Variation, lexicalization and grammaticalization in signed languages

Trevor Johnston, Adam Schembri

DANS LANGAGE ET SOCIÉTÉ 2010/1 (N° 131), PAGES 19 À 35
Éditions ÉDITIONS DE LA MAISON DES SCIENCES DE L'HOMME

ISSN 0181-4095
ISBN 9782735112319
DOI 10.3917/ls.131.0019

Article disponible en ligne à l'adresse
Variation, lexicalization and grammaticalization in signed languages

Trevor Johnston
Department of Linguistics at Macquarie University, Sydney, Australia
trevor.johnston@ling.mq.edu.au

Adam Schembri
Deafness, Cognition and Language Research Centre, University College, London
a.schembri@ucl.ac.uk

Introduction

It has been an observation of long-standing that there is considerable attested variation in most signed languages. This variation can be addressed in relation to ‘language external’ social variables (sociolinguistics) and ‘language internal’ forces of change (lexicalization and grammaticalization). The external factors can be further specified between social or inter-speaker constraints, and stylistic or intra-speaker constraints. Social factors include, for example, a signer’s age, region of origin, gender, ethnicity, and socio-economic status. Stylistic variation involves alternation between, for example, casual and formal styles of speech used by an individual speaker, often reflecting differing degrees of attention to speech due to changes in topic, setting and audience. The internal linguistic factors include phonological processes such as assimilation and reduction, on the one hand, and lexicalization and grammaticalization, on the other. Together, internal and external constraints form a complex interrelationship, with each influencing language use in distinctive ways.

The sociolinguistic study of spoken languages (SpLs) has long rested on two guiding principles: the “principle of quantitative modelling” and the “principle of multiple causes” (Young & Bayley, 1996). The first principle
refers to the need to quantify both variation in linguistic form and the relationship between a variant form and features of its surrounding linguistic environment and social context. The second principle reflects the long-standing assumption that no single linguistic or social factor can fully explain variation in language use. From one perspective, this is captured in the ‘apparent time hypothesis’ which suggests that variation in the linguistic system used by speakers of different ages at a single point in time can indicate a change in progress (Bailey, 2002). From another perspective, the use of variant forms in particular contexts is often indicative of differential patterns of grammaticalization of forms across a linguistic community (Chambers, 1995).

In this paper we focus on how variation in SLs reflects social factors interacts with linguistic factors that relate to lexicalization and grammaticalization. We first exemplify sociolinguistic variation at the levels of phonology, lexicon, and grammar, then we discuss variation and change from the perspective of lexicalization and grammaticalization.

1. Variation in SLs

In this section, we give illustrative examples of variation and change using data from systematic large-scale studies of sociolinguistic variation in ASL (American Sign Language), Auslan (Australian Sign Language) and NZSL (New Zealand Sign Language). With respect to the language external factors of region, age, gender, and ethnicity, these studies showed how these factors influence apparently random variation in rather systematic ways (Schembri & Johnston, in press).

1.1. Phonological variation and change

The first large-scale studies of phonological variation in ASL drew on a representative sample of the American deaf population and employed multivariate analyses of the data (i.e., an analysis which considers multiple variables simultaneously).

One study investigated variation in the sign deaf. Results showed that the factors that conditioned the phonological variants were linguistic, social and stylistic in nature (Bayley, Lucas, & Rose, 2000). Signers were less likely to use a citation form in nominal compounds, such as deaf^world, but

1. Section One is based on a much abbreviated version of a chapter by the authors in Pfau, R., Steinbach, M., & Woll, B. (eds) (forthcoming) Sign Languages. Berlin: Mouton de Gruyter.
2. Glossed signs can be viewed on the Langage et Société website <http://www.LetS.msh-paris.fr> (Auslan signs can also be viewed at www.auslan.org).
more likely to do so when deaf was part of a predicate, as in pro-3 deaf ‘She is deaf’. Social factors such as region and age were also important, e.g. signers in one region tended to use non-citation forms of deaf more than twice as often as signers in another, and older signers in some regions were found to be consistently more likely to use the citation form than younger signers.

Another study explored variation in ASL signs produced in citation form with the Bhandshape. The results showed that linguistic and social factors, in addition to phonological assimilation, were also at work in handshape variation. Phonological environment turned out to be significant, but it was not the most important linguistic factor. Instead, grammatical function was the strongest influence. Signers are more likely to choose an C handshape variant for wh-signs, for example, and a : variant for pronouns, whereas other lexical and function signs are more often realised in citation form. Social factors were also important, e.g., signers in some regions favouring the citation form, while those in others all disfavoured it.

A third study investigated location variation in signs produced on or near the signer’s forehead in citation form, but which often may be produced at locations lower than this. Again, analysis showed that grammatical function was the strongest linguistic factor, with noun, verbs and adjectives appearing more often in citation forms while prepositions and interrogative signs favoured lowered variants. Phonological environment was also important, with preceding signs made on or near the body having a significant influence on whether or not the target sign appeared as a lowered variant. The results also indicated that younger signers, men, and non-native signers all favoured lowered variants when compared to older signers, women and native signers. Regional and ethnic differences also emerged, with African-American deaf people and those from southern regions tending to use more citation forms than Whites and signers from the five other regions.

Auslan and NZSL sociolinguistic variation projects have also investigated phonological variation, focusing specifically on variation in the location parameter, replicating the ASL study reported above. For example, with respect to Auslan, Schembri et al., (2009) reported that variation in the use of the location parameter in these signs (e.g., name) reflects both linguistic and social factors, as has also been reported for ASL. Like the American study, the Auslan results provided evidence that the lowering of this class of signs reflects a language change in progress in the Australian deaf community, led by younger people and individuals from the larger urban centres. This geolinguistic pattern of language change (i.e., from larger to smaller population centres) is known as cascade diffusion, and is quite common.
The NZSL study found evidence of similar regional differences in the use of lowered variants, but age was not a significant factor in the dataset.

Furthermore, the results indicated that some of the particular factors at work, appear to differ in Auslan and NZSL when compared to ASL. First, the Auslan and NZSL studies suggested relatively more influence on location variation from the immediate phonological environment than is reported for ASL. Second, the Auslan and NZSL data suggested that location variation in this class of signs is an example of language change led by deaf women, not by deaf men as in ASL. This is typical of a language change known as change from below (i.e., one that is occurring without there being much awareness of this change in progress among the community of speakers or signers). Third, the Australian and New Zealand researchers showed that grammatical function interacts with lexical frequency in conditioning location variation (i.e., they found that high frequency verbs were lowered more often than any other class of signs), a factor not considered in the ASL study.

1.2. Lexical variation and change

Lexical variation presents the clearest examples of sociolinguistic variation in many SLs, with lexical choices often systematically associated with signers of a particular region, age, gender, religion, or ethnicity.

1.2.1. Region

Lexical variation has been noted in a wide range of SLs, and even signed varieties that are used across relatively small geographical areas can have multiple distinctive regional variants, e.g. the Netherlands SL has five regional dialects, with significant lexical differences between all regions but particularly between the south and the rest of the country (Schermer, 2004).

Lexical variation found in the two main regional varieties of Auslan (the northern and the southern dialects) is typical of that reported for many signing communities. The two dialects of Auslan differ in the signs traditionally used for numbers, colours, days of the week, and some other concepts. Indeed, the core set of vocabulary in certain semantic areas is actually different for every basic term in these dialects, as in the colour signs red, blue, green, yellow and black.

For ASL, Shroyer & Shroyer (1984) elicited data from signers in 25 states for 130 concepts which yielded 1,200 sign variants. Their data also suggested that, like BANZSL (British, Australian and New Zealand SLs con-
considered as one language) varieties, ASL regional variation was concentrated in certain semantic categories, particularly signs for food and animals.

What is of importance and interest here is not so much individual examples of lexical variation but the observation that the variation stems from similar sociolinguistic factors in different signing communities and manifests itself in very similar ways. In Britain, Australia, New Zealand and the United States, as in Europe, the establishment of residential deaf schools from the eighteenth to twentieth centuries accounts for much of the variation – and similarity – between signs in various regions. For example, signed communication was forbidden by some schools during the latter part of the nineteenth and for much of the twentieth century, leading to the creation of new signs by deaf children.

1.2.2. Age
The vast majority of deaf people have hearing families and the age at which they acquire SLs may be very late. Thus the intergenerational transmission of SLs is often problematic. This can result in considerable differences across generations.

One of the clearest manifestations of age difference is manifested in the fingerspelling of lexical items. Sutton-Spence, Woll, & Allsop (1990) analysed the use of the British manual alphabet in relation to four social factors: sex, region, age and communication mode used. There were no effects due to gender on the use of fingerspelling, but age was a significant factor. Similarly, a much smaller study of fingerspelling use in Auslan by Schembri & Johnston (2007) found that deaf signers aged 51 years or over made more frequent use of the manual alphabet than those aged 50 or younger.

With respect to ASL, Padden & Gunsauls (2003) reported that age and social class appeared to affect the use of fingerspelled proper versus common nouns, with older and working-class signers much more likely to fingerspell common nouns. They also stated that native signers fingerspelled more frequently, with university-educated deaf native signers using the most fingerspelling. This finding has been supported by the large-scale ASL sociolinguistic variation research mentioned above.

These age-related differences in fingerspelling usage undoubtedly reflect the educational experiences of older deaf people, many of whom were instructed using approaches that emphasized the use of fingerspelling. Language change is important here, as many older signers appear to prefer the use of traditionally fingerspelled items rather than the ‘new signs’ used by younger people.
With regard to lexical signs themselves, in the ASL sociolinguistic variation research mentioned above it was reported there were lexical variants for 24 of the 34 stimulus items that were unique to particular age groups. Similarly, variation in the NZSL numeral signs one to twenty is also systematically conditioned by social characteristics, especially age (McKee, McKee, & Major, 2006). Variation in numeral usage reveals diachronic change in NZSL, and increasing standardisation in this subset of the lexicon in younger age groups.

1.2.3. Gender
A well documented case of systematic variation in key vocabulary between males and females in an adult signing community can be found in Irish SL (Leeson & Greehan, 2004). For over a century, the Irish deaf community maintained distinct vocabularies associated with the different traditions of sign language use in the single-sex residential deaf schools in Dublin. Elsewhere evidence for gender-based variation has also been found. For example, Lucas, Bayley, Reed, & Wulf (2001) report that only 8 of the 34 stimulus items they studied did not show variants unique to either men or women. Gender differences in the use of fingerspelling in ASL were also reported by Mulrooney (2002), e.g. men were more likely to produce non-citation forms than women.

1.2.4. Ethnicity and religion
A distinct African-American variety of ASL exists which reflects the historical context of American deaf education, with specific schools having been established for African-American deaf children in some southern states during the period of segregation (Lucas, Bayley, Reed, & Wulf, 2001).

Generally, there are no clearly identifiable distinctions in the SL used by various ethnic groups in Australia, partly because the education of deaf children has, for the most part, never been segregated by ethnicity, unlike religion. Separate schools for Catholic deaf children were established in Australia which employed Irish SL as the language of instruction until the 1950s. As a result, an older generation of signers in some regions of Australia make some use of Irish SL signs and the Irish manual alphabet. Some Irish SL signs have been borrowed into regional varieties of Auslan (e.g., home, cousin) (Johnston & Schembri, 2007).
1.3. Grammatical variation and change

There also appears to be significant grammatical variation in SLs. However, there has been little research into morphosyntactic variation in ASL and BANZSL varieties, and there have not yet been empirical studies demonstrating whether there are consistent differences between individual signers due to gender, age, social class or region. Nonetheless, in many contexts, it appears that signers will vary in their choice and combination of the morphological, syntactic and discourse structures without any apparent change in communicative intent.

Schembri (2001) showed, for example, that native signers of Auslan varied in their use of ‘classifier’ handshapes to represent the motion of humans and vehicles. In his dataset, both the upturned Y handshape and the upright B may be used to represent a person moving, and a ~ handshape with the palm oriented sideways or downwards may represent vehicles. Schembri et al., (2002) and Johnston (2001) examined noun-verb pairs in Auslan, finding that not all signers made use of the same set of subtle differences in movement and other features sometimes used to distinguish signs referring to concrete objects from those used to indicate actions.

Similarly, Johnston & Schembri (2007) describe how Auslan signers have two major strategies available to them when producing sentences with indicating verbs. First, they may use an SVO constituent order of signs to represent actor versus patient roles (e.g., mother ask father ‘mother asks father’). Alternatively, they may convey this information by spatial modifications to the verb sign, using orders other than SVO (e.g., mother+left father+right left+ask+right ‘mother asks father’). The linguistic, stylistic and social factors that influence these types of choices have not yet been the focus of any research.

Like other SLs, both ASL and Auslan exhibit significant variation in the expression of subject. As part of the sociolinguistic variation in ASL and Auslan projects, variation in the presence of subject noun phrases was investigated in narratives (Lucas, 2001; Schembri & Johnston, 2006). The overall results of the two studies were remarkably similar: in Auslan almost two thirds (63%) of clauses had no overt subject noun phrase, almost identical to the figure in ASL (65%). Factors that conditioned an increased tendency to omit subject arguments included the use of a subject that identified a referent that was the same as the one in the immediately preceding clause; the subject having a non-first person referent; the use of role shift and the presence of some degree of English influence in the clause. In addition, in the ASL study social factors were relevant: women and older signers favoured overt subjects, whereas men and younger signers
did not. Unlike ASL, however, multivariate statistical analysis of the Auslan data suggested that social factors such as the signer’s age and gender were not significant.

Grammatical variation is not well understood because the linguistic factors that influence the types of variants (e.g., semantics, pragmatics) have not yet been the focus of much research in any signed language. They are also confounded by language contact and mixing with majority SpLs (and their written forms). Another major reason for the difficulty in researching grammatical variation has been the unavailability of large machine-readable corpora of SLs – a situation which is only now being addressed. Finally, the issue of grammatical variation cannot easily be separated from the broader issues of lexicalisation and grammaticalization, to which we now turn.

2. Lexicalization and grammaticalization in SLs
Linguistic variation may also be analysed with reference to lexicalization and grammaticalization. In this second perspective on variation, which is complementary to the sociolinguistic variation discussion in part one, the use in particular contexts of some specified phonological or lexical variants of individual signs or particular multi-sign constructions – called *synchronic contextual variation* (Heine, 2002) – is treated as indicative of differential stages of grammaticalization of these forms across a language community (Pagliuca, 1994).

2.1. Lexicalization in SLs
The signs uttered when communicating in a SL are not all of the same type. From one point of view – just as in SpLs – the conventionalised units of a SL can be divided into the two broad classes: an open class of content (or lexical) signs/words and a closed class of function (or grammatical) signs/words. Both these types of signs are roughly equivalent to the commonsense notion of *word* generally used to refer to the conventionalized free units of any language. Discovering the language-specific conventionalized semantic content of these signs is an empirical question, often requiring detailed textual analysis and fieldwork, but is nonetheless relatively straightforward.

From another point of view, however, there is a further word level distinction that needs to be made for SLs: a distinction between *fully-lexical* and *partly-lexical* signs. The need to make this second distinction stems from the fact that, unlike the phonemes of SpLs, the five basic formational components of signs in all SLs (handshapes, orientations, locations, movements, and non-manual facial expressions) can be individually meaningful as a function of iconicity or language-specific form-meaning conventionali-
larizations, or both. These components can each directly and componentially contribute to the meaning of a given sign form in predictable ways.

Partly-lexical signs are signs that, though conventionalized at the level of the meaningfulness of their components, do not have associated with them a meaning which is additional to or unpredictable from the value of those components when the sign is produced and used in various contexts. These types of signs have previously been called non-lexicalized signs (Johnston & Schembri, 1999) because they contrast with lexicalized signs whose meaning cannot simply be derived from that sign's form and/or its use in context. However, to avoid confusion of the term non-lexicalized or non-lexical sign (or word) with the term grammatical sign (or word) – in opposition to lexical or content sign (or word) – they are referred to here as partly-lexical signs in contradistinction to fully-lexical signs. In other words, a fully-lexical sign may be either a content word/sign or a function word/sign. Fully-lexical signs constitute the listable lexicon of a SL.

This basic observation regarding the degree of lexicalization of types of signs within SLs has been made many times within the literature, as Table 1 suggests. Naturally each researcher has focussed on different aspects of sign structure when identifying distinct types of signs so the two columns are not meant to imply the existence of two distinct or mutually exclusive categories of signs, nor that the labels identify the same entities in each column. What is of interest and relevance here is the widespread and repeated recognition of the existence of two broad categories of signs in SLs.

<table>
<thead>
<tr>
<th>Fully-lexical signs</th>
<th>Partly-lexical signs</th>
<th>Example citations</th>
</tr>
</thead>
<tbody>
<tr>
<td>“regular signs”</td>
<td>classifiers</td>
<td>Frishberg 1975</td>
</tr>
<tr>
<td>“regular signs”</td>
<td>classifier predicates</td>
<td>Liddell 1977</td>
</tr>
<tr>
<td>frozen signs</td>
<td>productive signs</td>
<td>Supalla 1978</td>
</tr>
<tr>
<td>fully specified signs</td>
<td>partially specified signs</td>
<td>Johnson &amp; Liddell 1986</td>
</tr>
<tr>
<td>lexical signs</td>
<td>non-lexical signs</td>
<td>Johnston &amp; Schembri 1999</td>
</tr>
<tr>
<td>standard signs</td>
<td>highly iconic structures</td>
<td>Cuxac 2000</td>
</tr>
<tr>
<td>plain verbs</td>
<td>indicating verbs</td>
<td>Liddell 2003</td>
</tr>
<tr>
<td>fully specified signs</td>
<td>depicting signs</td>
<td>Liddell 2003</td>
</tr>
</tbody>
</table>

Table 1:
Various categorizations that identify two major type of signs in SLs

Lexicalisation in SLs essentially occurs when a signed unit acquires a clearly identifiable and replicable citation form which is regularly and strongly associated with a meaning which is unpredictable and/or somewhat more specific than the sign's componential meaning potential, even
when cited out of context. Indeed, the meaning of the sign may or may not be quite unrelated to its componential meaning potential, i.e., it may be arbitrary. Of course, a sign may also be lexicalized ‘instantaneously’ through the assertion, and acceptance, of any form/meaning link within a linguistic community. However, an important observation is that the distinction between the two types of signs, whatever they may be called, is not so much categorical as gradient.

With respect to lexicalisation in sign languages, the notion of linguistic construction offers the opportunity to reconcile in one model these different types of signs and the processes by which gestures and highly iconic signs evolve into conventionalised lexical signs and thus, in turn, become available for grammaticalization. It is to this concept of constructions that we now turn.

2.2. Construction-based grammars and SLs
To summarize Diessel (2004), construction-based grammars take the Saussurian notion that the linguistic sign (word or morpheme) is a pairing of a specific form with a specific function and extend it to include grammatical (or multi-word) constructions. Constructions are the basic units of language. Importantly, constructions have general grammatical properties as well as idiosyncratic features. In this sense, they are not unlike the notion of construction in traditional grammar (e.g., the passive). The distinction is that in construction grammar, the notion has been generalized: all grammatical assemblies are constructions. Form comprises phonological, morphological and syntactic features; function includes semantic and discourse-pragmatic features. In other words, grammatical constructions are complex linguistic signs.

Moreover, construction grammar not only generalizes, it unifies. As in other functionally based theories of language and grammar, there is no categorical division between the lexicon and the grammar as ‘construction’ applies to all form-function pairs at whatever level in what is referred to as the grammar- or syntax-lexicon continuum, or simply lexico-grammar (Halliday, 1985; Langacker, 1998). Not only are words and grammatical constructions all linguistic symbols, there is a continuum between words and grammatical constructions, with idioms, famously, being the most obvious manifestation of this (Langacker, 2008).

Finally, rather than conceiving of linguistic rules as abstract and logically prior to their instantiation in actual utterances – a kind of algebra for combining words and morphemes, as in formal theories of language – the usage-based cognitive-functional approach which underpins construction
grammar conceives of linguistic constructions at all levels of structure as themselves meaningful linguistic symbols. In construction grammar, therefore, the productivity of language centres around the notions of pattern recognition (‘schemas’), frequency of usage, and analogy rather than rules the generate structures (Diessel, 2004; Tomasello, 2003).

The use of linguistic symbols in patterned ways involves constructions that can be differentiated along two continua: one of size or simplicity (from atomic to complex), and one of lexical specificity (from substantive to schematic or abstract) (Croft, 2001; Croft & Cruse, 2004). There is a continuum of constructions from highly canonical patterns, at one end, to highly idiosyncratic patterns, at the other end. Meaning derives both from individual symbols themselves (which would be characterised as substantive atomic constructions at the word or morpheme end of the continuum) and from the pattern or schema itself (complex and substantive or, at the most abstract, complex and schematic at the grammatical end of the continuum). There are also many types of mixed constructions. This is exemplified, for English, in Table 2.

<table>
<thead>
<tr>
<th>Construction type</th>
<th>Traditional name</th>
<th>English Examples</th>
</tr>
</thead>
<tbody>
<tr>
<td>Complex and (mostly) schematic</td>
<td>syntax</td>
<td>[SBJ be-TNS VERB—an by OBL]</td>
</tr>
<tr>
<td>Complex, substantive verb</td>
<td>subcategorization frame</td>
<td>[SBJ consume OBJ]</td>
</tr>
<tr>
<td>Complex and (mostly) substantive</td>
<td>idiom</td>
<td>[kick-TNS the bucket]</td>
</tr>
<tr>
<td>Complex but bound</td>
<td>morphology</td>
<td>[NOUN-s], [VERB-TNS]</td>
</tr>
<tr>
<td>Atomic and schematic</td>
<td>syntactic category</td>
<td>[DEM], [ADJ]</td>
</tr>
<tr>
<td>Atomic and substantive</td>
<td>word/lexicon</td>
<td>[this], [green]</td>
</tr>
</tbody>
</table>

**Key:** SBJ = subject, TNS = tense, OBL = oblique, OBJ = object, DEM = demonstrative, ADJ = adjective

Table 2: The syntax-lexicon continuum (Croft & Cruse, 2004: 255)

If we map the two dimensions in which constructions may be constituted (the atomic to the complex and the substantive to the schematic) on to the lexico-grammar of a SL such as Auslan we may be better able to account for the range of constructions that are possible, how they are interrelated, and the significance of their variable presence in the language. For example, from Table 3 we see that in many SLs the atomic units (i.e., the individual signs) are often not fully substantive, usually having one or more components or features unspecified.
<table>
<thead>
<tr>
<th>Construction type</th>
<th>Name or description</th>
<th>SL Examples (GLOSSES IN ITALICS)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Complex and (mostly) schematic</td>
<td>sign order or syntax</td>
<td>[Sbj VERB Obj]</td>
</tr>
<tr>
<td></td>
<td>(e.g., transitive plain verb)</td>
<td>(e.g., transitive plain verb)</td>
</tr>
<tr>
<td>Complex and (mostly) substantive</td>
<td>“traditional” idiom</td>
<td>[SORRY PRO-2 MISS TRAIN]</td>
</tr>
<tr>
<td>Complex and (mostly) schematic</td>
<td>depicting signs or ‘classifier’ handshapes</td>
<td>[DS(1):UPRIGHT.ENTITY.WITH.INHERENT.FRONT+PATH.LOCATION.MANNER.ORIENTATION]</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(no real citation form, save handshape)</td>
</tr>
<tr>
<td>Atomic and schematic</td>
<td>syntactic category</td>
<td>[PRO], [CONJ], [ADV], [NOUN], [VERB]</td>
</tr>
<tr>
<td>Atomic, schematic but bound</td>
<td>indicating verb (single ‘agreement’)</td>
<td>[BEGIN BODYANCHOR VERB-GOAL]</td>
</tr>
<tr>
<td></td>
<td>(citation form, fixed begin point, e.g., TELL)</td>
<td>(citation form, fixed begin point, e.g., TELL)</td>
</tr>
<tr>
<td>Atomic, schematic but bound</td>
<td>indicating verb (double ‘agreement’)</td>
<td>[SOURCE VERB-GOAL], [VERB+ORIENTATION]</td>
</tr>
<tr>
<td></td>
<td>(citation form 1st to 2nd person default, or variable)</td>
<td>(citation form 1st to 2nd person default, or variable)</td>
</tr>
<tr>
<td>Atomic and (mostly) schematic</td>
<td>layered or simultaneous morphology</td>
<td>[NOUN+LOC], [VERB+ASPECT], (e.g., plurals)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(e.g., aspect)</td>
</tr>
<tr>
<td>Atomic and (mostly) schematic</td>
<td>pointing or indexic signs</td>
<td>[PRO NON FIRST]+DIRECTION.LOCATION</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(no citation form, save handshape)</td>
</tr>
<tr>
<td>Atomic and (mostly) substantive</td>
<td>named indicating verb (direction and/or orientation unspecified)</td>
<td>[SOURCE.ORIENTATION=INVITE-GOAL.ORIENTATION]</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(has one or two ‘citation’ forms)</td>
</tr>
<tr>
<td>Atomic, (almost all) substantive verb</td>
<td>named indicating verb (very limited modification)</td>
<td>[SOURCE.PRO2 RETURN GOAL.PRO1]</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(has citation form)</td>
</tr>
<tr>
<td>Atomic and substantive</td>
<td>monomorphemic signs or fully-lexical signs (all values specified)</td>
<td>[PRO1], [BUT], [GREEN]</td>
</tr>
<tr>
<td></td>
<td>“lexical idioms”</td>
<td>(has citation form)</td>
</tr>
</tbody>
</table>

Key: DS = depicting sign, X-GLOSS = prefix or starting point of sign, GLOSS-X = suffix or end point of sign, GLOSS+X embedded simultaneous modification, X.X = two fused elements.

Table 3: Possible examples of the lexico-grammatical continuum in a SL (Auslan)

Moreover, there is a sense in which fully-lexicalized signs (constructions which are atomic and substantive) are ‘idiomatic’ in SLs in much the same way as multi-word idioms (constructions which are complex and substantive) can be idiomatic in SpLs and rarely so in SLs. (A common shared
feature of SLs is that most have been reported to have few genuine examples multi-sign idioms – idioms in the traditional sense – which are not calques from the surrounding SpL.) In other words, given the meaningfulness and/or iconicity of sign components, the idiosyncratic meaning of a fully-lexical sign suggests that lexicalization in SLs is similarly ‘idiomatic’. Lexical signs are in a sense idioms. It may well be that in SLs there is a large role for displaced or down-shifted idiomaticity within the linguistic system, rather than simply double articulation, in the creation of fully-lexical signs. This may have consequences for what is available as a source for grammaticalization in these languages.

2.3. Grammaticalization in SLs
Grammaticalization is a process of historical language change through which content words become grammatical words, clitics and ultimately morphemes, e.g., going to has changed to gonna when used as a future tense marker (both variants exist in modern English) (Givón, 2009; Hopper, 1998). Similarly, the Auslan sign bad-luck has acquired a context specific grammatical intensifying function (‘very’) when used immediately before or after adjectives or verbs, but only in the southern dialect (Johnston & Schembri, 2004). Its form in the two environments remains, as yet, unchanged. As another example, the modal can in BANZSL may be related to the lexical sign understand: very similar forms for both are used by some older BSL signers.

Grammaticalization also includes the process through which abstracted multi-word constructions become syntagms such that the pattern not only conveys meaning, it has grammatical consequences. For example, the order that the words/signs for an actor and undergoer appear relative to a verb may become fixed and meaningful in a language so that the order of, say, noun-1 verb noun-2, encodes the grammatical relations of Subject (noun-1) and Object (noun-2) respectively, in a nominative/accusative language, with or without the addition of other morphological markings such as case markers or agreement markers. This second process is sometimes also referred to as syntacticization (Givón, 1995).

Grammaticalisation theory is part of a broader cognitive-functional view of language in which language is understood to be the outcome of a long process of cultural evolution (Croft, 2000). Specifically, the grammatical dimension of language is regarded within this framework as emergent from frequency of usage and the result of historical and ontogenetic processes of grammaticalisation (Bybee & Hopper, 2001; Goldberg, 1995; LaPolla, 2006). Grammaticalization can most clearly be tracked
diachronically through the examination of written language. However, it can also be observed synchronically through cross-linguistic comparisons within language families, or within a language through variant word forms and multi-word constructions in particular contexts (*synchronic contextual variation*). In other words, the variation which is always present in language is rarely simply an index of social variables such as gender, class, etc, and the presence of similarly related variants of particular constructions often indicate on-going grammaticalization (Pagliuca, 1994), as with the case of the variable presence of *going to* and *gonna* in modern English, and variable forms of *can* and *cannot* in BSL.

With respect to SLs, the first stage in this general process of linguistic ‘emergence’ may even have its origins in gesture, rather than lexis (Johnston & Schembri, 2001; Wilcox, 2004). Some researchers have labelled this pre-grammaticalization pathway as *linguisticisation* (MacSweeny *et al*., 2008) – the process whereby some dimension of expressive communicative substance (*e.g.*, place of articulation or pitch and tone in vocalisation, or handshape, body location or facial expression in gesturing) becomes co-opted as a linguistic resource (*i.e.*, it becomes ‘emic’ within a communication system). Linguistic symbols or signs are then created that can themselves be lexicalised, much as we have described above, which then, in turn, become available for grammaticalisation as traditionally understood. In other words, it appears that grammaticalization pathways in SLs may also include a gestural substrate. We will not, however, pursue this topic further here and return to traditional grammaticalization pathways.

In the first part of this paper, we reported on research that has shown that variation in ASL and BANZSL does reflect regional, educational, social class, age and gender characteristics of its users. It is, however, more than likely that much observed variation in this language also reflects on-going grammaticalisation. The analysis of synchronic contextual variation is thus fundamental to fully understanding variation in SLs because SLs have no historical written texts, and few recordings or any significant historical depth. Consequently, these patterns of synchronic contextual variation can only be discerned by processing and analysing relatively large amounts of data and for this one needs machine-readable corpora with accurate metadata.

Until very recently these type of corpora have simply not existed for SLs so this type of research was impossible. In the past few years, however, a large machine-readable digital video corpus of BANZSL has begun to be created. Two Auslan corpora have already been created (a variation corpus and a general corpus), a NZSL variation corpus has been collected, and a
BSL corpus has begun to be collected. These corpora being enriched and made machine-readable through a long-term program of annotation. The existence of these corpora will provide an excellent case study of grammaticalization processes in SL because BANZSL has a rare, if not unique, attested historical depth and is used in three widely separated regions of the world.

With respect to synchronic contextual variation, for example, grammaticalisation theory predicts that one should find two types of related phenomena in a corpus of this type. First, there should exist pairs of sign variants in which one variant is clearly a lexical (or gestural) source and the other is an emerging or emergent grammaticalised form. Second, there should be evidence of (or indeed lack of evidence of) obligatory and arbitrary morpho-syntactic codings such as fixed orders and/or agreement or case morphology replacing clause-level constructions that are fully accounted for in terms of discourse, pragmatic and semantic principles, and iconicity.

In a corpus-based study of synchronic contextual variation, a specific linguistic feature in the corpus would be identified and tagged in the annotation, instances retrieved from the corpus, constructions in which they occur identified, and variations in form of target signs (e.g., phonological reduction) and meaning (e.g., a local construction-related sense) coded and analysed to reveal any change from lexical (or gestural) meaning to grammatical meaning. In other words, the research would identify and quantify features associated with the change of a lexical or content word into a grammatical or function word as described in grammaticalization theory: meaning loss, use in new contexts, loss of independent sign status, loss or reduction in form (Heine & Kuteva, 2002).

By way of example, a pilot study is being conducted on the Auslan corpus to collect evidence of the grammaticalization of forms and variants of lexical signs in the semantic area ‘finish’ (a well attested lexical source and grammaticalization pathway in SpLs, Heine & Kuteva, 2002). Approximately two hundred Auslan texts (narratives, recounts, conversations, etc.) are being annotated with unique identifying glosses (id-glosses, see Johnston, submitted) and linguistic annotations using ELAN, a multimedia annotation program (Hellwig et al., 2007). The annotations identify the linguistic environments (construction types) in which variants of certain

---

3. The creation of similar corpora of several European SLs (such as NGT, LSF, DGS and SSL) signals the beginning of a period of corpus-based research in SL linguistics which is certain to lead to a much better understanding of lexico-grammar and grammaticalization in these languages.
selected signs occur. In this study, clauses that express perfective aspect in some way are identified and annotated. All occurrences of signs expressing this meaning are being assigned an appropriate ID-gloss (e.g., FINISH, COMPLETE, HAVE, etc.) and any variation in their form annotated. ELAN is then used to view concordance patterns, extract collocations and colligations, and statistics on type/token frequency. These data are then exported to be manipulated in multi-variate analysis programs and then correlated with each other and with participants’ age, gender and region for evidence of the extent of grammaticalization. Manipulated corpus data of this type enables the identification of patterns that would otherwise be difficult or impossible to identify any other way. The hypothesis is that differential rates of grammaticalization across a geographically dispersed SL-using linguistic community are a significant part of the explanation with some regional variants correlating with different uses in different linguistic contexts, not just with the region itself.

3. Conclusion
In this article, we have explored some of the research conducted in the past few decades on sociolinguistic variation in deaf communities, with a particular focus on ASL, BSL, Auslan and NZSL. We have shown how, just as in SpL communities, variation is often not random, but is conditioned by social and linguistic factors. However, much work remains to be done as very little of these sociolinguistic corpora have been digitized and richly annotated using the latest multi-media annotation software such as ELAN. For example, other sociolinguistic variables need to be investigated, such as the use of mouth gestures, mouthing and fingerspelling, and stylistic factors need to be more fully explored. The influence of immigrant communities, and the impact of the many late learners and hearing and deaf second-language users on established SLs is also important.

We hope to have shown, however, that there is a complementary dimension to sociolinguistic variation that is only now opening up to systematic investigation to SL linguists through the use of large machine-readable corpora: the frequency and environments of use of particular variants as a manifestation of grammaticalization-in-progress. Pursuing such research questions will increase our knowledge about the sociolinguistics of SLs, as well as broaden our understanding of variation in language generally and in particular evidence of the extent of grammaticalization in SLs, given they are relatively young languages; and the similar lexical pathways or unique gestural pathways they exploit, given that they are face-to-face visual-gestural languages.
Though the origins of language remain unknown – and are in a fundamental sense unknowable – observations from the study of historical change in well documented languages, child language acquisition, pidginization and creolization, and the birth of SLs of deaf communities, increasingly reveal facts about variation and change in language that can help scaffold a hypothetical past which is more than mere speculation (Heine & Kuteva, 2007).

In sum, there is always variation in language, and variation may function as an index of social variables. In addition, however, some variation may actually reflect on-going language change, in particular lexicalization and grammaticalization. Only by taking all these observations into account can variation and change in SLs be properly understood.