

Linking a Web Lexicon of DSGS Technical Signs to iLex

Sarah Ebling*, Penny Boyes Braem**

* University of Zurich, Zurich, Switzerland

E-mail: ebling@cl.uzh.ch

** Center for Sign Language Research, Basel, Switzerland

E-mail: boyesbraem@fzgresearch.org

Abstract

A website for a lexicon of Swiss German Sign Language equivalents of technical terms was developed several years ago using Flash technology. In the intervening years, the backend research database was migrated from FileMaker to iLex. Here, we report on the development of a web platform that provides access to the same technical signs by extracting the relevant information directly from iLex. This new platform has many advantages: New sets of signs for technical terms can be added or existing ones modified in iLex at any time, and changes are reflected in the web platform upon refreshing the browser. Just as importantly, the new platform can now also be accessed through all major mobile operating systems, as it does not rely on Flash. We describe how information on the glosses, keywords, videos of citation forms, status, and uses of the technical signs is represented in iLex and how the corresponding web platform was built.

Keywords: Swiss German Sign Language, technical signs, iLex, web platform

1. Introduction

For more than twenty years, work has been carried out on a lexicon of Swiss German Sign Language (DSGS) (Boyes Braem, 2001). While the initial form of the lexicon was a FileMaker database, the lexicon has recently been migrated to iLex, a sign language lexicon and corpus software (Hanke and Storz, 2008). It is stored on a university server and is accessible by researchers working on specific projects. This DSGS iLex lexicon currently contains approx. 9000 signs arranged into different sub-lexicons.

Two sub-lexicons of signs together contain 685 technical terms in the domains of nutrition and economy (Boyes Braem et al., 2012). Before the database was migrated to an iLex form, these signs had been made accessible on the web using Flash technology. Here, we report on the development of a web platform that provides access to the same technical signs by extracting the relevant information from iLex. This new platform has many advantages: New sets of signs for technical terms can be added or existing ones modified in iLex at any time, and changes are reflected in the web platform upon refreshing the browser. Just as importantly, the new platform can now also be accessed through all major mobile operating systems, as it does not rely on Flash.

2. iLex

iLex is a client-server application developed specifically to support the creation and exploitation of sign language lexicons and corpora. As such, it is somewhat similar to ELAN, a common tool for creating sign language and spoken language corpora (Wittenburg et al., 2006). In an ELAN transcript, the glosses a user assigns to a sign segment take the form of strings in free-text fields with no underlying consistency check. By contrast, iLex requires the use of glosses that are standardized for a particular sign language by demanding in a first step that

the user choose from a set of previously introduced glosses. If an appropriate gloss is not available, a user can create a new one. All occurrences of a gloss in a transcript (gloss tokens) are linked back to their gloss type in the lexicon, and changes of the gloss type affect all gloss tokens in all transcripts. This is possible through an underlying relational database. The PostgreSQL database that is the backbone of iLex consists of 97 tables. Table 1 shows a selection of tables.

concepts
illustrations
illustration_instances
illustration_uses
meanings
metadata_descriptions
metadata_features
movies
projects
tags
terms
tiers
transcripts
types

Table 1: Selection of tables from iLex PostgreSQL database.

3. DSGS Signs for Technical Terms in the Previous Flash Web Platform

The DSGS signs for selected technical terms in the fields of nutrition (293 concepts) and economy (392 concepts) were analyzed and described in a two-year research project (Boyes Braem et al. 2012). The primary targeted users were DSGS signers, especially students in schools of higher education, who wanted more information in sign language about the meaning of specific technical

terms. The website has also been useful to DSGS interpreters looking for sign equivalents for terms in these technical fields. The resulting web lexicon provided the following information about each sign:

- **Gloss**, e.g., ALKOHOLSTEUER_1 [ALCOHOL-TAX_1]
- **Keywords** (German): Keywords are included as glosses often reflect only one meaning of a sign. Examples of keywords corresponding to the gloss ALKOHOLSTEUER_1 are “Alkoholsteuer” [alcohol tax] and “Steuer” [tax].
- Video of the **citation form** of the sign and of up to three widely used form variants
- **Domain**: “nutrition” or “economy”
- **Status** of the sign: “used” by more than three signers, “known” to be used by at least two signers, or “new” (as developed by a group of experienced DSGS signers who had knowledge of the corresponding technical field)
- **Definition** of the term in the form of videotaped DSGS and a written German back-translation of the DSGS signed text. A few definitions come in the form of illustrations as opposed to videos of DSGS signing. Definition of ALKOHOLSTEUER_1 (German back-translation): “Steuer, die im Kaufpreis von Alkohol inbegriffen ist [tax that is included in the price for alcohol]”.
- One to three **examples** of the meaning of the concept in videotaped DSGS and back-translated written German. Example (German back-translation): “Beim Kauf von Spirituosen (Bier, Wein usw.) wird mit dem Kaufpreis eine Steuer bezahlt. [When buying liquors (beer, wine, etc.), a tax is included in the price.]” For some concepts denoting kitchen utensils (nutrition domain), illustrations or photos of the utensil (e.g., a whisk) are provided instead of text examples.

4. Representation of Technical Signs in iLex

When migrating the DSGS lexical database from FileMaker to iLex, we introduced a sign type in the iLex lexicon for each **gloss** entry in the previous FileMaker database. To link the **videos of the citation forms** to the sign types, we introduced transcripts for the videos that hold token tags of the respective types. Figure 1 (at end of article) shows a transcript containing a token tag of the sign type ALKOHOLSTEUER_1 in a tier “Type/Subtype”. The tag spans the entire duration of the video. Token occurrences of this kind are excluded from any corpus statistics in iLex, as they do not represent meaningful examples of signs in context.

Associated with the technical signs in iLex is metadata information on the **domain** of the sub-lexicon (nutrition or economy). Information on the **status** of a sign was also stored as sign type metadata, using a closed vocabu-

lary that permits the assignment of only one of the three values “used”, “known”, and “new”. The **keywords** accompanying glosses were stored as concepts.

In iLex, technical terms are concepts assigned to specific projects. Terms may carry **definitions**. We stored the German back-translations of the DSGS definitions of our technical signs here. In addition, to link the German definitions with their DSGS video originals, we introduced a transcript for each DSGS video. In the transcript, the German definition is represented as a tag of a tier “Definition”, the length of the tag again being equal to the duration of the video. In addition, the transcript contains a second tier that records the sign type to which the definition is linked. Figure 2 (at end of article) shows the definition of the technical sign ALKOHOLSTEUER_1.

The same procedure was applied for the **examples** of usages of technical signs: Each example received its own transcript that holds the DSGS (video) and German (text) versions along with a link to the sign type.

It was noted in Section 3 that some definitions of technical signs come in the form of pictures as opposed to videos of DSGS signing. The same is true for examples, e.g., for concepts denoting kitchen utensils. iLex allows for such pictures to be linked to terms directly.

5. Web Platform

The new web platform that holds the technical signs is accessible at <https://linguistik-signlang.uzh.ch/>. Figure 3 (at end of article) shows the previously introduced entry ALKOHOLSTEUER_1 in the web interface.

Currently, users can consult the web lexicon of technical terms in four ways:

1. By searching for (sub-)strings of keywords
2. By restricting the search to one of the two domains nutrition and economy
3. By searching for signs with a given status, where possible values are “used”, “known”, and “new”
4. By searching for signs whose glosses start with a given letter

Combined searches, i.e., searches along more than one of the above parameters, are also possible.

The platform is implemented in PHP. Embedded in the PHP code are SQL statements that extract the information outlined in Sections 3 and 4 (gloss, keywords, video of citation form, status, definition, examples) from the iLex PostgreSQL database. More precisely, the information is queried from the 97 tables that make up the iLex database. To retrieve the information relevant for the web platform of technical signs, the 14 tables shown in Table 1 are consulted. Table 2 shows an SQL sub-query.

```

SELECT types.id, terms.name,
illustration_instances.file_path
FROM projects
JOIN terms ON projects.id=terms.project
JOIN concepts ON terms.concept=concepts.id
JOIN meanings ON concepts.id=meanings.concept
JOIN types ON types.id=meanings.type
LEFT JOIN illustration_uses ON
terms.id=illustration_uses.term
LEFT JOIN illustrations ON
illustration_uses.illustration=illustrations.id
LEFT JOIN illustration_instances ON
illustrations.id =
illustration_instances.illustration

```

Table 2: Sub-query to extract information from the iLex PostgreSQL database.

The database user issuing the statements has read-only rights to prevent possible SQL injection. Moreover, user input is handed over to the built-in PostgreSQL function `pg_escape_string` in PHP that masks a string. The PHP code with the SQL queries embedded in it will be made available to interested parties upon request.

6. Conclusion

In this paper, we have reported on the development of a web platform that provides access to technical signs in the domains of nutrition and economy by extracting the relevant information from a DSGS instance of iLex. The signs had previously been stored in a FileMaker database.

We have described how information on the glosses, keywords, videos of citation forms, status, and uses of the technical signs was migrated from FileMaker to iLex and how the corresponding web platform was built.

Currently, we are extending the search functionality offered in the web platform. A medium-term goal is to change the design of the website, which is still the same as that of the original (Flash-based) platform.

One of the advantages of the new linking of the web lexicon of signs for technical terms to iLex is making it easier to not only correct or modify existing terms but also add new technical fields to the web lexicon. We are in the process of expanding the website to include name signs for places that are already in the larger iLex lexicon. In the near future, as funding becomes available, we plan to add signs in the fields of sign language linguistics,

jurisprudence, medicine, natural hazards (such as earthquakes), language testing techniques, as well as name signs for well-known persons. The structure of the information presented on the website for these additional domains will in most cases be much the same as for the technical domains described above. As the information is based on iLex entries, however, this information will be backed up by the information and analysis necessary for all iLex entries, including metadata on the source of the information, more exact form descriptions (HamNoSys) as well as crosslinks to other lexical items in the larger lexicon and to examples in the annotated videos in the linked corpus. Finally, we expect all of the information from this website to be much more widely accessible as it is now available on tablets and mobile phones as well as on computers.

7. Acknowledgements

We gratefully acknowledge the work of Lorenz Nagele on the implementation of the PHP code. Johannes Graën's help in formulating and optimizing the SQL queries is much appreciated. We would also like to thank Thomas Hanke for his ongoing support of our work with iLex.

8. Bibliographical References

- Boyes Braem, P., Groeber, S., Stocker, H., and Tissi, K. (2012). Weblexikon für Fachbegriffe in Deutschschweizerischer Gebärdensprache (DSGS) und Deutsch. *eDITion: Fachzeitschrift für Terminologie*, 2, pp. 8-14.
- Boyes Braem, P. (2001). A multimedia bilingual database for the lexicon of Swiss German Sign Language. *Sign Language & Linguistics*, 4(1/2), pp. 133-143.
- Hanke, T., Storz, J. (2008). iLex: A database tool for integrating sign language corpus linguistics and sign language lexicography. In *Proceedings of the 6th Language Resources and Evaluation Conference (LREC)*, pp. 64-67, Marrakech, Morocco.
- Wittenburg, P., Brugman, H., Russel, A., Klassmann, A., and Sloetjes, H. (2006). ELAN: a professional framework for multimodality research. In *Proceedings of the 5th Language Resources and Evaluation Conference (LREC)*, pp. 1556-1559, Genoa, Italy.

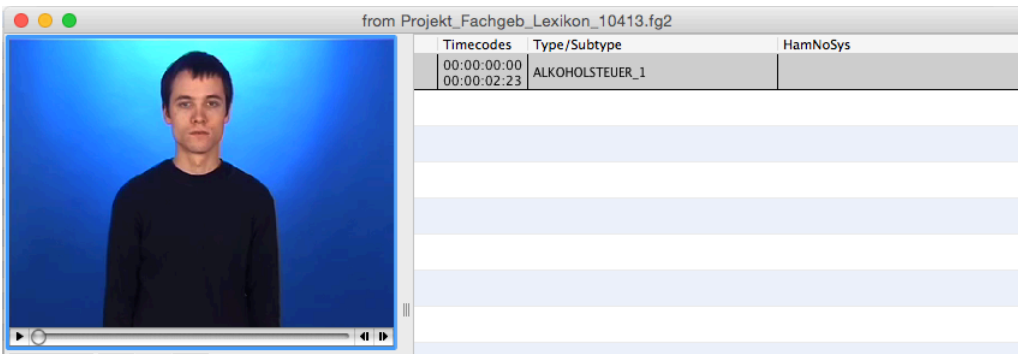


Figure 1: Link between videos of citation forms and sign types in the iLex lexicon: Sample transcript for ALKOHOLSTEUER_1 [ALCOHOL-TAX_1].

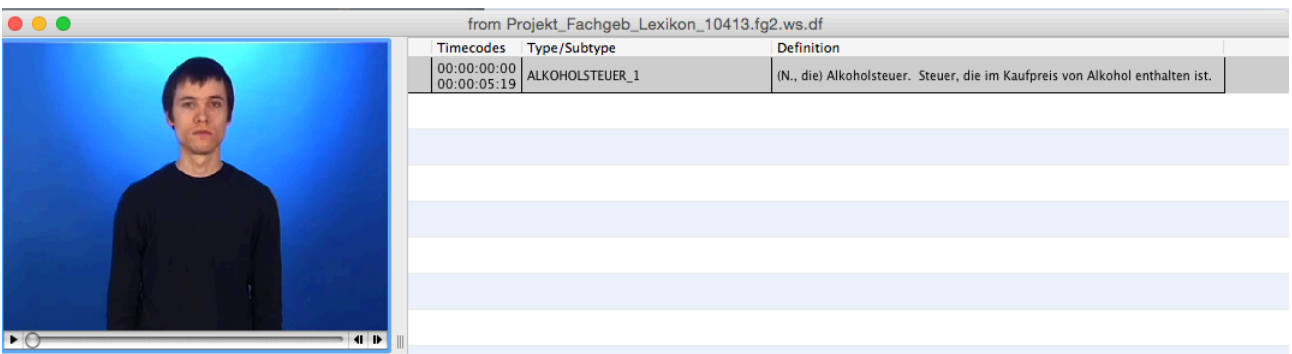


Figure 2: Link between videos of definitions and sign types in the iLex lexicon: Sample transcript for ALKOHOLSTEUER_1 [ALCOHOL-TAX_1].

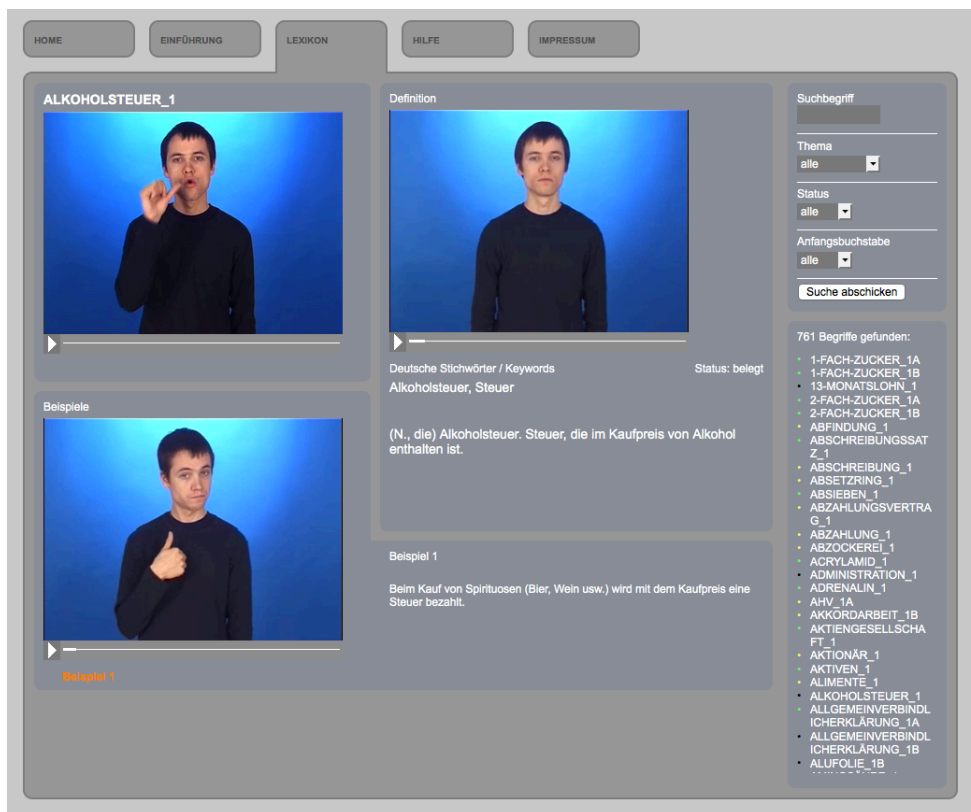


Figure 3: Screenshot of the web platform showing the entry ALKOHOLSTEUER_1 [ALCOHOL-TAX_1].