua-like technology, as demonstrated by the *Météo* system[1] of Chandioux (1976/1989). The integration of text and sign in Zardoz via a DCL script layer makes such applications particularly attractive, as sign translations can be dispatched by email to subscribers (small television stations, for instance, who don't possess a budget for a sign-language weather reader) as an ASCII script which can be animated locally using a stand-alone graphics package. The future of serious sign-MT research may well lie in the commercial viability of such on-line sign-MT bureaus.

8.1. Future Potential

[1] Though *Météo* is not a genuine interlingua system in any language-universal sense, it nevertheless clearly demonstrates the benefits accrued from restricted domain MT, the kind of task to which interlingua systems are most well suited.

In concluding, one should also not lose sight of the multimedia potential of this cross-modal technology. A language-configurable sign translation system with gesture articulation on a graphical doll display (as described in Conway & Veale 1994) overcomes the many limitations of using spliced video footage for sign generation, and will support native sign interfaces in applications as diverse as automated information-points, tutorial systems, sign-email, sign-teletext, and any interface where gestural communication is advantageous. Such techniques will have ready application outside the domain of sign language itself, inasmuch as a great deal of human extra-linguistic communication relies upon such gestural systems, or body-languages, for expression.

References

Adriaens, G. & S. L. Small. (1988). Word Expert Parsing Revisited in a Cognitive Science Perspective*. Lexical Ambiguity, S. L . Small, G. W. Cottrell & M. K. Tanenhaus (eds.),* San Mateo, CA: Morgan Kaufmann.

Chandioux, J. (1976). MÉTÉO: un système opérationnel pour la traduction automatique des bulletins météorologiques destinés au grand public, *Meta* **21**, pp 127-133.

Chandioux, J. (1989). Météo: 100 million words later. *In D.L. Hammond (ed.) American Translators Association Conference 1989.* Medford, NJ: Learned Information.

Conway, A. & T. Veale. (1994). A Linguistic Approach to Sign Language Synthesis, *in the proceedings of HCI'94, the Human Computer Interface Conference, Glasgow.*

Conway, A. & T. Veale. (1995). Building Signs: Representing Space and Structure in Automatic Sign Synthesis, *in the proceedings of ICLC'95, the 5th International Cognitive Linguistics Conference,* Albuquerque, New Mexico, July 1995.

Cunningham, P. & T. Veale. (1991). Organizational issues arising from the integration of the

Concept Network & Lexicon in a Text Understanding System, in *the Proceedings of the 12th International Joint Conference on Artificial Intelligence*. Morgan Kaufmann.

Emmorey, K. (1995).Interactions between Processing Spatial Information in Linguistic and NonLinguistic Domains, *in the proceedings of ICLC'95, the 5th International Cognitive Linguistics Conference,* Albuquerque, New Mexico, July 1995.

Grushkin, D. (1995). Metaphorical expressions of anger in ASL as a window on the culture of the American Deaf community, *in the proceedings of ICLC'95, the 5th International Cognitive Linguistics Conference,* Albuquerque, New Mexico.

Hobbs, J. (1978). Resolving Pronoun References, *Lingua* 44, p 311-338.

Holden, E. J. & G. G. Roy. (1992). The graphical translation of English text into Signed English in the hand sign translator system, *Computer Graphics Forum* **11**(3), pp357-366.

Klima, E. & U. Bellugi. (1979). *The Signs of Language.* Cambridge, MA: Harvard University Press.

Lakoff, G. & M. Johnson. (1980). *Metaphors We Live By*. Illinois: University of Chicago Press.

Lee, J. & T. L. Kunii. (1992). Visual Translation from Native Language to Sign Language, in *the proceedings of the IEEE workshop on Visual Languages,* September 1992, Seattle Washington.

Liddell, S. K. (1980). *American Sign Language Syntax*. Mouton: The Hague.

Liddell, S. K. & R. E. Johnson. (1986). American Sign Language compound formation processes, lexicalization and phonological remnants, Natural Language and Linguistic Theory **4**, pp 445-513.

Lyons, J. (1977). *Semantics*. London: Cambridge University Press.

Mellish, C. S. (1989). Some Chart-based Techniques for Parsing Ill-formed Input, *in the Proceedings of the 27th Annual Meeting of the Association for Computational Linguistics.*

Mitamura, T., E. H. Nyberg and J. G. Carbonell. (1991). An Efficient Interlingua Translation System for Multi-lingual Document Production, in *the proceedings of Machine Translation summit III*, Washington D.C., July 2-4, 1991.

Padden, C. A. & D. M. Perlmutter (1987). American Sign Language and the architecture of phonological theory, *Natural Language and Linguistic Theory* **5**, pp 335-375.

Patten, T. & J. Hartigan. (1993). Automatic Translation of English to American Sign Language, presented at *the 1993 National Conference on Deafness*, Columbus Ohio.

Sandler, W. (1989). *Phonological representation of the sign: Linearity and non-linearity in American Sign Language*. Providence, RI: Foris Publications.

Veale, T. & A. Conway. (1994). Cross-Modal Comprehension in Zardoz, An English to Sign-Language Translation system, presented at the *Fourth International Workshop on Natural Language Generation*, Maine, USA, 1994.

Veale, T. & P. Cunningham. (1992). Competitive Hypothesis Resolution in TWIG: A Blackboard-Driven Text-Understanding System, in *the Proceedings of the 10th European Conference on Artificial Intelligence*, Chichester: John Wiley.

Veale, T. & M. T. Keane. (1992a). Conceptual Scaffolding: Using metaphors to build knowledge structures, in *the Proceedings of the 10th European Conference on Artificial Intelligence,* Chichester: John Wiley.

Veale, T. & M. T. Keane. (1992b). Conceptual Scaffolding: A spatially founded meaning representation for metaphor comprehension, *Computational Intelligence* 8(3), pp 494-519.

Veale, T. & B. Collins. (1996). Space, Schematization & Metaphor in Sign: Sign Language Translation in the Zardoz System *in the proceedings of AMTA'96, The second conference of the Association for Machine Translation in the Americas, Montrèal.*

Veale, T. & B. Smyth. (1992). Krell: Knowledge Representation Entry-Level Language, the User Guide Version 1.0. *Hitachi Dublin Laboratory Technical Report*, HDL-TR-92-051.

Weaver, W. (1949). *Translation*. Reprinted in W. Locke & A. Booth (1955), *Machine Translation of Languages*, Cambridge, MA: MIT Press.

Werhli, E. (1992). The IPS system, *in the proceedings of COLING-92, the fourteenth international conference on Computational Linguistics*, pp 870-874.

Werhli, E. (1996). ITSVOX, System *in the proceedings of AMTA'96, The second conference of the Association for Machine Translation in the Americas, Montrèal.*

Wilks, Y. (1975). A preferential, pattern-Seeking, semantics for natural language inference, *Artificial Intelligence* 6, 53-74.

Yngve, V. (1957). A Framework for Syntactic Translation. *Mechanical Translation* 4(3), pp 59-65.

Yamada, K. (1996). A Controlled Skip Parser, *in the proceedings of AMTA'96, The second conference of the Association for Machine Translation in the Americas, Montrèal.*