ISOcat Data Categories for Signed Language Resources

Onno Crasborn[1] and Menzo Windhouwer[2]

[1] Radboud University Nijmegen, Centre for Language Studies
o.crasborn@let.ru.nl
[2] Max Planck Institute for Psycholinguistics, Nijmegen
menzo.windhouwer@mpi.nl

Abstract. As the creation of signed language resources is gaining speed world-wide, the need for standards in this field becomes more acute. This paper discusses the state of the field of signed language resources, their metadata descriptions, and annotations that are typically made. It then describes the role that ISOcat may play in this process and how it can stimulate standardisation without imposing standards. Finally, it makes some initial proposals for the thematic domain 'sign language' that was introduced in 2011.

Keywords: signed language resources, metadata, data categories, standardization.

1   Background and Context

Sign languages are unwritten languages world wide, despite some efforts to create a writing system called 'SignWriting' [1]. This is adopted by small groups of enthusiasts and occasionally by primary or secondary schools. Even in those cases, there is not yet a written tradition to speak of. There are few books and no newspapers in sign language, and few deaf people make their shopping lists in sign language. This is due in part to the fact that deaf people, if they go to school, learn to read and write in the surrounding spoken language, but also to the rapidly increasing technological possibilities of video. Where text messaging on mobile phones has since some time now been an attractive medium for deaf people, video connections over mobile phones are gaining popularity.

The implication for researchers is that recording language use requires making video recordings. Time-consuming manual annotations are necessary to make such recordings machine-readable, just as with speech. Where automatic speech recognition can at least provide a first pass in the transcription of speech, automatic processing by video recognition techniques is not yet available to linguists at this stage. Developments in this area are promising and can build on video processing algorithms are developed for all kinds of purposes (e.g. [2-3]). However, the results are not yet at such a level that spontaneous interaction recorded at moderate quality can be recognised at such a level that it would save linguistic researchers time. This quality cannot only be improved by better computer algorithms, it is also key that the quality and the level of detail of the annotation of training corpora is raised.

One step in this process is for linguists to agree on how to classify linguistic events in video recordings [4]. This holds both for corpora for language technology and for

linguistic data collections (which may in fact overlap). While there has been some discussion and a concrete proposal on sign-specific metadata categories [5], no such concrete proposal has been made for annotation of signing at any level. The meetings of the Sign Linguistics Corpora Network (SLCN) [6] in 2009 and 2010 have established that the time is right for proposals in this area, as the first large sign language corpora with their annotation conventions are being published online (e.g. [7-8]) and a large number of corpus creation projects are currently underway (see http://www.signlanguagecorpora.org).

The data category registry 'ISOcat' that will be discussed in this paper constitutes a platform for defining and referring to linguistic concepts and related metadata terminology. It offers a robust way of storing and referring to data categories at both metadata and annotation levels. It does not enforce a standard upon users, but offers facilities to refer in a standard way to those categories that are commonly agreed upon by the community or by a sub-community. Such agreements within a research community can evolve over time, and need not be ready at the start. Even for a single user working on larger sets of data, this can provide an advantage in enforcing consistent references to the same annotation value throughout a corpus or across multiple corpora.

2   ISOcat

The ISOcat Data Category Registry (DCR) is the ISO 12620:2009 [9] compliant registry of the ISO technical committee for "Terminology and other language and content resources" (ISO TC 37) [10]. The DCR plays a role in various flexible standards developed by this committee. For example, the Lexical Markup Framework (LMF; ISO 24613:2008 [11]), a standard for lexical resources, provides a meta model which can be adorned with ISOcat data categories to create an actually usable project or application specific model. But in fact any linguistic resource (schema) can be linked to data categories, and can thus make the semantics of the elementary descriptors explicit.

The DCR data model [12] distinguishes two major types of data categories: complex and simple categories. Complex data categories have a conceptual domain for which the values can be described in the registry as simple data categories. Each data category has an elaborate specification. The administrative section contains, among others, a technical persistent identifier, a human readable mnemonic, and various types of notes. The descriptive section then contains alternative names, definitions, examples, explanations, et cetera, in various working languages. An English name and definition is mandatory for every data category. Finally there is a linguistic section that contains the conceptual domain of a complex data category, which can be described specifically for various domains and object languages.

The persistent identifier is to be used by the linguistic resources to indicate which data category is meant. ISOcat uses so called 'cool URIs'.[1] For instance, http://www.isocat.org/datcat/DC-2523 is the persistent identifier for the metadata data category for geographical coordinates. Notice that the specification of the data category contains various names, possibly even in various languages, which can all be

---

[1] See Tim Berners-Lee, *Cool URIs don't change* at
 `http://www.w3.org/Provider/Style/URI.html`

used in the linguistic resource. However, it is also possible that the resource uses a very specific name or abbreviation for the data category which does not appear anywhere in the specification. This does not have to hamper semantic interoperability as long as the local category is semantically equivalent. This would then be indicated by its association through a persistent identifier with the ISOcat data category. For XML-based resources, ISO 12620:2009 defines a small XML vocabulary[2] of elements and attributes to embed data category persistent identifiers. For example:

```
<geoLocation

        dcr:datcat="http://www.isocat.org/datcat/DC-2523">
                52, 5
</geoLocation>
```

The dcr:datcat attribute in this example contains the persistent identifier and indicates that the geoLocation element in this resource is semantically equivalent to the complex geographical coordinates data category in the DCR.

Every linguist can register with ISOcat, at http://www.isocat.org/, and use, create and share data categories. The idea is that this grassroots approach will feed a standardised core, a coherent set of data categories reviewed and maintained by a group of international experts. Multiple of these expert groups, known as Thematic Domain Groups, have been established by ISO TC 37, e.g., one for metadata and one for morphosyntax, and more can be added upon need of the linguistic community. Recently, a Thematic Domain Group for sign languages has been introduced.

Linguists can use the ISOcat web interface, but a tool-oriented interface is also provided. This interface has been used to create a close integration with the ELAN annotation tool, which is used by many linguists to create sign language resources. In this tool, reference to ISOcat data categories can be made in the definition of controlled vocabularies; we will come back to this in section 5 below. This makes ISOcat an attractive platform for the development of standards for sign language annotation.

3   A Brief History of Signed Language Resources

3.1   Media and Metadata

Video annotation in the computer in software such as ELAN has only taken off in the last decade. Since sign language research started with the work of Tervoort [13] and Stokoe [14], film and later video recordings have been made on a variety of media, from reel-to-reel tapes to MiniDV cassettes. Analysis of these recordings took place by patiently winding back and forth the recordings, and making transcriptions with pen and paper or later in office software. By consequence, the resulting data sets have been limited in size, ranging from isolated sentences to short stories or short dialogues. Archives were not always well-ordered, as the number of tapes was such that researchers would remember what was where, but more importantly, there was little added value of creating proper metadata as the utterances contained on them would not be immediately accessible in the way that they are now through tools like ELAN.

---

[2] The latest version of this DC Reference vocabulary is available online at:
http://www.isocat.org/12620/

   Digital video slowly started to appear on the desktop in the 1990s, as individual files or as part of CD-ROM applications. Only some years later, DVDs became a proper successor to video tapes for home entertainment. For research purposes, longer video files than a few minutes could only be realistically created and used on the desktop since some ten years. Since then, digital video files have rapidly replaced video tapes as the medium of recording. Larger and more systematic corpora of video recordings for linguistic research and the development of language technology are now being created and published [6].

   As indicated above, the ordering of archives has become more important now that the sizes of data collections are growing and the media are less tangible. Metadata descriptions that formerly existed primarily on tape labels are being stored in a variety of formats, from Excel tables to formalized metadata documents like IMDI files. While the nature of the information is highly similar across all data collections (information about the signer, the project, the tasks or situations, etc.), the encoding of these data and the field names that are used are likely to vary. For use within research institutes, for instance, this should not be a problem, but upon collaborating with others, computer searches across data sets will be hard if not impossible.

   There are a number of metadata standards that stem from language technology or computer science more generally [15]. They range from very compact (Dublin Core[3]) to very detailed (IMDI[4]) and things in between (OLAC[5]). The European CLARIN project is developing a new strategy in the metadata description of resources that is potentially as elaborate as IMDI, but that allows for more flexibility by letting users decide which components are relevant to their data; hence the name 'component metadata' (CMDI[6]). The fields and values of CMDI are currently one of the most well-developed domains within ISOcat. In section 4.3, we will discuss the integration of sign language specific metadata fields in this approach.

3.2   Annotations

Where video annotation, including phonetic transcription, glossing and translation, has long taken place on paper and in office documents, nowadays most of these tasks take place in dedicated software like ELAN, iLex, or ANVIL. The actual tool that is used is likely to be guided by the research community of which one is a member, but the exchange of data is possible at least to some degree [16-17]. This is an important step, as fully standardized annotations across data sets would be hard to exploit if they were locked within the respective data sets. These tools lead to new possibilities for data coding and analysis, and to new views on how to go about annotating corpora in sign language research [8].

   An important first step in standardizing annotations is formed by publishing what one is using. This is done in increasing detail. Where the ECHO project used a few pages to summarize the conventions for a broad range of glossing, translation and phonetic transcription tiers [18], the glossing conventions alone are described in a few

---

dozen pages for the Auslan corpus [19]. The latter document emphasizes that glosses should be based on a lexicon ('ID-glosses') in order to ensure consistency. Sign lexicons themselves, as we will see below, are not standardized at all. While there may be significant overlap between languages in lexical forms, depending on the language pair of course, it will be hard to exploit these overlaps if the lexical annotations are not in some way interoperable.

3.3   Lexicons

Paper dictionaries of signed languages have been produced since the 1960s (e.g. Stokoe, Casterline & Croneberg [14] for American Sign Language, Brien [20] for British Sign Language). More recent productions have used CD-ROMs, DVDs or internet databases as media (e.g. NSDSK [21] for Sign Language of the Netherlands). They are often created for second language learners, and have so far not been based on substantial linguistic corpora. As these were simply not available until recently, there are rarely links between lexical databases and annotation files. The exception if formed by the German Sign Language dictionaries created in Hamburg in recent years (see [22]), where the iLex annotation program has a closely integrated lexicon module.

4   Sign Language-Specific Terminology

4.1   Introduction

Many data categories that are being developed for spoken languages will also apply to sign languages. This holds for many areas like semantics, morphosyntax, and pragmatics. However, we see a need for sign-specific standardised terminology in three different areas. We will discuss three areas where sign-language specific data categories can be useful: general concepts, metadata descriptions, and linguistic annotations for sign language specific linguistic concepts.

4.2   General Concepts

There are some general terms relating to sign language and deaf communication that would be good to define and standardise within ISOcat, so that they can be profitably used in descriptions of sign language resources, at either metadata or annotation level. The required definition that each data category contains may not only be of use among sign specialists, but also serves to make clear to other linguists what exactly is meant. The set includes the term 'sign language' itself, making explicit that this does not cover any type of visual communication, but specifically refers to the natural languages used in deaf communities. Other terms with a similar broad coverage include 'fingerspelling', 'sign-supported speech', 'CODA' (child of deaf adults), 'home sign' and 'auxiliary sign language'.

4.3   Metadata Descriptions

The metadata categories that were proposed in Crasborn & Hanke [23] were put together as additions to the IMDI metadata set, including actor properties such as

'hearing status' that may be specifically relevant to sign language resources. This 'sign language profile' has since been further discussed at an international workshop in November 2009, where need for only a few amendments was established.[7] They form excellent candidates for ISOcat metadata categories, whether in the thematic domain 'sign language' or 'metadata' or even both[8]. The IMDI sign language profile has been described in Crasborn & Hanke [23], an online manuscript, but had not been published or otherwise formalised. It has now been formalised in ISOcat within a CLARIN-NL project (IPROSLA[9]) in 2011. Below, we will briefly characterise the fields, which are part of the IMDI sections 'Content' and 'Actor'.

The selection of fields that a group of European experts agreed upon in 2003, and which was later adopted as a 'profile' in IMDI, is not everything one can possibly describe about a resource at metadata level. In fact, the starting point for that workshop was a much longer list [5]. Items from those list as well as any other elements can be more readily combined with standard items in CMDI than before in IMDI. ISOcat allows for the creation of one's personal data categories that in a metadata description can be combined with standardised fields. Sharing such fields that are needed to accurately describe a certain data category can lead to a gradual revision of the former sign language profile: it will be a task of the Thematic Domain Group to monitor such shared categories related to sign language, and to see whether they would be suited for standardisation. Important criteria will include explicitness of the definition, and relevance to the broader community.

4.3.1 Content Metadata

A number of properties were considered to often be important for sign language recordings in the 'content' area of IMDI. Some of them may have broader use, also applying to spoken language resources. While a language of a session can be specified, 'Language Variety' was suggested as an addition to cover the regional variation found in especially the lexicon of many signed languages. Moreover, a 'Description' for the language specified was added in order to allow for referring to use of sign-supported speech, which is really a (highly variable) performance issue involving the language use rather than a full-fledged language that should receive an ISO code. 'Elicitation Method' was seen as an open field that could be quite useful given the overlap in elicitation materials currently used for sign language corpora, such as the Tweety & Sylvester cartoon 'Canary Row'. Thus, the field name itself could be a complex data category in ISOcat, while the diverse values of the field could not be data categories, but free text.

Interpreting, too, is not something that will be unique to sign language resources, but the common use of interpreters in most deaf communities makes it more likely that interpreters also appear in sign language resources. 'Source' and 'Target' modality were distinguished, as well as 'Visibility' of the interpreter in the video recording, and whether an 'Audience' is present (as opposed to for example a

---

[7] We refer to the web site of the Sign Linguistics Corpora Network for a report on this workshop, `http://www.ru.nl/slcn/workshops/2_metadata`

[8] Data categories can be members of various thematic domains, but only one Thematic Domain Group can own the data category. When a data category is relevant for more domains the involved Thematic Domain Groups should work on harmonization of the specification.

[9] `http://www.ru.nl/sign-lang/projects/iprosla`

situation where the interpreter appears in a TV recording for an anonymous audience). For the modality, a distinction was made between 'sign language', 'speech', 'sign supported speech', 'text', 'fingerspelling' – terms that will all need to be explicitly defined when entered in ISOcat as simple data categories.

4.3.2 Actor Metadata

Most fields in the sign language profile were in the 'Actor' section: properties of language users, but also of researchers involved in the creation or publication of the resource. The properties are the fairly stable characteristics of persons, rather than referring to aspects of the behaviour of an actor in a particular movie.

Table 1. Metadata fields for actors in the IMDI sign language profile

Category	Field	Values or description
Deafness	Status	*Hearing, hard of hearing, deaf*
	Aid type	*None, conventional, cochlear implant*
Sign language experience	Exposure age	*(years; months)*
	Acquisition location	*Home from family, home from tutor, preschool teachers, teachers, family beyond home, friends*
	Sign teaching	Amount of experience with sign language teaching (*none, some, extensive*)
Family (both fields are needed for mother, father, and partner)	Deafness	*Deaf, hard-of-hearing, hearing*
	Primary communication form	*Sign, sign-supported speech, gesture, mix between signing and speaking, speech only, writing*
Education	Age	The age range during which the school was attended *(years; months – years; months)*
	School type	*Bilingual home programme, kindergarten, preschool, primary school, vocational training, college, university*
	Class kind	*Deaf, hard-of-hearing, deaf class in hearing school, individually integrated*
	Education model	*Bilingual, oral, mixed, sign monolingual, oral with interpreter*
	Location	City
	Boarding school	*Yes, no*

The actor properties fall into four categories: deafness, sign language experience, family, and education. A full list is presented in Table 1. The rightmost column presents a list of values (in italics) if this was suggested in the IMDI profile, and/or gives a short description of the field. For further details we refer to Crasborn & Hanke [23].

In the overview in Table 1, one can see that where the field names will always have a very specific interpretation and will need to be defined as a ISOcat data category, this is not always true for the actual values, which are sometimes self-evident and not in need of definition. This is the case for things like cities, or yes/no. 'Hearing', 'hard-of-hearing', and 'deaf', though, may well merit a specific definition in the context of their use in metadata descriptions of actors. These are matters for a Thematic Domain Group on sign language to further discuss, together with metadata experts.

4.4   Linguistic Annotations

The aspect of annotation that has received most discussion in the recent literature is glossing, as there are specific problems inherent in the representation of a signed language in the written code of a spoken language [24-19-25]. In many other areas of linguistic description, however, there will be many concepts where sufficient consensus in the linguistic literature has arisen. This does not necessitate agreement on the correctness of a specific theory: as long as concepts within a theory are sufficiently clear, they can be described as a data category with a specific label and a persistent identifier. For example, the theory of 'grand iconicity' by Cuxac [26] is not accepted by all linguists as a correct analysis of the grammar of sign languages, yet specific concepts from this theory may well be useful for the annotation of sign language data in a particular project. By using a standard reference to a data category such as 'personal transfer' (which others might characterise as 'role-taking' or 'constructed action'), this can be used in the annotation of videos in a way that is transparent to other researchers. Thus, it will facilitate the automated analysis of the corpus in question. In addition to theory-specific concepts, Schembri & Crasborn [4] suggest that the sign linguistics field could arrive at a shared theory for certain descriptions, that could lead to consistent and useable categories, just as IPA transcriptions form a usable basic level description of speech for a wide range of purposes independent of one's particular phonetic or phonological theory. This would be useful for phonetic descriptions of signing, the area where the modality-specific properties of sign language are most prominent and where ISOcat categories from other Thematic Domains will thus not apply.

4.5   Glossing

In principle, overlapping lexical items (same phonological form, same semantics), could be data categories in ISOcat as well. Pointing signs, for instance, exist in various varieties in all sign languages. Identical pointing signs could be described in ISOcat, so that they could be more easily compared across signed languages. Individual lexical databases would in turn refer to such a data category, adding instances and contextual examples from the own sign language, with further language-particular information. As with other annotations, establishing that two

forms are indeed similar both in their form and in their meaning and use is not a trivial task. We suggest that a focused effort to create a joint lexicon could well profit from ISOcat.

5   How to Move Forward

Referring to ISOcat data categories has recently been integrated in the ELAN multimedia annotation tool and the ARBIL metadata tool. These two are incorporated in the set of CLARIN-compliant standard tools, following an open source policy and using published XML schemas for their data files. ELAN has become a near standard in the annotation of sign language resources world wide. Within ELAN, data categories can be referred to in the creation of controlled vocabularies (CVs). These CVs are linked to 'linguistic types', which group annotation layers ('tiers'). Thus, users can straightforwardly select whatever fits with their own needs. The exchange of file templates that bundle specifications for types, tiers, and CVs can further promote standardisation of linguistic annotations.

In other words, the technical facilities are there. It is high time that linguists and language technologists come together within the Thematic Domain Group 'sign language', to make concrete proposals about ISOcat data categories that can then be proposed to the wider community for discussion. The nature of the ISOcat workflow implies that categories that have been created can be used right away, so that discussion need not only take place on the proposals on paper but can also be tested in the development of concrete corpora. For this to become a success, it will be important that corpora are actually published and shared. In that way, searches using ISOcat data categories as criteria can pay off directly, and the benefit of standardisation will not just be a long-term ideal.

References

1. Sutton, V.: SignWriting: On the occasion of its 25th anniversary November 1999. Sign Language & Linguistics 2(2), 271–282 (1999)
2. Cooper, H., Holt, B., Bowden, R.: Sign Language Recognition. In: Looking at People: Automatic Visual Analysis of Humans, Part D, pp. 539–562. Springer (2011)
3. Dreuw, P., Forster, J., Ney, H.: Tracking Benchmark Databases for Video-Based Sign Language Recognition. In: Proceedings of ECCV International Workshop on Sign, Gesture, and Activity (SGA), Crete, Greece (2010)
4. Schembri, A., Crasborn, O.: Issues in creating annotation standards for sign language description. In: Proceedings of the 4th Workshop on the Representation and Processing of Sign Languages: Corpora and Sign Language Technologies. ELDA, Paris (2010)
5. Crasborn, O., Hanke, T.: Metadata for sign language corpora. In: Background Document for an ECHO Workshop, May 8-9. Radboud University, Nijmegen (2003), `http://www.let.ru.nl/sign-lang/echo/docs/SignMetadata_May2003.pdf`

6. Crasborn, O.: The Sign Linguistics Corpora Network: towards standards for signed language resources. In: Calzolari, N., Choukri, K., Maegaard, B., Mariani, J., Odijk, J., Piperidis, S., Rosner, M., Tapias, D. (eds.) Proceedings of the Seventh conference on International Language Resources and Evaluation (LREC 2010), pp. 457–460. ELRA, Paris (2010)

7. Crasborn, O., Mesch, J., Waters, D., Nonhebel, A., van der Kooij, E., Bergman, B., Woll, B.: Sharing sign language data online. Experiences from the ECHO corpus. Int. J. Corpus Linguistics 12(4), 535–562 (2007)

8. Johnston, T.: From archive to corpus: transcription and annotation in the creation of signed language corpora. International Journal of Corpus Linguistics 15(1), 104–129 (2010)

9. International Organization for Standardization: Terminology and other language and content resources — Specification of data categories and management of a Data Category Registry for language resources. ISO 12620 (2009)

10. Kemps-Snijders, M., Windhouwer, M.A., Wittenburg, P., Wright, S.E.: ISOcat: Corralling Data Categories in the Wild. In: European Language Resources Association (ELRA) (ed.) Proceedings of the Sixth International Conference on Language Resources and Evaluation (LREC 2008), Marrakech, Morocco (2008)

11. International Organization for Standardization: Language resource management — Lexical markup framework (LMF). ISO 24613 (2008)

12. Kemps-Snijders, M., Windhouwer, M.A., Wittenburg, P., Wright, S.E.: A Revised Data Model for the ISO Data Category Registry. In: Proceedings of the 8th International Conference on Terminology and Knowledge Engineering (TKE 2008), Copenhagen, Denmark (2008)

13. Tervoort, B.T.M.: Structurele analyse van visueel taalgebruik binnen een groep dove kinderen. Noord-Hollandsche Uitgevers Maatschappij, Amsterdam (1953)

14. Stokoe, W.C., Casterline, D.C., Croneberg, C.G.: A dictionary of American Sign Language on linguistic principles, 2nd edn. Linstok Press, Silver Spring (1965, 1976)

15. Broeder, D., Declerck, T., Romary, L., Uneson, M., Strömqvist, S., Wittenburg, P.: A Large Metadata Domain of Language Resources. In: Lino, M., Xavier, M., Ferreira, F., Costa, R., Silva, R. (eds.) Proceedings of the 4th International Conference on Language Resources and Evaluation (LREC 2004), pp. 369–372. European Language Resources Association, Paris (2004)

16. Schmidt, T., Duncan, S., Ehmer, O., Hoyt, J., Kipp, M., Loehr, D., Magnusson, M., Rose, T., Sloetjes, H.: An Exchange Format for Multimodal Annotations. In: Kipp, M., Martin, J.-C., Paggio, P., Heylen, D. (eds.) Multimodal Corpora. LNCS (LNAI), vol. 5509, pp. 207–221. Springer, Heidelberg (2009)

17. Bird, S., Liberman, M.: A formal framework for linguistic annotation. Speech Communication 33, 23–60 (2001)

18. Nonhebel, A., Crasborn, O., van der Kooij, E.: Sign language transcription conventions for the ECHO project. Online manuscript (2004),
    `http://www.let.ru.nl/sign-lang/echo/docs/transcr_conv.pdf`

19. Johnston, T.: Auslan Corpus annotation guidelines. Online ms., Department of Linguistics, Macquarie University, Sydney, Australia (2011), `http://www.auslan.org.au/media/auslan-video/upload/attachments/AuslanCorpusAnnotationGuidelines19June2011.pdf`

20. Brien, D. (ed.): Dictionary of British Sign Language/English. Faber & Faber, London (1992)

21. NSDSK. Nederlandse Gebarentaal op CD-ROM. Communiceren met dove kinderen thuis en op school. Deel 1 (Version 1.00). Nederlandse Stichting voor het Dove en Slechthorende Kind, Amsterdam (1996)
22. Konrad, R., Schwarz, A., Prillwitz, S.: Fachgebardenlexica. In: Leuninger, H., Wempe, K. (eds.) Gebardensprachlinguistik, pp. 193–216. Signum, Hamburg (2001)
23. Crasborn, O., Hanke, T.: Additions to the IMDI metadata set for sign language corpora (2003), `http://www.let.ru.nl/sign-lang/echo/docs/SignMetadata_May2003.pdf`
24. Johnston, T.: Transcription and glossing of sign language texts: examples from Auslan (Australian Sign Language). International Journal of Sign Linguistics 2, 3–28 (1991)
25. Ormel, E., Crasborn, O., van der Kooij, E., van Dijken, L., Nauta, E., Forster, J., Stein, D.: Glossing a multi-purpose sign language corpus. In: Proceedings of the 4th Workshop on the Representation and Processing of Sign Languages: Corpora and Sign Language Technologies, pp. 186–191. ELDA, Paris (2010)
26. Cuxac, C.: La Langue des Signes Française (LSF) - Les Voies de l'Iconicité. Faits de Langues 15/16. Ophrys, Paris (2000)