Making an Online Dictionary of New Zealand Sign Language

2 authors:

Rachel Locker McKee
Victoria University of Wellington
65 PUBLICATIONS **295** CITATIONS

SEE PROFILE

David Mckee
Victoria University of Wellington
24 PUBLICATIONS **161** CITATIONS

SEE PROFILE

Some of the authors of this publication are also working on these related projects:

NZSL Vitality View project

Sociolinguistic Variation in NZSL View project

Making an Online Dictionary of New Zealand Sign Language*

Rachel Locker McKee *(rachel.mckee@vuw.ac.nz)*
and
David McKee *(david.mckee@vuw.ac.nz)*
Deaf Studies Research Unit, School of Linguistics and Applied Language Studies, Victoria University of Wellington, New Zealand

Abstract: The *Online Dictionary of New Zealand Sign Language* (ODNZSL),[1] launched in 2011, is an example of a contemporary sign language dictionary that leverages the 21st century advantages of a digital medium and an existing body of descriptive research on the language, including a small electronic corpus of New Zealand Sign Language. Innovations in recent online dictionaries of other signed languages informed development of this bilingual, bi-directional, multimedia dictionary. Video content and search capacities in an online medium are a huge advance in more directly representing a signed lexicon and enabling users to access content in versatile ways, yet do not resolve all of the theoretical challenges that face sign language dictionary makers. Considerations in the editing and production of the ODNZSL are discussed in this article, including issues of determining lexemes and word class in a polysynthetic language, deriving usage examples from a small corpus, and dealing with sociolinguistic variation in the selection and performance of content.

Keywords: SIGN LANGUAGE LEXICOGRAPHY, ONLINE DICTIONARIES, MULTIMEDIA DICTIONARIES, BILINGUAL DICTIONARIES, LEARNER DICTIONARIES, NEW ZEALAND SIGN LANGUAGE, VIDEO CONTENT, SIGN LANGUAGE CORPUS, POLYSYNTHETIC MORPHOLOGY, POLYSEMY, SOCIOLINGUISTIC VARIATION, SIGN LANGUAGE LINGUISTICS, USER PROFILE

Opsomming: Die maak van 'n aanlyn woordeboek van Nieu-Seelandse gebaretaal. Die *Online Dictionary of New Zealand Sign Language* (ODNZSL)[1] wat in 2011 bekend gestel is, is 'n voorbeeld van 'n hedendaagse gebaretaalwoordeboek wat gebruik maak van die 21ste-eeuse voordele van 'n digitale medium en 'n bestaande hoeveelheid beskrywende navorsing oor die taal, insluitende 'n klein elektroniese korpus van Nieu-Seelandse gebaretaal. Vernuwings in onlangse aanlyn woordeboeke van ander gebaretale het bygedra tot die ontwikkeling van hierdie tweetalige tweerigting-multimediawoordeboek. Video-inhoud en soekhoedanighede in 'n aanlyn medium is 'n groot vooruitgang om 'n gebareleksikon meer direk aan te bied en gebruikers op veelsydige maniere toegang te verleen tot die inhoud, maar los nogtans nie al die teoretiese uitdagings op waarvoor gebaretaalwoordeboekmakers te staan kom nie. Oorwegings by die redigering en totstandbrenging van die ODNZSL word in hierdie artikel bespreek, insluitende kwessies soos die

---

bepaling van lekseme en woordklas in 'n polisintetiese taal, die verkryging van gebruiksvoorbeelde uit 'n klein korpus en die hantering van sosiolinguistiese variasie in die keuse en werking van die inhoud.

Sleutelwoorde: GEBARELEKSIKOGRAFIE, AANLYN WOORDEBOEKE, MULTIMEDIA-WOORDEBOEKE, TWEETALIGE WOORDEBOEKE, AANLEERDERWOORDEBOEKE, NIEU-SEELANDSE GEBARETAAL, VIDEO-INHOUD, GEBARETAALKORPUS, POLISINTETIESE MORFOLOGIE, POLISEMIE, SOSIOLINGUISTIESE VARIASIE, GEBARETAALLINGUISTIEK, GEBRUIKERSPROFIEL

Introduction

Capturing the lexis of a signed language in a bilingual dictionary requires macro and micro decisions about issues of lemmatisation, ordering, variants, grammar, sense and usage. An even more fundamental challenge for sign language dictionary makers has traditionally been modality — attempting to describe a visual-gestural language in a static format. A decade ago, De Schryver (2003: 143) said this about electronic dictionaries:

> The arrival of the modern computer set in motion a series of lexicographers' dreams without equal in the history of dictionary making. Achieving the wildest of those electronic-dictionary vistas has the potential to result in reference works beyond all recognition.

In the 21st century, the electronic medium has indeed transformed vistas for makers and users of signed language dictionaries. The digital, online platform affords a vastly better fit between the dynamic nature of a signed language and possible dictionary formats. This article describes how the electronic medium was leveraged in making an online dictionary of New Zealand Sign Language (McKee et al. 2011). The aims, design and production of the dictionary (hereafter abbreviated as ODNZSL) are described, and some perennial issues for sign language dictionary makers are illustrated, including providing usage examples, determining word class and citation forms, and dealing with polysemy and mouthing. Sociolinguistic considerations of lexical variation and performing the lexicon on video are also discussed.

Aims and impacts of sign language dictionaries

The making of sign language dictionaries has been important in defining Deaf people as distinct language communities. In many countries, the production of a national sign language dictionary has been the foundational piece of language documentation leading to further attestation of the linguistic and cultural status of Deaf communities. Involvement in language documentation

activities raises consciousness of linguistic identity within the primary language community, and the artifact of a dictionary can support recognition of language minority status (Lucas 2003, Padden and Humphries 2005, Haualand and Allen 2009). Documenting the lexicon serves to distinguish the identity of a national Deaf community from — as well as aligning them with — signing communities in other countries or regions. However, promoting the impression of a unified national sign language via dictionary content can be misleading, and potentially prescriptive, in contexts where sign language use is actually locally variable (Lucas 2003; Reagan 2010; Schmaling 2012). In the African context for instance, Schmaling (2012: 273) observes that informants to national dictionary projects have often not been representative of the country's Deaf population, having been exposed to foreign forms of sign language via schooling and contact with overseas sign language users in development projects; as such, dictionary products may obscure the localised linguistic diversity in the wider spectrum of Deaf people who live outside capital cities and who are not formally educated. The *Dictionary of Southern African Signs* (Penn et al. 1992–1994) is noted by Schmaling as a work that purposefully addressed representativeness by documenting variants from at least eleven regional varieties identified within the South African Deaf community.

Sign language dictionaries create an avenue for outsiders to knowledge about the language. Common motives for sign language dictionary projects are to support the use of sign language as a medium of instruction in deaf education, and to increase Deaf people's participation in wider society through language recognition and the training of interpreters. Typical aims are captured in Ashipala et al.'s (1994, as cited in Lucas 2003: 325) statement about a Namibian sign language dictionary project: "We want people to know that NSL is a real language like Oshiwambo or English. We want Deaf Namibians to be proud of their Deaf culture and sign language, and we want hearing people to respect our culture and our sign language." In similar vein, the introduction to the print *Dictionary of New Zealand Sign Language* (Kennedy et al. 1997: ix) expresses goals of legitimising NZSL as a 'real' language through research and documentation, improving communication between Deaf people and hearing people by providing a bilingual learning resource, and helping to improve Deaf people's access to education and other social services.

ODNZSL (2011) is a third generation dictionary, following from two previous dictionaries that achieved initial documentation and language recognition. NZSL was made an official language of New Zealand by statute in 2006 (NZSL Act 2006), only 21 years after the first dictionary was made (Levitt 1986), and nine years after a larger dictionary appeared (Kennedy et al. 1997). Although a combination of policy, education, advocacy and research initiatives have all contributed to status change for NZSL, dictionary publication has played a key role in consolidating its public profile as a bona fide language used by a New Zealand community.

Development of NZSL dictionaries

The first NZSL dictionary was produced in 1986 by Dan Levitt, an American interpreter contracted to train the first cohort of sign language interpreters in New Zealand. In the absence of documented NZSL resources, language classes relied upon rosters of Deaf visitors as live language models. During class sessions, their signing was video-recorded and subsequently described and compiled in a dictionary. Levitt's photographic dictionary features many older signers, making it a valuable record of earlier signs that have since been replaced by modern variants. In making the 1986 dictionary, community members expressed a preference for photographic illustrations; they could not easily visualise signs disembodied from their users. The 1200 entries are alphabetically ordered by English gloss, in a uni-directional English–NZSL format. Variant forms are included. The Levitt dictionary promulgated the name 'NZSL' in the discourse of Deaf and hearing communities, and was an important artifact in bringing the existence of the language into the public domain.

The next dictionary (Kennedy et al. 1997, hereafter DNZSL) was a collaborative effort between the national Deaf Association and Victoria University. The project was led by Professor Graeme Kennedy, an editor of the *Oxford Dictionary of New Zealand English*, who had also supervised a PhD thesis describing the grammar and lexicon of NZSL (Collins-Ahlgren 1989). The methodology of the DNZSL underlies the content and record structure of the current ODNZSL, and is therefore outlined briefly here; (a more detailed account is found in the Introduction of DNZSL).

The 1997 dictionary contains 4,500 entries and was completed in six years, comprising 20,000 hours of work by paid staff, and uncounted hours of voluntary work by Deaf community members (Kennedy, p. xi). Kennedy took a systematic approach to content selection by establishing a taxonomy of semantic domains common to most languages, as well as a list of words (in English and Māori) pertinent to New Zealand life and to Deaf culture. Using this 'concept net' of approximately 7,000 words as a guide, Deaf groups were recorded conversing freely about suggested topics, and responding to lexical elicitation prompts. From the hours of resulting video recordings, signs were identified, re-filmed as citation forms, and then viewed by representative Deaf groups in three regions who attested whether the sign was recognised and/or used in their region, whether another variant existed, and what the sign meant in context. Validated signs were phonetically analysed and transcribed using the Hamburg Notation System (HamNoSys; Prillwitz et al. 1989), then drawn with the addition of movement symbols, and described in terms of their formation and usage, to create records for dictionary entries.

HamNoSys is a sign transcription system for describing the phonological parameters of a sign that were originally identified by Stokoe (1960) as handshape, place of articulation, movement, and orientation of the palm and fingers The set of HamNoSys symbols representing all permutations of these parame-

ters allows the phonetic features of any sign to be coded in a linear string, in the order of handshape, orientation, location (i.e. place of articulation), movement (as seen in figure 1). The DNZSL was one of few dictionaries at the time of publication to order its entries according to their handshape and location identity, rather than alphabetically by English translation (Zwitserlood 2010). DNZSL entries are grouped into 27 handshape categories, each of which includes up to eight related variants. A pictorial index of handshape categories, which also appear as section headers on dictionary pages, enables a user to identify the section in which a given sign should be found. Within each handshape section of the dictionary, signs are ordered by their place of articulation, starting from those made near the top of the head and moving progressively downwards in the signing space. Organisation by sign features is necessary to bi-directional access, allowing a user to search by the visual appearance of a sign if an English translation is not known. Privileging the formational features of signs in the dictionary layout also enhances users' receptive awareness about the form of signs, by having to attend to structural details that distinguish similar forms from each other. Nevertheless, this organisation is not familiar to new dictionary users, and can be challenging to use with precision. An English index at the back allows users to also search by English glosses. Figure 1 shows the elements and layout of an entry in the DNZSL.
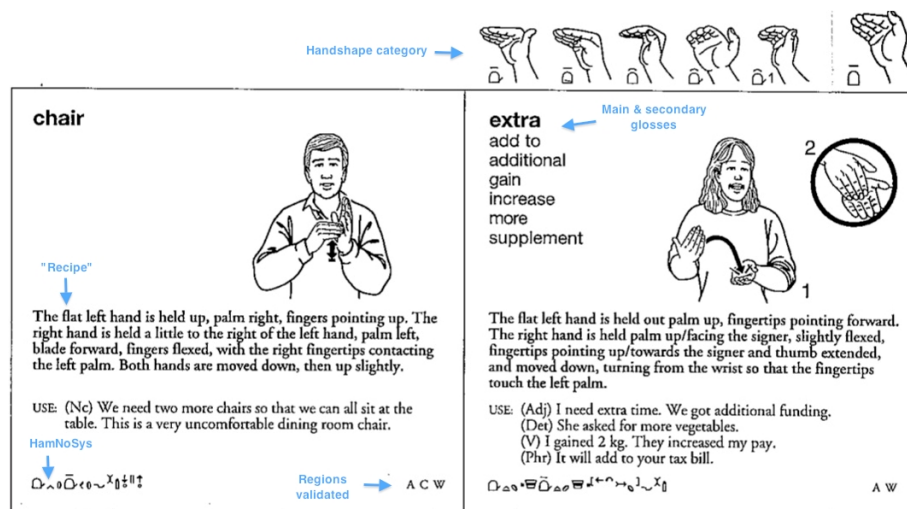

Figure 1: Entry structure in DNZSL (1997: 133)

"Online ease" factors

De Schryver (2003: 159) identified numerous advantages of electronic and online dictionaries described by lexicographers from 1980 onwards, and espe-

cially between 1990–2000. All of the 'online ease' factors that De Schryver lists are realised in the ODNZSL, and we comment on the relevance of these points (cited in italics) below.

"No space constraints, can handle huge quantities of data"
Migrating the existing NZSL dictionary database to an online interface allowed the addition of new data fields within records (such as sociolinguistic variation, semantic domain, usage examples) and enables the continuous addition of further entries. Schmaling (2012: 244) comments that while it is ideal for dictionaries to include all known variants, in practice this has often been unfeasible in sign language dictionaries because of limits on production resources and size. The Concise NZSL dictionary (Kennedy et al. 2002) containing the 2,000 most frequent signs was published in response to a need for a more portable volume than the original three-kilogram tome containing 4,500 signs on 800 pages. The online medium mitigates size and selection pressures, allowing for a more inclusive resource that is less constrained by the production, cost and usability parameters of a hard-copy text. According to Zwitserlood (2010: 454), most contemporary electronic sign language dictionaries have between 2,000–5,000 entries, and most print dictionaries have fewer than 2,000; so by current standards, both the print and online dictionaries of the NZSL are on the larger side, containing over 4,000 signs. Nevertheless, the ODNZSL represents only a portion of the established NZSL lexicon; its actual size is unknown, as is true for most signed languages (Johnston and Schembri 1999: 176). In light of the incomplete documentation of sign lexicons, online lexicons are vital to enabling ongoing description by future researchers.

"New types of information"
Moving to an online format allowed for re-design at all levels: the micro-structure of content within an entry, the macro-structure of dictionary organisation and user access to content, and the mega-structure of all elements in the dictionary (Hausmann and Wiegand 1989).

The micro-structure of entries was shaped by data fields established in the 1997 print dictionary; those that were transferred into the ODNZSL included a drawing, main gloss, secondary glosses, handshape category, grammatical and usage information. The ODNZSL entry components incorporated video clips of signs and usage examples (with English translation) and additional usage tags. The literal gloss of each example sentence not only shows the NZSL syntax, but allows users to click on any hyperlinked item in the sentence to go to the entry for that sign, providing cross-referencing.

The online format allowed for the addition of variation information (e.g., region, age, sign status) and hyperlinked tags explaining grammatical properties of signs. Since the website launch, Māori equivalents for all English main glosses have been added, to enable users to search for signs by a Māori word and to see lexical equivalence (or approximations thereof) between the three official languages of New Zealand.
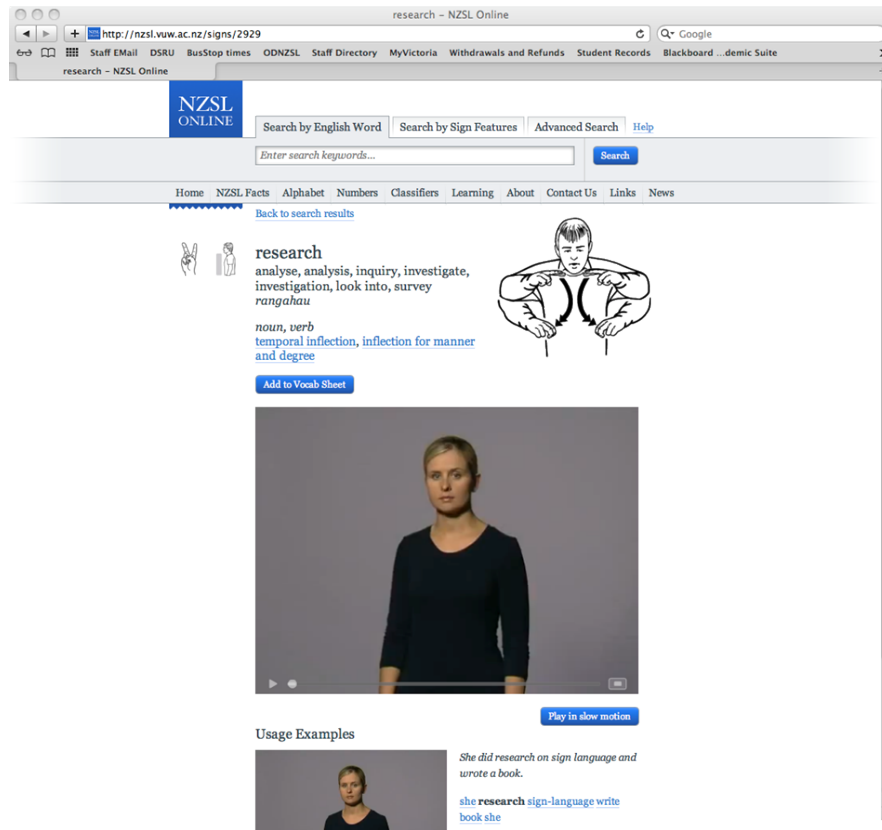
Figure 2: Entry display in ODNZSL (2011)

Relational searching enabled by the online medium considerably reduces the macro-structure issue of 'ordering'. Signs can be searched and displayed in ways that are more transparent than navigating a print volume organised according to phonological principles that are opaque to naïve users. Information about signs can be searched in various ways and combinations: via English word equivalents (which we also refer to in this article as main and secondary glosses), or by handshape and location features of a sign, or by 42 topic domains, or by tags for usage status (obscene, archaic, neologism, informal and rare). It was fortunate for the ODNZSL project that 1997 DNZSL editors made the (theoretically bold, at the time) decision to organise that dictionary according to sign form rather than alphabetically by English translations; hence, the immense task of phonetic analysis and coding of handshape and location features for each sign had been completed for all existing records. This HamNoSys coding was utilised as the basis for the 'search by sign features' in ODNZSL.
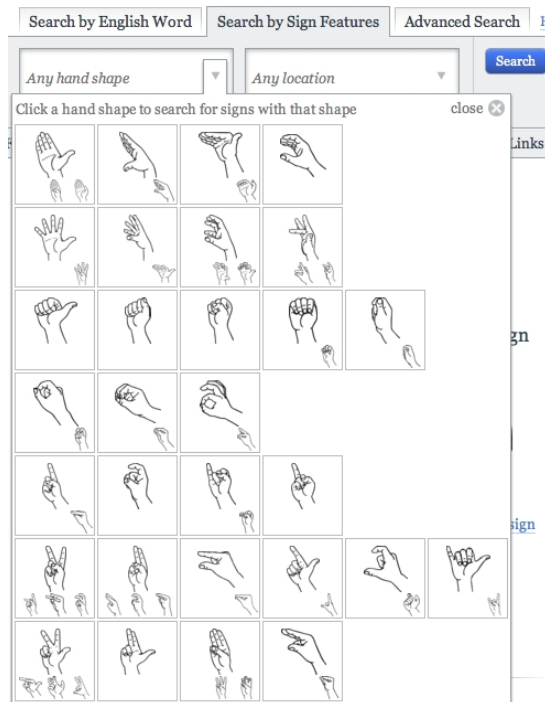
Figure 3: Handshape search menu

The dictionary needed to address both receptive and productive language needs of users, which requires being able to search by the visual form of a sign. Handshape and location features can be combined with topic tags in an advanced search, which is potentially helpful to a user who sees an unfamiliar sign used, and has an approximate knowledge of its form and domain, but not its meaning. Considering that users of the ODNZSL would be encountering a novel format for a dictionary, the design of 'search by sign' was informed by user consultation. As there is no equivalent of a conventional 'alphabet' for ordering signs, phonological features of handshape and location are used as the identifying parameters — which for users, is more analytically demanding than a 'search by word'. Handshape and location features need to be grouped and displayed in a logical manner in the 'search by sign' menus to facilitate user recognition of the most salient features. We examined recent online dictionaries of Danish Sign Language (Center for Tegnsprog 2008) and Finnish Sign Language (Kuurojen Liittory 2003), and their formats helped us to devise appropriate tables of NZSL handshapes and locations (see figures 3 and 4) that can be selected by clicking on pictures.
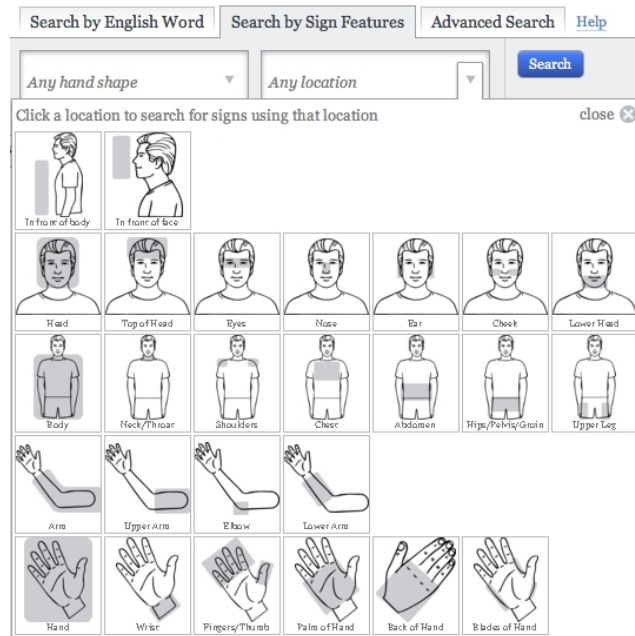
Figure 4: Location search menu

A handshape/location search yields a result list of forms that match the search parameters more or less closely; these are displayed by sign illustrations and hyperlinked glosses, which can be browsed directly from the results page, or refined further by returning to the sign feature menus. To determine the final format of these menus, adult NZSL students were given a pilot task to test how accurately they could identify salient handshapes and locations to find a set of target signs in the dictionary. Results indicated that users were able to make quite accurate choices, and pointed to cases of ambiguity which informed the final organisation of these menus (e.g., which handshapes to display at the top/category level, and which as sub-handshapes, and some that were not necessary distinctions for searching).

In its mega-structure, in addition to entry content, ODNZSL contains static information tabs that provide 'front' and 'back' matter about how to use the dictionary, history and grammatical aspects of NZSL, systems of numbers, fingerspelling, classifiers, attributions and acknowledgements, and links to NZSL-related organisations and learning resources. The front matter pages contain a video window that renders the information in NZSL translation, making the dictionary bilingual at the mega-structure level, and thus more accessible to Deaf NZSL users. A 'News' tab communicates relevant news about the dictionary to users, and offers a Facebook link, which has proven a useful means of dis-

seminating awareness of the dictionary. The homepage of the dictionary has a 'Show me a sign' display, which shows a sign illustration, gloss and entry hyperlink; these are randomly selected from within the database (excluding obscene signs after some early user feedback), and change at each log-in; this emulates a 'word of the day' feature in other online dictionaries.

"Video sequences, animation"
Signed languages are characterised by their visual-spatial production modality; transitions of movement and handshape are vital to distinguishing the form and meaning of signs. Analysing and representing these articulation features accurately in a dictionary is important both to the searchability of content, and to its value as a learning tool. The capacity for video content has been well utilised in recent online dictionaries of Flemish Sign Language (Van Herreweghe et al. 2004), Swedish Sign Language (Institutionen för Lingvistik 2009), Danish Sign Language (Center for Tegnsprog 2008), and Finnish Sign Language (Kuurojen Liittory 2003). Although the line drawings in the print NZSL dictionaries (1997 and 2002) were of high precision, it is transformative to be able to represent signs and their contextualised use through video clips. Learners particularly appreciate the slow motion playback, which allows the articulation of signs to be closely observed. The two-dimensional view that video affords required consideration of camera angles when filming certain signs to maximise the perceptible contrast in hand position or movement. Access to video information increases language learners' independence in building vocabulary knowledge. Whereas learners previously needed to ask a teacher or native speaker, "What is the sign for (…)?", to get a demonstration, they can now ask the dictionary for a (repeatable) video rendition of the vocabulary item. Lexical knowledge is also enriched by incidental discoveries in the dictionary about usage, variants and translations that a teacher might not have offered up. Finally, video content furnishes a repository of data for other researchers to refer to in cross-linguistic studies.

"Links with other software (e.g., learning exercises)"
The user survey indicated that learner exercises would be 'nice to have' in an online dictionary resource. ODNZSL contains a link to an interactive practice website for learners also developed at Victoria University. A vocabulary 'save and print' function allows users to select, caption and print customised sheets of sign illustrations from entries, which is valuable for learning and teaching purposes. Further learning links will be added as resources allow.

"Up-to-date and dynamic repository"
Following migration of print dictionary content to the online database, all records were checked and many were updated with clarifications of form, meaning, usage and grammatical tags. The capacity to continuously and rapidly update content is a great advantage (the server integrates any revisions

into entry display overnight), particularly as NZSL is still developing and being described. Community ownership of the dictionary is enhanced by the capacity to show a rapid editorial response to user feedback about content which is invited via an email link in the website. (Facebook and YouTube options for user feedback are currently being explored.)

"Cheap, if not free"

For both practical and principled (social impact) reasons it was decided to make ODNZSL openly available without registration or subscription. Free access ensures that the primary language community of Deaf people has equitable access to it as a resource, and can readily refer others to it. Open access softens the borders of traditional (and historically private) domains of sign language transmission, such as deaf schools, homes, clubs and community events, and more recently, NZSL classes. Its availability on the web and the portability of the medium helps to normalise it in the linguistic landscape. Importantly, NZSL becomes more available as a resource to hearing families and educators of deaf children who are not socially connected to a Deaf community. Since the launch in June 2011, informal feedback reveals an unexpectedly wide variety of users, including workmates of Deaf individuals, service providers in customer contact positions, teachers of various subjects who incorporate NZSL as an additional modality into classroom teaching, 'baby sign' classes for parents of hearing infants, schools, and even a children's rugby coach who consults it for translating instructions to a deaf player. These adventitious uses of the dictionary in wider society indicate the extended reach of a free, online resource.

The dictionary carries a Creative Commons license,[2] which permits content to be shared and adapted, with attribution, for purposes such as developing learning resources. Under this license, an independent software developer has created a free application that allows key elements of the dictionary to run on mobile devices, and is now working on a flashcard application. Enabling such initiatives extends the educational potential of the dictionary and harnesses a wider pool of creative expertise than our research unit has at its disposal. As lamented by many publishers of electronic reference tools, however, the downside of being free is the absence of sales revenue; securing funding for maintenance and development of the dictionary promises to be more difficult than attracting initial project funding.

Steps in making ODNZSL

The ODNZSL project was the initiative of the Deaf Studies Research Unit at Victoria University, working in partnership with educational and community stakeholders. A grant of NZD$750,000 for a three-year project was obtained from a government funding source.[3] The editorial team comprised a managing editor, two consulting editors, a database manager, and two research assistants.

All personnel worked on the project on a part-time basis. None had formal lexicography training, but most had previous relevant experience with sign linguistics, dictionary and corpus projects. A technical editor (database programmer), illustrator, translators and web designers were also contracted.

Effective dictionary design proceeds from a realistic understanding of the purposes and skills of users (Atkins and Rundell 2008). The ODNZSL project team included NZSL teachers and interpreters who brought experiential knowledge of likely user needs; we knew that the dictionary should serve both receptive and productive uses, and expected that the most frequent use would be language learners seeking NZSL equivalents for English words. Although we had a good sense of the dictionary audience, user needs were investigated at the outset of the project by a written survey of people in the following categories: teachers and teacher aides of Deaf children; adult learners of NZSL (in tertiary and community education); NZSL teachers (mainly Deaf); Deaf community members; NZSL interpreters; deaf students in schools; mainstream school students and teachers; parents/families of deaf children/adults (mainly non-Deaf). 301 survey responses were received, informing decisions about functionality and content that would meet the expectations and needs of the majority of users.

The first step was to migrate the 1997 DNZSL data to an online database in which content could be edited and re-formatted. Fortunately the ODNZSL technical editor worked on both projects, which expedited this transition. The lexicographical tool designed by the technical editor was Freelex,[4] an open source application that enables multi-user, online workflow, editing of data fields within records, quality assurance (QA) management, corpus search, and reports on user activity and record history.

Following migration, all record data was checked and updated as needed. New records were added, including further variants, neologisms and borrowings that had developed in NZSL in the 20 years since data was collected for the 1997 dictionary. In relation to a changing lexicon, it is relevant to note that during the past 25 years, provision of interpreting services and language recognition has lowered the barriers for NZSL users to participate in domains of higher education, white-collar occupations, the arts, media, and political advocacy. Deaf people's greater access to information and societal discourses has driven rapid growth in the NZSL lexicon, through coinages and borrowing from Australian, American and British Sign Languages.

All candidate signs to be added to the ODNZSL were validated by community groups in three regions. Although ODNZSL was intended as a general rather than a specialised dictionary, the two Deaf Education Centres were invited to submit signs relating to school curriculum subjects that they had already documented, for potential inclusion in the dictionary. In fact, many of the items submitted were very new coinages, nonce translations for English terms that had only contextually recognisable meaning, or coinages for concepts that could be translated by semantic extension of conventional signs

already in the dictionary (listed under an alternate English translation). As many of the suggested items apparently had marginal lexical status in NZSL, (some used only within a particular school or class), or already had conceptual equivalents in the established lexicon, approximately half of the suggested 'school signs' did not pass validation in a process of cross-school, and adult community consultation. Approximately 100 were added to the dictionary. For each new validated addition, a citation form was video-recorded for the purpose of making an illustration, and data fields for the record were populated.

During the revision of the whole database, some existing records from the previous dictionary failed QA for various reasons (e.g., duplicates or nearly identical phonetic variants) and are not displayed in the current ODNZSL. The total number of entries displayed in ODNZSL is 4070 with a further 2176 undisplayed records in the database. The total number of records in the database is 6,246.

The next main phase was to film citation forms and usage examples for each sign. The process of making usage examples from corpus data is discussed in more detail later in this article. A script for filming was created by typing example sentences in gloss form. Signs and usage examples (between one and four examples per entry) were filmed in five 'batches' of five-day blocks, over an 18-month period. Editing, uploading and QA of the clips, and further filming preparation was undertaken between batches of filming. Film personnel comprised a rotation of eight signing models (working in pairs for any given shoot), a cameraman, an assistant to log film shots, a director (the managing editor) to monitor accuracy of performance and consistency with dictionary illustrations, and a coach from the dictionary team to rehearse the signers. Over 11,000 clips were recorded, edited, quality assured, saved on a server in .mp4 format and linked as 'assets' to records in the database.

Each filmed usage example was translated into English, displayed beside the video window in the entry display. Below the translation in the entry is a literal syntactic gloss of the sentence in NZSL. To enable cross-reference to surrounding signs in the usage example, each sign gloss in the gloss sequence was tagged with the ID number of its corresponding record in the database, creating a hyperlink to the entry for that sign in the online dictionary. This feature was modeled on the Danish Sign Language Dictionary (Center for Tegnsprog 2008).

A web design company worked closely with the editing team to create a user interface for the website, and to integrate this with the database search engine. Ideas about features and functions of the website were drawn from review of recent online sign language dictionaries, in particular, the Danish and Finnish (Kuurojen Liittory 2003) dictionaries. At the final phase, a range of users (first and second language signers) trialed a set of search tasks within the website, and provided feedback on problems encountered; this feedback was used to tweak functionality prior to launch.

Corpus-based evidence

De Schryver's final 'online ease' factor is *"Rapid access to lexicographical evidence in corpora"*. The availability of electronic corpora has revolutionalised the sourcing of lexicographical evidence for dictionaries of major spoken languages. The majority of sign language dictionaries, however, have lacked any usage information let alone contextualised examples based on attested use. Content selection is often based on native speaker introspection or consultation rather than upon systematic analysis of a large body of data (Johnston 2003; Wilcox 2003; Zwitserlood 2010). This is a potentially unrepresentative basis for a dictionary, as perceptions about language use often differ from actual usage (Atkins and Rundell 2008).

The chief constraint on corpus-informed dictionaries of signed languages is the practical difficulty of developing machine-searchable corpora for visual languages without a conventional written form.[5] Rapid advances have been made recently in sign language corpus development, with the advent of software that enables searchable transcription and annotation of time-linked video files. ELAN (EUDICO Linguistic Annotator), originally developed for research on gesture, has become the standard tool for this, as it allows one to create, edit, visualise and search multi-tier annotations for video and audio data (Johnston and Schembri 2005; Leeson and Nolan 2008). A sophisticated multimedia relational database, 'iLex', has also been developed at the University of Hamburg for the purpose of sign corpus annotation and lexicography (Hanke and Storz 2008). Nevertheless, transcription of signed lexemes is done via a written gloss, and/or at sub-lexical level via a phonetic notation system — a hugely time intensive analytic process that poses issues such as the determination of lemmas, lexemes, and phrasal units in a signed language (Johnston and Schembri 1999), and the representation of non-manual, spatial and productive elements of the lexicon in use (Miller 2001; Brennan 2001). Sign language corpora are now being developed in numerous countries, although the scale remains relatively small in comparison to spoken languages, due to the labour intensive nature of transcription and the small number of people who have the analytic skills to do this work.[6]

Although it does not provide direct user access to a corpus, the ODNZSL is one of few signed language dictionaries to have utilised an electronic corpus as the main source of usage examples. The corpus is small and flawed in several respects, but its use is a step forward in providing lexicographical evidence for a sign language dictionary. Following the 1997 print dictionary of NZSL, the general editor, Graeme Kennedy, decided to create a first corpus of NZSL in order to analyse lexical frequency to guide compilation of a concise dictionary, and to inform teaching curricula (McKee and Kennedy 2006). The Wellington Corpus of NZSL (WCNZSL) consists of more than 50 hours of videotaped NZSL discourse by members of the Deaf community. 80 Deaf people

were videotaped taking part in 30 conversations and 43 monologues in dis-
course events that included conversations, narratives, speeches, meetings, dis-
cussions on more than 30 different topics, including school and education,
sport, work, health, childbirth, and politics. Over one year, videotapes were
transcribed into English glosses, creating a file consisting of 100,000 running
signs, comprising 7,222 types or lemmas (McKee and Kennedy 2006).

An example of the transcribed corpus is shown in figure 5. By contempo-
rary standards, this format is primitive in that the transcript is detached from
the original video sources and lacks annotation; it is also imperfectly lemma-
tised, which demands completely consistent glossing of every type and token
(see Johnston 2008). Non-manual syntactic markers and inflectional verb mor-
phology expressed in movement and space are not captured (in other words,
just 'bare' signs are represented). However, the text form of the corpus allowed
the use of a concordance tool, Wordsmith, to analyse lexical distribution (the
findings of which are reported in McKee and Kennedy 2006). Analysis revealed
that only 15 highly frequent signs (in the top 500) were missing in the original
1997 DNZSL, confirming that its coverage was quite representative. In the sub-
sequent Concise DNZSL (2002) the most frequent 2,000 signs were supple-
mented with less frequent, but nevertheless, important signs, such as names of
countries and cities, and vocabulary relevant to educational or health domains.

IX-2 KNOW IX-2 FAMOUS HORSE IX-loc AUSTRALIA IX-loc 1 MAN
FROM HERE HORSE-TROTTING NAME fs-SHANUE fs-DYE PRAM
KNOW IX-2 FAMOUS IX-3 MELBOURNE MELBOURNE CUP
MELBOURNE CUP WIN FIRST SCL-1-horse-leads FIRST 3 YEAR PAST 3
YEAR PAST IX-3 POS1 FRIEND IX-3 IX-1 LONG-TIME-AGO SMALL-
CHILD IX-3 FATHER IX-3 fs-WAS POS1 FRIEND IX-3 IX-1 IX-3 NMS-
nod IX-3 IX-3 BORN IX-3 IX-3 FATHER SISTER IX-3 FAMILY AREA
OLD GOOD TOGETHER GOOD EACH-OTHER IX-3 IX-3 fs-SON IX-3

Figure 5: Extract from Wellington Corpus of NZSL

Between 2005–2008, a further 14,000 tokens from 81 conversational data clips,
recorded in a sociolinguistic variation project, were added to the corpus. This
data was transcribed using ELAN annotation software, a screenshot of which is
shown in figure 6.

Text files from the two projects were combined to form a corpus linked to
the ODNZSL database. Figure 7 shows the appearance of this corpus within the
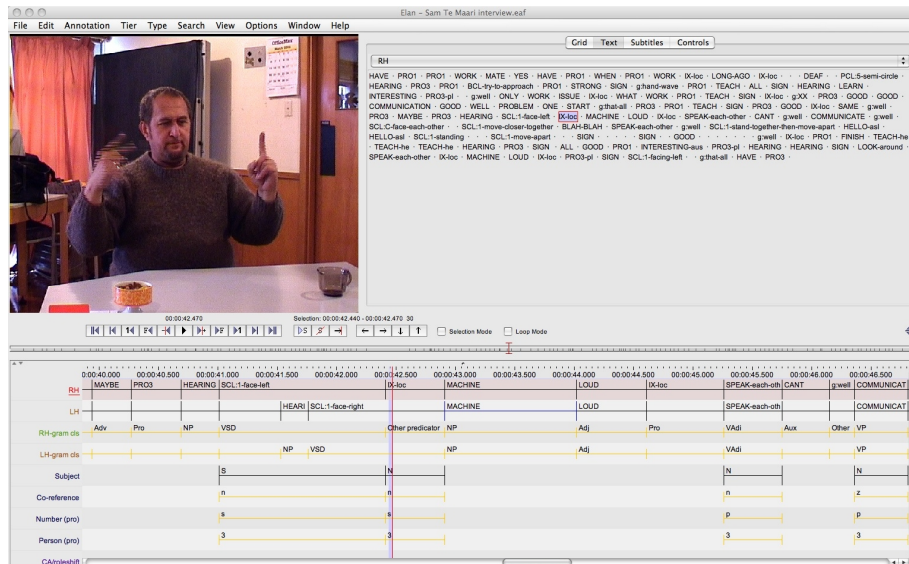editing database.

Figure 6: NZSL transcription and annotation in ELAN

As Atkins and Rundell (2008: 56) acknowledge, "Corpus creation is a pragmatic exercise, with compromise at all phases — design, data collection, and encoding." Although this corpus of 114,000 tokens is certainly compromised by its smallness (by spoken corpora standards) and imperfect sampling, it nevertheless represents a wide range of topics and speakers engaged in authentic discourse, and as such, provides valuable usage information about the lexicon. Corpus data informed the consideration of questions such as: Is the sign FIGHT used metaphorically as well as literally? Does the sign AFFAIR have a nominal or only a verbal sense? Which forms of HOT collocate with human vs. non-human subjects? In the editing process, the corpus was a valuable resource in making decisions about sense, word class, age and region-related usage (from metadata on signer identity), and provided the basis for example sentences, as explained in the next section.

The 1997 and 2002 NZSL print dictionaries include example sentences to illustrate sense and grammatical usage, but these are presented only in English, and so it is unclear whether they derive from authentic contexts of use in NZSL. Nor do they directly show a grammatical context in NZSL. It was therefore an advance for the ODNZSL to include filmed usage examples that are more authentic and informative with regard to both form and content. As the dictionary does not include definitions (meaning is indicated through English translation, sometimes supplemented by notes), usage examples are important for illustrating contextual features of both meaning and use.
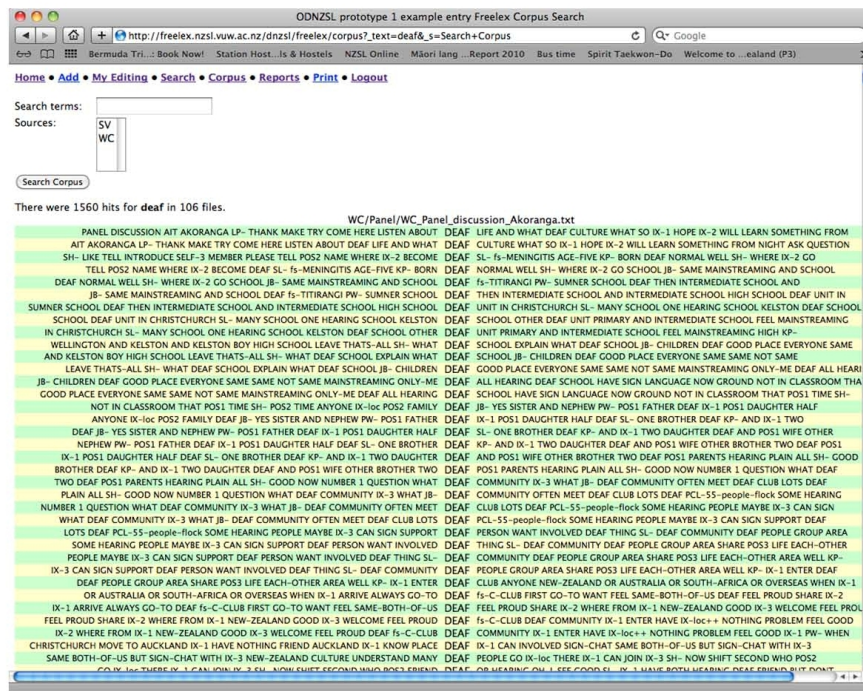
Figure 7: NZSL corpus in Freelex database (showing concordances of DEAF)

To create usage examples, the corpus was searched for tokens of each head-word, and up to three utterances were selected as the potential basis for examples. Some utterances in the corpus could be used verbatim, but most required editing. Our experience concurred with Atkins and Rundell's (2008) advice that most examples in a learners' dictionary will need to be modified from raw data in order to be accessible and maximally informative to various users. We adopted Atkins and Rundell's guiding principles that usage examples should be: (i) *Natural & Typical*: the word is shown in its most usual context, syntax and collocation; not idiosyncratic usage; not mixing registers or varieties; (ii) *Informative*: the sentence gives informative context (helps understand the sense of the word). (iii) *Intelligible*: the sentence contains no words that are more difficult than the headword; clear structure; succinct.

Our procedure for creating examples from corpus citations included removing identifying references to person, place and event, and offensive matter that was not essential context for the target word. In light of intelligibility for learners, extraneous clauses or unusual contexts that did not clarify typical sense or grammatical use were removed. As a general rule, each headword illustrated in an example needed to appear in an uninflected form, whereas it might have been inflected in the corpus data. Examples also had to be decontextualised

enough to stand alone as an intelligible utterance, which is not always the character of discourse among interlocutors who share real-world experience. Citations that required implicit knowledge of the immediate or wider conversational context were discarded or re-worked to make comprehensible examples.

With learners in mind, we tried to ensure that most signs within examples were relatively high-frequency and appear in the dictionary, enabling consistent cross-referencing. The imperative for clarity, and the fact that each example would have to be performed to camera and later translated into English, also led us to avoid certain structures particular to signed languages, including the following: 'constructed action', a quotative device for enacting reported speech or action, which can be convoluted to translate 'classifier' (polycomponential) constructions that depict objects and actions local to the discourse context, which are potentially wordy to translate, and not possible to cross-reference to dictionary entries as they utilise productive rather than conventional lexicon (Johnston 2001).

Due to the nature of certain signs, we made some exceptions to the guidelines; for instance, it is impossible to illustrate a usage context for an imperative such as 'Stop it!', or an exclamation such as 'You-idiot', without embedding it in constructed dialogue. Certain signs typically collocate with other signs that are low frequency, colloquial, or offensive. In these cases, we favoured the principle of 'natural, typical' usage. In some cases, examples were modified in the filming studio by the editor and sign model, if it became apparent that a sentence was too complex, or lacked sense out of its original context.

The translation phase (post-filming and QA of video examples) revealed some complications that in hindsight could have been minimised; for example, we wished that we had been more strict in applying the rule articulated for the Danish dictionary (Kristoffersen 2010) of avoiding spatially-encoded anaphoric pronominal reference, or null person reference (pro-drop). These structures are typical in signed language grammar, but difficult to deal with in English translation, which allows little ambiguity of person reference. Also not ideal were examples in which the sense of the target sign in NZSL, in the context of the example, had to be translated as a different word form than the main gloss displayed (e.g. main gloss — 'equal', while translation had to be 'equality'). In such cases, we made use of a secondary gloss in translation.

The goal of deriving all usage examples closely from corpus data was not achievable, as the Danish Sign Language dictionary makers also reported (cf. Kristoffersen 2010). We met corpus-related limitations: almost 50% of headwords were not found in the corpus, due to its limited size and some inconsistency of type labels for signs in the corpus. Another problem was that variants or synonyms for a single meaning were not consistently identified by distinct glosses in the corpus, making it impossible to reliably identify different sign forms in the corpus. Conversely, some polysemous or homophonous sign forms were glossed inconsistently according to meaning in context, rather than by a uniform label (see Johnston 2008, 2010 re. ID glossing issues in signed corpora). Unfortunately,

these transcription flaws became more apparent years after the fact, when trying to utilise the corpus in a dictionary. In the end, the usage examples reflect "pragmatic compromise" (Atkins and Rundell 2008), being corpus informed, but modified in the interests of achieving accessibility for users, and performability and translatability. Approximately 30% of examples were devised from scratch by a native speaker in the editorial team.

Nonetheless, a noteworthy benefit of corpus-based examples in the ODNZSL is their reflection of the Deaf cultural frames that underpin the lexicon in use. Fillmore (2003: 284) explains frames as the "background set of beliefs, practices, institutions, or ready-made conceptualizations available to the speakers of the language as the necessary underpinnings of the way they speak". The way signs are used by Deaf people may invoke frames that are particular to their shared experiences and understandings. For instance, in ODNZSL the usage example for 'family' is glossed in NZSL as, MY FAMILY ALL HEARING ME ONE DEAF (translation: "My family are all hearing. I'm the only Deaf person"). This statement is common in Deaf personal narratives, invoking mutually understood implications of this isolated status with regard to family relationships, language acquisition, and social identity. The example also shows a frequent collocation of the signs 'one/only' and 'Deaf' in NZSL users' description of themselves in social contexts with hearing people. Similarly, the entry for the sign 'oral/speaking' illustrates a distinctly Deaf worldview; the example is glossed as, HE ORAL SPEAK GOOD CAN HE (translation: "He can speak well"). While the translation is understandable on face value, it cannot convey the implicit reference to a social categorisation of Deaf individuals into those who speak more intelligibly or more often, and might therefore have a stronger affiliation with the hearing world, in contrast to those who don't, and may identify (and be identified) more strongly as NZSL community members. Culturally rich examples such as these are a valuable element of ODNZSL content. Nevertheless, we agree with Fillmore's (2003: 269) argument that dictionaries could usefully go further in making explicit the background frames that underpin native use of lexical items, especially for non-native users wanting to understand why a word exists and why a speaker would choose to use it in a context — in effect, providing ethnographic reference. Augmenting dictionary entries with encyclopedic-type definitions, where relevant, is not unfeasible in an online format (given 'no space constraints'), and is an interesting prospect for future development.

Performance of usage examples and sociolinguistic identities

Issues of identity and social structure often surface in lexicographical work on minority languages (Axelrod et al. 2003). Producing video content for the ODNZSL highlighted sociolinguistic considerations about form and style of language. Video clips are embodied linguistic performance by individuals who have social identities and styles that are immediately recognisable to their own

language community, and which will be perceived by learners and other dictionary users as prototypical models of the target language. Decisions had to be made regarding who should demonstrate the language: what is the ideal, or acceptable, embodiment of gender, age, ethnicity, regional origin, native language status, body stature, hair colour/style, sexual orientation? Given the impact of these personal factors on signing style and user perception of normative language models, we aimed for authenticity and diversity of representation, accepting the fact that there is neither a neutral nor ideal speaker of a language. To cover as many bases as possible with regard to sociolinguistic identities and style, the eight sign models comprise a mix of younger/older, native/non-native, female/male, Māori/non-Māori/Pasifika, northern/southern region individuals. Metalinguistic awareness of phonetic and grammatical form was also necessary in the signing models, as was the ability to perform scripted language 'naturally' on cue, in front of a camera for hours at a time. These requirements, and the level of scrutiny that the dictionary models would attract, narrowed the pool of eligible candidates for the job.

Also considered in filming was whether lexical variants more associated with certain social groups should be performed by a signer with matched identity characteristics. Although content of usage examples was not authored by the sign models, its animation (Goffman 1981) by a particular person may be perceived by viewers as representing a connection with the visible identity of the signer. For practical reasons, we only attempted to match variant to signer characteristics in the case of signs used mainly by older or younger people, signs with Māori cultural reference, and certain gender sensitive signs (e.g., anatomical and sexuality related).

Descriptive challenges

Word class tagging

Determining grammatical category (part of speech) in signed lexicons deeply challenges lexicographers and linguists (Johnston and Schembri 1999; Wilcox 2003, Zwiterserlood 2010). Indeed, Schwager and Zeshan (2008: 514) describe the dearth of systematic analysis of word class in signed languages as "shockingly neglectful". There are, however, methodological and theoretical reasons for ambiguity on this topic. In a typological survey of part of speech across signed languages, Schwager and Zeshan (2008: 513) describe "a substantial amount of systematic ambiguity or vagueness in many sign languages. For instance, in Indo-Pakistani Sign Language (IPSL) many signs tend to have rather general meanings that are narrowed down by the context of the utterance, and since many grammatical categories can remain unmarked, a sequence of signs can be difficult to categorise structurally, although its meaning is entirely clear semantically." This description applies to NZSL, in which a single sign may express various semantic roles; this is partly to do with the

motivated nature of many signs. For example in the ODNZSL, the sign glossed as BUILD (figure 8) may also translate in context as 'builder, carpenter, construct, nail' — senses that are motivated by the action of using a hammer, and are distinguishable by semantic and grammatical context, and/or by mouthing of a corresponding English word. In agreement with Johnston and Schembri's (1999: 150) conclusion that separate entries for nominal and verbal uses of signs could not be justified in an Australian Sign Language dictionary on the basis of form, many entries in ODNZSL have multiple word class tags and a correspondingly varied set of English word class translations.

Word class information displayed in ODNZSL is based on semantic and morphosyntactic evidence from usage in the corpus data, editorial intuition about the primary character of a sign as verbal, nominal, adjectival, etc., and to some extent, its potential equivalence in English. Since these judgments are not based on thoroughly principled analysis according to morphosyntactic criteria (as proposed by Schwager and Zeshan 2008, for example), we consider word class information to be only indicative about usage. Another source of word class ambiguity in signed languages relates to the polycomponential nature of many signs, as we discuss below.
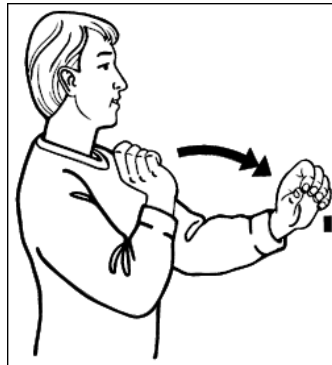

Figure 8: 'Build'

What qualifies as a lemma?

Defining signs, lemmas and lexemes for dictionary purposes is a problem previously examined by sign lexicographers (Brien and Turner 1994; Johnston and Schembri 1999; Brennan 2001). A key problem in defining the lexical status of signs is the polysynthetic capacity of signed languages to simultaneously combine meaningful elements in manual (hands), non-manual (face and body) and spatial dimensions to produce contextualised meanings. For example, figure 9 shows a sign conventionally glossed as 'reverse (a vehicle)'; this is actually a

polycomponential construction, or phrase, comprising several elements: a pro-form for 'vehicle' (right hand), a pro-form for 'ground' (left hand), a movement path of the vehicle's action (backwards), and manner (puffed cheek, indicating intensity of the agent's effort or focus). Although this construction occurs regularly and has a conventional meaning associated with it, in situated use, path and manner elements of the sign might vary to reflect details of a particular reversing event. Although this is a relatively straightforward example, such productive constructions do not easily fit morphological and lexical criteria for a fully lexicalised 'word' (Emmorey 2003; Schembri 2003).
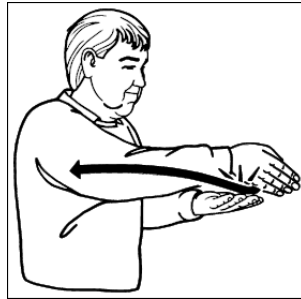

Figure 9: 'Reverse (a vehicle)' — a polycomponential construction

Johnston and Schembri (1999: 126) argue that many well-formed 'signs' observed in discourse, particularly polycomponential constructions, have reference local to a specific context but not established in wider use, and therefore have dubious status as entries in a dictionary. They state:

> A lexeme in Auslan is ... a sign that has a clearly identifiable citation form which is regularly and strongly associated with a meaning which is (a) more specific than the sign's componential meaning potential, especially when cited out of context, and/or (b) quite unrelated to its componential meaning (i.e., lexemes may have arbitrary links between form and meaning).

On the other hand, many productive (polycomponential) and morphologically inflected sign forms acquire lexical status through frequency of use and thus warrant inclusion in a dictionary (Brennan 2001). As Johnston and Schembri also note in their discussion of lexical status of signs, demonstrating semantic equivalence between languages can be an important objective of a bilingual dictionary that supports the inclusion of productive and inflected forms that are important for second-language dictionary users to discover.

Lexicographers of polysynthetic languages regularly debate the merits of entering verbs and nouns by their stem or by their regularly inflected forms. In relation to the morphologically productive African language isiZulu, for example, Prinsloo (2011) evaluates the arguments for lemmatisation by word forms that include regular affixes, versus by stems. He concludes overall that the

word approach is more accessible to most users, as searching by stem requires considerable morphosyntactic knowledge to deconstruct the components of complex forms that are encountered in texts as 'words'. Similarly, Kosch (2013) suggests that the decision for any given dictionary should be guided by consideration of users' expectations and grammatical knowledge.

In our editorial team's consideration of whether commonly inflected verb forms should appear as entries, the native signers (who are also sign language teachers) tended to be more in favour of their inclusion than the non-native signers, who were more inclined to separate lexicon from grammar. However, in light of the value for second language learners of seeing frequent and semantically stable uses of productive verb morphology, we did include some inflected forms that are regularly associated with distinct English translations. Figures 10 and 11 illustrate such 'sets' of inflected entries, based on the verbs LOOK and WALK, respectively.
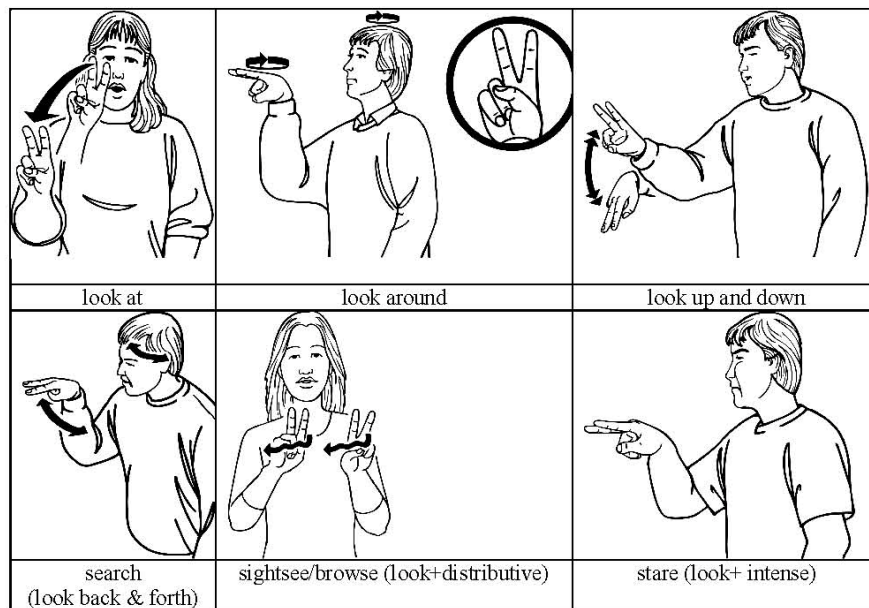

Figure 10: Entries for conventionally inflected forms of 'Look'

During filming of headword signs, sign models often found it difficult to produce certain signs (usually verbs and adjectives) devoid of inflectional morphology (e.g., manner, degree, location), suggesting that while these features are, in theory, optional, in natural usage they are closely fused with the lexicon. Arbitrating on citation form for each item in a dictionary has something in common with the orthography decisions required in making dictionaries of

languages with an oral tradition (cf. Axelrod et al. 2003), and is a consultative process. As Axelrod et al. observe, native speaker perspective in this process is vital, but does not always guarantee agreement on judgments about 'accurate' or 'neutral' forms. We acknowledge the potential for ongoing editorial revision in this regard.
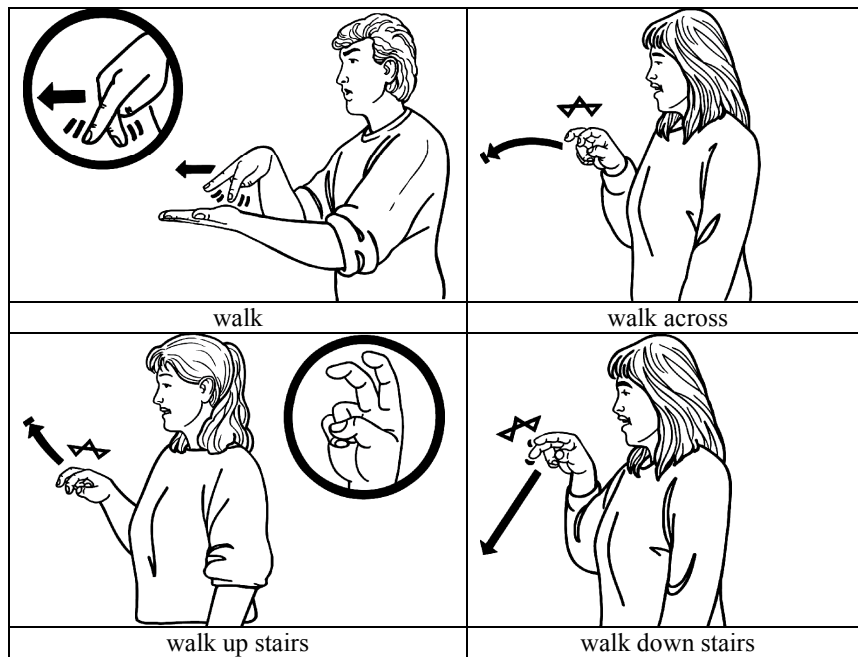

walk	walk across
walk up stairs	walk down stairs

Figure 11: Entries for conventionally inflected forms of 'Walk'

Polysemy, mouthing and functional equivalence

NZSL has much polysemy and some homophony in its lexicon, which requires decisions about how to present multiple form–sense relationships. In addition to semantic context, an important means of disambiguating meanings of polysemous and homophonous signs in NZSL is the simultaneous mouthing of spoken English words, which stems from Deaf people's exposure to that lexicon in their everyday lives. Co-articulation of mouthed words with manual signs occurs in many signed languages, and the status of mouthing in relation to a sign lexicon is the subject of theoretical debate, which is beyond the scope of this article (e.g., Boyes-Braem and Sutton-Spence 2001, Ebbinghaus and Hessman 1996). Suffice to say that mouthing is significant in lexicographical decisions about polysemous signs, yet is not often discussed explicitly in relation to the practicalities of making of sign language dictionaries. An exception

is the online Danish Sign Language Dictionary project, which outlined their rationale (Kristoffersen and Boye Niemalä 2008) for deciding whether to treat word mouthings that are regularly associated with signs at the level of phonological form, or at the semantic level of contextualised meaning. Kristoffersen and Niemalä explain that if mouth patterns are treated as a formational element of a sign (i.e., as a stable, and contrastive 'phonological' feature), the result would be multiple dictionary entries for one manual form paired with alternate mouthings. For example, a manual sign with the prototypical meaning 'house' might have seven entries corresponding with seven dwelling-related Danish words that are regularly mouthed to extend the semantic range of this manual sign. On the other hand, if mouthing is treated at the level of meaning (i.e., not a conventional part of its form), the result is to make a single entry for a polysemous sign, with a listing of frequent senses that are expressed in combination with mouth patterns. Native speaker consultants on the Danish dictionary favoured treating mouthing at the level of meaning, as did the NZSL native speaker editors (for the mostpart); i.e., one sign that is typically produced with several semantically related mouthings has a single entry.[7] An example in the ODNZSL is the entry for the sign BUILD (figure 8), which lists related senses that might be specified in discourse by mouthing and/or semantic context. As not all secondary glosses displayed in an entry are necessarily typical mouthings (but rather, alternate English translations), this is not an entirely satisfactory solution for informing users how signs are potentially combined with mouthing — which varies considerably in situated discourse.

A practical difficulty with representing polysemous items in ODNZSL is that video clips require a decision as to which, if any, mouthing the signer should show with the citation form of the headword. Our compromise was to film such citation forms with no mouthing (although this is not very typical in NZSL), and to show contextually appropriate mouthing in the usage examples. For instance, the entry LEARN/SCHOOL (secondary glosses — examination, lesson, schooling) (http://nzsl.vuw.ac.nz/signs/5724) displays a citation form with no mouthing, but two usage examples with natural mouthing of 'learn' and 'school' respectively.

Homophonous signs — a single manual form that has unrelated senses (and correspondingly unrelated word mouthings) — are treated as separate forms. For example, signs for LIVE (http://nzsl.vuw.ac.nz/signs/565) and TRUE (http://nzsl.vuw.ac.nz/signs/440) are homophonous, and are displayed as separate entries. This departs from the usual practice of organising a lexicon by contrasting forms, rather than meanings, thereby entering a headword once with all senses (related and unrelated) listed. Given the bilingual and pedagogical aim of this dictionary to present form–meaning correspondences for mainly English-speaking users, and considering search and display possibilities in the online medium (i.e., more 'space'), it seemed preferable to display distinct senses of homophones as individual entries. All analytic decisions about how signs should be arranged in a dictionary are, of course, theoretically

debatable; our decisions were influenced by native speaker intuition about the form–meaning correspondences, and by consideration of users' ability to navigate content without the need for extensive explanatory text, or over-cluttering the layout of individual entries. A resulting flaw in the ODNZSL macrostructure is that identical polysemous and homophonous forms which appear in separate entries are not cross-referenced to each other, which should be rectified in future.

Sociolinguistic variation: recording lexical variation and change

As unwritten languages of limited diffusion, signed languages tend to have considerable lexical and sub-lexical variation, which may be both socially patterned and idiosyncratic. Natural variation has often been intentionally or inadvertently disregarded in the making of sign language dictionaries, especially those produced by non-linguists for standardising purposes (Johnston 2003; Schmaling 2012). While lexical variation and innovation is not generally problematic for sign language users, it can be challenging for dictionary makers to capture, and sometimes for users (both native and non-native speakers) to accept in documented form. As Collins-Ahlgren (1994) commented about preparing an earlier dictionary of NZSL, whatever is contained between the covers will inevitably be regarded by second language users as the prescribed standard, and by the primary language community as an approved subset of the language — which may be contentious among members who have not contributed directly to its content or do not see their lexical preferences reflected. It is for such reasons of face validity of process and content that the preface to the DNZSL (Kennedy et al. 1997) lists, by name and region, the 90-plus Deaf informants who participated in elicitation and validation processes.

Sign language dictionaries have a role as a repository of evidence of language change (Wilcox 2003). This archival function was taken into account in making the ODNZSL and its predecessor; both document diachronic and synchronic variation in NZSL by including lexical and phonological variants. A search for the verb 'to die' for example, yields four variants, and three variants for the numeral 'nine'. Research on sociolinguistic variation (McKee and McKee 2011) shows that the lexicon of older signers is distinct from that of younger signers and that regional variation exists. From 1979, an Australian-based signing system was introduced into deaf education, which re-lexified a substantial proportion of NZSL from that time onwards. Variants recorded in the ODNZSL reflect both external and internal processes of change. External influences include vocabulary introduced from above through the education system, and from below via spontaneous borrowing by Deaf individuals travelling and living abroad, and via immigrants in the NZSL community. Internal change and innovation is prompted by new cultural referents and experiences, and in some cases by phonological and morphological processes. An example of morphological change is seen in early and modern variants of the concept

'open-minded' — originally a two-sign calque from the English phrase (shown in figure 12), which has evolved by relocating the sign OPEN to the location of MIND (forehead), thus incorporating two morphemes into a form that can be regarded as a compound lexeme (shown in figure 13).
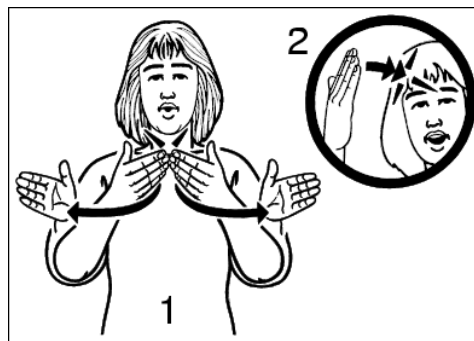

Figure 12: 'Open minded' — earlier form

Figure 13: 'Open minded' — contemporary form (morphologically reformed)

The boundaries of sign language lexicons are quite permeable compared to most spoken languages; individual borrowing and coining are commonly accepted strategies for filling lexical gaps. Editorial calls thus had to be made regarding when a foreign sign (usually from Auslan, ASL or BSL) qualifies as an established borrowing that should appear in a NZSL lexicon, and at what point neologisms are used widely and regularly enough to warrant an entry (e.g., recent coinages for 'Facebook', 'upload'). In practice, it was not possible to formulate rules to resolve these questions, which are a matter of varying perception. Our course of action was to include all signs that were recognised by the editorial team and community informants to be in use more widely than

one social grouping or region of country, and especially if they fill a lexical gap in NZSL (such as 'digital', or 'culture') rather than competing with an established sign. In the case of borrowings, language origin is displayed in the entry only when it is clear; however there are many untagged signs that have almost certainly been borrowed into NZSL by earlier generations, or by secondary routes (e.g., American and Irish signs that have entered NZSL via Auslan borrowings), making their origin now opaque. In the case of neologisms, users can provide feedback via the online feedback form and the Facebook page, which may inform future editorial revisions.

Conclusion

The ODNZSL illustrates the potential of the online medium to enable a more dynamic and authentic representation of a signed language lexicon and to offer users an interactive reference tool. A key element that enriches both descriptive and instructive value is the addition of video content, in addition to line drawings of signs, both of which can be downloaded for offline use.

The relational nature of online searching circumvents macrostructure issues of how to order entries and facilitates bi-directional searching, via the visual features of signs, or by word equivalents, topic domains or other tags. The 'search-by sign' function (using handshape and location features) promotes receptive use by learners who want to identify unknown vocabulary; this feature, and NZSL translation of front matter, also increases accessibility of the dictionary to Deaf NZSL users.

Editorial advantages of the online medium include the capacity to edit records easily, with overnight updating of the database, and to maintain a collaborative workflow in a multi-user team. New records can be created as new data comes to hand, and the microstructure of entries can potentially be adapted to display different fields from records in the database.

The electronic medium does not resolve longstanding lexicographical issues for sign language researchers, such as: deciding what qualifies for entry status (i.e., treatment of signs that are marginally lexicalised); determining morphologically unmarked citation forms; assigning word class; achieving lexico-pragmatic equivalence between the words of different languages (in this case NZSL, English and Māori); accounting for mouthing; and representing variation (in content selection, and in its embodied performance). The medium does, however, promote consultative and recursive editing processes, and allows for revision of content decisions in light of new evidence and ideas.

Dictionary content was not generated directly from an electronic corpus, but a transcribed corpus of NZSL informed selection, senses, and usage information. Although the format and scale of the corpus used to support the ODNZSL has been surpassed by recent advances in sign language corpus methods, the development of usage examples from a body of authentic language data is a significant advance in sign language dictionary making practice. Reference to corpus evidence in illustrating the use of signs not only

informs about typical linguistic contexts, but also offers glimpses into the cultural discourse of the NZSL community.

User feedback indicates that the availability of a user-focused online dictionary has widened societal access to knowledge about NZSL. It is hoped that this tool will support positive social change for the NZSL community, as well as contributing to improving the standard of contemporary online sign language dictionaries.

Notes

1.   NZSL Online is found at: http://nzsl.vuw.ac.nz/.
2.   See: http://creativecommons.org/licenses/by-nc-sa/3.0/.
3.   The ODNZSL project was funded by the New Zealand Tertiary Education Commission's 'Encouraging and Supporting Innovation' Fund, 2008–2011.
4.   Freelex is Open Source lexicography software by Dave Moskovitz. Download at: http://www.matapuna.org/.
5.   Several writing systems have been created to capture the phonetic or lexical form of signs, but their use tends to be limited to linguists or educators who adopt them for technical or pedagogical purposes. They are not widely known by signers. See Zwitserlood (2010) for a recent summary of sign 'writing' systems.
6.   An overview of recent developments in sign language corpus work is found in Crasborn et al. (2008). Also see the ongoing work of the Sign Linguistics Corpora Network hosted at Radboud University, Nijmegen: http://www.ru.nl/slcn/.
7.   Boyes-Braem (2001) reports the opposite decision by Deaf consultants for a lexical database of Swiss-German Sign Language, in which identical manual signs with contrasting mouthings are treated as separate entries. A different approach to polysemous signs with multiple mouthings is taken in the iLex project which is building a lexicographical database from a corpus of German Sign Language (DGS): the manual form is treated as the type (or lemma), and combinations of the manual form with different German word mouthings are treated as sub-types (lexemes) of the lemma. In this analysis, distributional properties of the manual sign can be analysed either independently, or in its various relationships with mouthing components in discourse (Reiner Konrad, 2012 Sign Corpus Linguistics Summer School, University of Hamburg, unpublished lecture notes.)

References

Ashipala, S.P., D. Profelius, M.N. Haikali, N. Israel, F.T. Linus, H.H. Nilenge, T.F. Haiduwah and R.N. Hashiyanah. 1994. The Development of a Dictionary of Namibian Sign Language. Erting, C.J., R.C. Johnson, D.L. Smith and B.D. Snider (Eds.). 1994: 342-346.

Atkins, B.T.S. and M. Rundell. 2008. *The Oxford Guide to Practical Lexicography.* Oxford/New York: Oxford University Press.

Axelrod, M., J. Gómez de Garcia and J. Lachler. 2003. The Role of Literacy and Collaboration in Documenting Native American Languages: A Report from the Jicarilla Apache Dictionary Project. *Sign Language Studies* 3(3): 296-321.

Boyes-Braem, P. 2001. A Multimedia Bilingual Database for the Lexicon of Swiss German Sign Language. *Sign Language and Linguistics* 4(1–2). *Special Issue: Sign Transcription and Database Storage of Sign Information*: 133-144.

Boyes-Braem, P. and R. Sutton-Spence (Eds.). 2001. *The Hands are the Head of the Mouth: The Mouth as Articulator in Sign Languages.* Seedorf/Hamburg: Signum Verlag.

Brennan, M. 2001. Encoding and Capturing Productive Morphology. *Sign Language and Linguistics* 4(1–2). *Special Volume: Sign Transcription and Database Storage of Sign Information*: 47-62.

Brien, D. and G. Turner. 1994. Lemmas, Dilemmas and Lexicographical Anisomorphism: Presenting Meanings in the First BSL–English Dictionary. Ahlgren, I., B. Bergman and M. Brennan (Eds.). *Perspectives on Sign Language Usage: Papers from the Fifth International Symposium on Sign Language Research* 2: 391-407. Durham: The International Sign Linguistics Association (ISLA).

Center for Tegnsprog. 2008. *Ordbog over Dansk Tegnsprog.* (The Danish Sign Language Dictionary.) http://www.viittomat.net/http://www.tegnsprog.dk/.

Collins-Ahlgren, M. 1989. *Aspects of New Zealand Sign Language*. Unpublished Doctoral Thesis. Wellington, New Zealand: Victoria University of Wellington.

Collins-Ahlgren, M. 1994. Issues in Preparing and Presenting a Dictionary of New Zealand Signs. Erting, C.J., R.C. Johnson, D.L. Smith and B.D. Snider (Eds.). 1994: 335-341.

Crasborn, O., T. Hanke, E.D. Thoutenhoofd, I. Zwitserlood and E. Efthimiou (Eds.). 2008. *Construction and Exploitation of Sign Language Corpora. Proceedings of the 3rd Workshop on the Representation and Processing of Sign Languages.* Paris: ELRA. http://www.lrec-conf.org/proceedings/lrec2008/workshops/W25_Proceedings.pdf.

De Schryver, G.-M. 2003. Lexicographers' Dreams in the Electronic-Dictionary Age. *International Journal of Lexicography* 16(2): 143-199.

Ebbinghaus, H. and J. Hessman. 1996. Signs and Words: Accounting for Spoken Elements in German Sign Language. Edmonson, W. and R. Wilbur (Eds.). 2008. *International Review of Sign Linguistics*. *Volume 1*: 23-56. Mahway, NJ: Lawrence Erlbaum Associates.

Emmorey, K. (Ed.). 2003. *Perspectives on Classifier Constructions in Sign Languages.* Mahwah, NJ: Lawrence Erlbaum Associates.

Erting, C.J., R.C. Johnson, D.L. Smith and B.D. Snider (Eds.). 1994. *The Deaf Way: Perspectives from the International Conference of Deaf Culture*. Washington D.C.: Gallaudet University Press.

Fillmore, C.J. 2003. Double-Decker Definitions: The Role of Frames in Meaning Explanations. *Sign Language Studies* 3(3): 263-295.

Goffman, E. 1981. *Forms of Talk.* Oxford: Basil Blackwell.

Hanke, T. and J. Storz. 2008. iLex — A Database Tool for Integrating Sign Language Corpus Linguistics and Sign Language Lexicography. Crasborn, O., T. Hanke, E.D. Thoutenhoofd, I. Zwitserlood and E. Efthimiou (Eds.). 2008: 64 -67.

Haualand, H. and C. Allen. 2009. *Deaf People and Human Rights*. Helsinki: World Federation of the Deaf.

Hausmann, F.J. and H.E. Wiegand. 1989. Component Parts and Structures of General Monolingual Dictionaries: A Survey. Hausmann, F.J., O. Reichmann, H.E. Wiegand and L. Zgusta (Eds.). 1989–1991. *Wörterbücher. Ein internationales Handbuch zur Lexikographie / Dictionaries. An International Encyclopedia of Lexicography / Dictionnaires. Encyclopédie internationale de lexicographie*: 328-360. Berlin: Walter de Gruyter.

Institutionen för Lingvistik. 2009. *Svenskt teckenspråkslexikon*. (Swedish Sign Language Dictionary.) www.ling.su.se/teckensprakslexikon.

Johnston, T. 2001. The Lexical Database of Auslan (Australian Sign Language). *Sign Language and Linguistics* 4(1–2). *Special Issue: Sign Transcription and Database Storage of Sign Information:* 145-169.

Johnston, T. 2003. Language Standardization and Signed Language Dictionaries. *Sign Language Studies* 3(4): 431-468.

Johnston, T. 2008. Corpus Linguistics and Signed Languages: No Lemmata, No Corpus. Crasborn, O., T. Hanke, E.D. Thoutenhoofd, I. Zwitserlood and E. Efthimiou (Eds.). 2008: 82-87.

Johnston, T. 2010. From Archive to Corpus: Transcription and Annotation in the Creation of Signed Language Corpora. *International Journal of Corpus Linguistics* 15(1): 104-129.

Johnston, T. and A. Schembri. 1999. On Defining Lexeme in a Signed Language. *Sign Language & Linguistics* 2(2): 115-185.

Johnston, T. and A. Schembri. 2005. The Use of ELAN Annotation Software in the Auslan Archive/Corpus Project. Paper presented at the Ethnographic E-research Annotation Conference, University of Melbourne, Victoria, Australia, 15–16 February 2005.

Kennedy, G., R. Arnold, S. Fahey and D. Moskovitz. 1997. *A Dictionary of New Zealand Sign Language*. Auckland: Auckland University Press with Bridget Williams Books.

Kennedy, G., D. McKee, R. Arnold, P. Dugdale, S. Fahey and D. Moskowitz (Eds.). 2002. *A Concise Dictionary of New Zealand Sign Language*. Wellington: Bridget Williams Books.

Kosch, I.M. 2013. *Expectation Levels in Dictionary Consultation and Compilation.* Paper presented at the Seventeenth Annual International Conference of the African Association for Lexicography (AFRILEX), which was hosted by the Department of African Languages, University of Pretoria, Pretoria, South Africa, 2–5 July 2012.

Kristoffersen, J.H. 2010. *From Utterance to Example Sentence: A Lexicographic Approach to Exploitation of Corpus.* Poster presented at Sign Linguistics Corpora Network (SLCN) Workshop 4: Exploitation. Deaf Centre, Friedrichstraße 12, 10969, Berlin, 3–4 December 2010. Retrieved electronically 31 July 2012, from: http://www.ru.nl/slcn/workshops/4_exploitation/workshop-posters/.

Kristoffersen, J.H. and J. Boye Niemalä. 2008. How to Describe Mouth Patterns in the Danish Sign Language Dictionary. De Quadros, R.M. (Ed.). 2008. *Sign Languages: Spinning and Unraveling the Past, Present and Future. TISLR9, Forty Five Papers and Three Posters from the 9th. Theoretical Issues in Sign Language Research Conference, Florianopolis, Brazil, December 2006*: 230-238. Petrópolis/RJ. Brazil: Editora Arara Azul. http://www.editora-arara-azul.com.br/EstudosSurdos.php.

Kuurojen Liittory. 2003. *Suvi Suomalaisen viittomakielen verkosanakirja*. (Finnish Sign Language Dictionary.) http://suvi.viittomat.net.

Leeson, L. and B. Nolan. 2008. *Digital Deployment of the Signs of Ireland Corpus in Elearning.* Unpublished conference paper, Language Resources and Evaluation Conference, Marrakesh, Morocco, June 2008.

Levitt, D. 1986. *Introduction to New Zealand Sign Language.* Auckland: National Foundation of the Deaf.

Lucas, C. 2003. The Role of Variation in Lexicography. *Sign Language Studies* 3(3): 322-340.

McKee, D. and G. Kennedy. 2006. The Distribution of Signs in New Zealand Sign Language. *Sign Language Studies* 6(4): 372-390.

McKee, R. and D. McKee. 2011. Old Signs, New Signs, Whose Signs? Sociolinguistic Variation in the New Zealand Sign Language Lexicon. *Sign Language Studies* 11(4): 455-527.

McKee D., R. McKee, S. Pivac Alexander, L. Pivac and M. Vale (Eds.). 2011. *Online Dictionary of New Zealand Sign Language*. Deaf Studies Research Unit, Victoria University of Wellington, New Zealand. http://nzsl.vuw.ac.nz/nzsl.

Miller, C. 2001. Some Reflections on the Need for a Common Sign Notation. *Sign Language and Linguistics* 4(1–2). *Special Issue: Sign Transcription and Database Storage of Sign Information*: 11-28.

Padden, C. and T. Humphries. 2005. *Inside Deaf Culture.* Cambridge, Mass.: Harvard University Press.

Penn, C., D. Ogilvy-Foreman, D. Simmons and M. Anderson-Forbes. 1992–1994. *Dictionary of Southern African Signs for Communicating with the Deaf.* 5 volumes. Pretoria: Human Sciences Research Council.

Prillwitz, S., R. Leven, H. Zienert, T. Hanke and J. Henning. 1989. *HamNoSys version 2.0. Hamburg Notation System for Sign Languages: An Introductory Guide.* Hamburg: Signum.

Prinsloo, D.J. 2011. A Critical Analysis of the Lemmatization of Nouns and Verbs in isiZulu. *Lexikos* 21: 169-193. Online at: http://lexikos.journals.ac.za/pub/article/view/42.

Reagan, T.G. 2010. *Language Policy and Planning for Sign Languages.* Washington DC: Gallaudet University.

Schembri, A. 2003. Rethinking 'Classifiers' in Signed Languages. Emmorey, Karen (Ed.). 2003. *Perspectives on Classifier Constructions in Sign Languages*: 3-34. Mahwah, NJ: Lawrence Erlbaum.

Schmaling, C. 2012. Dictionaries of African Sign Languages: An Overview. *Sign Language Studies* 12(2): 236-278.

Schwager, W. and U. Zeshan. 2008. Word Classes in Sign Languages: Criteria and Classifications. *Studies in Language* 32(3): 509-545.

Stokoe, W.C. 1960. Sign Language Structure: An Outline of the Visual Communication Systems of the American Deaf. *Studies in Linguistics: Occasional Papers No. 8.* Buffalo: University of Buffalo: Department of Anthropology and Linguistics.

Van Herreweghe, M., S. Slembrouck and M. Vermeerbergen. 2004. *Digitaal Vlaamse Gebarentaal–Nederlands/Nederlands–Vlaamse Gebarentaal woordenboek.* http://gebaren.ugent.be.

Wilcox, S. 2003. The Multimedia Dictionary of American Sign Language: Learning Lessons about Language, Technology and Business. *Sign Language Studies* 3(4): 379-392.

Zwitserlood, I. 2010. Sign Language Lexicography in the Early 21st Century and a Recently Published Dictionary of Sign Language of The Netherlands. *International Journal of Lexicography* 23(4): 443–476.