Moving Heads and Moving Hands: Developing a Digital Corpus of Irish Sign Language.
The 'Signs of Ireland' Corpus Development Project

Lorraine Leeson
Centre for Deaf Studies, School of Linguistics, Speech and Communication Sciences, Trinity College Dublin
leesonl@tcd.ie

John Saeed
Centre for Language and Communication Studies, School of Linguistics, Speech and Communication Sciences, Trinity College Dublin
jsaeed@tcd.ie

Deirdre Byrne-Dunne
Alison Macduff
Cormac Leonard
Centre for Deaf Studies, School of Linguistics, Speech and Communication Sciences, Trinity College Dublin

Abstract

This paper outlines the establishment of the first digital corpus of Irish Sign Language (ISL) using a software programme called ELAN. The Signs of Ireland comprises 40 signers making it the largest digital annotated corpus of a signed language in Europe. This paper describes the way in which such software enhances sign linguistic research, and outlines some of the limitations that arise, in great part, because of the lack of a standardized notation system for signed languages, because of the need for human consistency when working on annotation, and the fact that you will 'get out what you put in' when working with a digital corpus: that is, the decisions made regarding the annotations influence analysis results.

1. Introduction

Irish Sign Language is an indigenous language of Ireland. It is used by some 5,000 Irish Deaf people as their preferred language (Matthews 1996) while it is estimated that some 50,000 non-Deaf people also know and use the language to a greater or lesser extent (Leeson 2001). The Signs of Ireland corpus is part of the Languages of Ireland programme at the School of Linguistic, Speech and Communication Sciences, TCD. It comprises data from Deaf Irish Sign Language (ISL)

users across Ireland in digital form, and has been annotated using ELAN, a software programme developed by the Max Planx Institute, Nijmegan. The corpus is housed at the Centre for Deaf Studies, a constituent member of the School.

While technology has opened the way for the development of digital corpora for signed languages, we need to bear in mind that traditionally sign linguists have worked painstakingly with VHS material, and pen and paper in hand when beginning to gloss data from signed languages. The fact that signed languages are articulated in three dimensional space, using not only the hands and arms, but also the head, shoulders, torso, eyes, eyebrows, nose, mouth and chin to express meaning (e.g. Klima and Bellugi 1979 for American Sign Language (ASL); Kyle and Woll 1985, and Sutton-Spence and Woll 1999 for British Sign Language (BSL); and McDonnell 1996; Leeson 1996, 1997, 2001; O'Baoill and Matthews 2000 for Irish Sign Language (ISL)) leads to highly complex, multi-linear, potentially dependent tiers that need to be coded and time-aligned.

As with spoken languages, the influence of gesture on signed languages has begun to be explored (Armstrong, Stokoe and Wilcox 1995, Stokoe 2001; Vermeerbergen and Demey (forthcoming)), while discussion about

what is linguistic and what is extra-linguistic in the grammars of various signed languages continues (e.g. Engberg-Pedersen 1993, Liddell 2003, Schembri 2003). While these remain theoretical notions at a certain level, decisions regarding how one views such elements and their role as linguistic or extra-linguistic constituents plays an important role when determining what will be included or excluded in an annotated corpus. Such decisions also determine how items are notated, particularly in the absence of a written form for the language being described.

2. ELAN

Originally developed for gesture researchers, ELAN has become the standard tool for establishing and maintaining signed language corpora. ELAN (EUDICO Linguistic Annotator) is an annotation tool that allows you to create, edit, visualize and search annotations for video and audio data. It was developed with the aim to provide a sound technological basis for the annotation and exploitation of multi-media recordings. ELAN is specifically designed for the analysis of languages, sign languages, and gesture, but it can be used by everybody who works with media corpora, i.e., with video and/or audio data, for purposes of annotation, analysis and documentation. *(Source: ECHO Project -* *http://www.let.ru.nl/sign-lang/echo/index.html?http&&&www.let.ru.nl/sign-lang/echo/data.html)*

3. Collecting the Data

In summer 2004, Deirdre Byrne-Dunne collected data from some 40 signers aged between 18 and 65 from across the Republic of Ireland. The corpus community comprised male and female signers, all of whom had been educated in a school for the Deaf. None were sign language teachers, as we wished to avoid the collection of data from signers who had a highly conceptualized notion of 'correct' or 'pure' ISL. The fact that Deirdre is herself Deaf, and an established member of the Irish Deaf community, meant that the potential for 'Observor's Paradox' (Labov 1969) while not reduced, took on a positive spin: knowing who the interviewer/ recorder of data was, and knowing their status as a community member, lent itself to the informants opening up and using their 'natural' signs rather than a variety that they might have assumed a university researcher would 'expect' or 'prefer'.

It also meant that the informants who knew Deirdre, either as a former class-mate or from within the Deaf community, code-switched to use lexical items that would not typically be chosen if the interlocutor was unknown. For example, some 'school' signs were used (BROWN). And in other instances, informants, telling stories that they had self-selected, referred to Deirdre during the recounting of their tales, often noting that 'you know'/ 'you remember'. As a result, the data collected is remarkable and we take this opportunity to sincerely thank the participants for their contribution.

We have touched on the fact that data collected included self-selected narratives. We also asked participants to tell 'The Frog' story, which is a picture sequence format telling the story of a young boy who, with his dog, searches for his frog, which has escaped from a jar. Informants were also asked to sign the content of the Volterra picture elicitation task, a series of 18 sets of paired pictures showing a series of situations that aim to elicit transitive utterances. Both the 'frog' story and the Volterra picture elicitation task have been used widely in signed language specific descriptions and in cross-linguistic comparisons, including ISL (e.g. Leeson 2001 for ISL; Johnston, Vermeerbergen Schembri and Leeson (forthcoming) for Australian Sign Language, Flemish Sign Language and ISL; Volterra et al. 1984 for Italian Sign Language; Coerts 1994 for Sign Language of the Netherlands).

Funding permitting, we would like to expand the data on file to include renditions of Chafe's Pear Story and Aesop's fables, dialogues, and interviews with Deaf ISL users regarding how they view ISL in order to record the current status and usage of ISL. We would ideally also like to supplement this with register specific data, such as descriptions of occupational activities to elicit the range of register specific vocabulary available within the community at present.

4. Codification

When annotating a corpus, there is a myth that the annotators are neutral with respect to the data and that they simply 'write down what they see'. While this is untrue for anyone working on a corpus linguistics project, it is doubly untrue for those working on a signed language. As ISL does not have a written form, there is no standard code for recording it. While some established transcription keys exist (HamNoSys, Sign Writing, Stokoe

Notation), none of these are compatible with ELAN and none are fully developed with respect to ISL.

The other problem with these transcription systems is that they are not shared 'languages' – that is, in the international sign linguistic communities, these transcription codes are not common place, and to use one in place of a gloss means limiting the sharing of data to an extremely small group of linguists. However, glossing data with English 'tags' is problematic too. Pizzutto and Pietrandrea (2001) point out the dangers inherent in assuming that a gloss can stand in for an original piece of signed language data. They note that "It is often implicitly or explicitly assumed that the use of glosses in research on signed [languages] is more or less comparable to the use of glosses in research on spoken languages … this assumption does not take into account, in our view, that there is a crucial difference in the way glosses are used in spoken as compared to signed language description. In descriptions of spoken (or also written) languages, glosses typically fulfill an ancillary role and necessarily require an independent written representation of the sound sequence being glossed. In contrast, in description of signed languages, glosses are the primary and only means of representing in writing the sequence of articulatory movements being glossed" (2001: 37). Later, they add that " … glosses impose upon the data a wealth of unwarranted and highly variable lexical and grammatical information (depending upon the spoken/written language used for glossing)." (ibid: 42).

Thus, the glossing of signed data is fraught with potential problems – even when a team is working very consistently and cross-referencing work in a diligent manner, as is the case here. The Signs of Ireland project appears to be unique in that *all* annotated data is verified by a Deaf research assistant who holds a masters degree in applied linguistics. All three annotators have masters degree qualifications in linguistics/ communications and all hold Deaf Studies specific qualifications, making them uniquely qualified to work with this data.

4.1 You Get Out What You Put In
While one of the most positive features of ELAN is the fact that the stream of signed language data runs in a time-aligned fashion with the annotations, the problem remains that any search function is restrained by the consistency and accuracy of the annotations that have been inputted and second-checked by the Signs of Ireland team.

For example, several ISL signs may be informally glossed in the same way, but the signs themselves are different, for example, WHAT (1), which is articulated using two hands, both taking an 'L' handshape, and having contact at c. locus. This is considered the 'citation form' of the sign:

In contrast, WHAT (2) is articulated on one hand, with the palm facing the signer. The middle finger wriggles a little in articulation. This is considered to be an informal variant- for example, it would not usually be taught in an ISL class:

The fact that both of these signs are glossed in the same way demonstrates that any frequency count that would subsequently be carried out using ELAN would not distinguish between the two on the basis of the gloss, WHAT, alone. The same is true of signs such as HEARING [1] as used by older signers ("L" handshape at chin) and HEARING [2] ("x" handshape at chin) as used by younger signers, and DEAD (citation form), compared with varieties that appear to be influenced from British Sign Language:

HEARING [1]

HEARING [2]

This example raises another issue – and one which we have avoided to date- that of tagging items according to grammatical function. We have not done this because we do not yet know enough about the grammatical function of items in ISL to accurately code to that level, but despite this, our annotations do reflect assumptions about the nature and structure of certain items. We also take very seriously the concerns of linguists who have recorded the impact of early codification of signed languages like Flemish Sign Language (VGT) (Van Herreweghe and Vermeerbergen 2004).

4.2 Beginning to Annotate a Signed Text
Despite the fact that we wanted to avoid making assumptions about word class and morpho-syntax, the act of annotating a text means that certain decisions have to be made about how to treat specific items. For example, it is known that non-manual signals, articulated on the face of the signer, provides information that assists in parsing a message as for example, a question or a statement, or in providing adverbial like information about a verbal predicate (e.g. Leeson 1997; O'Baoill and Matthews 2000 for ISL, Sutton-Spence and Woll 1999, Brennan 1992, Deuchar 1984 for British Sign Language; Liddell 1980 for American Sign Language). When it comes to annotating such features, we have to decide if we treat non-manual features as dependent tiers, relative to the manual signs that they co-occur with, or as independent tiers containing information that may be supra-segmental in nature.

At the lexical level, there were decisions to be made as to what constitutes a word in ISL. While established lexical items that have citation forms in dictionaries/ glossaries of ISL were 'easy' to decide on, there was the issue of how to determine if a sign was a 'word' or a 'gesture' or part of a more complex predicate form, often described as classifier predicates. The fact that some signers used signs related to their gender or age group challenged the annotators – they had to determine whether a sign that was new to them was a gendered variant (Le Master 1990, 1999-2000, Leeson and Grehan 2004), a gendered generational variant (Le Master ibid, Leonard 2005), a mis-articulation of an established sign (i.e. a 'slip of the hand' (Klima and Bellugi 1979), an idiosyncratic sign, a borrowing from another signed language (e.g. BSL), or a gesture. Our team's experience and qualifications helped the decision making process here. All decisions were recorded in order to provide a stable reference point for further items that challenged that shared characteristics with items that were discussed previously.

The use of mouth patterns in signed languages provide another challenge for annotators dealing with signed languages. Mouthings and mouth gestures have been recognized as significant in signed languages, and while mouthings are often indicative of the language contact that exists between spoken and signed languages, mouth gestures are not (for example, see Boyes Braem and Sutton-Spence 2001, Sutton-Spence, forthcoming).

 Given that the Signs of Ireland corpus will, in the first instance, be used by researchers looking at the morpho-syntax of the language, we opted to not annotate the mouth in a very detailed manner. Instead, we have provided fairly general annotations following from those listed in the ECHO project annotations list.

5. Using Space to Describe Space: Topographic Space
More complicated were data streams where signers describe locations or objects in detail. Signers use space to describe space. Working from a cognitive linguistics perspective, Leeson and Saeed (2003, 2004, 2005, forthcoming) have reported on aspects of the use of space in ISL to encode information about relative location, signer perspective, the 'blending' of space and the use of iconicity in narratives.

For example, Leeson and Saeed (2005, forthcoming) discuss aspects of such complex use of signing space in ISL. They report on a current affairs presentation by a male signer in his early 30s who is discussing the early stages of the war in Iraq. The signer establishes locative relationships (where Turkey and Kuwait are positioned relative to Iraq) and then builds on the fact that the location of Iraq is active for interlocutors when it comes to expressing the proposed simultaneous invasion of Iraq from the north and south. Thus, the non-dominant hand does not even need to maintain an explicit reference to Iraq. Leeson and Saeed note that Iraq is currently 'gapped' (Talmy 1996) insofar as it is not explicitly referenced, pragmatically it is active as the discourse theme.

In this same example, the simultaneous nature of the invasion of Iraq is expressed simultaneously in ISL. Leeson and Saeed describe how the signer maps the invading forces onto the loci previously established for Turkey and Kuwait, again using topographical space (Sutton-Spence and Woll 1999) to locate the relative locations of these countries. The facing of the fingertips in the CL5+open classifier demonstrates a movement away from Turkey/Kuwait, towards a central locus, that of Iraq. The movement of this classifier ends at the locus for Iraq.

They also report the use of iconicity encoded in the relative locations of Turkey, Iraq and Kuwait. The signer maps real world relative locations onto signing space, creating a conceptually complex representation.

There is a blending of spaces involving locations and entities in the real world, a projected map of these, and elements of the signer's real space. The signer uses parts of his body to represent entities moving around the war zone, while his eye gaze is directed towards the projected map, representing the viewpoint of the signer himself as narrator/viewer of the events.

We see other examples of this kind of complex structures in the Signs of Ireland data, for example, in the story of a Deaf woman discussing an incident with her motorbike. She recounts how a garden hose got entangled in the back of her bike when leaving for work early one morning. She recounts how a truck driver tried to tell her that there was a problem while she was stopped at some traffic lights – but she was didn't understand what he was saying and was afraid that he was following

her. Eventually, the driver stops her at another junction and points out the problem.

The interesting thing about this piece is the extent of iconicity, which is encoded in mapping the real world' onto signing space. The space then becomes active for the signer as she was on the morning in question (i.e. recounting the event as it unfolded), and the signer's c. locus is also commensurate for the signer as narrator, the signer as herself in the present time, the truck driver, and later on, herself at work and a colleague/s at work whom she recounts the incident for. The relative location and size of the signer vis-à-vis the other participants introduced into signing space reflects the real-world locations of the signer and these participants on the day of the event that she is recalling. For example, the stills below demonstrate the relative positionality of the signer and the truck driver.
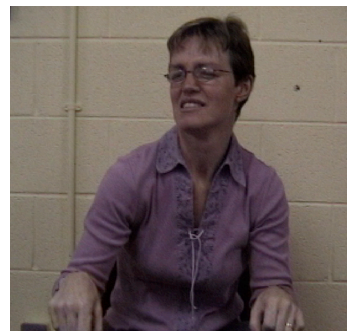
Signer as 'HERSELF-ON-MOTORBIKE'


Signer as 'TRUCKDRIVER'

She also encodes aspects of the size and shape of her motorbike, including the location of the mirror and handlebars. She further encodes information about the size and shape of the garden hose that is discussed in the story, especially where she discusses rolling up the hose and carrying it over her arm as she continued her journey by motorbike to work.

WRAPPING-HOSE-AROUND-ARM-AND-SHOULDER

PUTTING-HOSE-OVER-SHOULDER
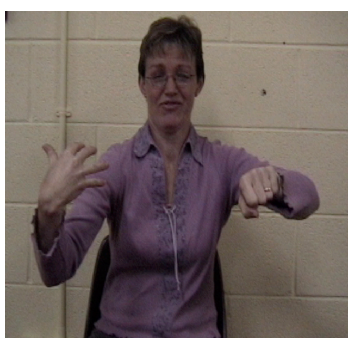
DRIVING-MOTORBIKE-WHILE-CARRYING-HOSE-OVER-RIGHT-SHOULDER

The major issue here is that the complex use of signing space to represent these multiple facets of how one conceptualizes an event and encodes it to represent that conceptualization draws on what have been considered core linguistic (lexicalized items, verb structure, etc.) in nature. But it also includes those items that have traditionally been considered as extra linguistic such as 'cognitive iconicity' and gesture (Wilcox 2004a, 2004b).

6. Deciding on the Morpho-syntax

Coming back a step, in transcribing data, one of the big questions concerns morphological structure. The question then becomes, 'what is a morpheme in a signed language?' Following from this we can ask if all signs have decompositional morphemes or if iconicity and gesture actually govern structure (see for example Liddell 2003 for discussion of whether morphology in ASL is as consistent as previously claimed, Wilcox 2004a, 2004b for discussion of the role of cognitive iconicity in motivating structure, and Sallandre (forthcoming) for an overview of the French tradition regarding Highly Iconic Structures (HIS)).

While many sign linguists argue for aspects of grammar such as verb agreement embedded in the morpho-syntax of signed languages (Engberg-Pedersen 1993 for Danish Sign Language, Meier 1998 for Israeli Sign Language, Sutton-Spence and Woll 1999 for British Sign Language, Sandler and Lillo-Martin 2006 for a range of signed languages, and McDonnell 1996, 2004 for Irish Sign Language), some researchers have rejected a morphologically driven approach to the description of such structures. Instead, scholars like Scott Liddell (2003) have embraced a 'strong' version of cognitive linguistics, arguing that iconicity drives what has been described as verb agreement for person and location. Liddell (2003: 97-98) proposes that what he calls 'indicating verbs' differ from citation form verbs in that they "… are capable of being meaningfully directed in space towards entities, directions or places." He goes on to note that "if requested to produce an instance of an indicating verb, signers will often produce a form that is not meaningfully directed at anything. This makes sense since the verb is being produced out of any discourse context – other than the request to produce a verb".

The fact that such 'indicating verbs' (Liddell (ibid) or 'agreement verbs' (McDonnell 1996) occur extensively in discourse describing events of motion and location, where the relative location of referents in space is fluid, not fixed, creates problems when deciding on how to annotate for such locations in space and the motion that occurs between such points. Our annotations insinuate that we adopt the principle that some kind of co-referencing between the signer's locus (c.) and that of a referent established at another point in signing space occurs, and we draw on the referencing format developed by Engberg-Pedersen (1993) for Danish Sign Language, in this respect. We note here that such annotations infers also that we take as given that there is some form of agreement morphology embedded in the structures that we analyse, reflected in the use of the + symbol to annotate movement of a verb form between locations (e.g.

c+TELEPHONE+f indicates an onset point at the c locus, from where the verb moves forward to an offset point that is to the front of the signer. It indicates agreement between first and non-first person, and might be translated as '(I) (will) phone (you)'. Here c is commensurate with a location that is morphologically meaningful as is f. Again, the annotation infers an underlying acceptance of how space is used in a signed language, so the notion of 'writing down what you see' is challenged.

Such loci for first and non-first person are tracked across discourse stretches (e.g. see Janzen 2005). While we have not yet explored the frequency and consistency of such reference tracking in this corpus in detail, we can say already that the establishment of loci and the tracking of referents in signing space in a consistent manner across a signed text is the most consistent feature that we note in the corpus to date. This is regardless of the age of the signer, the region that they live in or their gender. Thus the mapping of 'real world' locations onto signing space and partitioning out that space seems fundamental to what signers do. The Signs of Ireland corpus will allow us to test the boundaries of how space is used in a manner that simply was not possible using VHS and pen and paper.

6.1 How do we view "classifiers"?

The idea that signed languages make use of 'classifier predicates' has been fraught with controversy. That signed languages encode information about the size and shape of referents, about animacy, about path of motion, direction of movement and noun class (e.g. two-legged, vehicular, animate) is well established (see Emmorey (ed) 2003 for an excellent overview of this subject) but the nature of such structures is still open to much debate. Are these structures truly linguistic in nature? Or are they remnants of gesture that have become highly conventionalized within a given signed language? For example, Vermeerbergen and Demey (forthcoming) describe a set of gestures used by Flemish speakers when carrying out a task that are remarkably similar in terms of articulation to some simultaneous structures typical of many signed languages (Liddell's 'buoys' (cf Liddell 2003).

Further, the issue is compounded by the modality that signed languages are expressed in. Given that signed languages are visual-spatial languages that make use of three dimensional signing space and use the hands, arms, torso, head and facial features as articulators, it makes sense that aspects of the real world are mapped onto signing space. Thus an object in the distance that moves as a vehicle is encoded as a flat inanimate entity moving in a specific direction, at a relative speed at a relative distance from the signer – encoded from the signer's visual point of view.

We have noted that these challenge us in terms of how we might consistently annotate data, but again indicates that when we adopt established, conventionalized approaches to annotation, we are to a large extent buying into the theories that are in place. For example, we annotate classifiers as 'CL' even though we are extremely aware of the debate that rages over their status (e.g. Schembri 2003) and though more and more researchers follow Liddell (2003) in referring to certain simultaneously established lists as 'buoys', we have not done so in this corpus. This also demonstrates how ongoing description of aspects of signed languages can impact on annotation and thus mark a corpus as 'dated' if it does not embody such changes. The tension between being maximally accurate in annotating and aiming for a longitudinally stable corpus that can be enlarged over time means that we will have to take account of the need to review annotations and perhaps adjust them over time to best reflect our understanding of what is happening in ISL. Again, this reflects just how annotations are anything but neutral.

7. Deaf Sign Language Users: The Encoding of Visual Motivation

Aspects of being Deaf – or perhaps, more accurately – being a visual language user – is encoded in ISL too. Thus, from a Deaf perspective, what is unseen is not discussed, and actions arising without volitional involvement on the part of the signer are typically encoded with averted eyegaze (Leeson and Saeed 2003). For example, if a signer is approached from behind, and someone taps them on the shoulder, the information encoded only includes what the signer can see or feel, in this case, a hand tapping the signer's shoulder. While we can describe these constructions as passives (focus on the semantic Undergoer rather than the Actor), from a transcription point of view, what is most important is that (1) the hand tapping the shoulder is encoded and (2) that eyegaze is averted for the duration of the initial point of contact.

Another example of the visual nature of being Deaf is found in the following example [MOTORBIKE STORY] where the signer notes that she felt a strange vibration when she sat on her bike. As a result, she looked in her rear view mirror, but as it was dark, she could not see anything significant. Here, the signer is focusing on the fact that she is a visually motivated person – she doesn't hear the sound of the hosepipe caught in the wheel of her bike – she feels it. But, because it is dark, she cannot see it and drives on.

LOOKING-TO-CHECK-SOURCE-OF-PROBLEM

The data in the corpus is rich in this type of reflection of the experience of deafness, but again, it is only when one moves beyond the physical manifestation of grammar and considers world view, pragmatics and semantics that one gets to this in signed language data.

8. Beyond the Sign – Identifying Clause and Sentence Boundaries

The question of determining what constitutes a clause, or indeed a sentence, in a signed language is one that researchers have debated in some depth. Johnston, Vermeerbergen, Schembri and Leeson (forthcoming) tackled this issue for three signed languages – Auslan (the signed language of Australia), Flemish Sign Language (VGT) and ISL. They found that non-manual features such as eyeblink, which has been suggested as an indicator of clause boundaries (Wilbur 1994), was not always possible to identify due to visual restrictions (i.e. the signer may be in shadow on camera and so eyeblinks cannot be identified accurately). Instead, Johnston et al. (ibid.: 7) report that,

"In order to avoid as much as possible the kinds of variable coding and interpretation of the data that appears almost inevitable between data sets from different signed languages (and even within a data set from a single signed language), we applied one basic criterion for 'clause-hood': the nucleus of each clause

should be some kind of predicating element, often a verb (Van Valin & LaPolla, 1997). In other words, each individual verb represented a separate clause, even if there were no explicit and separate signs for the various arguments of the verb …"

They also point out that some of the data that they analysed contained predications without any verbs and/or without a copula (ibid.: 24):
"For example, some descriptions of a spatial arrangement or existence only use signed arguments that are located through the placement of signs in the signing space or through body shifts left or right during the production of certain signs (e.g., COWBOY-lf INDIAN-rt 'There's a cowboy here and an Indian there') with no verb."

Such issues also arise in the ISL data here, though, unlike the rather more straightforward elicited data described in Johnston et al., the ISL data is more complex given that signers layer timeframes, event schemas, participants and point of view predication.

In annotating the data, boundaries were not marked but pause structure (e.g. longer holds on the phrase-final sign) and the co-occurrence of certain non-manual features (e.g. eyeblink or return of eye-gaze to camera), suggest a pattern of phrase marking in the ISL data that we have worked with thus far. This is one of the features that we will analyse in detail when the corpus is complete.

9. Implications for Technical Work

The question of how the work on a corpus can assist technical development is an open one, which raises more questions in turn. For example, if considering machine translation, we can ask how do you write re-write rules when we don't 'know' the grammar of ISL (i.e. ISL is under-described)? How do you create an idealized standard when, from what we see, there is so much variation? How do you capture the role of iconicity and gesture and encode that as 'grammar'? How do you capture the multi-dimensional interaction between non-manual and manual articulators? How do you parse for individual conceptualization of the world, encoded in a three-dimensional signing space where relative location and the path of movement is meaningful? These are the questions challenging IT people working with signed languages today.

10. Applications of the Corpus: A Hub for Research

Beyond serving as a body of data for sign linguists, the Signs of Ireland corpus has the potential to be a hub for research activity focused on ISL. It has potential for those working towards developing a text-to-sign model of ISL. It also has potential as a tool for students learning ISL and those training as interpreters or teachers of the language. Beyond that, this corpus represents a significant archive of data from an endangered language that was severely suppressed from the 1940s until the late 1990s. (Some would argue that in education, ISL continues to be suppressed. Indeed, we argue that the degree of variation that exists in the data examined to date reflects the institutional suppression of the language.)

This corpus then, provides a snap shot of the degree of variability that exists among Deaf male and female signers aged 18-65 in the Republic of Ireland. If this corpus leads to greater understanding of the complexity and beauty of ISL, then it will have served a significant purpose. It is our hope that it will do much more than that.

11. Summary

While discussing progress to date at a recent Signs of Ireland meeting, we were reflecting on the question of whether the data supported the notion of a 'standard' ISL. One of the researchers noted that 'the only thing that the signers have in common is that they all have 2 hands and they all have heads': variation is rampant in ISL. Understanding that this variation is a direct result of language policy that banned use of ISL in the classroom and in public places (Leeson and Grehan 2004) helps us to understand the sociolinguistics more: there is no extensive regional variation because there were only 2 major schools for the Deaf and until the introduction of the notion of 'mainstream education', most profoundly deaf children attended these schools. However, there was segregation on the basis of gender, and further, on the basis of perceived ability to speak and lipread. This led to very extensive gendered variation in the language and class cohort variants. The latter is yet to be fully recorded and discussed.

The act of annotating this corpus brings us closer to the language. Despite the breadth of expertise of the Signs of Ireland team as ISL users, we have all been amazed by the complexity and beauty of the data. The systematicity underlying how events are encoded and the conventionalization of cognitive iconicity reflects the fact that human languages can develop and thrive even when they are discouraged and actively suppressed. We hope that this corpus will reflect this miracle of language survival, and that the Signs of Ireland corpus will function as a cornerstone for ISL researchers in the future who will have significant synchronic and diachronic data to work with.

12. Acknowledgements

13. References

Armstrong, D.F., W.C. Stokoe and S.E. Wilcox, *Gesture and the Nature of Language*, Cambridge University Press, Cambridge, 1995.

Boyes Braem, P. and R. Sutton-Spence (eds), *The Hands are the Head of the Mouth: The Mouth as Articulator in Sign Languages*. International Studies on Sign Language and Communication of the Deaf, 39, Signum, Hamburg, 2001.

Coerts, J. Constituent order in Sign Language of the Netherlands, In M. Brennan and G.H. Turner (eds) *Word-order Issues in Sign Languages – Working Papers*, The International Sign Linguistics Association (ISLA),Durham, 1994, pp 47-71.

Emmorey, K. (ed) *Perspectives on Classifier Constructions in Sign Languages*, Lawrence Erlbaum and Associates, New Jersey, 2003.

Engberg-Pedersen, E., *Space in Danish Sign Language*, Hamburg, Signum Verlag, 1993.

Janzen, T. *Perspective Shift Reflected in the Signer's Use of Space,* Centre for Deaf Studies, University of Dublin, Trinity College, Dublin, 2005.

Klima, E. and U. Bellugi, *The Signs of Language*, Harvard University Press, Harvard, 1979.

Kyle, J.G. and B. Woll, *Sign Language: The Study of Deaf People and their Language*, Cambridge University Press, Cambridge, 1985.

Himmelmann, N.P., Language Documentation: What is it and what is it good for?, In Gippert, J., N.P. Himmelmann and U. Mosel (eds.), *Essentials of Language Documentation,* Mouton de Gruyter, Berlin and New York, 2006, pp 1-30.

Himmelmann, N.P., The challenges of segmenting spoken language In Gippert, J., N.P. Himmelmann and U. Mosel (eds.), *Essentials of Language Documentation,* Mouton de Gruyter, Berlin and New York, 2006, pp 253-274.

Johnston, T., M. Vermeerbergen, A. Schembri and L. Leeson, 'Real data are messy': Considering cross-linguistic analysis of constituent ordering in Australian Sign Language (Auslan), Vlaamse Gebarentaal (VGT), and Irish Sign Language (ISL). In: Perniss, P., Pfau, R., Steinbach, M. (Eds.) *Visible Variation: Comparative Studies on Sign Language Structure*. Mouton de Gruyter, Berlin. (41 blz.) (in press).

Labov, W, *The Logic of Nonstandard English*, Georgetown Univeristy 20th Annual Round Table, Monograph Series on Languages and Linguistics, No. 22, 1969.

Leeson, L., *The Marking of Time in Signed Languages with Specific Reference to Irish Sign Language*, Unpublished M. Phil Dissertation, Centre for Language and Communication Studies, University of Dublin, Trinity College, 1996.

Leeson, L., *The ABC of ISL.* Booklet to accompany the TV series. RTE/Irish Deaf Society, Dublin, 1997.

Leeson, L. *Aspects of Verb Valency in Irish Sign Language*, Unpublished Doctoral Dissertation, Centre for Language and Communication Studies, TCD, Dublin, 2001.

Leeson, L. and C. Grehan, To the Lexicon and Beyond: The Effect of Gender on Variation in Irish Sign Language. In Van Herreweghe, M. and M. Vermeerbergen, *To the Lexicon and Beyond: The Sociolinguistics of European Sign Languages,* Gallaudet University Press, Washington DC, 2004, pp 39-73.

Leeson, L. and J.I. Saeed, *Exploring the Cognitive Underpinning in the Construal of Passive Events in Irish Sign Language (ISL).* Paper presented at the 8th International Cognitive Linguistics Association Conference, Spain, July 2003.

Leeson, L. and J.I. Saeed, The Windowing of Attention in Simultaneous Constructions in Irish Sign Language. In T. Cameron, C. Shank and K. Holley (eds.) *Proceedings of the Fifth High Desert Linguistics Conference*, 1-2 November 2002, Albuquerque, University of New Mexico, 2004.

Leeson, L. and J.I. Saeed, *Conceptual Blending and the Windowing of Attention in Simultaneous Constructions in Irish Sign Language (ISL)*, Paper Presented at the 8th International Cognitive Linguistics Association Conference, Seoul, Korea, July 2005.

Leeson, L. and J.I. Saeed, Conceptual Blending Blending and the Windowing of Attention in Simultaneous Constructions in Irish Sign Language (ISL), In Vermeerbergen, M., L. Leeson and O. Crasborn (eds), *Simultaneous Constructions in Signed Languages – Form and Function*, John Benjamins, Amsterdam, (forthcoming).

Le Master, B. *The Maintenance and Loss of Female and Male Signs in the Dublin Deaf Community*, PhD. Dissertation, Los Angeles, University of California, 1990.

Le Master, B., Reappropriation of Gendered Irish Sign Language in One Family, *Visual Anthropology Review*, 15(2), 1999-2000, pp. 1-15.

Le Master, B. What Difference Does Difference Make? Negotiating Gender and Generation in Irish Sign Language. In S. Benor, M. Rose, D. Sharma and Q. Shang (eds), *Gendered Practices in Language*, Centre for the Study of Languages and Information Publication, Stanford, 2002.

Leonard, C. 2005, Signs of Diversity: Use and Recognition of Gendered Signs among Young Irish Deaf People, Deaf Worlds, Vol. 21 (2), pp. 62-77, 2005.

Liddell, S.K. *Grammar, Gesture, and Meaning in American Sign Language*, Cambridge University Press, Cambridge, 2003.

Matthews, P.A. *The Irish Deaf Community- Volume 1*, ITE, Dublin, 1996.

McDonnell, P, *Verb Categories in Irish Sign Language*, Unpublished Doctoral Dissertation, Centre for Language and Communication Studies, University of Dublin, Trinity College, 1996.

McDonnell, P (ed) *Deaf Studies in Ireland: An Introduction*, Doug McLean, Glouc., 2004.

Meir, I., *Thematic Structure and Verb Agreement in Israeli Sign Language*, PhD Dissertation, Hebrew University of Jerusalem, 1998.

Mosel, U., Fieldwork and Community Language Work. In Gippert, J., N.P. Himmelmann and U. Mosel (eds.), *Essentials of Language Documentation,* Mouton de Gruyter, Berlin and New York, 2006, pp 67-86.

O'Baoill, D. and P.A. Matthews, *The Irish Deaf Community, Volume 2 – The Linguistics of ISL*, ITE, Dublin, 2000.

Pizzuto, E. ans P. Pietrandrea, The Notation of Signed Texts, *Sign Language and Linguistics*, Vol 4, No. 1/2, pp29-45, 2001.

Sallandre, M., Simultaneity in French Sign Language Discourse, In Vermeerbergen, M., L. Leeson and O. Crasborn (eds), *Simultaneous Constructions in Signed Languages – Form and Function*, John Benjamins, Amsterdam, (forthcoming).

Sandler, W. and D. Lillo-Martin, *Sign Language and Linguistic Universals*, Cambridge University Press, Cambridge, 2006.

Schembri, A., Rethinking 'Classifiers' in Signed Languages, In K. Emmorey (ed), *Perspectives on Classifier Constructions in Sign Languages*, Lawrence Erlbaum and Associates, New Jersey, 2003, pp. 3-34.

Seifart, F. Orthography Development, In Gippert, J., N.P. Himmelmann and U. Mosel (eds.), *Essentials of Language Documentation,* Mouton de Gruyter, Berlin and New York, 2006, pp 275-299.

Stokoe, W. C., *Language in Hand: Why Sign Came Before Speech*, Gallaudet University Press, Washington DC, 2001.

Sutton-Spence, R. Mouthings and Simultaneity in British Sign Language, In Vermeerbergen, M., L. Leeson and O. Crasborn (eds), *Simultaneous Constructions in Signed Languages – Form and Function*, John Benjamins, Amsterdam, (forthcoming).

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

Talmy, L., The Windowing of Attention in Language. In M. Shibitani and S.A.Thompson, (eds.): *Grammatical Constructions- Their Form and Meaning.* Oxford: Clarendon Press, 1997, pp235-287.

Tsunoda, T., *Language Endangerment and Language Revitalisation – An Introduction*, Berlin and New York, Mouton de Gruyter, 2005.

Wilbur, R., Eyeblinks and ASL phrase structure, Sign Language Studies 84: 221-240,1994

Van Herreweghe and Vermeerbergen 2004, Flemish Sign Language: Some Risks of Codification, In Van Herreweghe, M. and M. Vermeerbergen, *To the Lexicon and Beyond: The Sociolinguistics of*

European Sign Languages, Gallaudet University Press, Washington DC, pp. 111-140, 2004.

Van Valin, R. D. and R.J. La Polla, *Syntax-Structure, Meaning and Function,* Cambridge University Press, Cambridge, 1997.

Vermeerbergen, M., L. Leeson and O. Crasborn, Simultaneous Structures in Signed Languages: An Introduction, In Vermeerbergen, M., L. Leeson and O. Crasborn (eds), *Simultaneous Constructions in Signed Languages – Form and Function*, John Benjamins, Amsterdam, (forthcoming).

Vermeerbergen, M. and E. Demey, Sign + Gesture = Speech + Gesture? Comparing Aspects of Simultaneity in Flemish Sign Language to Instances of Concurrent Speech and Gesture. In Vermeerbergen, M., L. Leeson and O. Crasborn (eds), *Simultaneous Constructions in Signed Languages – Form and Function*, John Benjamins, Amsterdam, (forthcoming).

Volterra, V., A. Laudanna; S. Corazza; E. Radutzky and F. Natale, Italian Sign Language: the order of elements in the declarative sentence. In F. Loncke (ed) *Recent Research on European Sign Languages*, Svets and Zeitlinger, Lisse, 1984, pp. 19-48.

Wilbur, R.B, Eyeblinks and ASL Phrase Structure, In *Sign Language Studies* 84, pp 221-240, 1994.

Wilcox, S, Gesture and Language: Cross-linguistic and Historical Data from Signed Languages, *Gesture* 43, 2004a, pp 43-73.

Wilcox, S. Cognitive Iconicity: Conceptual Spaces, Meaning and Gesture in Signed Languages, *Cognitive Linguistics* 15-2, 2004b, pp 119-147.

6/9/2006 v final