Exploring Aspects of Spontaneous Signing in the DGS Corpus

Maria Kopf, Reiner Konrad, Gabriele Langer, Marc Schulder, Lutz König
Institute of German Sign Language and Communication of the Deaf
University of Hamburg, Germany
{maria.kopf, reiner.konrad, gabriele.langer, marc.schulder, lutz.koenig}@uni-hamburg.de

Abstract

Most use of sign language is spontaneous, unplanned, embedded in a one-to-one situation and transient. General sign language corpora aim at such naturalistic data. Thus it can be expected that they include phenomena of spontaneous language similar to the ones described for spontaneous speech in vocal languages: that is, (dis)fluencies such as pauses, hesitations, errors, false starts and repairs as well as discourse markers. In this paper we explore which of the known phenomena of spontaneous language from previous research on vocal and sign languages could be identified in the DGS Corpus using the annotations at hand. We describe our search strategies, consider additional annotation tiers for spontaneous language, and provide examples for the phenomena identified.

Keywords: spontaneous signing, disfluency, annotation, DGS Corpus, German Sign Language

1. Introduction

In our ongoing research on German Sign Language (*Deutsche Gebärdensprache*, DGS) with the *DGS Corpus* (Prillwitz et al., 2008), we come across passages that appear to be semantically incorrect. For example, a current price is given in Deutschmarks instead of euros[1]. Other passages at first glance look like annotation errors, for instance the manual sign EARRING1A^ (used for meanings like 'woman' or 'girl') being accompanied by the mouthing "junge" (boy).[2] Upon closer inspection, we can determine these to be either errors or slips of the hands, a typical phenomenon in spontaneous signing (Keller et al., 2003). This begs the question: what other phenomena of spontaneous signing can we find in our data, and how do we deal with them in our annotations? And conversely, how can we use our existing annotations to detect different phenomena of spontaneous signing?

Most sign language use is spontaneous face-to-face interaction embedded in the present communication situation, transient and largely unplanned. Other times, signing is informed by careful planning, review, recording, correcting and rerecording, produced for permanent use – a conceptually "written" form of signing (Koch and Oesterreicher, 1985).

Most linguistic sign language corpora, while recorded, aim at a more naturalistic form of signing for their data, closer to spontaneous signing than to edited signed text. Data collection formats are preferred in which sets of signers spontaneously sign with each other on topics of interest or in "staged communicative events" (Himmelmann, 1998, pp. 185–186). Thus, the data to be dealt with in these sign language corpora can be seen as language-in-interaction and be expected to contain similar characteristics and challenges as spontaneous speech events of vocal language[3], studied under the terms *disfluencies, turn taking* or *repairs*.

With this article we present a first attempt to identify phenomena of spontaneous signing that occur in the *DGS Corpus*. To do these phenomena justice, we propose an additional annotation scheme to supplement the basic annotation of the corpus. We do not offer a comprehensive typology in the way Shriberg (1994) does for English, as we only looked for phenomena in stretches of monologue-like signing where one signer holds the floor (turn).

After an overview of related work and introducing the *DGS Corpus* (section 2) we discuss how annotation schemes could be expanded or supplemented to permit more in-depth analyses of spontaneous signing (section 3). We then illustrate a number of phenomena with examples from the corpus and discuss how they (and the search for them) differs from their counterparts in vocal language (section 4). We conclude with a discussion of our findings and outlook to future steps (section 5).

2. Background

2.1. Spontaneous Language

Spontaneous language use has been studied from different fields with different frameworks and goals. The study of phenomena of spontaneous language

[1] https://ling.meinedgs.de/release4/html/2025500_en.html#t00154443

[2] https://ling.meinedgs.de/release4/html/1212402_en.html#t00033228

[3] We use the term vocal language for audio-phonetic languages (Sallandre and Garcia, 2013), commonly also referred to as *spoken languages*.

Proceedings of the 12th Workshop on the Representation and Processing of Sign Languages, pages 262–274
16 May 2026.

has a long history with psycholinguistics under the term *disfluency*, focussing on language production. Conversation analysis examines it from the point of discourse structure, calling it *talk-in-interaction*. Computational linguistics analyses it for the purpose of automatic speech processing and language acquisition studies it for the sake of improving fluency in L2 speakers. See Kosmala (2022) for an overview and intriguing combination of the different frameworks. One of the most fundamental works in the study of disfluency is Shriberg (1994), which presents a comprehensive classification of disfluency types for American English. In 1999 the *Disfluency in Spontaneous Speech (DiSS)* workshop series was established from which a broad canon of literature has grown. While multimodal accounts are quite established, studies on sign languages have not yet been part of the workshop series.

In studies approaching spontaneous language from the discourse analysis framework, sign languages are used to compare phenomena cross-linguistically and cross-modally (cf. Floyd et al., 2016). In the field of sign language linguistics the interest on discourse management and interactional phenomena is growing as well. Corpus-based studies comparing vocal and sign languages (here French Belgian Sign Language (*Langue des Signes de Belgique Francophone*, LSFB) are compiled in Degand et al. (2019). For a brief overview of additional studies see Lutzenberger et al. (2024). Most literature, though, focusses on individual phenomena. The palm-up gesture as discourse marker has received significant attention (Lepeut and Shaw, 2025; McKee and Wallingford, 2011; Notarrigo and Meurant, 2014). Gabarró-López (2020b) gives an overview of studies on discourse markers in sign languages, covering American Sign Language (ASL), Venezuelan Sign Language (*Lengua de Señas Venezolana*, LSV), Spanish Sign Language (*Lengua de Señas Española*, LSE) and LSFB. Gabarró-López (2020a) adapted a vocal languages annotation protocol for discourse markers for use with LSFB. Notarrigo and Meurant (2019) study repetitions in spontaneous signing, also for LSFB.

For DGS we found a small number of studies dedicated to spontaneous signing. Hansen (2012) discusses several aspects of discourse studies such as coherence, cohesion and context. Volk (2016) reports on discourse functions of the palm-up gesture in DGS, as does Kaul (2024) who analyses data from the *Public DGS Corpus*. Mpadanes (2020) studies repairs in DGS, also based on *DGS Corpus* data. Based on annotations of filmed monologues, Oviedo et al. (2018) studies filled and unfilled pauses in the context of fluency and perceived disfluency. Keller et al. (2003) studies articulation errors and their repairs, based on a collection of signed stories chosen to evoke slips of the hands.

2.2. Sign Language Annotation

Besides documenting and archiving, one of the first steps in building sign language corpora is basic annotation, in which each manual sign is identified by labelling it with an ID-gloss (Johnston, 2010) or directly linking token to type entries in a lexical database (Hanke and Storz, 2008). Further annotations depend on the goals and the resources of each project and are documented in annotation guidelines. See Hodge and Crasborn (2022) regarding good practices in annotation.

Phenomena of spontaneous language which are annotated in several corpora are repetitions, false starts and gestures (Kopf et al., 2022). For repetitions the distinction between grammatical and non-grammatical forms is decisive for annotation. In grammatical forms the repetition is used to mark aspect or plural; in non-grammatical forms the sign is repeated for other reasons, e. g. to intensify or gain time for language planning. For the focus of this article only non-grammatical forms are of interest. False starts are marked with a suffix in the Auslan Corpus and STS Korpus (Johnston, 2024; Mesch and Wallin, 2021). In the Auslan Corpus a false start is identified by a repair sequence in the direct or near context. The STS Korpus additionally annotates pauses due to hesitations and lengthenings in the form of a hold. The way in which gestures are annotated differs in how exactly gestures are defined. One gesture that nearly all corpora annotate with a special ID-gloss is the palm-up gesture (Kopf et al., 2022).

An aspect that occurs when annotating spontaneous sign language is the use of the articulators, i. e. the hands, for non-linguistic purposes such as scratching one's nose (Crasborn et al., 2020; Konrad et al., 2022). For the purpose of this article we exclude such uses.

Another challenge when annotating spontaneous signing lies in the multimodality of language. As Braem and Volterra (2023, p. 32) state: "the facial, gestural, and other prosodic behaviors produced simultaneously with either speech or signing had traditionally usually been relegated to a kind of limbo with other phenomena that were considered 'paralinguistic' or 'nonlinguistic'". But in the study of spontaneous language these "nonlinguistic" behaviours are crucial. The difference between repair sequences and rhetorical stylistic devices in signed statements of the form "*A, no, B*" can often only be recognised by subtle differences in facial expressions and head movement. During pauses in the manual production, whole sections of the uttered statement can be produced on only the face. Research on disfluency in speaking and signing does recognize this and adjusts methods and theories (again, see Braem and Volterra, 2023). Still, the disregard of non-manual activity is evident today in

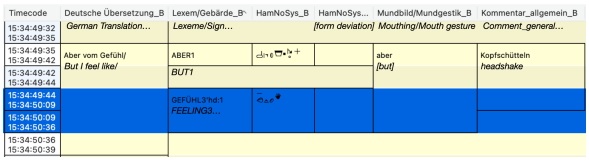

Figure 1: iLex transcript with tiers for translation, token, mouthing/mouth gesture, and comment.

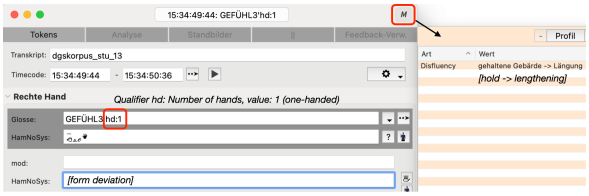

Figure 2: iLex token-tag window with metadata.

sign language annotation as the main focus in corpus annotation continues to be on manual signs, in part due to resource restrictions. While mouthings and mouth gestures are often annotated, other non-manuals are not part of a basic annotation.

2.3. DGS Corpus

The *DGS Corpus* (Prillwitz et al., 2008) is an annotated reference corpus of DGS. It contains monologues and dyadic discourses that were captured in the context of various collection tasks (Nishio et al., 2010). The recordings from its first data collection campaign (2010–2012), which the analyses in this article are based on, cover 560 hours of language production by 330 participants from all parts of Germany. Over 50 hours of the *DGS Corpus* have been released publicly as the *Public DGS Corpus* in a number of different formats (Konrad et al., 2024a), which includes the fully annotated research dataset *MY DGS – annotated* (Konrad et al., 2024b).

The corpus contains annotations and translations. Translations into German are done by sign language interpreters; broken-off utterances are marked by a slash. Using the iLex annotation tool (Hanke and Storz, 2008), annotators time-align the translations sentence by sentence, then segment and lemmatise tokens (token-type matching). Mouthing and mouth gesture are also annotated in this first annotation pass, including a note on the deviation from the type's citation form. Gestures and non-manual signals like head nods and shakes are annotated with special gesture types when occurring on their own or less formally in a comment tier when they co-occur with a manual sign.

In the lemma revision process, tokens with form deviation are reviewed and either the deviation is specified using the Hamburg Notation System (HamNoSys) (Hanke, 2004) or tokens are grouped into form features called 'qualifiers' (see Loos and Konrad, 2022). Additional information can be specified via metadata entries.

In addition to its basic annotation, the *DGS Corpus* labels so-called non-tokens for lexicographic work (Langer et al., 2016). Non-tokens are sign uses that should not count as evidence for a sign type or its use by a particular signer. In addition to metalinguistic references and contrastive untokens, non-tokens include various cases of slips of the hands, false starts, aborted signs, malformed or deviated sign forms, gestures, and trying out or searching for signs. Therefore, offering an overlap between non-tokens and phenomena of spontaneous signing.

Figure 1 shows a snapshot of an iLex transcript[4] with timestamps and the tiers "German Translation", "Token", "Type citation form" (HamNoSys), "Token deviation" (HamNoSys), "Mouthing/Mouth gesture", and "Comment". The signer interrupts their signing (the translation ends with a slash) and the second sign is held with a duration of nearly one second (42 frames). Figure 2 shows the token-tag window of the second sign with the qualified type gloss and the metadatum *Disfluency*.

3. Additional Annotations

In order to spot tokens or translation tags of interest for this study we implemented the metadata category *Disfluency* (cf. section 2.3). For annotating the corresponding sequences we introduced two tiers in the transcripts: *Disfluency_types* and *Disfluency_parts*. Tags in *Disfluency_types* can span over more than one token, tags in *Disfluency_parts* are used to label tokens within repetitions and repairs. Parts of a repetition are the sign's first instantiation and its repetition. Usually one sign is repeated, but in example 1[5] a combination of two signs meaning "I participated in all the cities" is part of a non-adjacent repetition (see link in footnote to view the example). Figure 3 shows the annotation of this in iLex.

Parts of a repair are *reparandum*, *interregnum*, and *reparans* as introduced by Shriberg (1994) (cf. Kosmala, 2022, p. 51; Piu and Bove, 2007, p. 399), *interregnum* being optional. In example 2[6] the signer recounts Diana, Princess of Wales, caring greatly about children (phrased in German and DGS as *"having a heart for children"*). He first signs HEART2B, PERSON1 and FOR1 accompanied

[4] Public transcript matching fig. 1: https://ling.meinedgs.de/release4/html/1182343_en.html#t00050035

[5] https://ling.meinedgs.de/release4/html/1414312_en.html#t00073006

[6] https://ling.meinedgs.de/release4/html/1413485_en.html#t00035311

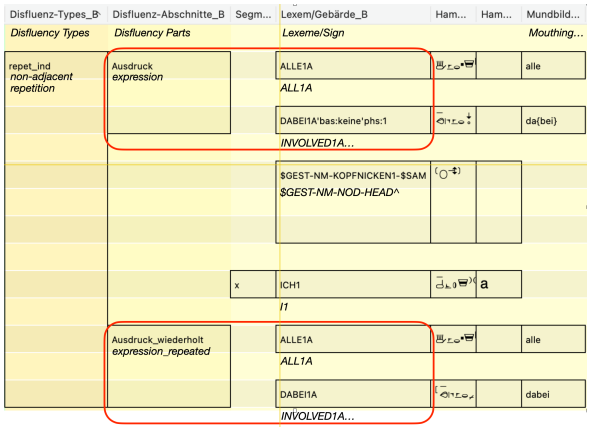

Figure 3: Disfluency annotation for a repetition.

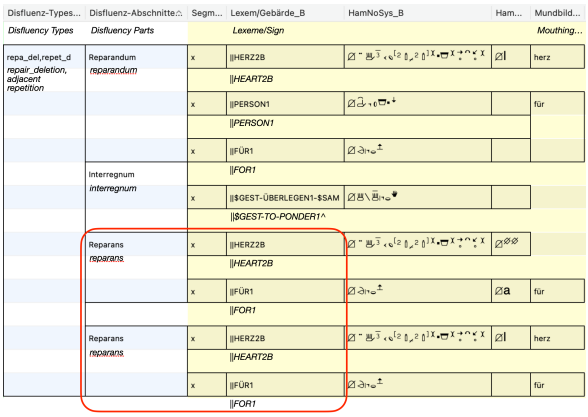

Figure 4: Disfluency tiers for repair and repetition.

by the mouthing “herz für” (heart for), then hesitates while gesturing, and repeats the sequence HEART2B FOR1 without PERSON1 two times. Its disfluency annotation is shown in fig. 4.

As Crible (2018, p. 78) shows for vocal languages, one sequence can display several disfluency types. In our current work we did not opt for macro-labels of sequences but allowed the selection of more than one type in the *Disfluency_types* tag, as for example 2, where we tagged a repair with deletion (see section 4.1) and an adjacent repetition (see section 4.2).

We agree with Crible (2017) that “[i]t is not [certain] whether these annotation systems could be applied to languages other than their original”. Nevertheless, we hope to offer ideas and perspectives which can be tested for other sign languages.

An added benefit of this work is the ongoing quality control it provides for corpora, as annotation artefacts or missing annotations are detected while looking for disfluencies.

4. Selected Phenomena

We follow Shriberg in her definition of (dis)fluencies as any “stretch of linguistic material [which] must be deleted to arrive at the sequence the speaker ‘intended’” (Shriberg, 1994, p. 1, as quoted in Crible, 2017). We do not only concentrate on material that interrupts the fluency of the sign stream but also include material that helps retaining it (“fluencemes” in the terminology of Götz (2013)). Comparing (dis)fluency typologies from Shriberg (2001), Kosmala (2022), Notarrigo and Meurant (2019), and Blache et al. (2017), we decided to concentrate on the following phenomena: *Pauses, Repetitions, Elisions, Lengthening, Repairs,* and *Articulation Errors* (sections 4.1 to 4.6). Based on Crible (2018) we added *Discourse Markers* as a further phenomenon (section 4.7).

We now describe how we searched for these phenomena of spontaneous DGS using several strategies: the database query capabilities of iLex, existing annotation, previous work with the corpus data and frequency counts. Queries resulted in lists of candidates that were manually searched for cases of the phenomena. Each phenomenon is introduced and we describe the subphenomena we encountered, illustrated by examples.

4.1. Repairs

A repair is a sequence where the signer detects a problem, halts the production and then continues. The halting can occur after the sign is completed, creating a pause, or during sign execution, resulting in a broken off sign or a sign with a reduced movement or missing repetition. The continuing part can be a correction of the prior utterance (in the form of a deletion, insertion or substitution), but it also can be a complete new phrase, as is the case for false starts. Following Shriberg (2001) we focus on five kinds of repairs: repairs of articulation errors, repairs including a deletion, insertions, substitutions, and false starts.

To detect repair sequences we first used several queries detecting break-offs in the translation tier (indicated by a slash) at the beginning, middle or end of a translation tag, also in combination with tokens marked as aborted in the gloss tier with a HamNoSys-symbol.

Not all shortened sign forms are aborted signs and marked as such with the symbol for break-offs. Shortened sign executions with missing repetitions are marked through a qualifier indicating the number of repetitions or absence of repeated motion (phases). We used queries to find all cases where a sign with a repeated motion in its citation form is signed without repetition (*phase:1*) or without movement at all (*phase:0*) and signed again with repetition within the next five signs.

Another query to find self-repair candidates identified repetition of a sign after a break-off by looking

at translation tags with a break-off slash and checking whether the last sign of that tag re-occurred within the first four signs of the next translation tag. Cases with significant pauses between translation tags were filtered out, as this pause implies a pause in the production, followed by a resumption or change of topic.

A deletion was illustrated in example 2 (section 3), where the signer deletes PERSON1 in the repair sequence. A false start can be seen in example 3[7] where the signer interrupts his signing by signing NO1A and then restarts with a different phrasing.

In example 4[8] an incorrect sign is substituted, the signer having made a mistake in naming a district of Hamburg—not Blankenese, but Sülldorf—and corrects this with a shake of the head while signing OR4B before showing the manual alphabet sign for S.

In addition to substitutions of incorrect signs we found repairs with substitutions of lexical variants. In example 5[9] the signer talks about a person interviewing a woman. She starts with the onset of the sign WOMAN1A, which involves tracing the shape of the breast. Then she aborts the execution of this sign by blending smoothly to a full execution of the lexical variant WOMAN4A. The signer uses WOMAN1A fourteen times in the corpus, and WOMAN4A only three times. It is not clear why in example 5 she changes to the less frequently used variant. Possible explanations are the attempt to replace WOMAN1A by another lexical variant because of an ongoing discussion about the political correctness of the sign WOMAN1A or a commonly observed pattern of signing multiple lexical variants in a row for concepts with rich regional variation, possibly as a strategy to ensure understanding across dialectal boundaries.

Example 6[10] (also seen in fig. 5) shows the signer starting to sign ENGLAND4, stop and change to ENGLAND3, both times breaking off the execution, then fully produce ENGLAND4. In the entire conversation segment of about 16 minutes, the signer uses ENGLAND3 six times and ENGLAND4 seventeen times. Shortly before, his conversation partner used ENGLAND3, so this is most likely an attempt by the signer to adapt his signing to his interlocutor. (Later on in the conversation the signer uses ENGLAND3 once, then switches back to ENGLAND4 and even later he fully signs both signs in a row.)

...	**ENGLAND4***	**ENGLAND3*** **ENGLAND4**	
	$INDEX1	$INDEX1* $INDEX1*	...
...	**e{ngland}**	**england**	
	It's from **England**.		

Figure 5: Repair with repetitions and break-offs in example 6[10]. Shown are the three signs meaning 'England', the first two broken off. Below are glosses for dominant and non-dominant hand, mouthing, and translation. For "e{ngland}", only "e" was mouthed.

4.2. Repetitions

Repetitions are cases where basic meaningful elements appear more than once. Notarrigo and Meurant (2019) give a comprehensive analysis and typological description of repetitions based on corpus data of spontaneous signing in LSFB. They distinguish three form groups of repetition: adjacent, non-adjacent (with one or more elements in between), and framing (non-adjacent with the function to highlight a central element). They report that in sign language linguistics literature adjacent repetition is identified as a disfluency marker indicating hesitation or repair. The authors looked for repetitions of manual signs in the stretch of one signer's turn, excluding interactional factors. Although nearly half of the repetitions had a pragmatic function, there were only few cases where repetition indicates hesitation or repair.

In this work, we limit our analysis to utterances of single signers, excluding aspects of immediate interaction with the addressee. However, as is discussed for example 6, wider interactional contexts could still play a role. To find adjacent repetitions we used a query looking for two neighbouring tokens with the same gloss name irrespective of numbers and letters for variants or extensions for qualifiers.

A case of adjacent repetition can be seen in example 7[11], where the signer recounts her school days. TEACHER2 is signed twice. There is no pragmatic or grammatical reason for this repetition.

A non-adjacent repetition can be seen in example 8[12]. The signer talks about the British royal family moving into a castle. The sign TO-MOVE1

[7] https://ling.meinedgs.de/release4/html/1210997_en.html#t00065404

[8] https://ling.meinedgs.de/release4/html/1179224_en.html#t00164519

[9] https://ling.meinedgs.de/release4/html/1248862_en.html#t00043921

[10] https://ling.meinedgs.de/release4/html/1248090_en.html#t00050405

[11] https://ling.meinedgs.de/release4/html/1418889_en.html#t00021205

[12] https://ling.meinedgs.de/release4/html/1248505_en.html#t00032409

is signed two times with the sign TO-LIVE4 in between. There is no manual or non-manual marking which would indicate a repair or emphasis.

Another pattern, which may be called lexical doubling, is seen quite often in the corpus: Two consecutive lexical variants for one concept, often sharing one mouthing, where each part is also a stand alone sign for the concept. Examples of this pattern are FAMILY1 FAMILY3 as in example 9[13] and BEAUTIFUL1A BEAUTIFUL3 as in example 10[14]. This kind of redundancy might simply be a strategy to ensure understanding when several lexical or even regional variants exist for the same concept, as already discussed in section 4.1.

In other cases, the second sign might have a visual or illustrating function due to its iconic motivation, e. g. providing information on the visual or spatial properties of the object, setting it up in space or providing a hold (ground) to be referred to subsequently. In example 11[15] the signer describes the equipment of a bus. First signing BUS2 (iconic motivation: using a steering wheel), followed by BUS1A (tracing the shape of the vehicle with two C-handshapes). The latter variant can be located anywhere in the signing space, in contrast to the fixed place of articulation (PoA) of the first. The two variants are followed by an index and the signs CURTAINS2 and CURTAINS3, which are executed at the same PoA as BUS1A, indicating the curtains were in the bus. Here the visual space is used to draw the whole picture of the bus with its curtains.

A case with lexical variants and a modified form can be seen in example 12[16]. The signer is retelling a picture story where a cage is picked up. He uses three strategies to sign the concept 'cage'. A one-handed modification of the sign CAGE4 in front of the face, as if he were inside the cage looking through the bars, followed by an unmodified CAGE4, and then the lexical variant CAGE2.

4.3. Elisions

Elisions are cases where an element of a word or the whole word is not realised (Blache et al., 2017; Crible, 2018). As morphemes are often executed simultaneously in sign languages, elisions are quite hard to detect. Looking at the phonological parameters of a sign, a sign executed in front of the body instead of the head could be seen as an elision of the PoA, but there still is a PoA, it just changed to a less marked PoA (discussed as "lowering" in the literature). The same is true for handshapes and orientation, one always has a handshape and orientation when signing, but it might deviate from the citation form. However, a sign with a path movement in its canonical form can be executed without movement. So to detect elisions we concentrated on movement using the already existing annotation with the movement qualifier *phase*, described in section 4.1, e. g. LOUD1C in example 13[17] in which the signer describes the learning environment at her boarding school.

The deletion of a whole word is easier to detect, as it is part of a repair sequence (cf. section 4.1). To find cases where the movement is not completely deleted but only reduced we searched for tokens being considerably shorter than the average duration of their type. To detect outliers, we calculated the z-score[18] of each sign's duration and inspected tokens with a score of $z < -1$.

With this detection strategy we found many cases with a shortened movement. This might in part be an artefact of annotation, which allows changing the playback speed or even looking at individual frames. That way, annotators may detect handshapes, often between two clearly produced signs, and annotate them as tokens with deviations, which are unlikely to be visible to the naked eye. Still, they might be seen as a reduced sign originating in a semi-realised planning event, as in the case of example 14[19] where the signer talks about playing squash. She signs WORKOUT1 with a shortened movement, followed by TO-PLAY2 without any markers for a repair.

Both strategies will detect cases that should not be considered aspects of spontaneous signing, but are deliberate markings to indicate an interruption in a process, what Liddell and Johnson (1989) call "unrealized-inceptive aspect". One such case is example 15[20], where the signer describes the passage of time with the sign PROCEEDING1^, which is aborted to show the interruption when she got pregnant, and resumed afterwards. Therefore, candidates for elision have to be double checked for their status as true elisions.

[13] https://ling.meinedgs.de/release4/html/1431428_en.html#t00011118

[14] https://ling.meinedgs.de/release4/html/1210997_en.html#t00061437

[15] https://ling.meinedgs.de/release4/html/1429709-16344227-16451821_en.html#t00015207

[16] https://ling.meine-dgs.de/release4/html/1413240_en.html#t00000402

[17] https://ling.meine-dgs.de/release4/html/1212402_en.html#t00064531; note that in public data, qualified types are indicated by an asterisk, while the gloss is LAUT1C'phs:0 in the internal corpus.

[18] Z-score, or standard score, indicates the number of standard deviations that a value is above or beyond the population's mean value. It is defined as $z = \frac{x-\mu}{\sigma}$, x being the value, μ the mean, and σ the standard deviation.

[19] https://ling.meinedgs.de/release4/html/1212402_en.html#t00085710

[20] https://ling.meinedgs.de/release4/html/1418889_en.html#t00055607

4.4. Lengthening

Lengthenings or prolongations are executions of words with longer durations of phonemes (or syllables) than usual. Other phenomena can also be prolonged, e. g. filled pauses or discourse markers (Crible et al., 2024). Oviedo et al. (2018) describes lengthening as one form of filled pauses.

To search for lengthened signs, we again used the z-score of sign durations, as we did for shortenings (section 4.3), this time searching for longer than usual instances. Long outliers can be more extreme than short ones, so we searched for signs with $z > 5$ for our initial investigation. We found lengthenings of two forms: either the sign is repeated multiple times[21] or the hand is held in the first or last PoA of the sign.

In example 16[22] the signer describes different vocations that are part of the education in a school in Leipzig. He repeats $LIST1 several times while trying to remember the vocations. He looks at his interlocutor to start with, but after the sign LEIPZIG1A his gaze wanders to the top left corner of the room while he is thinking. He continues signing different vocations, his gaze still to the top left corner. Arriving at number four of his list he again lengthens the sign with a reduplication followed by a filled pause ($GEST^ in the current public release, internally already revised to $GEST-TO-PONDER1^). Finally he remembers the fourth vocation and his gaze is back on the interlocutor again.

In example 17[23] the signer seems to have difficulties recalling where the last sign language festival he attended took place. After repeating the signs for 'sign language festival', he holds the sign A-MOMENT-AGO1A. During the whole sequence of repeating and holding, his head is bowed, he looks down, and the tip of his tongue is visible in the corner of his mouth. Once he recalls the location he nods and continues to sign IN1 BERLIN1A.

In contrast to the described unintended forms of prolongations, lengthening can also be used intentionally to indicate an intensifier. Example 18[24] shows the signer prolonging LONG-TEMPORAL4A as well as their mouthing of "lang" (long) to indicate that the referenced time span was particularly long.

4.5. Pauses

Pauses can be described as interruptions of information production. Gabarró-López and Meurant (2016, p. 85) define them as "periods of no signing at all" and include stops with crossed hands and relaxed hands in the neutral signing space or along the body.

Most authors differentiate between unfilled or silent pauses and filled pauses. Unfilled pauses are often defined by a timespan of silence, i. e. the interruption of the sound signal in vocal languages or missing or "neutral" (Oviedo et al., 2018) activity of the hands and face in sign languages, e. g. Oviedo et al. (2018) specify a minimum of 160 milliseconds for DGS. Filled pauses are stretches filled with "conventional and neutral phonetic form[s], (e. g. 'euh' in French)" (Crible, 2018, p. 72). For DGS, Oviedo et al. (2018) describes four kinds or filled pauses: a) crossed hands accompanied with non-manual activity, b) a longer transitioning between two signs (due to a break in or slow execution of the movement), c) non-morphological prolonging of sign segments (e. g. a prolonged hold at the beginning of a sign), and d) gestures without meaning.

In the *DGS Corpus* silent pauses are not annotated and the time length of tokens are determined by the rules of narrow transcription, i. e. transition movements are not part of the token. This is important to keep in mind when calculating time between tokens. For detecting stretches without signing, other than between utterances, we used a query that checks whether there are long periods within a translation tag without any sign token and without inputs from an interlocutor. To find the most clear cut cases we started out with pauses of a full second. In future passes, this period can be reduced to match the approach of Oviedo et al. (2018).

A part of filled pauses is covered by existing annotations and therefore already annotated. Relevant types include $GEST-OFF1^[25] (the palm-up gesture), $GEST-TO-FOLD-ONES-HANDS[26] (used when the signer folds their hands within a turn), $GEST-NO-IDEA1^[27] (a gesture meaning 'no idea'), and five variants with the gloss name $GEST-TO-PONDER (gestures used while or for thinking, three of which occur in the public corpus and can thus be found in its type index.[28])

We found both kinds of pauses, filled and unfilled. Within filled pauses we detected the four kinds described by Oviedo et al. (2018): crossed hands with non-manual elements, long transitions, lengthening (see also section 4.4), and gestures.

[21] Multiple repetitions without breaks are annotated as one token in the *DGS Corpus*.

[22] https://ling.meinedgs.de/release4/html/1414563_en.html#t00105700

[23] https://ling.meinedgs.de/release4/html/1414312_en.html#t00090244

[24] https://ling.meinedgs.de/release4/html/1178133_en.html#t00000845

[25] https://ling.meinedgs.de/release4/types/type16636_en.html

[26] https://ling.meinedgs.de/release4/types/type72381_en.html

[27] https://ling.meinedgs.de/release4/types/type60946_en.html

[28] https://ling.meinedgs.de/release4/ling/types_en.html

In example 19[29] the signer recounts a football tournament. As he tries to remember which team he played against, he folds his hands, looks up and furrows his brow.

When working with corpus data it is apparent that not all non-activity of the hands are pauses in a strict sense. Sometimes manual signing is halted while the face, head or torso is used in a CA to add information to the utterance or the head and facial expression are used to comment on the signing, as in example 20[30], and thus should not be interpreted as a pause at all, so just checking for activity of the hands is not enough. Even cases with hardly any manual and non-manual activity must be viewed in context. While they can be indicators of active thinking as in example 21[31] (pause between ATHLETICS1 and TO-SWAP3A) they might also be re-enactments of thinking or hesitating in a constructed action (CA), although our observations indicate that in most such cases they are explicitly marked.

4.6. Articulation Errors

Articulation errors, or slips of the hands, are cases of incorrect signing in an otherwise correct production. For vocal languages it is estimated that on average within 1,000 words or 10 minutes there is one slip of the tongue (Fromkin, 1980). For sign languages, Keller et al. (2003) show that slips of the hands follow the same patterns found in vocal languages regarding the types of errors (e. g. anticipations, perseverations, substitutions, blends, fusion, etc.), monitoring behaviour (repairs) and affected units (from phonological parameters up to full signs and even phrases). The only difference between them and slips of the tongue lies in their distribution: complete signs are affected more than complete spoken words, whereas morphemes and phrases are less affected in sign languages than in vocal languages.

Articulation errors are already a subcategory of repairs (see section 4.1), but as only half of all slips of the hands are repaired by the signers themself (Keller et al., 2003) there is a need for a distinct category, annotating articulation errors that occur outside of repair sequences. To find articulation errors, we relied on pre-existing work that marked non-tokens during lexicographic analysis (see section 2.3). Developing detection strategies like those described in previous sections proved challenging, as there are many valid reasons for deviations from the citation form that would have to be filtered out. And clearly distinguishing articulation errors from phenomena such as assimilations, signing style or lax signing also proves to be extremely difficult. Newkirk et al. (1980) suggest that slips of the hand change a parameter of the sign completely while juncture assimilations substitute only part of the parameter, e. g. the handshape is changed only towards the end of the sign. The distinction between different linguistic phenomena should be refined for future work on articulation errors.

A phonological articulation error can be seen in example 22[32] where the signer talks about her final exam. While signing the two signs GRADUATION-OR-COMPLETION2 and TO-EXAMINE1 meaning 'final exam', a metathesis of the handshapes happens. The wrong execution of TO-EXAMINE1 is aborted and the sign is repeated in it's correct form. The two examples from section 1 represent a lexical substitution (Deutschmark instead of euros) and a lexical anticipation (girls anticipated instead of boys). Both examples also show a mismatch between the manual sign and the mouthing.

4.7. Discourse Markers

There exist numerous approaches for defining the category of discourse markers (DM), however, they still are, in the words of Crible (2018, p. 34) a "slippery linguistic category". For our purposes we define them as signs that are used on a metadiscursive level to structure discourse.

In previous research, some signs and gestures, especially palm-up, have been discussed with regard to discourse organisation. Palm-up is attested for many sign languages with a variety of functions. Volk (2016), Debreslioska et al. (2025), and Kaul (2024) show that palm-up gesture has, among others, discourse regulating functions in DGS, which is also attested for several other sign languages (e.g. McKee and Wallingford, 2011; Lepeut and Shaw, 2025). Some authors also discuss individual signs that function as DM, such as HEY and WELL in ASL by Hoza (2011), FINE in ASL by Metzger and Bahan (2001), different kinds of buoys in LSFB by Gabarró-López (2019), as well as SAME and palm-up in LSFB by Gabarró-López (2020b), or ENCAMBIO (INSTEAD-OF) and ES-DECIR (THAT-IS-TO-SAY) in LSE by Villameriel (2010). Other studies try to detect all discourse markers in a given piece of signing, such as Pérez (2006) for LSV, Gabarró-López (2025) for LSFB and Villameriel (2008) for LSE.

Furthermore, as Crible (2018) showed in her bottom-up research on French and English, there are language-specific as well as universal patterns

[29] https://ling.meinedgs.de/release4/html/1209006_en.html#t00024826

[30] https://ling.meine-dgs.de/release4/html/1419797_de.html#t00091547

[31] https://ling.meinedgs.de/release4/html/1414123_en.html#t00015311

[32] https://ling.meinedgs.de/release4/html/1212402_en.html#t00123445

of DMs, so not all described DMs might apply to DGS. To our knowledge there is no comprehensive study of DMs in DGS, however, some DGS signs have been mentioned serving the function of DMs, e. g. palm-up (Debreslioska et al., 2025), BEISPIEL (EXAMPLE) and BEDEUTUNG (MEANING) (Hansen, 2012), and DONE (Jahn et al., 2021).

Crible (2018, p. 82) notes "a highly pervasive and prominent use of DMs in spoken language" and reports a relative frequency of DMs around 55 occurrences per thousand words. Thus, in a corpus of spontaneous signing, signs that function as DMs can be expected to be among the most frequent elements. Consequently, we scanned the 100 most frequent signs of the *DGS Corpus* for possible candidates. Of those 100, we identified over twenty that we suspect of having a discourse structuring function. Among them are signs that previously had been identified as DM (e. g. palm-up, DONE), signs that have a comparable semantic origin as the DM signs mentioned for other sign languages (THEN1A[33], EQUAL1A[34], various kinds of buoys) and several signs that suggested themselves during lexicographic analysis or annotation as having DM functions, e. g. MEANING1, illustrated by example 23[35].

This can only be a first glimpse into an area not thoroughly investigated so far and further analysis on discourse markers in DGS is highly desired. The possible candidate signs identified could be a good starting point for such an analysis to detect, verify, and describe their functions as DMs.

5. Discussion and Future Work

Errors and slips of the hands in signing are hard to identify. In order to spot a mistake one has to rule out that the signing is intentional and correct. DGS includes a large amount of lexical variation, a great amount of liberty of how to execute a specific sign, and flexibility in syntactic structures. Therefore, unless a mistake is clearly marked by a self-correction as in example 4, one always has to consider the possibility that occurrences such as repetitions have an intentional linguistic or communicative purpose. In contrast to vocal languages repetitions of lexical variants (lexical doubling) as well as pairing an iconically motivated sign with a conventional sign seem to be sign language specific communicative strategies that might not only be found in spontaneous but also in planned signing.

As the examples given in section 4 show, phenomena of spontaneous signing go far beyond repairs of errors. Basic annotation that focusses on manuals can only be used to a limited extent to detect them. For more comprehensive searches, detailed annotation of non-manuals (body posture, head movement, facial expression, eye gaze, etc.) is also required. Building on this, a dedicated annotation of phenomena as we discuss in section 3 may be undertaken. Scaling this sufficiently for frequency-based analyses as, for example, Crible (2018) did for two vocal languages, is, unfortunately, outside the scope of the DGS-Korpus project. However, future projects will hopefully use the *DGS Corpus* to explore such matters.

Aside from manual annotations, another potential source of information to guide researchers may be kinematic data gained from automatic pose recognition models (Schulder and Hanke, 2020). While the semantic interpretation of signing with computer vision methods, especially at the level of nuance relevant to linguistic analysis, remains an open challenge, pose information may already be helpful, e. g. for spotting unfilled pauses.

Spontaneous language also offers the potential to revisit the assumptions and paradigms of current annotation schemata. One such undertaking in the related field of gesture studies is the *MultiModal MultiDimensional* (M3D) labelling system (Rohrer et al., 2025), which exhibits a very broad view on annotating linguistic phenomena.

Finally, while the separate presentation of phenomena gives the impression that they can be clearly distinguished and delineated, this is not necessarily the case. We found many instances in which several phenomena occurred in parallel or were intertwined. Crible (2018) and Shriberg (2001) use the term *complex* for these cases. Accounting for such phenomenon complexes in annotation schemes will require further deliberation.

6. Acknowledgements

This document has been produced in the context of the joint research funding of the German Federal Government and Federal States in the Academies' Programme, with funding from the Federal Ministry of Education and Research and the Free and Hanseatic City of Hamburg. The Academies' Programme is coordinated by the Union of the Academies of Sciences and Humanities.

[33] https://ling.meinedgs.de/release4/types/type14891_en.html

[34] https://ling.meinedgs.de/release4/types/type13890_en.html

[35] https://ling.meinedgs.de/release4/html/1246100_en.html#t00053300

7. Bibliographical References

Philippe Blache, Roxane Bertrand, Gaëlle Ferré, Berthille Pallaud, Laurent Prévot, and Stéphane Rauzy. 2017. The corpus of interactional data: A large multimodal annotated resource. In Nancy Ide and James Pustejovsky, editors, *Handbook of Linguistic Annotation*, pages 1323–1356. Springer Netherlands, Dordrecht.

Penny Boyes Braem and Virginia Volterra. 2023. Through the signed language glass: Changing and converging views in spoken and signed language research. In Terry Janzen and Barbara Shaffer, editors, *Signed Language and Gesture Research in Cognitive Linguistics*, pages 23–48. De Gruyter Mouton.

Onno A. Crasborn, Inge Zwitserlood, Els van der Kooij, and Richard Bank. 2020. Annotation conventions for the Corpus NGT. Annotation convention, Centre for Language Studies & Department of Linguistics, Radboud University, Nijmegen, Netherlands.

Ludivine Crible. 2017. Discourse markers and (dis)fluency in English and French: Variation and combination in the DisFrEn corpus. *International Journal of Corpus Linguistics*, 22(2):242–269.

Ludivine Crible. 2018. *Discourse Markers and (Dis)fluency: Forms and functions across languages and registers*, volume 286 of *Pragmatics & Beyond New Series*. John Benjamins Publishing Company, Amsterdam.

Ludivine Crible, Ivana Didirková, Christelle Dodane, and Loulou Kosmala. 2024. Towards an inclusive system for the annotation of (dis)fluency in typical and atypical speech. *Clinical Linguistics & Phonetics*, 38(4):381–398.

Sandra Debreslioska, Anna Kuder, and Pamela Perniss. 2025. The palm-up open hand gesture across language modalities: A comparison of German and DGS (German Sign Language). *Journal of Pragmatics*, 249:99–119.

Elisabeth Degand, Laurence Meurant, Gaëtanelle Gilquin, and Anne Catherine Simon, editors. 2019. *Fluency and Disfluency across Languages and Language Varieties*. Number 7 in Corpora and Language in Use. Presses universitaires de Louvain, Louvain-la-Neuve, Belgium.

Simeon Floyd, Elizabeth Manrique, Giovanni Rossi, and Francisco Torreira. 2016. Timing of visual bodily behavior in repair sequences: Evidence from three languages. *Discourse Processes*, 53(3):175–204.

Victoria Fromkin. 1980. Introduction. In *Errors in linguistic performance: slips of the tongue, ear, pen, and hand*, pages 1–12. Academic Press, New York, NY.

Sílvia Gabarró-López. 2019. Describing buoys from the perspective of discourse markers: A cross-genre study of French Belgian Sign Language (LSFB). *Sign Language & Linguistics*, 22(2):210–240.

Sílvia Gabarró-López. 2020a. Are discourse markers related to age and educational background? A comparative account between two sign languages. *Journal of Pragmatics*, 156:68–82.

Sílvia Gabarró-López. 2020b. Discourse markers, where are you? Investigating the relationship between their functions and their position in French Belgian Sign Language conversations. *Sign Language Studies*, 20(2):231–263.

Sílvia Gabarró-López. 2025. Discourse markers in French Belgian Sign Language (LSFB) dialogues and their translation into French: A corpus-based study. *Languages*, 10(9):243.

Sílvia Gabarró-López and Laurence Meurant. 2016. Slicing your SL data into basic discourse units (BDUs). Adapting the BDU model (syntax + prosody) to signed discourse. In *Proceedings of the LREC2016 7th Workshop on the Representation and Processing of Sign Languages: Corpus Mining*, pages 81–88, Portorož, Slovenia. European Language Resources Association (ELRA).

Sandra Götz. 2013. *Fluency in Native and Nonnative English Speech*, volume 53 of *Studies in Corpus Linguistics*. John Benjamins Publishing Company, Amsterdam.

Thomas Hanke. 2004. HamNoSys – representing sign language data in language resources and language processing contexts. In *Proceedings of the LREC2004 Workshop on the Representation and Processing of Sign Languages: From SignWriting to Image Processing. Information techniques and their implications for teaching, documentation and communication*, pages 1–6, Lisbon, Portugal. European Language Resources Association (ELRA).

Thomas Hanke and Jakob Storz. 2008. iLex – a database tool for integrating sign language corpus linguistics and sign language lexicography. In *Proceedings of the LREC2008 3rd Workshop on the Representation and Processing of Sign Languages: Construction and Exploitation of Sign Language Corpora*, pages 64–67, Marrakech, Morocco. European Language Resources Association (ELRA).

Martje Hansen. 2012. Textlinguistik: Gebärdensprache im Kontext. In Hanna Eichmann, Martje Hansen, and Jens Heßmann, editors, *Handbuch deutsche Gebärdensprache: sprachwissenschaftliche und anwendungsbezogene Perspektiven*, number Band 50 in Internationale Arbeiten zur Gebärdensprache und Kommunikation Gehörloser, pages 199–224. Signum, Seedorf.

Nikolaus P. Himmelmann. 1998. Documentary and descriptive linguistics. *Linguistics: An Interdisciplinary Journal of the Language Sciences*, 36(1):161–196.

Gabrielle Hodge and Onno Crasborn. 2022. Good practices in annotation. In Trevor Johnston, Julie A. Hochgesang, and Jordan Fenlon, editors, *Signed Language Corpora*, pages 46–89. Gallaudet University Press, Washington, D. C.

Jack Hoza. 2011. The discourse and politeness functions of HEY and WELL in American Sign Language. In Cynthia B. Roy, editor, *Discourse in signed languages*, number 17 in Sociolinguistics in deaf communities series, pages 69–95. Gallaudet University Press, Washington, D.C.

Elena Jahn, Gabriele Langer, and Cornelia Loos. 2021. Are we DONE yet? a corpus-based analysis of DONE as a clause-boundary diagnostic in DGS. Poster presented at Formal and Experimental Advances in Sign Language Theory (FEAST 2021).

Trevor Johnston. 2010. From archive to corpus: Transcription and annotation in the creation of signed language corpora. *International Journal of Corpus Linguistics*, 15(1):106–131.

Trevor Johnston. 2024. Auslan Corpus annotation guidelines. v2. Technical report, Monash University & Macquarie University.

Kim Josephine Kaul. 2024. *Palm-up Geste. Ein Vergleich zwischen Laut- und Gebärdensprachen*. Term paper, Georg-August-Universität, Göttingen, Germany.

Jörg Keller, Annette Hohenberger, and Helen Leuninger. 2003. Sign language production: slips of the hand and their repairs in German Sign Language (DGS). In *Cross-linguistic perspectives in sign language research. Selected papers from TISLR 2000*, volume 41 of *Internationale Arbeiten zur Gebärdensprache und Kommunikation Gehörloser*, pages 307–333, Amsterdam, Netherlands. Signum Verlag.

Peter Koch and Wulf Oesterreicher. 1985. Sprache der Nähe — Sprache der Distanz. Mündlichkeit und Schriftlichkeit im Spannungsfeld von Sprachtheorie und Sprachgeschichte. *Romanistisches Jahrbuch*, 36(1):15–43.

Reiner Konrad, Thomas Hanke, Amy Isard, Marc Schulder, Lutz König, Julian Bleicken, and Oliver Böse. 2024a. Corpus à la carte – improving access to the Public DGS Corpus. In *Proceedings of the LREC-COLING 2024 11th Workshop on the Representation and Processing of Sign Languages: Evaluation of Sign Language Resources*, pages 390–399, Turin, Italy. ELRA Language Resources Association (ELRA) and the International Committee on Computational Linguistics (ICCL).

Reiner Konrad, Thomas Hanke, Gabriele Langer, Dolly Blanck, Julian Bleicken, Ilona Hofmann, Olga Jeziorski, Lutz König, Susanne König, Rie Nishio, Anja Regen, Uta Salden, Sven Wagner, Satu Worseck, Oliver Böse, Elena Jahn, and Marc Schulder. 2024b. MY DGS – annotated. Public Corpus of German Sign Language, 4th release. Dataset, University of Hamburg.

Reiner Konrad, Thomas Hanke, Gabriele Langer, Susanne König, Lutz König, Rie Nishio, and Anja Regen. 2022. Public DGS Corpus: Annotation conventions. Project Note AP03-2018-01, Universität Hamburg, DGS-Korpus project, IDGS, Universität Hamburg.

Maria Kopf, Marc Schulder, Thomas Hanke, and Sam Bigeard. 2022. Specification for the harmonization of sign language annotations. EASIER project deliverable D6.2, Universität Hamburg.

Loulou Kosmala. 2022. Rethinking (dis)fluency within the scope of interactional linguistics and gesture studies. *Studia Universitatis Babeș-Bolyai Philosophia*, 67(2):49–66.

Gabriele Langer, Thomas Hanke, Reiner Konrad, and Susanne König. 2016. "Non-tokens": When tokens should not count as evidence of sign use. In *Proceedings of the LREC2016 7th Workshop on the Representation and Processing of Sign Languages: Corpus Mining*, pages 137–142, Portorož, Slovenia. European Language Resources Association (ELRA).

Alysson Lepeut and Emily Shaw. 2025. Interactive functions of palm-up: Cross-linguistic and cross-modal insights from ASL, American English, LSFB and Belgian French. *Languages*, 10(9):239.

Scott K. Liddell and Robert E. Johnson. 1989. American Sign Language: The phonological base. *Sign Language Studies*, 64(1):195–277.

Cornelia Loos and Reiner Konrad. 2022. Detailed annotation and qualifiers. Project Note AP03-2021-01, DGS-Korpus project, IDGS, Universität Hamburg, Hamburg, Germany.

Hannah Lutzenberger, Lierin De Wael, Rehana Omardeen, and Mark Dingemanse. 2024. Interactional infrastructure across modalities: A comparison of repair initiators and continuers in British Sign Language and British English. *Sign Language Studies*, 24(3):548–581.

Rachel McKee and Sophia Wallingford. 2011. 'So, well, whatever': Discourse functions of *palm-up* in New Zealand Sign Language. *Sign Language & Linguistics*, 14(2):213–247.

Johanna Mesch and Lars Wallin. 2021. Annoteringskonventioner för teckenspråkstexter. version 8. Technical report, Avdelningen för teckenspråk, Institutionen för lingvistik, Stockholms universitet, Stockholm.

Melanie Metzger and Ben Bahan. 2001. Discourse analysis. In Ceil Lucas, editor, *The Sociolinguistics of Sign Languages*, pages 112–144. Cambridge University Press, Cambridge.

Elena Georgia Mpadanes. 2020. Approaching repairs in German Sign Language conversations: an analysis of data from the DGS corpus. Master's thesis, University of Hamburg, Hamburg.

Don Newkirk, Edward S. Klima, Carlene Canady Pedersen, and Ursula Bellugi. 1980. Linguistic evidence from slips of the hand. In *Errors in linguistic performance: slips of the tongue, ear, pen, and hand (12th International Congress of Linguists, 1977; Working Group on Speech Errors)*, pages 165–198, Vienna, Austria. Academic Press.

Rie Nishio, Sung-Eun Hong, Susanne König, Reiner Konrad, Gabriele Langer, Thomas Hanke, and Christian Rathmann. 2010. Elicitation methods in the DGS (German Sign Language) Corpus project. In *Proceedings of the LREC2010 4th Workshop on the Representation and Processing of Sign Languages: Corpora and Sign Language Technologies*, pages 178–185, Valletta, Malta. European Language Resources Association (ELRA).

Ingrid Notarrigo and Laurence Meurant. 2014. Nonmanuals as markers of (dis)fluency. In *Proceedings of the LREC2014 6th Workshop on the Representation and Processing of Sign Languages: Beyond the Manual Channel*, pages 135–142, Reykjavik, Iceland. European Language Resources Association (ELRA).

Ingrid Notarrigo and Laurence Meurant. 2019. Conversations spontanées en langue des signes de Belgique francophone (LSFB) : fonctions et usages de la répétition. *Lidil*, 60.

Alejandro Oviedo, Thomas Kaul, Katharina Urbann, Reiner Griebel, and Leonid Klinner. 2018. Zur bedeutung von Pausen in der Deutschen Gebärdensprache – eine linguistische Annäherung. *DAS ZEICHEN*, 108:116–125.

Yolanda Pérez. 2006. Marcadores manuales en el discurso narrativo en la lengua de señas venezolana. *Letras*, 48(72):349–363.

Marie Piu and Rémi Bove. 2007. Annotation des disfluences dans les corpus oraux. In *Actes de la 14ème conférence sur le Traitement Automatique des Langues Naturelles. REncontres jeunes Chercheurs en Informatique pour le Traitement Automatique des Langues (Posters)*, pages 397–406, Toulouse, France. ATALA.

Siegmund Prillwitz, Thomas Hanke, Susanne König, Reiner Konrad, Gabriele Langer, and Arvid Schwarz. 2008. DGS Corpus project – development of a corpus based electronic dictionary German Sign Language / German. In *Proceedings of the LREC2008 3rd Workshop on the Representation and Processing of Sign Languages: Construction and Exploitation of Sign Language Corpora*, pages 159–164, Marrakech, Morocco. European Language Resources Association (ELRA).

Patrick Louis Rohrer, Ulya Tütüncübasi, Júlia Florit-Pons, Ingrid Vilà-Giménez, Núria Esteve-Gibert, Ada Ren-Mitchell, Stefanie Shattuck-Hufnagel, and Pilar Prieto. 2025. Multidimensional labeling of gesture in communication: the M3D proposal. *Corpus Pragmatics*, 9(4):411–433.

Marie-Anne Sallandre and Brigitte Garcia. 2013. Epistemological issues in the semiological model for the annotation of sign languages. In Laurence Meurant, Aurélie Sinte, Mieke Van Herreweghe, and Myriam Vermeerbergen, editors, *Sign Language Research, Uses and Practices: Crossing Views on Theoretical and Applied Sign Language Linguistics*, pages 159–178. De Gruyter Mouton.

Marc Schulder and Thomas Hanke. 2020. OpenPose in the Public DGS Corpus. Project Note AP06-2019-01, DGS-Korpus project, IDGS, Hamburg University, Hamburg, Germany.

Elizabeth Shriberg. 1994. *Preliminaries to a Theory of Speech Disfluencies*. PhD dissertation, University of California, Berkeley.

Elizabeth Shriberg. 2001. To 'errrr' is human: ecology and acoustics of speech disfluencies. *Journal of the International Phonetic Association*, 31(1):153–169.

Saúl García Villameriel. 2008. La interpretación de los marcadores del discurso de la lengua de signos española. In *Actas del III Congreso de la Federación Española de Intérpretes de Lengua de Signos y Guías-Intérpretes (FILSE)*.

Saúl García Villameriel. 2010. EN-CAMBIO y ES-DECIR: origen de dos marcadores discursivos de la lengua de signos en el español oral. In *Actas del IX congreso de lingüística general*, Valladolid, Spain.

Elisabeth Volk. 2016. Discourse functions of palm-up in German Sign Language (DGS). Poster presented at the 5th Meeting of the Formal and Experimental Advances in Sign Language Theory Colloquium (FEAST 2016), Venice, Italy.