Information Types and Use Cases

Gabriele Langer, Anke Müller, Felicitas Otte, Sabrina Wähl

DOI (latest version): 10.25592/uhh.fdm10232

Version 0	2021-08-06	Draft: Some sections not finished
Version 1	2022-05-24	First public version

Zusammenfassung

Dieses Arbeitspapier ist Teil des lexikographischen Manuals des DW-DGS[1], in dem wir unsere allgemeinen Entscheidungen bezüglich des Typs des Wörterbuchs und dessen Inhalt darlegen. Wir beschreiben die verschiedenen Informationseinheiten, die in einem Wörterbucheintrag gegeben werden können. Diese Informationseinheiten sind Elemente der Mikrostruktur des Wörterbuches. Zusätzlich werden die Möglichkeiten für die Benutzer:innen, Informationen im DW-DGS zu suchen, beschrieben (Zugangsstrukturen). Dieser Teil beinhaltet auch eine Vorstellung der potentiellen Nutzer:innen-Gruppen des DW-DGS. Wir skizzieren verschiedene Nutzungsmöglichkeiten und zeigen auf, welche Informationen und Zugangsmöglichkeiten für welche Nutzer:innen-Gruppe besonders nutzbringend sind.

Abstract

These project notes are part of the lexicographic manual for the DW-DGS[2]. In these project notes, we describe our general decisions concerning dictionary type and content. We describe the different information types given in the entries (elements of the microstructure) and the ways a user can find desired information in the DW-DGS (access structures). The latter includes a closer look at the potential user groups of the DW-DGS. Different use cases of the dictionary and which information types and access structures are especially profitable to which user group are discussed.

[1] Vollständiger Name des DW-DGS: *Digitales Wörterbuch DGS. Das korpusbasierte Wörterbuch DGS – Deutsch.*
[2] Full Title of the DW-DGS: *Digitales Wörterbuch DGS. Das korpusbasierte Wörterbuch DGS – Deutsch.*

Table of Contents

1 Corpus-Based Methodology

The corpus-based and to some extent corpus-driven and largely corpus-bound dictionary of DGS, the DW-DGS, will be the first corpus-based dictionary of its kind.[3] Sign language corpora of considerable sizes have become available only recently and can now be used as a basis for lexicographic analysis and description. Sign language corpora differ from corpora of spoken languages (in their written form) in several aspects: They use videoformat to capture the visual and spatial nature of Sign languages (SLs). In addition, SLs lack an established writing system and thus machine-readable representations of the language in question. Therefore special challenges with regards to accessing and analyzing the data occur. Tokenization and lemmatization are done manually by token-type-matching via glosses that are time-aligned to the language data in video format. This means that search and analysis routines cannot be based on an adequate written representation of the signed text. Tools such as automatic PoS taggers are not available yet and search and analysis routines require previous manual annotations. Apart from these challenges, it is now possible to conduct corpus-based analyses of signs' meanings and uses in context for the first time. Not everything that is commonplace and easily done in the lexicographic work of well-resourced languages with a written tradition such as German or English can be realized with an SL corpus, yet. However, the sign language corpus provides us with new insights on signs' properties and uses that can now be observed systematically in the data, such as information on different meanings and uses in natural signing (used for sign sense discrimination), frequency effects, collocations, and regional distribution. Naturally, these new possibilities are explored and put to use in the compilation of the dictionary. Pieces of information can now be based on and evidenced by corpus data, facilitating a new level of quality to dictionary work. Consequently, the corpuslinguistic approach of our work influences what information types can and are included in the dictionary entries. The DW-DGS contains some information types which are included in a sign language dictionary for the first time. Consequently, new ways of structuring and presenting this information needed to be developed. Due to deeper and broader information on the lemma signs, the entries need to provide adequate structures for the more complex contents. While the user can profit from a broader range of information types, the appearance and complex structures of the dictionary are new and might be unexpected and intimidating to the user at first glance. We expect that putting some effort into educating users on how to use the dictionary to their best gain will be necessary.

2 Lexicographic Concepts and Terminology

This section contains a summary of lexicographic concepts and terminology following Svensén (2009) with some minor adaptions. These terms and concepts are necessary for understanding the later sections of these project notes.

Linguistic activities:

- *Reception* (*'passive', decoding*): reading, listening, watching
- *Production* (*'active', encoding*): writing, speaking, signing

[3] We are well aware that the *Corpus-Based Dictionary of Polish Sign Language* is the first general sign language dictionary that is based on a large corpus and bears *corpus-based* in the name. From the description of their dictionary writing process, it becomes clear that while they used the corpus as starting point, source, and inspiration, they also added signs and senses based on introspection. Thus, it remains somewhat unclear to what extent the dictionary is corpus-based and which pieces of information derive from corpus analysis and which from introspection. The Polish team also had a much broader approach to senses and only defined macrosenses. "The division of meanings into macrosenses was performed based on the linguistic intuitions of the lexicographers." (see https://www.slownikpjm.uw.edu.pl/en/page/pomoc#metodologia).

Lexicographic function and information needs:

- Reception/decoding: given form → finding certain content
- Production/encoding: given content → finding certain form

Object language(s): language(s) the dictionary sets out to describe
Metalanguage[4]: language used for descriptions, explanations, comments in the dictionary, labels, headings

L1: first/native language of user
L2: second/foreign language of user

Monolingual dictionary

- One object language (user's native language/L1 or a foreign language/L2)
- Metalanguage usually identical with object language

Bilingual Dictionary

- Two object languages (typical case)
  - one: user's native language (L1)
  - other: user's foreign language (L2)
- Meta-language is usually user's native language

Bilingual bidirectional dictionary

- bilingual dictionary aimed at users from both language communities of the two object languages; users have one of the object language as their L1

Entry of a bilingual dictionary:

- *Source Language*: language of the lemma/headword (word, expression) in the entry
- *Target Language*: language of the counterpart (*equivalent*) of the lemma sign given in in the entry

Dictionary functions:

- reception of L2 text
- production of L2 text
- translation of existing L2 text into L1 text
- translation of existing L1 text into L2 text

Dictionary types:

- monodirectional dictionaries
  - L1 → L2 dictionaries (native language = source language)
    - Production of foreign text
    - Translation of native language text into foreign text
  - L2 → L1 dictionaries (native language = target language)
    - Reception of foreign text
    - Translation of foreign text into native language
- Bidirectional dictionary
  - L1→L2/L2→L1
    - Can be used for any of the scenarios listed above

[4] Svensén (2009, 12) calls this *description language.*

3 User Groups

Major aims of the DGS-Korpus-Projekt and its dictionary are the documentation of DGS, as well as advancing the field of corpus linguistics and lexicography of sign languages in general and specifically for DGS. Apart from this more academic motivation, the dictionary is intended as a reference tool for DGS signers and has to serve the practical purpose of satisfying the information needs of its users. Since our dictionary will be the first and probably for a long time the only corpus-based dictionary of DGS, it is devised to offer valuable information to diverse user groups. The following groups are potential users of the DW-DGS:

- Proficient DGS signers (native signers and other signers that consider DGS as their native language or use it as their primary means of communication, that is, L1)
- Deaf Students
- Advanced and intermediate learners of DGS (L2), usually with German as their first language (L1)
- Novel learners and interested people with only little knowledge of DGS (L2), with German as their first language (L1)
- Hearing people with a professional contact to the deaf[5] community and hearing parents of deaf children
- DGS teachers (language professionals)
- DGS interpreters (language professionals)
- Linguists interested in DGS and signed languages in general (language professionals)

We are aware that these different groups have diverse and partially contrary information needs. For example, the novel learner might be overwhelmed by the wealth of information that is presented in an entry for one single sign and may favor a reduced view, while the advanced learner or the native signer might find this information especially interesting and valuable. The dictionary cannot address all these needs in the best possible way for all groups at the same time in one layout. The in-depth analyses that the corpus data allow for favor a deeper coverage which is better suited for native signers, advanced learners, and language professionals (DGS teachers, interpreters). However, we also try to offer helpful information and access structures to the less experienced user and novel learner. For example, the subject index will be especially valuable to the novel learner.[6]

In the following text when talking of user groups we think of expected prototypical users as representatives of these groups. In this regard *deaf* may be used as synonym to the group of users with DGS as their L1 and German as their L2 and *hearing* may be used as synonym to the group of users with German as their L1 and DGS as their L2 (Learners of DGS in the widest sense). We are well aware that the reality is much more diverse than this simplified view and that the status of German and DGS as L2 and L1 respectively may be more fluid and less clear for individual German deaf persons (see below in the next section) but for the purpose of these project notes this may suffice.

4 Metalanguage of the Dictionary

The metalanguage of a dictionary is the language of description and used for headers, captions, menu items, labels, comments, and the like. Ideally, the metalanguage of a dictionary should

[5] We use *deaf* to be inclusive of all deaf/Deaf persons.

[6] We are still in the process of discussing whether the final product will provide a layout choice for beginners and novel learners with a reduced range of information in order to reduce complexity and to provide a low-threshold access.

be the native language of the user. For bidirectional bilingual dictionaries, user groups with two different L1 languages are involved. Svensén states that "[…] the dictionary's description language […] must be intelligible to both user groups." (2009, 18).
In the DW-DGS we rely on written German as metalanguage. This decision takes into account that German deaf persons are to some extent bilingual and have some knowledge of (written) German. German as metalanguage thus is intelligible to user groups with both German and DGS as their L1. Most German deaf persons have hearing parents and have grown up with German as the language of communication in their original families. German is the surrounding majority language. It is taught in schools and it is the language deaf people encounter in official documents and commercially written texts. It is also used for everyday written notes and texts, e.g. in text messaging (Power & Power 2004). German is also the language that deaf persons typically use most in writing as no writing system is established for writing DGS. Thus, we can assume that written German is intelligible not only to users with German as L1 but also to most users with DGS as L1, at least to some extent.
We are aware that only some deaf users might be in full command of written German while for others reading texts in an only partly known spoken language is difficult. Notes and comments should be given in easy language with regard to vocabulary and complexity of syntax. Labels used in the dictionary will be explained not only in German but also in DGS in the front/back matter of the dictionary.
The second reason for using German and not DGS as metalanguage is of a more practical nature. Using a sign language as the metalanguage in an electronic sign language dictionary throughout (e.g. for all labels, headings, notes, explanations, as well as definitions) would either require the use of some widely-known writing system, or some form of direct visual representation such as movies or animations. A well-established and widely-known writing system for DGS and sign languages in general is not available to date. Using movies or animations throughout is not very practical either. It is difficult to imagine a complex sign entry containing different information kinds without any written elements that guide the eye and structure the entry. The content of written text within a complex entry is much more easily scanned and received than the content of a movie. Written text is static and therefore always available to the scanning eye while the content of a movie is fleeting and has to be watched all the way through before its content is available. Using DGS as metalanguage in the dictionary throughout the macro-structure and microstructure, for explanatory comments as well as for labels, notes, headings and guiding elements is therefore not very practical or user-friendly.[7] These practical considerations and the overall usability, contributed to the decision to use written German as the metalanguage of the DW-DGS.
Dictionary definitions in monolingual dictionaries are another place where metalanguage is involved.[8] Ideally, for a monolingual dictionary the definitions are provided in the same language as is being described by the definitions. Such definitions serve two functions at the same time: First, they identify and explain different meanings and uses of the lemma sign in an abstract way, and second, by doing this, they provide an additional linguistic context where the lemma sign is put into a meaningful relation with other words/signs of the object language of similar or related meaning, such as synonyms, antonyms, hypernyms, co-hyponyms, and other associated named units of a prototypical frame the lemma sign might be used in. Such relations and connections are useful to the learner to build up and strengthen their vocabulary and build associations between related words/signs of the language. A definition provided in a language other than the described language can fulfill the first function but not the second one.

[7] It is common practice for sign dictionaries (even monolingual ones) to use the written majority language for informational, guiding, and structural written elements within the dictionary macrostructure and microstructure (comp. Kristoffersen & Troelsgård 2012, 312).
[8] See for example Piotrowski 1989, 74-75 on this point.

It not unusal for minority languages without their own writing system that the more widely known majority language is used as metalanguage.[9]
With regard to the definitions, the following additional points have been taken into consideration for the decision to use written German instead of DGS as metalanguage for definitions:

- Without written DGS, definitions in DGS would have to be recorded and presented as movies.[10]
- Easy editing: Lexicographic definitions are a very dense, self-contained, and controlled text type. Definitions are usually re-worded, changed, and corrected several times during the article writing process. Written text is easily changed, while changing a recorded signed text requires much more effort, time, and technical as well as human resources.
- There are no established conventions for dictionary definitions in DGS as a text type. Signed definitions would have to be newly developed for lexicographic purposes. This would be a lengthy process, best led by members of the language community/native signers that have had a training in linguistics and lexicography.
- For a true monolingual treatment of definitions, signed definitions would have to be developed in DGS without a first version in written German. Avoiding inference from a written language while creating a signed dictionary definition is difficult. Detailed and complex texts are not easily recorded in the studio without some written representation or notes as memory aid for the signing model.
- It is an open question whether the integration of films with signed definitions alone would be as practical for finding the desired information in a complex entry as written definitions are. The scanning eye and brain of a user might be overwhelmed by a large number of movies presented, the contents of which are available only when starting and watching the movies.

5 Object Languages of the Dictionary

The DW-DGS has two object languages: DGS and German.
DGS is the main object language. It is the lesser-used and lesser-resourced language of the language pair and receives the focus of attention in the dictionary.
For any linguistic unit, pattern or example of the object language DGS given in the dictionary, there has to be some adequate representation easily accessible to the user. For this purpose we use the integration of videos and micons for the representation of signs and signed utterances at different places in the entry.[11] In the DW-DGS micons are small movie thumbnails of one variant form (usually the main variant that serves as citation form for the entire lemma sign) in combination with the entry number of the respective lemma sign. A click on the movie thumbnail shows the corresponding movie in the movie display area, a click onto the entry number serves as a link to the addressed entry. A sense preview box is triggered by mouse-over and lists the signposts (see below) of the respective entry to give an idea of the meaning of the

[9] Popkema notes that for languages with a strong dialectal differentiation and without a standard dictionary makers face the issue of having to choose a dialect to be used as metalanguage in the dictionary (2010, 42). This is avoided when choosing another language as metalanguage.
[10] Recorded movies or animated moving images (e.g. produced by avatar technology and the like). We do not discuss the possibility of animated movies here, because at the moment they are not available to us in an easy-to use, fast and feasible way.
[11] We have ruled out the options of using a notation or a writing system that is not widely established, as well as the use of glosses to represent signs. For the rationale and discussion of these possible alternative options and more details see AP11-2021-01 (Otte et al. 2022).

sign.[12] Micons or groups of micons are used throughout the microstructure and access structures to represent DGS units in a compact way. Examples from the corpus are provided as movies only.
German is the second object language of the DW-DGS. It is represented in its written form.

6 Monolingual Focus of the DW-DGS

The DW-DGS contains monolingually oriented descriptions of DGS signs and their meanings and uses. We take the actual use of the signs in the corpus data (tokens of the lemma sign in context) as a starting point for the analysis. Different meanings and uses are then summarized and described as dictionary senses. This is called word sense discrimination (WSD) in lexicography. Each sense is identified by a sense description, i.e., the dictionary definition, and illustrated by examples taken from the corpus. Collocational patterns discovered in the data are selected and provided in the entries. Also, whenever possible, (near) synonyms and antonyms are provided for each sense.[13]
The focus of our corpus-based lexicographic analysis and description of the lemma signs is done in the same way as in a monolingual dictionary with one exception: In the DW-DGS the dictionary definitions are not provided in DGS – the object language, instead written German, used for description as metalanguage.[14] [15]

7 Bilingual Features of the DW-DGS

While the DW-DGS is primarily focused on DGS from a monolingual perspective, it also contains some bilingual information, in order to provide access via German words and facilitate bilingual uses such as translations. This makes the DW-DGS some kind of hybrid between a monolingual and bilingual dictionary.[16] In addition to a monolingually oriented treatment of the signs, we also provide German translation equivalents [EqW] for the sign senses identified, thus fulfilling the function of a bilingual dictionary DGS→German. The translation equivalents are then the basis from which the German side of the dictionary (the German index) is generated. The German index is primarily meant as an access point to find the desired sign entries and is not to be taken as a list of fully edited German entries proper. For more information on the German index see section 10.4 below.
Other bilingual elements – that is integrating elements of the second object language German and information thereof in the dictionary– are:

- German translations of the DGS examples [BT]
- labels for diasystematic information on the use of the German word [EqL]

[12] For details and rationale of this mode of representation used throughout the DW-DGS in micro- and macro-structure see AP11-2021-1 (Otte et al. 2022). The sense preview box only appears with micons referring to an entire lemma sign, not to micons referring to a single sign sense. Micons referring to a lemma sign appear without any mouth activity.
[13] For more information on Word Sense Discrimination in the DW-DGS see also AP10-2016-02 (Langer 2020) and Langer et al. (2018a).
[14] While we consider this a feasible and practical solution, nevertheless the inclusion of signed definitions remains a desired and valuable information type for monolingually oriented sign language dictionaries in the future.
[15] In lesser researched, primarily oral or minority languages, the metalanguage used in dictionaries is often one that is more widely used and understood than the object language (e.g., Latin in the middle ages, Russian for indigenous Northern Eurasian Languages).
[16] It could be called a *bilingualized monolingual dictionary* (compare Hannay 2003, 152-153; Svensén 2009, 20) or a *bridge dictionary* (Svensén 2009, 20) since the definitions are given in German (the second object language). Such a hybrid approach is not unusual for sign language dictionaries. See for example for Auslan, the printed *AUSLAN dictionary* (Johnston 1989) or the online Danish Sign Language Dictionary (*Ordbog over Dansk Tegnsprog*).

- some indications of reflexive or transitive use and syntactic patterns, given through extended equivalent choices including prepositions and reflexive or object pronouns (***sich*** *bemühen,* ***sich*** *beschweren,* ***sich*** *einsetzen* ***für****,* ***etwas*** *verbessern,* ***etwas*** *wiederherstellen,* ***bei jemandem*** *übernachten)*

7.1 German Deaf Signers as 'Foreign' Language Learners or Users of German?

The prototypical situation for the use of a bilingual dictionary has been described as follows: One object language of the bilingual dictionary is the user's native language (their L1) and the other object language is a foreign language to them (their L2).[17] Minority language users surrounded by the majority language have a different relationship to the majority language than the typical foreign language learner. For instance, in many cases they can be assumed to be functionally bilingual to some extent. The same can be said for deaf signers and their relationship to the surrounding German language.

Most deaf people are to some extent bilingual and have some knowledge of German. Some may consider German as their only or their second native language. However, because of the lack of direct access to spoken German, for many deaf persons German remains a language that they are not fully proficient in. While some command of written German (at least in reception) can be assumed for most German signers, the relationship of German deaf signers to the German language might be somewhat different than that of a typical language learner to their foreign L2. On the one hand, German can be seen as the natural choice for written metalanguage information in a DGS dictionary, on the other hand, the command of German often remains incomplete in German signers and thus their information needs concerning German can be assumed to be similar to the ones of a language learner. For this reason, we use German as metalanguage of the DGS but in at the same time treat the relationship of deaf signers to German as that of a foreign language (L2) learner, knowing that this might not be a fully adequate assumption.

Furthermore the label 'L1' likewise does not describe the typical case here. An L1 DGS signer in our discussion of user groups is a person that uses primarily DGS in their everyday life and for whom DGS is the most accessible and expressive means of communication, the language of their emotional home and the language of choice. It is not necessarily the first language this person acquired, as many deaf children have hearing parents and are exposed to DGS not from the very beginning.

8 Dictionary functions and Use Cases

Depending on which language is the native language (L1) and which is the foreign or second language (L2) of the user, i.e. DGS or German, and which occasion for consultation (use case) is given different kinds of information are needed and could be provided in a bilingual dictionary of DGS – German. Naturally, information on L1 units need not be very detailed, while information on L2 units need to be more explicit and detailed, especially for production purposes (see DU3/4, HU3/4 below). Information already known to the native speaker/signer can appear to be superfluous or irritating and make it more demanding to find the sought new information. However, some information that is superfluous for certain user groups at specific occasions of consultation cannot be avoided when the dictionary is to allow for bidirectionality, that is, serve users of various groups – both with DGS and with German as their respective L1 languages. This is also a common practice in bilingual lexicography (comp. Svensén 2009, 18).

[17] Compare for example Svensén 2009, 14.

8.1 Functions of DW-DGS as a Monolingual Dictionary

There are two kinds of monolingual dictionaries depending on the user groups they are intended for: Monolingual dictionaries for native speakers/signers and monolingual dictionaries for learners (learners' dictionary).[18] Because native speakers/signers are much more competent in the object language than learners these two groups have different needs regarding the kind, depth and detail of information provided.

The following possible dictionary functions are a selective summary of the most prominent ones for a monolingual dictionary – adapted for a DGS monolingual dictionary[19]:

Use case	Dictionary function		Information sought in the dictionary
DU5	reception of DGS text (understanding)	reception of L1	meaning of DGS (unknown) sign*
HU5		reception of L2	
DU6	production of DGS text	production of L1	a) how to use a sign correctly (form; meaning; restrictions, style/usage; grammar: modifications, constructions; iconicity) b) whats the sign to use when… / for… c) what other signs can be used that express the same; signs with similar/ related meanings
HU6		production of L2	
DU7	correction of DGS text	correction of L1	a) verification of a sign's existence b) verification of correct use of a sign
HU7		correction of L2	
DU8 HU8	acquiring and strengthening vocabulary	vocabulary building	a) vocabulary related to a certain topic, subject field, theme b) signs of a specific region c) geographic names d) proper names
DU9 HU9	metalinguistic interest	search for specific kinds of information	signs with specific properties or of a specific class; dictionary grammar

[18] Learner's dictionaries can be specified further: L2 a foreign language vs. L2 as second language and with regard to the language competence of the learners: dictionaries for beginners, intermediate and advanced learners.

[19] This is a selective summary of different functions mentioned in literature for monolingual dictionaries of spoken languages (Tarp 2013, Engelberg & Lemmnitzer 2004, 93-105, Herbst & Klotz 18-21, Atkins & Rundell 2008, Svensén 2009, 13). The dictionary functions are described here with an adaptation to the case of DGS. Note that by giving this summary we do not claim that the DW-DGS can serve all of these uses in the same way. This listing is meant to provide a frame of reference for the further discussion of information types and their functions to the user.

DU10 HU10	pleasure, enjoyment, fun, interest and language awareness	non-purposeful browsing through the dictionary	any sign, sense, fact, map, or feature that catches the eye or attention; mediostructure: hyperlink cross-references, signposts, maps, examples

*unknown signs: regional variants, technical terms, old words/signs, foreign words

We expect that native signers will use the dictionary in text production mainly to reassure their knowledge about known signs (DU6/DU7) and in text reception they probably will encounter only few cases of unknown signs (mostly regional signs, old signs, or signs for technical terms) that they might need or want to look up (DU5). From our experience L1 DGS signers are especially interested in different regional signs used for the same concept. The DW-DGS may also be used to browse through and explore one's own language for the pure fun and enjoyment of it while discovering new and interesting facts about it (DU10). German signers have never seen and used a complex dictionary of their own language and it might take some time and education to get them accustomed and skilled in the use of the dictionary of their language.
Learners (L2 users) are confronted mostly with the core vocabulary of the language in reception (HU5) and for production tasks (HU6/HU7) they need to be given more detailed information on the meaning and use of a sign.

8.2 Functions of DW-DGS as a Bilingual Dictionary

Depending on the dictionary function and taking source and target language into consideration, we can distinguish the following eight major use cases for bilingual uses of the DW-DGS (adaption of Svensén 2009, 18 to our dictionary).

Users with DGS as L1 and German as L2 (abbreviation *DU* stands for assumed typical *deaf user*):

Use case	Dictionary functions		Information sought in the dictionary
DU1	reception of German text (understanding)	reception of L2	meaning of German word
DU2	translation of existing German text into DGS (understanding)		meaning of German word, finding corresponding/best matching DGS equivalent
DU3	production of German text	production of L2	finding adequate German word; instructions on how to use it correctly
DU4	translation of existing DGS text into German		finding corresponding/best matching German equivalent; instructions on how to use it correctly

Users with German as L1 and DGS as L2 (abbreviation *HU* stands for assumed typical *hearing user*):

Use case	Dictionary functions		Information sought in the dictionary
HU1	reception of DGS text (understanding)	reception of L2	meaning of DGS sign
HU2	translation of existing DGS text into German (understanding)		meaning of DGS sign, finding corresponding/best matching German equivalent
HU3	production of DGS text	production of L2	finding adequate DGS sign; instructions on how to use it correctly
HU4	translation of existing German text into DGS		finding corresponding/best matching DGS equivalent; instructions on how to use it correctly

9 Information Types

In the following, we describe the information types that are part of the microstructure as developed for the pre-release entries.[20] Which information types are included in a particular entry depends on the properties of the lemma sign described, the data available and the entry type.[21] All information types that can appear in the entries are listed in the table (see appendix). Here additional information and specific aspects of certain information types is given (codes refer to the codes for the information types listed in the appendix).

9.1 Information Types on Entry Level

Information types on the entry level are described here following the order they appear in in the microstructure of the entry.

[number] – Entry Number – [EN]

Each entry has an ID number, which is an integral number and uniquely refers to this entry and respective lemma sign. The number functions as the address of the entry. It does not carry any meaning by itself, but it is essential to ensure the unique identifiability of each entry. Particularly in places in the dictionary where multiple entries are listed, some micons and thus entries would be hard to keep apart from thumbnail movie or large movie alone. For example, when there are several cross-references to signs with same or similar form given in an entry [CRI, CRS], these signs look (almost) the same formally on thumbnail and on video; without the entry number the user would not be able to reliably distinguish between the listed entries. Thus, the entry number supports clear identification and navigation between entries.

FORM – Form – [V1/V2]

The citation forms of the main and additional variants, that are covered by the scope of the entry, are shown by means of movies. These form illustrations do not show any mouth activities

[20] The layout of the final product may differ to that of the pre-release entries and this may also affect some elements of the microstructure (e.g. ordering and spatial distribution of elements, use of labels, structure indicators, presentation of information types) but not the content of the entries as such.

[21] For more information on different entry types see AP10-2021-01 (Wähl et al. 2021).

and are thus an abstraction of the range of forms within the scope of the entry, as signs in context usually occur with some mouth activity, be it a mouthing or a mouth gesture.
The main variant [V1] is obligatory for each entry. The variants are identified by numbers added to the entry number, separated by a dot. 465.1 is the main variant of the sign described in entry 465; 465.2 is the second variant and so on. Next to each variant label is a play button that plays a studio recording of a signing model producing the variant in question.

KOMMENTAR – Comment on Form [CF]
A movie representing the form of each variant shows one specific execution of the sign variant's form, i.e. a clean, isolated execution of its selected citation form [V1/V2]. Sometimes additional information is given on the range of differences in form below the variant level. These are forms the specific variants may take on in actual use, for example dropping oder adding a second hand, minor differences in handshape, location or movement and the like. This information is included in the information type labeled KOMMENTAR.
This information is helpful in use cases that focus on the reception and understanding of DGS (DU5/HU5, HU1, HU2). When a person stumbles across an unknown sign in a DGS text, the first step is finding that sign in the dictionary to get more information on it. Information on possible form variation, as given in this section, helps the user determine whether the target sign is within the scope of a specific entry. The same goes for the corrective use cases DU7/HU7.

KOMMENTAR – Comments – [CU/CO]
The information type KOMMENTAR is rather flexible in terms of its contents. If there is something noticeable about the sign in question, e.g. regarding restrictions on use, this information will be added here. Comments include information on age variation, possible origins of the sign, more detailed information on frequency, and other details that do not fit into any other part of the entry.
Information on age variation is most helpful to hearing users, specifically for use cases HU2, HU3, and HU4. In the receptive use case HU2, the search for a corresponding German equivalent might be helped by a comment saying that the sign is more likely to be used by a young person rather than an older person, as the user can than choose an equivalent that shows the same tendency in German. The overall impression of the translation is then a closer match to the DGS text than it would have been otherwise. In the productive use cases DU6, HU6, HU3, and HU4, the user is helped in choosing the sign that fits the situation best. The user can avoid confusion, as an elderly person will not accidentally use youth slang DGS. In terms of the monolingual perspective, this is the type of information, that a regular user of DGS might just find generally interesting (DU10/HU10). They might start thinking about who uses a certain sign in their social surroundings and realize that there are differences there. The information thus help shape the metalinguistic awareness of L1 DGS users as well as advanced learners (DU9/HU9).
For a few signs, (speculative) information on its origin might be added. So far, this has been the case when signs saliently incorporate fingerspelling or an aspect of the fingerspelling of the word (e.g. entry 600). This has been particularly relevant for signs denoting months. The information on sign origin might be helpful to users expanding their vocabulary (DU8/HU8), as the user might be able to use the origin to remember what the sign looks like. It also might be interesting to deaf people, again, as a way of learning about metalinguistic connections within their language, in this case specifically about sign formation and perhaps even development (DU9).
The comment also might contain information about the frequency of individual sign variants. While the overall frequency of the lemma sign instantiations in the corpus is shown by the symbols in BELEGLAGE [FR] (see below), this does not say anything about the proportions of

the variants in relation to each other. If there are stark differences between the variants, this information can be added here (e.g. entry 421). Productive use cases (DU6, HU6, HU3, and HU4) benefit the most from this information, as the user can avoid using unusual variants of a sign.
Lastly, there may be various other comments on the sign that seem important enough to be mentioned in this information type. One example can be found in entry 364, where it is clarified in the comment, that a lot of instances of this sign are genderneutral and it is only the context that gives information about the gender of the person in question, if this information is given at all. Information like this is important to all hearing users of the dictionary. In receptive uses (HU1, HU2, DU5, HU5), it becomes clear that the sign is not necessarily gendered and the user is told where to look for information of the person's gender (context clues). In productive uses (DU6, HU6, HU3, HU4), the user is informed that this sign can be used independently of the referent's gender and can also be used in context where gender is not specified.

BELEGLAGE – Frequency – [FR]
The DW-DGS is corpus-based and the three boxes in BELEGLAGE indicate the overall frequency of the sign in the corpus data. This information is directly generated from corpus data. All tokens that belong to the scope of the entry as determined by lemma establishment and that are not marked as non-counting non-tokens (see Langer et al. 2016) are counted.

Symbol	threshold of corpus tokens
□□□	0 - 24
■□□	25 - 49
■■□	50 - 99
■■■	> 99

This information is especially valuable for users, when having to decide which of possible competing signs to choose in production (DU6, HU6, HU3, HU4). It may be also helpful for learners, sign language teachers, and other language professionals to determine which signs are more basic and which may belong to the more advanced vocabulary (DU8, HU8, DU9, HU9).

GRAMMATIK – Grammatical comment – [GrE]
As grammatical categories have not yet been solidly established for the description of sign languages, we are careful with using grammatical labels to describe signs. Some of the clearer phenomena are labeled in the grammatical comment. These include directional verbs (e.g. entry 358) and numeral incorporation (e.g. entry 361).
This kind of information is particularly useful to hearing users, as they might not be aware of the potential grammatical modifications of a given sign. In the receptive use cases (HU1, HU2), learning about these modifications will help identify the sign; while in the productive use cases (HU3, HU4), the user is informed that sign's form could or should be modified to fit the context that they might want to use the sign in.
In terms of monolingual use cases by DGS signers, the grammatical information might still have the same function described above, if the signer is not familiar with the sign in question e.g. due to its regional distribution (DU5, HU5, DU6, HU6). If the signer is familiar with the sign in question, the comment might still be interesting as a way of learning about the categorization of that sign's modification. The user might be looking at multiple signs with the same kind of modification and consciously realize for the first time that the modifications of these different signs work in the same way (DU9/HU9).

REGIONAL – Information on regional distribution – [REt/REm]
Regional distribution for a lemma sign including all its variants is given when there is a noticeable regional distribution attestable for the sign and all its senses. An examples can be found in entry 1546. In such cases, a written description of the regional distribution [REt] is given in the head of the entry, and an additional link to the distributional map is provided [REm]. There are also cases where only one variant form exhibits regionality as in 59, and 360.
DGS shows a considerable amount of variation, especially regional variation. Giving information on a sign's distribution thus becomes an important task for language documentation. DGS signers are generally aware of regional variation, but may not know all the possible signs there are, whereas learners usually start with one variant and encounter further variation over the time of learning the language.
Information on regionality of signs promotes the language knowledge for all signers as they learn more about the regional distribution of signs (DU8/HU8). The distributional maps may also support sign language teachers who can use e.g. the maps in their classes. Another possible use case is an interpreter preparing for a job in a region other to the one she/he usually works in (DU5/DU6). Using the common regional signs in an interpretation situation may reduce misunderstandings.
Accordingly, a bilingual use case is interpreting from DGS to German - a bilingual situation where regional signs may be of importance (HU1-HU4).
Additionally, deaf and hearing users can enhance or reassure their knowledge concerning the different uses of a sign's senses within different regions. This is useful for both language perception and production and furthermore expands the user's metalinguistic knowledge on DGS.

The central part of an entry is the list of senses section. Three different units are contributing to the list that is sign senses, phrases and multisign names.

BEDEUTUNG – Sense – [S]
For each lemma sign one or more senses of the sign are listed. A sense is one specific meaning or use of that entry's lemma sign described further in the entry. Each sense is identified by the section label BEDEUTUNG followed by # and an identifying sense number, e.g. 440#1 refers to the first sense of entry 440.
A sense section can be collapsed showing only the signpost [SP] or expanded showing all information types addressed to the particular sense.

PHRASE – Phrase – [Ph]
A phrase is a combination of the entry's lemma with another sign that forms one idiomatic expression with a specific established meaning or use. Phrases are placed in the sense list below the senses. They are marked by the label PHRASE followed by # and an identifying number. Phrases can be considered sublemmas as they are separate multiword expressions (MWE) listed in the entry. Examples of phrases are 440#10 and 649#7.
A phrase section can be collapsed showing only the signpost [SP] or expanded showing all information types addressed to the phrase.

MEHRTEILIGER NAME – Multisign name – [MN]
A multisign name is a proper name that consists of a multiword expression (MWE) for which additional information is provided. Multisign names are sublemmas that are placed in the sense list. They are marked by the label MEHRTEILIGER NAME followed by # and an identifying number. An example of a multisign name, here for the city of Bremerhaven, is 437#4.
A multisign name section can be collapsed showing only the signpost [SP] or expanded showing all information types addressed to the multisign name.

ZUSAMMENSETZUNGEN – Compound-like construction – ZEs/ZSe
Compound-like constructions are multisign elements whose makeup mirrors that of a German compound, in that each occurring sign semantically matches one of the components in the corresponding German compound. The German compound is usually also the mouthing of the DGS compound-like construction. As the meaning of the compound-like construction usually is compositional, transparent and corresponding to the mouthed German compound, no definition is provided in the entry. Compound-like constructions are marked by the label ZUSAMMENSETZUNGEN. When the compound-like construction cannot be related to any of the senses of a sign in particular a compound-like construction is listed as a sublemma at the bottom section of entry. For an example see (Kranken|haus) in entry 441.
Compound-like constructions are represented by several micons in sequence [ZEs]. The German compound is also listed [ZSe]. A vertical line indicates which parts of the mouthed compound are mirrored by the DGS signs of the compound-like construction. One of the compound's slots may be filled by more than one sign indicated by a frame around the sign options.

VERWANDT – Cross reference to related signs – [CRV]
Indicated by the label VERWANDT ('related') signs that may be considered related to the lemma sign are listed. Related signs share crucial formational elements that are interpreted as not being accidental but are best explained in terms of iconicity and modification – that is, they usually share properties of the underlying image and often are semantically close to each other. The general idea is that one of these signs may have derived from the other or that they share a common conceptual and iconic origin (see Langer at al. (2020): AP10-2016-1, section 2.1.4).
The information on related signs helps users build up a vocabulary knowledge (HU8) or become aware (DU8) of morphological, iconic and structural connections between signs. The cross-reference to related signs invites users to compare signs and reflect on iconic and formational processes and relations. See for example entry 458.

FORMGLEICH – Cross reference to signs with identical citation form – [CRI]
Indicated by the label FORMGLEICH cross-references to signs sharing the citation form in one of their variants are listed. See for example entries 350, 357, and 359. See entry 485 for indication of which form variant is identical to which form variant of another entry.
For deaf users, the cross-references to signs with the same form can be helpful after a form search, especially when the found entry does not seem to fit the expectation. The desired match may quickly be found within signs with the same form (DU3/DU4). For hearing users, the [CRI] likewise gives quick access to all possible meanings and entries related to one citation form (HU1/HU2).

FORMÄHNLICH – Cross reference to signs with similar citation form – [CRS]
Indicated by the label FORMGLEICH cross-references to signs with a similar citation form or form variant are provided. This information is useful especially for learners, to be or become aware of significant small form differences of signs leading to semantically distant signs (HU1/HU3/HU4). It also helps searching by form, e.g. if the first choice is not successful, the desired match may quickly be found within signs of similar form (HU1). The information may also be useful for DGS course instructors looking for examples of confusables (DU8/DU9) as well as their students (HU8/HU9), or for persons seeking inspiration for poetic production (DU10).

KONKORDANZ – Concordance – [CD]
Via a clickable button at the very bottom of an entry the user can open the concordance for the entry. This is a KWIC view of all of tokens included in the Public DGS Corpus that are matched to the sign described in the given entry. The concordance can be used to get more examples of the lemma sign being used in context and to see more form variations/modifications in different linguistic contexts.

9.2 Information Types on Sense Level

The middle part of the entry contains the list-of-senses section. Here senses of a lemma sign are listed, followed by MWE sublemma entries of multisign names and phrases containing the lemma sign. All three types are indicated by their category label followed by a # with a entry-internal running id for identification and addressing purposes. To each of these units (senses, phrases and multisign names) several information types are included in the microstructure to provide further information.

9.2.1 Three Units Contained in the Sense List

BEDEUTUNG – SENSE – [S]
Different contextual meanings and uses found in the corpus data are distinguished from each other, summarised and described as different senses in the dictionary (Word Sense Discrimination – WSD). In the entry each sense included is listed in the middle part of the table identified by the label BEDEUTUNG followed by the sense number (#[number]). The ordering of the senses within an entry is determined individually for each entry with the following guiding principles: First, groups of closely related senses or subsenses usually are kept together following each other in the list in a logical sequence, and second, more common/frequent senses are listed first followed by the more uncommon/infrequent ones.[22]
Senses of an entry can either appear in the collapsed overview form – displaying only the signpost information [SP] of the sense – or they can be expanded and display all information types given for this particular sense.

PHRASE – PHRASE – [Ph]
In the DW-DGS, a *phrase* is a MWE having a meaning which is non-compositional and not transparent. A phrase is usually composed of two elements, though it is possible that there are phrases with more constituents. With regard to microstructure the dictionary treats phrases similar to senses. The information types provided for senses are also given for phrases. Additional information on the FORM is provided for phrases: A studio recording shows the whole phrase including a typical mouthing [PhM] and a set of micons represents the signs contained in the phrase [PhMic]. Examples for phrases are 440#10 and 649#8.

MN- MEHRTEILIGER NAME – Multisign name
Multisign names are proper names that are composed of more than one sign. With regard to microstructure multisign names are treated similar to senses. The information types that are given for senses, are also given for multisign names. Additional information on FORM is provided for multisign names: A studio recording shows the whole multisign name including a typical mouthing [MNM] and a set of micons represents the signs contained in the multisign name [MNMic].

[22] As we do not analyse and tag all tokens of a type for senses commonality/frequency here is not a robust count but more an impression of the lexicographer gained during sample tokens analyses and scanning of the corpus data.

Examples for multisign names are 17#2 Rotes Kreuz (red cross) and 437#4.

9.2.2 Information Types Addressed to a Sense

{no label} – Signpost – [SP]
In the DW-DGS we use signposts[23] for providing an overview and quick access to multi-sense entries. Each sense is given a signpost which is placed at the top of each sense section of the microstructure and is shown in the menue-like overview of the collapsed sense-section of the entry.[24] The fully collapsed view is given e.g. when opening a multi-sense entry from the DGS index (e.g. 193, fully collapsed). With a click on the label BEDEUTUNG each sense can be expanded and collapsed individually by the user for navigating and exploring. When coming from different places in the dictionary (e.g. from a synonym cross-reference or the German index) the respective addressed sense will already be expanded in the entry when opened (e.g. 193#10, sense 10 expanded).
Signposts are used for disambiguation and do not contain full definitions. They just hint at the senses' meaning/use. Often a signpost includes the most common equivalent, a combination of equivalents and/or sometimes a short paraphrase. It may also contain a semantic category, a domain label or some other disambiguating context. In the collapsed entry signposts support the scanning eye to browse through the semantic information and choose the sense of interest for further investigation.
Signposts are especially useful for users that are looking for a specific meaning of a sign, for example in the following use cases:
DU4: Having a sign and a specific meaning of the sign in mind, the deaf user needs to find and identify the corresponding sense to look up matching German equivalents.
HU5/DU5: To extend one's vocabulary repertoire a user might want to learn about a sign's range of meanings. For this they can look at the summarizing signpost menue for an overview and expand individual senses step by step for more information about meanings and uses.

FORM – Form of phrase or multisign name – [PhM, PhMic, MNM, MNMic]
This information type is provided only for the sublemma categories of phrases and multisign names. A movie with a studio recording of the whole phrase or multisign name is provided. The movie is opened in the movie display area by klicking on the arrow icon in the FORM section. The form of the MWE is presented as studio recording in which a signing model signs the phrase/multisign name with the most common mouthing [PhM, MNM**]**.
To further illustrate the form of a phrase or multisign name, its is represented by a string of micons, one micon for each sign that contributes to the phrase or multisign name. In some cases one part of the multisign pattern can be expressed alternatively by several lexical variants. In that case, all signs are shown that occur in the corpus in this combination with the respective meaning. Other possible lexical variants not evidenced in the corpus are not included. Micons of alternative signs for one part are grouped together by a frame. The micon that refers to a

[23] Signposts in dictionaries are guiding elements with the function of providing quick cues to disambiguation of the senses of long polysemous entries and thus providing orientation for guiding the user effectively to the particular sense that they are looking for. Signposts were invented for and are primarily used in monolingual learners' dictionaries. They are either placed together as a menue at top of the entry or given separately as signposts before each full definition (cf. Atkins & Rundell 2008, 444 on *short definitions*).
[24] The signpost information is optimized for contrasting several senses of a lemma sign in the collapsed entry. Apart from this – their original function in the entry – they are also re-used in two other functions in the dictionary: First, they are also listed as a summary of the signs' range of meaning in a *sense preview box*. This box appears in places where micons represent the whole entry when the mouse hovers over the entry number of the micon. The box can be found at the DGS search page or in non sense-specific cross-references e.g at FORMÄHLICH, here the sense preview box helps to select the sign of interest among similar-looking signs on basis of its meaning.

form of the lemma sign of the entry that is viewed, is marked with a point in a red box instead of the usual ID number.
The display of the part of the MWE with micons is particularly useful to hearing users, as it explicates which signs contribute to the multisign construction and allows them to get more information on the individual components by following the linking to their respective entries. Examples can be found in entry 420 phrase #6 and entry 17 multisign name #2.

MUNDBILD – Typical mouthing/mouth gesture – [Mtht, MthM]
In this section, the most common mouth activities associated with the described sense, phrase, or multisign name are given.
A movie of the sign with one typical mouthing or mouth gesture is provided. It can be activated in the movie display area by clicking the "play"-button. The movie shows a studio recording of a signing model producing a sign with the most frequent mouth activity occurring with this sense. In the case of a mouth gesture, the movie shows what the mouth gesture looks like. Following the play-button, typical mouthings are presented in writing without capital letters so as to distinguish it from information on the second object language German, since mouthings are considered information on the form of a sign in actual use. Mouth gestures are indicated by *[MG]* in this section. The list of mouth activities given in this section is non-exhaustive and based on frequency observations in the DGS corpus. This information is given for all senses, phrases, and multisign names, but examples can be found in 193#12 (sense), 440#10 (phrase), and 17#2 (multisign name) respectively.

ERKLÄRUNG – Explanation (dictionary definition) – [**Def**]
An explanation [ERKLÄRUNG] of the specific meaning/use is provided for each individual sense – paraphrased and described in written German. These explanations differentiate the senses a sign can have and indicate the contexts of the sign's use in each sense. It is important to recognize and understand a sense in question for reception as well as for production and to identify it to access further information provided on it. The explanation is a means to clearly but also abstractly describe the sense. It is complemented by further information such as synomyms and antonyms [BEDEUTUNGSGLEICH, ENTGEGENGESETZT] and examples [BEISPIELE] that illustrate the repective sense. These three information types directly complement and mutually contextualize each other and thus together contribute to the clarification of a sense. Explanation [ERKLÄRUNG] and German equivalents [DEUTSCH] each also promote and solidify the understanding of the repective other information type.

DEUTSCH – German translation equivalents – [EqW]
German translation equivalents are listed in German words and phrases separated by small bullets under the label DEUTSCH. The most prominent and best-matching structural equivalents are linked to the German index and marked in blue as opposed to default black of additional equivalents not listed in the German index.
Translation equivalents provide a quick access to the meaning of a DGS sign, especially for learners of DGS in reception (HU1). When translating DGS into German (HU2, HU4) the list of translation equivalents help to find a suitable word for translation in a given context. German equivalents also help in case of DU3 (production of German text for signers with DGS as L2) even when this use case not optimally served. For production of L2 or translation into one's L2, more information on the L2 item is needed than just some of the semantics. In the DW-DGS we do not provide full coverage of the German words and their properties as object language. This shortcoming is made up for by links to the German index, which again link the German equivalents to their respective entries in the DWDS, an online Dictionary of German. Here, deaf users with German as L2 can look up information on the German equivalent concerning grammar and word formation, e.g. grammatical gender of nouns, forms of the cases etc. In the

DW-DGS the focus is on the semantics and some additional syntactic behavior which aids the identification of the sense of the German equivalent, e.g. by adding reflexive pronouns or hints on the syntactic construction with respect to the animacy of the object referent (e.g. *jemanden* vs. *etwas*), or constructions requiring sense-specific prepositions, e.g. *Geduld haben mit*, *übertragen auf*. Another kind of information the DW-DGS gives on German translation equivalents is some hints on register, style and usage (see EqL below).

DEUTSCH – disambiguator – [EqD]
The disambiguator is addressed to the single translation equivalent. German translation equivalents are often polysemous and therefore ambigious. In general, a list of equivalents often disambiguates each element of the list, so no further disambiguator is necessary in the entry. In other cases an equivalent needs an additional disambiguating context to specify which meaning or nuance of the polysemous German word is equivalent to the sign's sense or to indicate the scope or domain of used. For example, the sense 606#5 refers to movements of the mouth that are visible when articulating words (of a spoken language) or mouthings of a signed language. Instead of establishing two senses, these are lumped to one, giving a hint on the use domain of the respective German translation equivalent: *Mundbild (Gebärdensprache), deutlich artikulieren (Lautsprache).* Examples for disambiguating a German polysemous word see: 84#6, 485#2. Entry 417#3 exemplifies the default case not showing disambiguators in the entry, but in the German index where they are necessary for sense discrimination and disambiguation (comp. Svensén 2009, p. 261 f.).
This information is particularly relevant to use cases HU2 and DU4, as it aids the selection of the right German word from the list of equivalents in a given context when translating from DGS.

DEUTSCH – diasystematic label – [EqL]
The diasystematic label is addressed to the German equivalent and indicates that it is marked i.e. it belongs to a certain register, style, subject field, or is a metaphorical meaning extension and the like. The label is also used to give hints on the grammatical markedness as {pl.} (only used in plural form).
The label makes explicit what is intuitively known to most users with German as L1. It is valuable information for deaf users with incomplete command of German – when producing German texts or translating into German (DU3 /DU4). The information is relevant in order to find a matching of register or style, and to become aware of properties of the German equivalents that may not be shared by the DGS expression.
The following labels are used in the DW-DGS to provide information on the usage of the German word (equivalent):
abwert. (derogatory),
fachspr. (technical term),
fam. (informal, intimate),
form. (formal),
gehob. (literary),
idiom. (idiomatic)
metaph. (metaphorical, figurative)
ugs. (colloquial)
reg. (regional)
pl. (plural form)
The diasystematic labels also appear with the equivalent in the German index.
Examples of label use can be found with listed equivalents for 443#7, 136#3, 409#1, 628#4.

ANMERKUNG– Additional information on usage – [CS]
This information type contains comments on particularities in the sign's use that are specific to a particular sense. Most commonly, we describe some minor differences in form. This might mean that only one of the form variants within the entry's scope is used for this sense or it might mean a different modification like a repetition of the sign's movement. Other usage patterns can also be described here, such as the use in certain age groups or usage within role shift. This information is presented in written German. Examples can be seen in 361#1 and 417#3.

GRAMMATIK– Additional information on grammar– [GrS]
Grammatical information specific to the described sense can be given under the label GRAMMATIK. Mentioned here are sign modifications in terms of location to indicate arguments, either by moving the sign's location as a whole or, in the case of directional verbs, by starting and ending the sign in locations that have referents assigned to them. For the latter, we describe the types of arguments in written German and add an arrow in between to indicate the movement of the sign. In entry 351 several sense contain this information type.
This kind of information is particularly useful to hearing users, as they might not be aware of the potential grammatical modifications of the sense. In the production use cases (HU3, HU4), the user is helped in finding a modification suitable for the context that they might want to use the sign in. For L1 DGS signers, the comment might still be interesting as a way of learning about the categorization of that sign's modification. The user might be looking at multiple signs with the same kind of modification and consciously realize for the first time that the modifications of these different signs work in the same way (DU9, HU9).

BEISPIELE – authentic examples – [BM]
We decided to use so called *authentic examples*, i.e. examples taken directly from the corpus in opposition to edited or competence examples.[25] The example movie [BM] may be played in the movie display area by clicking on the play button on the left to example context and translation.
Authentic examples have the overall advantage that through them it is possible to show a wide range of signers within the dictionary. Another advantage is that these examples show language in use which is as natural as corpus data may possibly be. For DGS learners this variation of signers and thus signing styles may be initially somewhat challenging, but it is also a good perception practice for encounters with deaf people outside the DGS classes, where signing is often adapted to the needs of novel learners (HU1).
The examples support the understanding of the senses by illustrating them. Deaf users may watch the examples in order to ensure they have found the correct sense of a sign they know and thus the correct translation equivalents they can use for it in a German text (DU3).
Another case is that a signer has found the German word but there may be different signs that can be used. Watching the examples may clarify further which sign is the most suitable one (DU1, DU2).

BEISPIELE – context to the example – [BC]
The context to the example is an optional information given when the example cannot stand alone and needs context information. The context is written German text and is presented in square brackets. The context can be a key word or even a full sentence.[26]
In the DW-DGS authentic examples are shown outside the context they were produced in and it happens regularly that some explanation is necessary to understand the example itself, e.g

[25] For more information on the topic see Langer et al. 2018b.
[26] For more information on the example context see Langer et al. 2018b.

resolve unclear references. The example context provides the necessary information to understand the example properly. It is also easier – especially for learners – to understand an isolated example clip when being primed on the topic or context of the example first. This information is beneficial for all user groups.

BEISPIELE – German translation of example – [BT]
DGS examples are translated into written German. The translation equivalent of the target sign is written in bold face to highlight it.
Deaf users can have a look at the chosen translation equivalent and see directly how it would be used in a German utterance. Thus they can learn more about the use of the translation equivalent in question and may even apply it for language production of a written German text (DU3).
DGS L2 users, especially learners, are supported in the understanding of the example as they either know the translation before watching the example or can read it afterwards and compare the signed example and the translation. They also see an example for translating the respective sense of the sign (HU2).

BEISPIELE – Link to MEINE DGS – [BL]
Many examples shown in the DW-DGS are also published in their larger context in the Public DGS Corpus (MEINE DGS – MY DGS). Where this is the case two red buttons appear below the movie display area area whenever the play button for the example is clicked. The buttons are links to the respective movie the example can be found in in to *MY DGS* and *MY DGS – annotated* [27].
Clicking *Ganzes Gespräch in MEINE DGS* will lead to the movie on the website MY DGS but not to the exact time where the example appears. Clicking *Transkript in MEINE DGS – annotiert* leads to the respective segment within the transcript where the example can be found. This function allows the user to view the wider context of the example. It shows directly what the data basis for the DW-DGS is – the DGS Corpus. It reassures the user that the information given is based on language data which is even in parts publicly available.
In the annotated transcript of the Portal *MY DGS – annotated*, learners of DGS can look up other signs of the example they may not know, click to the types list, scroll to the top of the types entry and find cross reference(s) to the corresponding dictionary entry or entries (if existent), thus finding more information on the other signs occurring in a specific example (cf. Müller et al. 2020).

LANGFORM – Long form – [LFMic, LP]
Some senses may be expressed by either a multisign unit (the long form) or a single sign (the short form) that is one of the parts of the multisign unit. An example of a long form can be seen in entry 441#3. The sign 441 (main sense: ‘sick’) can mean health insurance company (literally from DGS: sickness fund) on its own, but also may combine with another sign to express the same meaning.
Long forms are represented by series of micons [LFMic]. One of the long form’s slot may be filled by more than one sign, the alternatives are then represented next to each other grouped by a surrounding frame.
In addition to the micons the German compound equivalent [LP] corresponding to the long form is given. Long forms usually mirror German compounds in their makeup. A vertical line indicates which parts of the German compound are represented by which signs of the long form.

[27] For more Information see Müller et al. 2020.

BEDEUTUNGSGLEICH – Synonymous or near synonymous sign – [SynMic]
Whereever possible synonyms or near synonyms are provided for a particular sense. These are signs that have the same meaning as or a very similar meaning to the sign described in the sense. These synonyms fulfill three roles: First, they function as a kind of secondary definition[28] to the entry's sign. L1 signers for whom the written German definition might not be easy to understand can look at the DGS signs listed as (near-)synonyms and still gauge the meaning of the sign. This corresponds to use case DU5 as these are deaf users looking at signs that they are not familiar with e.g. due to regionality. Hearing users that prefer to learn about a sign's meaning by looking at DGS rather than a written German definition would profit in the same way (HU5). Second, by listing signs with the same meaning, they give the user the opportunity to learn these signs as well e.g. when trying to expand vocabulary (DU8, HU8). In use cases HU1 and HU2, the hearing user looking up an unfamiliar sign might choose to look at the synonyms as well and will then have an easier time understanding the rest of the text. Third, the synonyms function as cross-references to other dictionary entries. This is particularly helpful for all translational use cases working towards a translation into DGS (DU2, HU4). The list of synonyms serves as a set of potential alternative translations. Similarly, if the user is correcting a DGS text (DU7, HU7), the synonyms are a quick way to look at alternatives to signs that might not quite work in the given context. This is helpful on its own, but is supported by the cross-references as the user is then able to jump to the entry of an alternative translation they are considering and can thus learn more about whether the sign is a good fit or not.

BEDEUTUNGSGLEICH – Regionality of synonym – [SynR]
If a sign is listed as a synonym to another sign, but the two differ in their regionality, the label *regional* is added underneath the micon of the synonymous sign. The user may then get more information on the regionality of the synonymous sign by clicking on it and checking the information in the sign's entry [REt, REm]. The use cases for this information type are the same as for [SynCM], listed below.

BEDEUTUNGSGLEICH – Constrastive Map of synonyms – [SynCM]
DGS is known for having several regional variants for some concepts, especially for specific sets of concepts such as signs denoting months, weekdays, and people (woman, man, girl, boy etc.). Here signs tend to differ depending on the region. Beyond just listing the signs used in other regions as synonyms, it thus makes sense to give bundled information on what sign is used where. This is done via a contrastive map of synonyms. The different signs are represented by different colors and the German regions are filled with the color that corresponds to the most common sign in that region. Additionally, the number of attested signers using this sign are given for each region and a pie chart indicates how many different signs have been used in the particular region by how many people. The maps are thus very detailed, but still well suited for a quick overview over the general distribution of signs. As a visual representation it is more accessible to deaf users than a written description would have been. For blind users, a machine-readable table containing the same information is provided.
An overview of existing regional variants for the concept is an information that deaf users take a general interest in (DU10). In terms of translating German to DGS (DU2, HU4), this information might be helpful to chose the DGS sign that is either the most widely used one or the one that matches the region of the intended addresse of the DGS text. When producing DGS text (HU3) or doing vocabulary work (HU8) the user learning DGS is helped in finding the

[28] Synonyms or antonyms can support the differentiation and disambiguation of the respective sense. They can function as elements of (cumulative) synonym definitions in DGS. For (cumulative) synonym definitions see Svensén (2009, 214-217).

appropriate sign used in their region. This way, a DGS learner in Northern Germany is unlikely to accidentally choose a sign from Southern Germany.

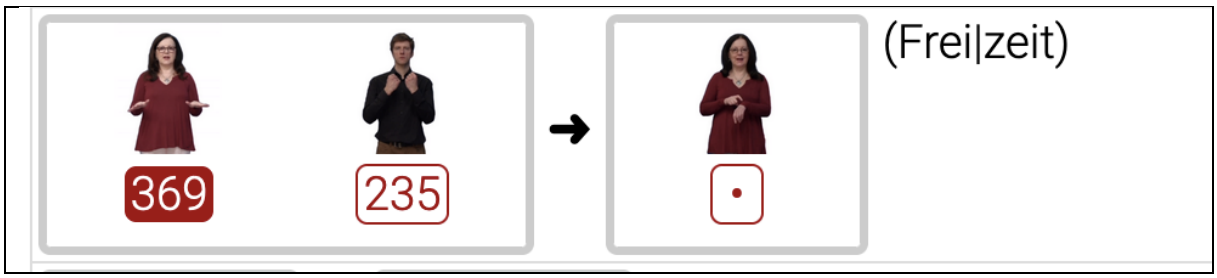
ENTGEGENGESETZT – Antonymous sign – [AntMic]
Marked by the label ENTGEGENGESETZT signs with opposite or complementary meaning (antonyms) are listed for a particular sense. Just like the synonyms, antonyms are pieces of information directly displayed in the entry and at the same time they function as cross-references to the corresponding entries. The use cases for antonyms are also very similar to those of syonyms.
First, DGS antonyms also contribute to the information on the sign's meaning. A user that prefers to learn about a sign's meaning in DGS rather than written German (DU5, HU5), can use both the synonyms and the antonyms to gain an impression of the sign's meaning. Second, the antonyms can be used to expand the user's vocabulary (DU8, HU8). When creating translations into DGS (DU2, HU4), looking at antonyms in addition to the target sign may help with the remainder of the translation.

ENTGEGENGESETZT – Regionality of antonym – [AntR]
When a regional sign is listed as an antonym to another sign, the label *regional* is added underneath the micon of the antonymous sign. The user may then get more information on the regionality of the antonym by clicking on it and checking the information in the sign's entry [REt, REm].

ENTGEGENGESETZT – Constrastive map of antonyms – [AntCM]
Contrastive maps of a cluster of several competing regional signs for a concept is provided in in the antonym spot when possible. These maps are of the same type as contrastive maps of synonyms [SynCM]. They are included in an entry, if the antonyms to the sense described in the entry form a cluster of competing regional signs. For example, entry 127#1 describing in the sense 'woman, adult person' contains a map with the different signs for 'woman' in the synonym section and a map with the different signs for 'man, adult person' in the antonym section.
In reception of DGS (use cases DU5, HU5, HU1, HU2) antonyms might help understand the remainder of a text. So, looking at antonyms when looking up a specific sign makes sense and specifically looking at antonyms that match the sign in regionality increases the probability of that sign appearing in the remainder of the text. Also, when expanding vocabularly (DU8, HU8), it usually makes sense to focus on vocabularly from one region. The map helps determine which signs fit the bill. Due to the visual nature of the map and the cross-references to many different entries, these cluster maps are also a fun feature for people browsing through the dictionary without a specific aim (DU10, HU10).

HÄUFIGE KOMBINATIONEN – Collocational patterns – [CollMic]
The dictionary includes information on collocational patterns and semantic preferences patterns for signs. These are presented as series of micons with alternate signs being shown within the same frame. Semantic preferences are indicated in written German, either within the frame of alternate micons or, if no micons are given, in their own frame. The order the signs appear in is expressed through arrows between the micons. These might either be unidirectional if the signs always appear in a set order or they may be bidirectional if the order is undetermined. Examples of information on collocational patterns can be found in 354#1, 327#3, 193#2.
Collocates help the user distinguish between different senses and give an impression of the sign's usage in context. As a kind of short examples collocational patterns are especially helpul in L2 DGS production (HU3/HU4) as they provide templates of linguistic contexts that are common with the respective sign.

ZUSAMMENSETZUNGEN – Compound-like constructions – [CPDMic, CPDeq]
Compound-like constructions have been described above (see ZEs, ZEe). When a compound-like construction can be related to a particular established sense of a sign it is listed with that particular sense. The information provided consists of the sign combination in form of micons [CPDMic] and the German compound in written German [CPDeq].
The micons used for the compounds are ones that have the sign model articulating the mouthing. For each sign involved in a compound-like construction a separate micon is displayed. For example, the compound-like construction mirroring *Frei|zeit* (spare time) in entry 354#1 includes two signs for 'free' matched to *frei* (free), each of which may be used in the compound-like construction, and one sign for 'time' matched to *zeit* (time). The section in this case thus includes three micons.
The compound-like DGS construction is mirrowing a German compound in the sequence of parts. The German compound suggests the structure of the sign string and is mouthed along the sign production. It is represented by written German [CPDeq]. The two (or more) parts of the German compound that are mirrored by the sequence of signs are separated by a vertical line. The written German word also represents the mouthing that accompanies the signs and functions at the same time as translation equivalent. It is also included in the German index.



The information on compound-like constructions can be helpful in a number of use cases. When encountering an unknown sign during the reception of DGS (DU5, HU5), it might be helpful to spare a glance at the compound-like constructions in case the occurrence in the given DGS text is part of such a construction. Being aware of the connection of the two (or more) signs will facilitate a better understanding of the text. Information on compound-like constructions might also help in correcting a DGS text (DU7, HU7), as the user may be seeking to confirm that a second form of a compound does in fact exist. A user who is expanding their vocabulary (DU8, HU8), can profit from learning about the compound-like constructions a sign may appear in alongside learning the sign itself.
The access to DGS compound-like constructions via the German equivalent in the German index is particularly helpful to users coming from a German text and seeking to either understand (DU1) or translate it (DU2, HU4) or for L2 text production in DGS (HU3).

REGIONAL – Regional distribution text and map – [RSt, RSm]
When there is a prominent regional distribution attestable for a particular sense in contrast to the other senses, e.g. senses 414#1 and 414#4 information on regionality is included on the sense level. The same sign (414) is used in different regions for different senses.
In such cases a written description of the regional distribution may be found in the field REGIONAL [RSt]. In addition a link is provided that leads to the corresponding distributional map [RSm].
For this information type the use cases described for regional distribution on sign level [REt, REm] apply accordingly.

SACHGRUPPEN – Subject Areas – [Sub]
Each sense is attributed with one or several subject areas. Subject areas can be used as a quick, concise indication of what contexts a sign might be appropriate for. In the entry subject areas are realized as hyperlinks. In the respective subject area that lists other signs used in comparable situations are listed. This information is particularly helpful to learners that want to produce a DGS text on a specific topic (HU3, HU4).
Some subject areas such as *Eigenschaften* (properties) and *Tätigkeit, Handlungen* (actions) not only let the user know about possible contexts of the sign but also may indicate how to use the sign. A sign in the subject area property is likely to be a descriptive sign and should be used accordingly. This is helpful to use cases concerned with L2 DGS production (HU3, HU4), but also might be interesting from a metalinguistic perspective (DU9, HU9). The latter use case is also served by the subject area *Wendung/pragmatische Funktion* (idiom, pragmatic function), which includes many of the less straightforward sign uses, e.g. senses like 'Bless you!'.
The subject areas also serve as an access structure (see section 11.5 Subject Index for a description and the use cases of the access structure). The list of subject areas in the entry also serves as a link to the subject index and thus can be used to quickly navigate to other signs that may be used in a similar context.

10 Access Structures

Access structures allow users to navigate directly to the parts of the dictionary that match their informational needs on a specific occasion, that is primarily entries and senses.

10.1 Macrostructure of DGS entries

The macrostructure is the primary access structure of a printed dicitionary. It involves a clearly defined conventional order in which entries are listed. For dictionaries of spoken languages, the most used method is the alphabetic ordering of the lemmas. For sign languages, we have neither a written representation of the citation form, nor a conventionally fixed orthography, nor a conventional ordering (alphabet) of writing symbols. Printed sign dictionaries with a more monolingual perspective, such as the Auslan (Johnston 1989) or BSL dictionary (Brien 1992), usually order the sign entries in accordance to certain paramaters of sign form (such as handshape, height of place of articulation, number of hands, and the like) in a given fixed order. The macrostructure enables the user to determine the place of an entry according to the sign's form in the sequence of entries to facilitate a search via sign form. In electronic dictionaries the macrostructure – that is the method of ordering dictionary articles – becomes less important when a search function via sign form is provided. However, an adequate macrostructure may be desirable for political reasons and to encourage browsing through entries, because they place entries in neighboring positions that are similar in the formational element used for ordering. In the DW-DGS pre-release entries, the DGS index offers four different entry orderings to choose from: First, by number (*Nummer*): Entries are displayed in numerical order by entry number, which is a randomly assigned number without further significance except for being unique and unambiguous. Second, by handshape (*Handform*), third by number of involved hands (*Händigkeit*), and fourth by Location (*Lokation*).

10.2 Search Function via Sign Form

A search function via sign form will be provided in the final product. Such a search function facilitates finding the entry of a specific sign the user has encountered or in mind, as in the use cases HU1, HU2, DU3, and DU4.

10.3 Graph

The graph (Graph) is a visual structure in which all of the DW-DGS entries are represented. The entries are visualized as little dots, which are supplemented by micons when the user hovers over a dot. The entry dots are connected by differently colored lines, which represent different relations between the signs described in the entries. These relations include synonyms (SynMic), antonyms (AntMic), collocations (ColMic), MWEs (Ph, MN, ZEs, ZEe), compound-like constructions (CPDMic), related signs (CRV), and having the same (CRS) or a similar form (CRI).

The main purpose of the graph is to serve as an overview. As the user of a printed dictionary can look at the book(s) and immediately get an impression of the size of the dictionary, we would like our online users to be able to look at the graph and realize how many entries and cross-references between them we have. The user can choose to show or hide each category of cross-references individually. Seeing the different categories presented in this graph hints at the different aspects that we consider in the dictionary: semantics of a sign, form of the sign, use in context etc.

Together with the DGS index, the graph serves as a visual, non-text-based access way to the dictionary. The text on this page is kept to an absolute minimum: only the filter function of the different categories of cross-references uses one word German descriptions. As such, it is an access way suitable for users that are less comfortable navigating the more text-based German index and subject index.

The filter function that allows the user to modify the graph to only show certain kinds of connections gives the graph a playful element. Through this, it is well suited to aimless exploration of the dictionary and let's the user stumble across entries without specifically looking for them. This might be particularly interesting for deaf signers that use DGS often and are just looking to get an impression of the dictionary rather than asking it to perform any specific task for them.

10.4 German Index

The German index (Index Deutsch) is generated from the translation equivalents listed in the entries. When a polysemous German word is given as a translation equivalent, a particular sense of the German word is distinguished by disambiguating specifications where necessary or sensible. The German index is automatically compiled and groups all DGS equivalents with the same German base word together with disambiguating hints. This results in a structured alphabetical list of entry-like German words with some degree of sense discrimination provided as needed for the parallelisation with the corresponding signs and their sense structures.

It has to be noted that no specific lexicographic effort goes into the description of German words for their own sake as German entries. Links to external lexicographic online resources (DWDS) are provided in the German index where possible in order to give more information on the German words and their uses. This feature primarily benefits the native signers: Having found German equivalents to a specific sign, they might want to know how to use the German equivalents. Wanting to learn more about a German word, the signer could also look it up in the German index, getting some indication of its senses through the matched sign equivalents provided in the index. When further information on the German words is needed, it is only one click away.

Lemma selection on the German side of the dictionary is limited to the equivalents resulting from the translation equivalents given for the signs. Thus, the equivalents are limited to what the DGS-entries require and provide. This is not corpus-based and does not treat German from a monolingual perspective in the same way as the DGS side is treated, e.g. no elaborated entries proper for the German words in the index are provided. However, within the limits of the project resources, we consolidate the German index list to entry-like groups that are headed by a German word and indicate different senses matching DGS-sign senses by disambiguating

elements. The equivalents also include indications of syntactical information (see below) whenever needed to distinguish between different uses of the German word.

Example of entry-like grouping of equivalents in the German index:

behandeln	behandeln (Medizin)
	behandeln (Umgang)
einfach	einfach (kurzentschlossen, unbürokratisch)
	einfach (mühelos, unkompliziert)
	einfach (nicht komplex)
	einfach (schlicht)

It is interesting that deaf members of the focus group and other deaf people confronted with the dictionary for the first time often specifically ask for the search option via German words / equivalents. This expresses their expectation of having an access via their L2 to the entries of their L1 sign language[29], a search function that is realized by the German index.

10.5 Subject Index

The subject index (Index Sachgruppen) lists subject areas and the signs' senses that have been attributed to them. It consists of three columns: the subject area, the matching senses' SignPost, and the micons representing the sign entry and its relevant sense. We have implemented a two-level hierarchy in the subject areas, as there are cases of larger subject areas, e.g *Bildung, Lernen* (education, learning) that are best sorted into smaller sets to be helpful to the user, e.g. *Bildung, Lernen: Schule* (education, learning: school) and *Bildung, Lernen: Schulfächer* (education, learning: school subjects). As can be seen in these examples, these hierarchies are visualized simply by adding a colon between the larger set and the smaller subset.

The subject index is an useful access point to novel learners of DGS, DGS teachers, interpreters, and competent DGS signers. Novel learners of DGS can use the groupings by subjects to practice vocabulary, as DGS classes commonly teach signs by topic. Likewise, DGS teachers preparing classes can use the subject index to find vocabulary that they might want to add to a lesson on a certain topic.

Two user groups with a similar aim but in different contexts are hearing people that start interacting with deaf people in certain situations and interpreters preparing for assignments. A hearing person meeting a deaf person in a specific situation (at work, at a sports club etc.) might want to learn some signs specifically for this context. Subject areas such as *Arbeit und Beruf* (work and occupation) and *Sport* (sports) are a good place to start this kind of vocabulary work. Interpretators are also likely to want to prepare for interactions in a specific contexts depending on their assignment. For example, if an interpreter is booked for an appointment at a school, they might want to use the subject group *education, learning: school subjects* to brush up on the signs for the school subjects that might come up in conversation.

Lastly, feedback from the deaf community shows us that the subject index is also a preferred access point for native DGS signers. When looking for a specific sign in the dictionary, some signers use the subject index rather than the German index. They feel that there are more likely to find the sign they are looking for when looking at groups of signs rather than having to navigate in the alphabetically orded list of the German index.

[29] This is in line with the observation Popkema (2010, 89) describes. A common search strategy of native users of a lesser used language, that displays much variation and has no orthography, is to use the dominant (L2) language as source for searching for a word of their own (L1) language as target in a bilingual dictionary e.g. to check its spelling.

11 References

11.1 Lexicography

Atkins, B.T. Sue / Rundell, Michael (2008). *The Oxford Guide to Practical Lexicography*. Oxford, New York: Oxford University Press.

DeCesaris, Janet (2012). On the nature of signposts. In: Fjeld, R. V. / Torjusen, J.M. (eds.): *Proceedings of the 15th EURALEX International Congress*. Oslo. Department of Linguistics and Scandinavian Studies, University of Oslo, 532–540. URL: https://euralex.org/wp-content/themes/euralex/proceedings/Euralex%202012/pp532-540%20DeCesaris.pdf (visited on 2022-05-20).

Dziemianko, Anna (2019). Homogeneous or Heterogeneous? Insights into signposts in Learners' Dictionaries. In: *International Journal of Lexicography*, 2019, Vol. 32, No. 4, 432–457. https://doi.org/10.1093/ijl/ecz011 (visited on 2022-05-20).

Engelberg, Stefan / Lemnitzer, Lothar (2004). *Lexikographie und Wörterbuchbenutzung*. 2. ed. Tübingen: Stauffenburg Verl.

Hannay, Mike (2003). Types of bilingual dictionaries. In: van Sterkenburg, Piet (ed.): *A Practical Guide to Lexicography.* Amsterdam: Benjamins, 145-153.

Herbst, Thomas / Klotz, Michael (2003). *Lexikografie.* Paderborn: Schöningh.

Kristoffersen, Jette H. / Troelsgård, Thomas (2012). The electronic lexicographical treatment of sign languages: The Danish Sign Language Dictionary. In: Granger, S. / Paquot, M. (eds.): *Electronic Lexicography*. Oxford University Press, 293–315.

Langer, Gabriele / Hanke, Thomas / Konrad, Reiner / König, Susanne (2016). "Non-tokens": When Tokens Should not Count as Evidence of Sign Use. In: Efthimiou, Eleni et al. (eds.): *Proceedings of the LREC2016 7th Workshop on the Representation and Processing of Sign Languages: Corpus Mining*. 10th International Conference on Language Resources and Evaluation (LREC 2016), Portorož, Slowenien, 28.5.2016. Paris, Frankreich: European Language Resources Association (ELRA), 137-142. URL: https://www.sign-lang.uni-hamburg.de/lrec/pub/16013.html (visited on 2022-05-20).

Langer, Gabriele / Müller, Anke / Wähl, Sabrina. (2018a). Queries and Views in iLex to Support Corpus-based Lexicographic Work on German Sign Language (DGS). In: Bono, Mayumi et al. (eds.): *Proceedings of the LREC2018 8th Workshop on the Representation and Processing of Sign Languages: Involving the Language Community. 11th International Conference on Language Resources and Evaluation (LREC 2018)*, Miyazaki, Japan, 12.5.2018. Paris, Frankreich: European Language Resources Association (ELRA), S. 107-114. URL: https://www.sign-lang.uni-hamburg.de/lrec/pub/18026.pdf (visited on 2022-05-20).

Langer, Gabriele / Müller, Anke / Wähl, Sabrina / Bleicken, Julian (2018b). Authentic Examples in a Corpus-Based Sign Language Dictionary – Why and How. In: Čibej, Jaka et al. (eds.): *Proceedings of the XVIII EURALEX International Congress: Lexicography in Global Contexts (EURALEX 2018*), Ljubljana, Slowenien, 17.-21.7.2018. Ljubljana, Slowenien: Ljubljana University Press, 483-497. URL: https://euralex.org/publications/authentic-examples-in-a-corpus-based-sign-language-dictionary/ (visited on 2022-05-20).

Müller, Anke / Hanke, Thomas / Konrad, Reiner / Langer, Gabriele / Wähl, Sabrina (2020). From Dictionary to Corpus and Back Again – Linking Heterogeneous Language

Resources for DGS. In: Efthimiou, Eleni et al. (eds.): *Sign Language Resources in the Service of the Language Community, Technological Challenges and Application Perspectives. Proceedings of the 9th Workshop on the Representation and Processing of Sign Languages*. 12th International Conference on Language Resources and Evaluation (LREC 2020), Marseille, France, 2020. Paris, France: European Language Resources Association (ELRA), 157-164. URL: https://www.sign-lang.uni-hamburg.de/lrec/pub/20025.html (visited on 2022-05-20).

Piotrowski, Tadeusz (1989). Monolingual and Bilingual Dictionaries: Fundamental Differences. In: Tickoo, Makhan L. (ed): *Learners' Dictionaries: State of the Art*. Singapore: SEAMEO Regional Language Centre. 72-83.

Popkema, Anne Tjerk (2010). State of the Art of the Lexicography of European Lesser Used or Non-State Languages. In: Dykstra, Anne / Schoonheim, Tanneke (eds.): *Proceedings of Euralex 2010*, Leeuwarden/Ljouwert, The Netherlands, 6.-10.07.2010. Leeuwarden/Ljouwert, The Netherlands: Fryske Akademy, 65-98. URL: https://www.euralex.org/elx_proceedings/Euralex2010/000_Euralex_2010_03_Plenary_POPKEMA_State%20of%20the%20Art%20of%20the%20Lexicography%20of%20European%20Lesser%20Used%20or%20Non-State%20Languages.pdf (visited on 2022-05-20).

Power, Mary R. / Power, Des (2004). Everyone Here Speaks TXT: Deaf People Using SMS in Australia and the Rest of the World. In: *Journal of Deaf Studies and Deaf Education*, 9(3), 333–343.

Svensén, Bo (2009). *A Handbook of Lexicography*. *The Theory and Practice of Dictionary-Making*. Cambridge: Cambridge University Press.

Tarp, Sven (2013). 19. Lexicographic functions. In: Gouws, Rufus Hjalmar et al (eds.): *Supplementary Volume Dictionaries. An International Encyclopedia of Lexicography: Supplementary Volume: Recent Developments with Focus on Electronic and Computational Lexicography*. Berlin, Boston: De Gruyter Mouton, 460-468. https://doi.org/10.1515/9783110238136.460 (visited on 2022-05-20).

11.2 Dictionaries and Language Ressources

Brien, David (ed.) (1992). *Dictionary of British Sign Language/English*. London: Faber and Faber.

Johnston, Trevor (1989). *AUSLAN Dictionary. A dictionary of the sign language of the Australian deaf community*. Victoria: Australian Print Group.

Centre for Sign Language, University College UCC (2008-2017). *Ordbog over Dansk Tegnsprog*. [The Danish Sign Language Dictionary]. URL: http://www.tegnsprog.dk (visited on 2022-05-20).

DWDS – Digitales Wörterbuch der deutschen Sprache. Das Wortauskunftssystem zur deutschen Sprache in Geschichte und Gegenwart. Berlin-Brandenburgischen Akademie der Wissenschaften. URL: https://www.dwds.de/ (visited on 2022-05-20).

Łacheta, Joanna / Czajkowska-Kisil, Małgorzata / Linde-Usiekniewicz, Jadwiga / Rutkowski Paweł (eds.) (2016). *Korpusowy słownik polskiego języka migowego / Corpus-based Dictionary of Polish Sign Language*. Warsaw: Faculty of Polish Studies, University of Warsaw. URL: https://www.slownikpjm.uw.edu.pl/en (visited on 2022-05-20).

11.3 Project Notes

Konrad, Reiner / Hanke, Thomas / Langer, Gabriele / König, Susanne / König, Lutz / Nishio, Rie / Regen, Anja (2020). *Public DGS Corpus: Annotation Conventions / Öffentliches DGS-Korpus: Annotationskonventionen*. Project Note AP03-2018-01. Version 3. Hamburg, Germany: DGS-Korpus project, IDGS, Hamburg University. https://doi.org/10.25592/uhhfdm.1860.

Langer, Gabriele (2020). *Vorgehen bei der Analyse für die Artikelschreibung (Wörterbuch).* Project Note AP10-2016-02. Version 3. Hamburg, Germany: DGS-Korpus project, IDGS, Hamburg University. https://doi.org/10.25592/uhhfdm.9577.

Langer, Gabriele / Müller, Anke / Wähl, Sabrina / König, Susanne / Hanke, Thomas / Konrad, Reiner (2020). *Lemmatisierungsregeln (Vorläufige Wörterbuch-Einträge).* Project Note AP10-2016-01. Version 4. Hamburg, Germany: DGS-Korpus project, IDGS, Hamburg University. https://doi.org/10.25592/uhhfdm.1899.

Schulder, Marc / Blanck, Dolly / Hanke, Thomas / Hofmann, Ilona / Hong, Sung-Eun / Jeziorski, Olga / König, Lutz / König, Susanne / Konrad, Reiner / Langer, Gabriele / Nishio, Rie / Rathmann, Christian (2020). *Data Statement for the Public DGS Corpus*. Project Note AP06-2020-01. Version 1. Hamburg, Germany: DGS-Korpus project, IDGS, Hamburg University. https://doi.org/10.25592/uhhfdm.1855.

Otte, Felicitas / Müller, Anke / Langer, Gabriele / Wähl, Sabrina / Hanke, Thomas (2022). *Sign Representation in the DW-DGS.* Project Note AP11-2021-01. Version 1. Hamburg, Germany: DGS-Korpus project, IDGS, Hamburg University. https://doi.org/10.25592/uhhfdm.10222.

Wähl, Sabrina / Langer, Gabriele / Müller, Anke / Otte, Felicitas / Hanke, Thomas / Konrad, Reiner (2021). *Lemma Selection and Entry Types.* Project Note AP10-2021-01. Version 1. Hamburg, Germany: DGS-Korpus project, IDGS, Hamburg University.

Appendix 1: Table of Information Types

Description of the different information types presented in the microstructure of the entries.

Information types addressed to lemma sign and its variants (at entry level)

	Information type	Label in Entry	Content	Form, medium and presentation, typography or symbols	Comment	Information content	Example
EN	Entry number	{no label}	Unique address of lemma sign, ID number	Integral number, heading each entry 193			all entries 361
V1	Form: main variant	FORM	variant form	studio reproduction of manual form without mouthing (movie) … .1 [object language DGS]	main or first variant chosen as citation form [obligatory information]	form	all entries 361.1
V2	Form: additional variants	FORM	variant form	studio reproduction of manual form without mouthing (movie) … .[number >=1] [object language DGS]	additional variants if there are any	form	361.2
CF	Form: comment on form	KOMMENTAR	description of range of minor form differences in actual use	note in written German [meta-language]		form	361

CU	Use: comment on use/usage	KOMMENTAR	comments concerning the use of a sign or one of its variants	note in written German [meta-language]	e.g. use by age group; restrictions in use of variants	use, usage	430, 1
CO	Other comments	KOMMENTAR	other comments	note in written German [meta-language]	any additional information that is occasionally given on sign e.g. assumed origin of sign	additional information	381
Fr	Frequency	BELEGELAGE	indication of frequency (number of tokens) in the corpus	three small squares: from empty to filled □□□ ■□□ ■■□ ■■■ [symbols]		frequency	all entries 193
GrE	Grammatical comment	GRAMMATIK	grammatical properties	written label indicating the grammatical property of the sign [meta-language]	e.g. directional sign or number incorporation	grammar	351, 361
REt	Regional distribution: text	REGIONAL	regional distribution of the sign or one of its variants (corpus & feedback data)	note in written German [meta-language]	description of main area of use	distribution	e.g. 269
REm	Regional distribution: map	REGIONAL	regional distribution of the sign (all	map via link (**Karte**) [link: meta-language]		distribution	e.g. 269 (map)

			variants) (corpus & feedback data)				
S	Sense	BEDEUTUNG #	dictionary sense	identified by sense number after hash # [structure indicator]	various information types addressed to sense (see table below)	sense	all entries 193
Ph	Phrase	PHRASE #	a multisign phrase with an idiomatic meaning and sometimes a specific mouthing (MWE)	identified by sense number after hash # [structure indicator]	sublemma- various information types addressed to phrase (see table below)	MWE phrase	440#10
MN	Multisign name	MEHRTEILIGER NAME #	multisign name (MWE)	identified by sense number after hash # [structure indicator]	sublemma - various information types addressed to multisign name (see table below)	MWE multisign name	17#2
ZEs	compound-like construction: signs	ZUSAMMENSETZUNGEN	multisign unit (MWE), mirroring the structure of the German compound that is matches the mouthing	series of Micons (with mouthings) alternative signs of one part are grouped by frame, consecutive parts are indicated by arrow [object language DGS]	sublemma; parts of the compound-like construction can consist of a set of alternative signs; run-on when not matched to any specific sense of the sign	MWE	354
ZEe	compound-like	ZUSAMMENSETZUNGEN	German compound the	German compound word in brackets with	German compound word = typical	**addressed to**	354

	construction: mouthing/ translation equivalent		multisign unit mirrors (equivalent)	indication of borders between parts [object language German and as mouthing part of the object language unit of DGS]	mouthing and base for multisign unit	**compound-like MWE;** equivalent	
CRV	Cross-reference: related sign	VERWANDT	cross-reference to an iconically related sign	Micon (no mouthing) + Tooltip with signposts [object language DGS]		cross-reference	82, 193
CRI	Cross-reference: identical sign form	FORMGLEICH	cross-reference to a different sign with the same citation form or an identical variant form (homonyms)	Micon (no mouthing) + Tooltip with signposts [object language DGS]	for more than one variant: variant of same form indicated see 485	cross-reference	193, 485
CRS	Cross-reference: similar sign form	FORMÄHNLICH	cross-reference to a different sign with a similar form or a similar variant form	Micon (no mouthing) + Tooltip with signposts [object language DGS]		cross-reference	193
CD	Concordance	{no label}	concordance of tokens in context in public corpus	button linking to concordance view: [German as meta-language] Konkordanz im Öffentlichen Korpus	function: shows further examples of use as included in the public corpus	link to other product	most entries 409

Information types addressed at sense/phrase/multisign name (at sense level):

	Information type	Label in Entry	Content	Form, medium and presentation, typography or symbols	Comment		Example
SP	Signpost	{no label}	written indication or summary of sense	written German [metalanguage]	guiding element with the function of quick disambiguation / orientation	used for: summary in collapsed entries; in subject index; as tooltip with micons	all entries 193
PhM	Form: movie	FORM	form of phrase	movie triggered by arrow icon ▶: studio recording of complete phrase with the usual mouthing/mouth gesture [object language DGS]	given for each **phrase**		phrase 420#6
PhMic	Form: set of micons	FORM	signs contributing to phrase (MWE)	series of Micons (with mouthings), alternative signs of one part are grouped by frame, consecutive parts are indicated by arrow [object language DGS]	given for each **phrase;** sublemma; parts of the phrase can consist of a set of alternative signs		phrase 420#6

MNM	Form: movie	FORM	form of multisign name	movie triggered by arrow icon ▶: studio recording of complete multisign name with the usual mouthing/mouth gesture [object language DGS]	given for each **multisign name**		multisign name 17#2
MNMic	Form: set of micons	FORM	form of multisign name	series of Micons (with mouthings) alternative signs of one part are grouped by frame, consecutive parts are indicated by arrow [object language DGS]	given for each **multisign name;** sublemma; parts of the phrase can consist of a set of alternative signs		multisign name 17#2
Mtht	Mouthing/ mouth gesture	MUNDBILD	typical mouthings and mouth gestures associated with this sense	written words (without capital letters) [mouthing as part of object language form DGS] and rough indication of mouth gestures: [MG] [German abbreviation as meta-language]	**addressed to sense:** non-exhaustive list based on frequency of corpus data		all senses 193#12
PhMth	Mouthing/ mouth gesture	MUNDBILD	typical mouthing(s) or mouth gesture for whole phrase	written German (without capital letters) [mouthing as part of object language form DGS]	**addressed to phrase:** approximation as exact form of mouthing is not		phrase 440#10

				and rough indication of mouth gestures: [MG] [German abbreviation as metalanguage]	analysed and given		
MNMth	Mouthing	MUNDBILD	typical mouthing(s) for whole multisign name	written German (without capital letters) [mouthing as part of object language form DGS]	**addressed to multisign name**	no extra recording for multisign name with mouthing	multisign name 17#2
MthM	Mouthing/ mouth gesture	MUNDBILD	one typical mouthing or and mouth gesture	movie triggered by arrow icon ▶: studio recording of main variant with one mouthing/mouth gesture typical for this sense	**addressed to sense**		all senses 193#12
Def	Explanation	ERKLÄRUNG	explanation of meaning/use covered by this sense	written German [metalanguage]	dictionary definition		all entries 193#12
EqW	Equivalent: word	DEUTSCH	German translation equivalents	German words and phrases separated by small bullets; links to German index for most prominent equivalents [object language German]		also used for generating the German index and linked to the German index	all entries 434#1
EqD	Equivalent: disambiguator	DEUTSCH	disambiguator to German polysemous equivalent	written German in round brackets [metalanguage]	**addressed to equivalent;** either disambiguator of German	disambiguators for equivalent differentiation	disambiguator in entry: 84#6, 485#2

					polysemous word or eqivalent differentiation		disambiguator only in index: 417#3
EqL	Equivalent: diasystematic label[30]	DEUTSCH	information that German equivalent word is marked – i.e. belongs to a certain register, style, subject field, or is a metaphorical meaning extension and the like	Abbreviation/written German in curved brackets [metalanguage]	**addressed to equivalent**; either information on use restrictions of German word and/or differentiation of its senses	also appears in German index e.g. {form.}, {ugs.}, {metaph.}, {idiom.}	443#7, 136#3, 409#1, 628#4
CS	Additional information	ANMERKUNG	additional comments to sign use, relevant to this sense only	written German [metalanguage]			361#1, 417#3
GrS	Grammatical information	GRAMMATIK	indication of argument roles for directional signs	written German (with arrow) [metalanguage]	addressed at sense		351 (all senses)
Bi	Example: structure indicator	BEISPIELE	example	identified by example number ❶❷... [structure indicator]	structure of whole example:		

[30] Comp. Svensén 2009, p. 315-332. *Marking, esp. 317/318.*

					❶ ▶ [context] translation		
BM	Example: authentic example	BEISPIELE	authentic example	movie triggered by arrow icon ▶: sequence taken directly from corpus data [object language DGS]			193#12
BC	Example: context	BEISPIELE	indication of topic or previous context of an example in the corpus data	written German in square brackets [metalanguage]	optional; addressed to BM		193#12
BT	Example: translation	BEISPIELE	translation of example	German translation of the DGS example with **equivalent** in boldface letters [object language German]	addressed to BM		193#12
BL	Example: links to *MEINE DGS* (Public Corpus)	BEISPIELE	larger signed context of example [object language DGS]	2 red buttons [meta-language German]: Ganzes Gespräch in MEINE DGS Transkript in MEINE DGS – annotiert	addressed to BM	link to other product	409#2 Example 1
LFMic	Long form: MWE	LANGFORM	multisign unit (MWE), often mirroring the structure of a German compound	series of Micons (with mouthings) alternative signs of one part are grouped by frame, consecutive parts are indicated by arrow (→)	**addressed to sense;** simple sign and multisign unit share sense and the mouthing:		

			used as mouthing	[object language DGS]	additional multisign form is given		
LP	Long form: parts of the long form	LANGFORM	German word with a straight slash between the parts of long form	German (compound) word in brackets with indication of borders between parts [object language German]	German compound word = base for multisign unit; listed in German index	**addressed to long form;** equivalent not listed twice in index	
SynMic	Synonyms: signs	BEDEUTUNGSGLEICH	synonyms, near synonyms and regional lexical variants	micon + movie (with mouthing) [object language DGS]			414#1
SynR	Synonyms: regional	BEDEUTUNGSGLEICH	indication of synonyms regionality	label *regional* [German as metalanguage]	addressed to synonym		414#1
SynCM	Synonyms: contrastive map	BEDEUTUNGSGLEICH	contrastive distribution map of set of lexical variants and their regional distribution	link to map is thumbnail map with label *Regionale Verteilung* [German as metalanguage]	regional variants that follow map are lexical variants contained in the map		14#1 (map) 127#2 (map)
AntMic	Antonyms: signs	ENTGEGENGESETZT	antonyms and complementary signs	micon + movie (with mouthing) [object language DGS]			193#4
AntR	Antonyms: regional	ENTGEGENGESETZT	indication of antonyms regionality	label *regional* [German as metalanguage]	addressed to antonym		193#5
AntCM	Antonyms: contrastive map	ENTGEGENGESETZT	contrastive distribution map of set of	link to map is thumbnail map with label *Regionale*	regional variants that follow map are		planned, not implemented yet

			lexical variants and their regional distribution	*Verteilung* [German as metalanguage]	lexical variants contained in the map		
ColMic	Collocations	HÄUFIGE KOMBINATIONEN	collocational patterns and semantic preferences	series of Micons (with mouthings) alternative signs of one part are grouped by frame, consecutive parts are indicated by arrow (→) / two-headed arrow (↔) indicates that neigbours can appear as left or right neighbour [object language DGS]	neighbours signs in collocational patterns or semantic preference can consist of a set of alternative signs; selectional preference may contain hint on category of group or category may replace micon of neighbour [category: German as meta-language]		354#1, 91#1, 327#3, 193#2 (↔), 193#1 (with headshake)
CPDMic	Compound-like constructions	ZUSAMMENSETZUNGEN	multisign unit (MWE), mirroring the structure of German compound used as mouthing	series of Micons (with mouthings) alternative signs of one part are grouped by frame, consecutive parts are indicated by arrow (→) [object language DGS]	sublemma / indication; parts of the compound-like construction can consist of a set of alternative signs	**addressed to sense;** part (target sign) is realization of given sense within MWE	354#1

CPDeq	compound-like construction: mouthing/ translation equivalent	ZUSAMMENSETZUNGEN	German compound the multisign unit mirrors (equivalent)	German compound word in brackets with indication of borders between parts [object language German and as mouthing part of the object language unit of DGS]	German compound word = typical mouthing and base for multisign unit; listed in German index	**addressed to compound-like MWE;** equivalent	354#1
RSt	Regional distribution: text	REGIONAL	regional distribution of one sense or a group of related senses (corpus & feedback data)	note in written German [metalanguage]	description of regionality	distribution	one sense: 1018#1 group of related senses: 329#1,2
RSm	Regional distribution: map	REGIONAL	regional distribution of one sense or a group of related senses (corpus & feedback data)	map via link (**Karte**) [link: metalanguage]		distribution	one sense: 1018#1 (map) group of related senses: 329#1,2 (map)
Sub	Subject area	SACHGRUPPEN	subject area	written German with link to subject area in subject index; separated by small bullets [metalanguage]	function: onomasiological access to vocabulary via subject index		485#2