



Article

Numeral Incorporation as Grammaticalization? A Corpus Study on German Sign Language (DGS)

Felicitas Otte , Anke Müller , Sabrina Wähl  and Gabriele Langer 

Institute of German Sign Language and Communication of the Deaf, University of Hamburg, 20354 Hamburg, Germany; anke.mueller@uni-hamburg.de (A.M.); sabrina.waehl@uni-hamburg.de (S.W.); gabriele.langer@uni-hamburg.de (G.L.)

* Correspondence: felicitas.otte@uni-hamburg.de

Abstract: Numeral incorporation describes the merging of a numeral sign with a lexical sign to create a single sign with a compositional meaning, e.g., “three weeks.” As a phenomenon of simultaneous morphology, numeral incorporation is unique to sign languages. While researchers disagree on the exact morphological structure of the construction, it has, thus far, mainly been described as a synchronic, phonological phenomenon. Using the DGS corpus, a language resource on German Sign Language, we explore the possibility of numeral incorporation resulting from a language change process, specifically a grammaticalization process. Our dataset comprises tokens belonging to nine different signs that may occur in numeral incorporations. We find a cline of three constructions in the corpus, which shows a progression from free morpheme to cliticized morpheme to bound morpheme (affix). A comparison of the usage frequency of the three constructions in different age groups reveals that signers use more incorporations the younger they are. Following the apparent time approach, these observations are taken as indicators of diachronic language change. We describe to what extent the properties of numeral incorporation fit with the grammaticalization hypothesis and conclude that while the emergence of numeral incorporation is an instance of language change and shows some aspects seen in grammaticalization, the gradual change fails to exhibit some crucial aspects of grammaticalization and, thus, should not be regarded as an example thereof.



Citation: Otte, Felicitas, Anke Müller, Sabrina Wähl, and Gabriele Langer. 2023. Numeral Incorporation as Grammaticalization? A Corpus Study on German Sign Language (DGS). *Languages* 8: 153. <https://doi.org/10.3390/languages8020153>

Academic Editors: Renata Enghels, Marlies Jansegers and Tom Bossuyt

Received: 15 February 2023

Revised: 1 June 2023

Accepted: 4 June 2023

Published: 19 June 2023



Copyright: © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

Keywords: German Sign Language; numeral incorporation; grammaticalization; corpus study

1. Introduction

Numeral incorporation is a very common phenomenon in sign languages and denotes the integration of a numeral sign into a lexical sign by merging features of both into a new sign form. In German Sign Language (DGS), for example, temporal signs such as WEEK1A “week” can incorporate handshape features of a restricted set of numerals to denote the meaning “quantity of unit,” e.g., “quantity of weeks.” In the past, numeral incorporation has been described as a synchronic, morpho-phonological phenomenon (Chinchor 1982; Mathur and Rathmann 2011; Semushina and Mayberry 2019). In this study, we want to explore the possibility of numeral incorporation being the result of diachronic language change, specifically, a grammaticalization process. Our study is based on the recently compiled DGS corpus, which contains videos of signed language data. As there is no suitable diachronic data available for DGS, we employ the synchronic data of the DGS corpus to compare the use of numeral incorporation across different age groups (using the apparent time approach). Following the usage-based framework of construction grammar, we can deal with heterogenous data showing multiple and, at times, competing constructions at work, e.g., sequential and simultaneous constructions.

Our results confirm that numeral incorporation is a field with semantic and phonological constraints, but the data also indicates ongoing language change: frequency of use and phonological change may lead to an increase in use of numeral incorporation in

terms of quantity, but also in terms of attracting more lexical signs to the pool of signs allowing for numeral incorporation. For some signs, the language change process leading to numeral incorporation is incomplete and will need research on future data to confirm the entrenchment—or abandonment—of the construction.

2. Theoretical Background

In this section, we provide some background on the phenomena relevant to our study. We start with an introduction to the numeral systems in sign languages and DGS, in particular (Section 2.1), continue with a description of numeral incorporation (Section 2.2), and, lastly, provide some background information on grammaticalization (Section 2.3.).

2.1. Numeral Systems in Sign Languages

Before describing the specifics of the DGS numeral system, we provide a brief introduction to sign language phonology, which will be helpful in following the subsequent descriptions. Signs can be generally described as consisting of four different parameters: handshape, location, hand orientation, and movement. As Brentari (2012, p. 22) described it, parameters are “different phonemic groups of features.” Thus, they constitute contrastive properties in the form of a sign. Specifically, a parameter can only be specified for one (possibly complex) value at a time. In a fully specified sign, all parameters are specified in the phonemic representation and are realized in the articulation of a sign. The phonological status of these parameters is shown by minimal pairs, as illustrated in Appendix A, Figure A1. Changing the value of one parameter of an individual sign, e.g., the hand orientation of MILK2C “milk,” leads to a different sign, e.g., DIFFICULT1 “difficult.”

Sign languages differ in their numeral systems with respect to the underlying count base, as well as to the motivation of forms and productive strategies to form numerals higher than 10, and whether they employ one or two hands. Sagara and Zeshan (2016, pp. 28–33) give a typological overview of the cardinal number system of 30 sign languages. Numeral systems are often a mixture of lexical and productive word forms.

Most of the larger or urban sign languages have a base of 10, which is similar to the majority of spoken languages. This may be due to the number of fingers of the human hands (Zeshan et al. 2013, p. 360). Numeral signs can reflect different iconic motivations, e.g., form iconicity with respect to the written number (which Sagara and Zeshan (2016, p. 28) call orthographic iconicity) or to the initial letter of the written spoken language numeral, or different kinds of a finger-for-number analogy for the numbers 1 to 10. In a two-handed system, the second hand is added to enumerate 6–10, while in a one-handed system, the numbers six to nine or 10 can be articulated on one hand as well, e.g., by touching a certain finger with the thumb.

The numeral system of DGS is two-handed, and the numerals for the numbers 1–10 are “based on ‘number-for-number iconicity’ (Taub 2001), that is, the number of extended fingers corresponds to the numerical value to be expressed” (Sagara and Zeshan 2016, p. 29). The selection of fingers results in specific handshapes serving to distinguish the numerals from 1–10. These are called numeral handshapes. DGS numerals from 1 to 10 vary based on which fingers are extended (e.g., “one” can be expressed by either an extended index or an extended thumb) and the hand orientation (see Konrad et al. 2022, pp. 14–16, for more detail). Variants in hand orientation either have the palm directed away from or towards the body. Two-handed signs (from 6–10) may exhibit mixed orientations; that is, with one of the hands directed away from the body and the other directed towards the body. All numerals from 1–10 share one orientation feature: the hand is upright, and the extended fingers point upwards. For pictures of the numeral handshapes 1–10, see Appendix B, Figure A2, which include variants in handshape but leave out variance in palm orientation (palm directed away or towards the body).

Numerals for numbers higher than 10 are usually complex numerals; that is, they are formed by syntactic or morphosyntactic rules (von Mengden 2010, p. 28). In DGS, and in sign languages in general, there are two basic categories: numerals expressed in a single,

morphologically complex sign, or in a sequence of signs. Numerals of the first group are expressed in a regular way by combining a specific movement and hand orientation with selected extended fingers to express values from 1 to 9. Sagara and Zeshan (2016, p. 29) describe this strategy as “combining a numeral handshape with a numeral movement pattern.” The teens, i.e., the numbers from 11 to 19¹, are formed combining the numeral handshapes from 1 to 9, with a meaningful movement signifying the addition of 10 (see Appendix C Example (1), for an example featuring a TEEN sign). Multiples of 10 are likewise expressed, combining numeral handshapes from 2 to 9, with another movement signifying the multiplication with 10 to produce numerals from 20 to 90. The multiples of hundred and thousand are also expressed applying the same strategy. The numeral handshapes from one to nine/ten can be attached to a numeral movement to signify 100 up to 900, and 1000 up to 10,000. Figure 1 illustrates the numerals 100, 200, and 900. Tens, hundreds, and thousands differ with regard to their movements.

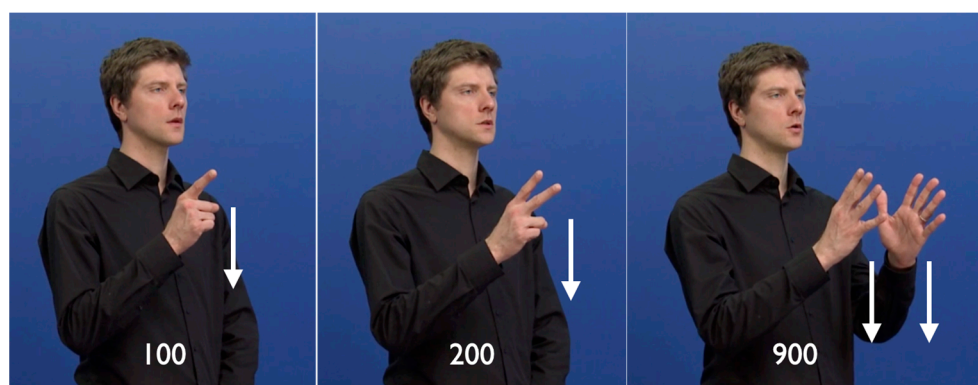


Figure 1. Productive strategy for multiples of 100.

Millions, on the other hand, are complex numerals that are expressed sequentially; that is a numeral from 1 to 10 followed by a sign for million, meaning the multiplication of millions by the preceding number. All numbers that express high values by combining lower values are expressed through a mixture of morphosyntactic strategies. They are not relevant for this paper but should be mentioned for completeness. A number like 47 is signed sequentially in DGS: first the numeral for seven, followed by the numeral for 40 (additive strategy). This order mirrors the practice in spoken German. Example (2) in Appendix C illustrates the use of such a numeral.

While the term complex numerals denotes all of the above described numerals higher than 10, from now on, we restrict the use of the term to complex numerals of the first kind, that is to complex DGS numerals with simultaneous morphology encompassing the teens, the multiples of ten, hundred, and thousand.

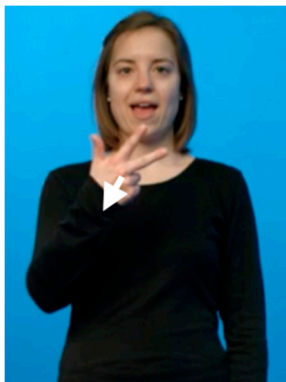
Another remark on terminology may be in order. The terms “numeral sign” and “numeral” refer to the same concept, the linguistic expression of a number. If we want to refer to the handshape parameter of numerals only, we speak of numeral handshapes as opposed to numeral signs or numerals (with all parameters specified).

2.2. Numeral Incorporation


Numeral incorporation is the simultaneous combination of a numeral sign and a lexical base sign. As a linguistic structure, it is common to many sign languages (Sagara and Zeshan 2016, pp. 31–32). Lexical signs that allow numeral incorporation usually belong to semantic fields such as time units, calendric terms, measurements, currencies, and school grades. However, the specific set of signs that do incorporate numerals depends on the language (Sagara and Zeshan 2016, p. 32; Ktejik 2013, p. 208; Mathur and Rathmann 2011, p. 62; Chinchor 1982, p. 77). Another restriction pertains to which numerals can be incorporated. This limitation seems to be of a phonological nature, depending on the phonological features of the numeral or the numeral system of a sign language

(see Sagara and Zeshan 2016 for a typological overview). In DGS, the numerals 1 to 10 can be incorporated,² whereas Semushina and Mayberry (2019, pp. 109–10) report the incorporation of the numerals from 1 up to 15 in Russian Sign Language (RSL), at least for one sign. Phonological features of the incorporating lexical base sign play a role as well (Semushina and Mayberry 2019; Mathur and Rathmann 2011), e.g., restrictions on the handshape parameter. We will elaborate on this a little bit later in this section.

The following example in Figure 2a,b shows a sequential expression with the meaning of “three weeks.” This construction stands in contrast to the numeral incorporation with the same meaning shown in Figure 3. We will use this example to illustrate the morpho-phonological process behind numeral incorporation. As the exact nature of the process is still under discussion, we describe several perspectives.


	Handshape	“3-handshape:” fist with extended thumb, index, and middle finger
	Location	Centralized in front of the body
	(palm) Orientation	Towards the body
	Movement	Slight forward movement

(a)

	Handshape	“Index-handshape:” fist with extended index finger
	Location	Centralized in front of the body
	(palm) Orientation	Away from the body
	Movement	Linear movement to the right (left when produced with the left hand)

(b)

Figure 2. (a). Phonological parameters and illustration of the DGS sign THREE1A. (b). Phonological parameters and illustration of the DGS sign WEEK1A.



Handshape	“3-handshape:” fist with extended thumb, index, and middle finger
Location	Centralized in front of the body
(palm) Orientation	Away from the body
Movement	Linear movement to the right (left when produced with the left hand)

Figure 3. Phonological parameters and illustration of the incorporated form “three weeks.”

Figure 2a,b show the fully specified (non-incorporated) signs signed in sequence (the numeral THREE1A³ “three” and the lexical sign WEEK1A “week”) and describe each sign’s phonological parameters.

When signed sequentially, both signs realize all of their phonological parameters; that is, they are fully specified signs with their own specific movements, locations, orientations, and handshapes. In a numeral incorporation, in contrast, there is only one sign in which the numeral and the lexical base sign are merged. As a phonological parameter can be realized only through one value at a time, some values in the incorporated form are determined by the numeral, while others are determined by the base sign and as a consequence; some of the individual signs’ features are deleted. The result is a morphologically complex form (Figure 3) that takes on the handshape of the numeral but retains the specifications of the base sign in the other three parameters⁴.

As for the analysis of numeral incorporation, various explanations have been proposed for different sign languages. In an early analysis of numeral incorporation in American Sign Language (ASL), Chinchor (1982) describes the sign formation process of numeral incorporation as the combination of two signs. These are not both fully specified signs, though, as the numeral sign is regarded as unspecified for location and movement (Chinchor 1982, p. 130). Chinchor, furthermore, claims that the base sign (in its unmodified non-incorporating form) has to have a handshape matching the numeral “one” in one of its allomorphs. This rule has also been stated by Mathur and Rathmann (2011) as a tendency in both ASL and DGS.

Mathur and Rathmann (2011) describe numeral incorporation as a morphological process where a numeral and a sign with the above-mentioned semantic properties become one sign. Both signs have “a fixed phonological realization” (Mathur and Rathmann 2011, p. 57) in their fully specified form. In the process of merging, the numeral lends its handshape and loses its features of placement and movement, while the lexical base sign loses its handshape parameter and lends its other parameter values to the resulting construction. Mathur and Rathmann (2011) place this process within the group of non-concatenative word formation, as the articulation of the relevant features is simultaneous rather than a sequence of morphemes. Which signs take part as base signs in the process could be specified in a list in the lexicon (comparable to lists of irregular verbs with vowel change as for instance the past tense morpheme in English). However, there are phonological reasons for why some signs incorporate numerals and others do not. Thus, it seems that if the lexical sign has a marked handshape⁵, the handshape cannot be deleted; if the numeral has a complex, distinctive movement feature, this movement likewise cannot be deleted. In two-handed signs, symmetry constraints may prevent incorporation (Mathur and Rathmann 2011, pp. 67–69). Semushina and Mayberry (2019) attest and refine these general findings for RSL.

A contrary morphological analysis is proposed by Liddell (1996), who describes incorporated forms as consisting of a bound root as a lexical base and a numeral morpheme. The bound root is specified for all phonological parameters, except for the handshape, while the numeral morpheme consists of only a handshape and, as such, is also a bound morpheme. The two morphemes are then combined to form one sign with four specified phonological parameters. This analysis is based on the fact that some incorporated forms do not have a non-incorporated, independent base sign with four phonological parameters, e.g., the ASL sign O'CLOCK ("o'clock") is always specified for a specific time, i.e., it always contains a numeral handshape. Thus, it does not exist independently as a fully specified base sign. Aiming at finding a theory that accounts for these, as well as the signs with fully specified independent base signs, Liddell argues that incorporated forms as the one shown above for "week" are not based on the fully specified sign, but they are likewise formed with a bound root, which matches the fully specified sign in all parameters but a handshape (as it has an empty handshape slot). This solution entails the assumption of two different signs in cases where an independent lexical base exists, though the lexical base looks the same as the bound root with the numeral handshape for "one." Liddell (1996, p. 208) explains these forms as having developed "through semantic extension," e.g., the combined sign meaning "one week" extended its meaning to also mean "week."

Out of these analyses, our own understanding of numeral incorporation most closely resembles Mathur and Rathmann (2011), as their analysis features a fully specified sign as a lexical base that is combined with a numeral handshape. Our reasoning for this will be discussed in Section 3.3.

2.3. Grammaticalization

Grammaticalization can be roughly described as a phenomenon of language change through which a linguistic item becomes more grammatical, e.g., a lexical content word like "go" becomes used in another function, e.g., as a future tense marker, and changes its behavior subsequently. However, when is a language change process also grammaticalization? Lehmann (1985) and Heine and Kuteva (2007) have devised some useful criteria, which we will refer to in detail in Section 5.2.

When it comes to theories of grammaticalization, the most prototypical example is the grammaticalization of a single element, as e.g., the emergence of a new morpheme that arises from the reanalysis of a linguistic construction. Reanalysis is a parameter of grammaticalization that is also emphasized by Hopper and Traugott (2003, pp. 50–63). Reanalysis describes a process of structural reinterpretation of an item in a given context. For example, the original structure of the "going-to" construction was a lexical verb (to go) with a to-infinitive sub-clause expressing an intention. The sentence "I am going to get groceries" in this original interpretation could likewise be expressed as "I am walking in order to get groceries." The structure of the construction was then reanalyzed to instead consist of a temporal marker for future (be going to), followed by a lexical verb ("get," in the above example). The reanalysis culminates in the usage in new contexts that do not allow for an interpretation in the original structure: "I am going to think about it" cannot be phrased as "I am walking in order to think about it," as the latter sentence expresses a different meaning. As is apparent from this example, it is not only the individual item that undergoes a grammaticalization process, but rather an entire construction. Himmelmann (2004, p. 31) emphasizes this aspect of grammaticalization: "The unit to which grammaticization properly applies are constructions, not isolated lexical items." The use of the grammaticalized item in new contexts is described as a "host-class expansion" by Himmelmann (2004, p. 32). That is, the grammatical item may combine with more and broader classes than in the source construction. Heine and Kuteva (2007, pp. 35–36) call this effect of a grammaticalization process the parameter of extension.

For languages with a long, written tradition, there is ample material to document and research instances of language change and grammaticalization. As a result, cross-linguistically similar processes have been attested that led to the identification of "clines"

(Hopper and Traugott 2003, pp. 6–9), suggesting historical processes moving in a certain direction. The knowledge of these helps identify instances of grammaticalization where only synchronic data is available and internal reconstruction is the method of choice when grouping and classifying senses and functions of a linguistic item in question.

While phonological changes have been attested in sign languages, e.g., ASL (Frishberg 1975), and different historical stages of language use have been documented in historical dictionaries or historical descriptions of sign languages, these typically encompass single lexical signs, but not grammar or language in use. Even though historic records of sign language use are scarce, there are some findings of grammaticalization in sign languages (Pfau and Steinbach 2011; Janzen 2012). Pfau and Steinbach (2011, p. 684) note that the method of internal reconstruction is useful in sign language research, though less reliable than a diachronic comparative approach. Hypotheses can be built based on inferences from patterns of the synchronic grammar, assuming grammaticalization processes to be usually unidirectional and taking into account what is known about common grammaticalization paths (Pfau and Steinbach 2011, p. 684). Typically, grammaticalization in signed languages proceeds “from lexical element to free grammatical marker,” e.g., development of aspectual and tense markers from verbs and adverbs, and less often or scarcely “from free grammatical marker to grammatical affix” (Pfau and Steinbach 2011, p. 689), which may be due to characteristics of signed language morphology, or the fact that sign languages tend to be relatively young languages (Janzen 2012, p. 836). In contrast to spoken languages, manual or facial gestures are a well-attested source of grammaticalization, even at times directly from gestures to functional elements, thereby skipping a step of lexicalization (Pfau and Steinbach 2011, pp. 689–93; Janzen 2012, pp. 829–34).

In the absence of historical data, synchronic language corpora become the database of choice. Representative, well-curated corpora allow for a thorough comparison of variant constructions. Following Bybee (2006, 2011), we regard the manifold manifestations of language use as a base for a reconstruction of cognitively organized rules that are informed by language use. Frequency of use is a relevant aspect for causing changes, e.g., frequency of co-occurrence of linguistic items as a prerequisite for chunking (Bybee 2011, pp. 70–71) most often leads at least to phonetic reduction. Phonetic reduction is a case of language change but not necessarily a grammaticalization process. The more frequently a construction is used, the more likely it is to become part of the cognitive linguistic rules; that is, the more likely it is to become entrenched (Bybee 2011, pp. 77–78). In the case of competing constructions, it is thus the one that is most frequent that is likely to become entrenched and, in consequence, lead to language change by abandonment of the other constructions, and “as long as frequency is on the rise, changes will move in a consistent direction” (Bybee 2011, p. 77). DGS lacks corpora with historic data of language use. However, a synchronic corpus has recently become available. Therefore, we can apply the apparent time hypothesis in our investigation of this potential grammaticalization process. According to Bailey et al. (1991), people within a generation tend to preserve the linguistic stage of their youth. Thus, a gradient age distribution in a language community can be taken to indicate language change. The oldest people in this approach represent the earliest language stage in the dataset, while the youngest people represent the most recent time period. This theoretical concept, used primarily in sociolinguistics in the wake of Labov (1963), attested its usefulness as an analytical tool through an evaluative test with real-time data (Bailey et al. 1991). The apparent time approach is the method of choice where historical linguistic data are lacking and has been used in sign language research e.g., by Hanke et al. (2017) and Dachkovsky (2022). Where possible, it should be complemented by diachronic, real-time approaches.

3. Materials and Methods

3.1. The DGS Corpus

The DGS corpus⁶ is compiled and curated by the DGS-Korpus project and serves as a reference corpus for a corpus-based dictionary of DGS (the DW-DGS). A total of

50 h of annotated and translated data are available to the public via three portals: one addressing the language community (*MY DGS*), and two addressing linguistic researchers (*MY DGS—annotated* and *MY DGS—ANNIS*) (see Hanke et al. 2020; Isard and Konrad 2022). For our study, we used the annotated parts of the reference corpus, which includes more data and more detailed annotations than the Public DGS Corpus.⁷

The DGS corpus is based around 560 h of natural and near-natural signing from 330 participants rooted in the deaf community. The participants use DGS fluently and on a daily basis. They were recorded in pairs throughout Germany from 2010 to 2012. Participants were asked to complete different tasks, such as conversations about given topics and retellings of stories. The corpus is balanced with regard to region, gender, and age of the participants (see Schulder et al. 2021). Four age groups were sampled based on the age of the participants at the time of the recording, ranging from the youngest group born between 1994 and 1981 (age group 18–30) to the oldest group born in 1950 and before (age group 61+). In between, there are age groups 31–45 (of people born between 1980 and 1966) and 46–60 (of people born between 1965 and 1951). At the time of conducting our study in August 2022, the corpus data consisted of 92 h of lemmatized material with 671,851 tokens.

3.2. Corpus Annotation

The DGS data are annotated using the iLex database and annotation tool (Hanke and Storz 2008), which facilitates type-token matching and is searchable via SQL. Each type has a unique type id, but also an (id-)gloss name roughly hinting at the meaning of a sign.⁸ The annotation conventions (Konrad et al. 2022) specify all aspects of the annotation and provide an overview of different sign categories. We will briefly describe parts of the annotation process and the glossing conventions that are relevant to our study. The DGS-Korpus project is still ongoing, and the descriptions below match the conventions from the third release of the Public DGS Corpus.

In a first pass of annotation, translations are added on the sentence level; the signs are segmented and lemmatized (type-token matching); and the mouthings and mouth gestures⁹ are added (Hanke et al. 2023, p. 200). Any modifications and forms not exactly matching the citation forms are marked with the letter “a” (for “Abweichung”—deviation¹⁰). In the second step of annotation, these items are reviewed and categorized, possibly following a group discussion if the classification of the form is uncertain. As a result, information on the specific form of a given token is added, e.g., on a sign being one- or two-handed or a sign showing repeated movement (Hanke et al. 2023, p. 204). In preparation specifically for the study at hand, some of the annotations of numeral signs and numeral incorporating signs were revised by the annotation team.

During the lemmatization process, each sign token is matched with a type that has a unique, persistent ID, a gloss name, and a HamNoSys notation¹¹. The gloss name not only hints at the meaning of the sign but also includes a number to distinguish different signs with the same gloss name (e.g., lexical variants) and may include a letter to indicate a phonological variant (e.g., *YEAR1A* and *YEAR1B*). Numeral signs constitute their own category of signs and are marked by the prefix “\$NUM.” This prefix is attached to numerals of different kinds, such as ordinals, cardinals, divisions, etc., but also to some numeral incorporating signs, which do not have an independent base form (such as *\$NUM-CLOCK1A*)¹². In the DGS corpus, the numerals 1 to 10 are annotated using four different glosses: *\$NUM-ONE-TO-TEN1A*, *\$NUM-ONE-TO-TEN1B*, *\$NUM-ONE-TO-TEN1C*, and *\$NUM-ONE-TO-TEN1D*. These phonological variants differ with respect to hand orientation (away from the body/towards the body/one hand oriented in either direction). For our analysis and for readability in this paper, we subsumed and coded these types as *\$NUM-ONE-TO-TEN*.

Some of the forms deviating from the citation form of a sign type are annotated by adding qualifiers to the respective sign type gloss¹³. These qualifiers categorize and label recurring formational patterns of deviations from the assumed citation form across different

signs that may be of phonetic, phonological, or morphological nature. Most relevant to this study is the qualifier “q” (quantification),¹⁴ which is added to either a numeral to specify the value of the numeral (possible values range from 0 to 10) or to a numeral incorporating sign to likewise specify the value of the incorporated numeral (possible values range from 1 to 10).¹⁵ The qualifier values may simply consist of a number but may also contain additional letters. The letters identify variants of the numeral handshape, e.g., the letter “d” in “1d” means a handshape with extended thumb, whereas “1” refers to the extended index finger.

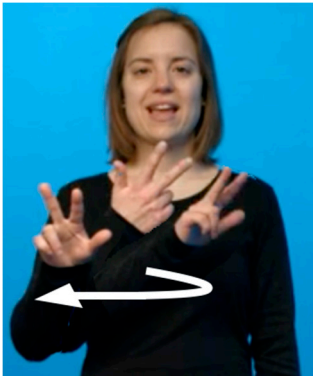
For example, a non-incorporated form meaning “three weeks” looks like this in glosses: \$NUM-ONE-TO-TEN1A’q:3d WEEK1A. These refer to two distinct signs, as indicated by the blank space between the glosses (see also Figure 2a,b for the signs’ form). The second sign is WEEK1A, which matches its citation form and is, thus, not marked in any other way. The first sign is one of the phonological variants of the 1–10 numerals and is specified for the value “3d,” which means that the produced handshape is the 3-handshape involving the thumb. For comparison, the glossing of an incorporated form looks as follows: WEEK1A’q:3d. Only the sign WEEK1A is articulated, and it is in this case the one specified by the value “3d,” indicating that the sign is produced with the very same handshape as the numeral sign in the non-incorporated construction (see also Figure 3 for the sign’s form).

All sign types (see Appendix D) and all examples (see Appendix C) can be found in the Public DGS Corpus. Note that the annotations in the Public DGS Corpus differ in some respects from the annotations of the reference corpus: Qualified forms are only indicated with an asterisk instead of a full specification of the modification, thus subsuming all qualified and deviating forms into one category. Signs of the category \$NUM, on the other hand, end by specifying the handshape according to the numbers 1–10.

3.3. Concept of Study

From the morpho-phonological complexity of the incorporating construction arises the question of how this structure may have developed and whether grammaticalization is involved. A balanced, annotated corpus with access to metadata information gives us the opportunity to take the synchronic variation, as well as sociolinguistic properties of participants such as region, gender, and age, into account. With a comparably large corpus, we have considerably more natural language data than was the case for any of the previous studies on numeral incorporation, some of which are based on elicited material, some on interviews, and some on both methods. This gives us the opportunity not only to attest the synchronic variation, but also to analyze it with respect to age groups, thus estimating trends of diachronic change according to the apparent time approach.

The annotated DGS corpus data suggest a three-partite variance of constructions when it comes to the combination of a numeral from 1 to 10 and a sign serving as a base sign. Additionally to constructions without incorporation, as in Figure 2a,b, and incorporated constructions, as in Figure 3, there is also an in-between structure that consists of an individually signed numeral and a base sign with numeral handshape incorporation (Figure 4).



Handshape	“3-handshape”: fist with extended thumb, index, and middle finger
Location	Centralized in front of the body
(palm) Orientation	Towards the body → away from the body
Movement	Small arch with a wrist turn, followed by a linear movement to the right

Figure 4. Phonological parameters and illustration of the cliticized form ‘three weeks’.

In this construction, the handshape parameter of the lexical base sign is modified in a process of progressive handshape assimilation: The numeral’s handshape perseveres throughout the subsequent lexical sign. Handshape assimilation is a regular and frequently occurring phonological process in sign languages (see Quer et al. 2017, pp. 53–54) and is often observed in quick or casual signing. Interestingly, an brief examination of the corpus data hints at a tendency towards regressive assimilation (handshape anticipation) in DGS rather than progressive assimilation. The assimilation to numeral handshapes thus stands out and indicates a morphosyntactic environment prone to cliticization. Additionally, there is a reduction of phonological material on part of the numeral sign. Though there are still two signs, they are less distinguishable as the numeral has lost its own movement value (slight forward movement), with the result of a smooth movement along both signs, resulting in a twist of the hand that comes from the change of hand orientation.

We will call the sequential production of two signs as shown in Figure 2a,b a **phrasal construction**, the incorporated form as shown in Figure 3 an affixation or **affixed form**, and the in-between construction we will call the **cliticized form** (Figure 4). It can be regarded as an incorporated form that is immediately preceded by the numeral whose value is identical to the incorporated one. Whenever we mention incorporating forms, we are referring to both affixed and cliticized forms. The existence of three variant constructions can be regarded as an indicator of ongoing language change, assuming that, synchronically, competing constructions show up during an ongoing language change process; for example a grammaticalization process.

Based on these observations, we propose a cline of three different constructions that show the progression towards grammaticalized numeral incorporation:

- Phrasal construction: In this construction the two elements are articulated independently from, but adjacent to, one another. The numeral keeps the location and orientation of its unmodified form, and the base is articulated with the handshape of its unmodified form. This is reflected in the corpus through the annotation of a numeral and a lexical base sign (see Appendix C Example (3) for an example featuring the phrasal construction).
- Cliticization: The numeral is still articulated as its own distinguishable sign with its own location and orientation. However, the noun loses its handshape and takes on the numeral handshape instead. The movement of the numeral sign is lost; instead, the movement of the lexical base is extended to accommodate the twist that is needed to match the hand orientations of both signs. This is reflected in the corpus through the annotation of a numeral and a modified lexical base sign (see Appendix C Example (4) for an example featuring the cliticized construction).
- Affixation: Here, the numeral only exists as the numeral handshape of the base sign. The two elements have become one sign. This is reflected in the corpus through the annotation of a single modified lexical base sign (see Appendix C Example (5) for an example featuring the affixed construction).

The existence of a cliticized form in the data speaks towards a process that starts with two independent signs, as proposed by [Mathur and Rathmann \(2011\)](#), rather than a bound morpheme and a bound root, as [Liddell \(1996\)](#) argues. For our study, we look at those instances of numeral incorporation that are based on independent base signs because we see the necessity of considering the relation between the independent sign and the incorporated form in order to explore the diachronic development of the incorporated form. This leaves open the issue of how to account for those incorporated forms that have no independent base sign, like the ASL sign O'CLOCK ("o'clock"). In DGS, there are also several signs without fully specified independent forms, including \$NUM-CLOCK1A "o'clock" (see [Appendix C Example \(6\)](#)). Synchronically, there is a syntactic paradigm with a suppletive form consisting of \$NUM-CLOCK1A, plus numeral incorporation for the numerals 1 to 10. From 11 onwards, signs such as CLOCK1 are used, which form phrasal constructions with any numeral. We will come back to the topic of numeral incorporations without a base sign, and their implications for the process as a whole in the conclusion ([Section 6](#)).

On a synchronic level, we agree with [Mathur and Rathmann \(2011\)](#) that the incorporated form is created through the combination of a base sign with all four phonological parameters and a fully specified numeral, and that, in this morphological process, the handshape parameter of the base sign gets deleted and is replaced by the handshape of the numeral. We call the form with complete incorporation "affixation" in the sense of "adding a morpheme to a structure" that is nevertheless being realized simultaneously with the base sign. The numeral morpheme within such an incorporated construction may furthermore be reanalyzed as consisting of a handshape only (as in [Liddell 1996](#)), and we hypothesize that, in fact, a process of reanalysis occurred diachronically. If we look at the numeral system of DGS, we see a model for such a reanalysis in the complex numerals. We suggest that these were reanalyzed as a meaningful handshape morpheme affixed to a meaningful base morpheme consisting only of a centrally located, oriented movement, thus yielding numerals as 20, 30 to 90, or 100, 200, and up to 900, and so on, applying the strategy of combining a numeral handshape with a specific movement. Indeed, [Sagara and Zeshan \(2016, p. 29\)](#) call this strategy "numeral incorporation" because they put this simultaneous morphological process and numeral incorporation in lexical signs as described here in the same category (see also [Zeshan et al. 2013, p. 363](#)). We suggest that this view reflects cognitive processes (of reanalysis) of the language community at large. As a result of the reanalysis we propose, the numeral handshape morphemes "one of a X," "two of a X," and so on become available to be attached not only to elements in the numeral system to form complex numerals, but also to other lexical signs that denote typical quantifiable units. In other words, the numeral sign system becomes a model for a new set of constructions in which the numeral morpheme is affixed to other signs, such as temporal expressions.

3.4. Conducting the Study

At the preparative stage of our investigation, we deliberately excluded cases of attested numeral incorporation where no independent base sign can be found, as we were looking for signs that could occur in both incorporating and non-incorporating constructions. We also excluded all types that had fewer than five tokens with numeral incorporation in the corpus in order to avoid idiosyncrasies and other exceptional uses. A first query for signs meeting these two conditions gave us the following list of signs:

- WEEK1A, WEEK1B and WEEK1C (in the following subsumed and coded as WEEK1) "week."
- MONTH1 "month."
- YEAR1A and YEAR1B (subsumed and coded as YEAR1) "year."
- DAY2 "day."
- HOUR2A, HOUR2B, and HOUR2C (subsumed and coded as HOUR2) "hour."
- EURO1 "euro" (currency).¹⁶
- OLD8B "age."

Following our idea that numeral morphemes might have been abstracted from complex numerals, we included two incorporating signs from the numeral system, namely the hundreds and the thousands, as an assumed model for reanalysis and to check if they are used in the corpus as expected:

- \$NUM-HUNDREDS1 “hundreds.”
- \$NUM-THOUSANDS1 “thousands.”

These last two form a category we call “complex numerals,” while the longer list above is split into two semantic categories: the “temporals” (the first five items) and the remaining two signs that we group under “miscellaneous.”

The signs listed so far are the lexical hosts of our three constructions. In order to assess the frequency of complete numeral incorporation, we also need to look at the competing constructions: the phrasal construction and cliticization. As numeral incorporation occurs only with numerals ranging from 1 to 10 in DGS, we are only looking for occurrences with quantifying values in that range, be they quantified by a numeral sign or a numeral hand-shape. Our categorization for the annotation values “phrasal construction,” “cliticization,” and “affixation” is thus as follows:

1. Phrasal construction: a numeral from 1 to 10 followed by a lexical sign of the above list without the qualifier q.
2. Cliticization: a numeral from 1 to 10 followed by a lexical sign of the above list with qualifier q (specified for the same number value).
3. Affixation: a lexical sign of the above list with qualifier q and without a preceding numeral from 1 to 10.

You can see the corresponding constructions in glosses in Table 1.

Table 1. Annotation of numerals plus lexical sign for three construction types.

Category		Annotation Scheme and Example ‘Three Weeks’
Phrasal Construction	Scheme	\$NUM-ONE-TO-TEN’q:n BASE-SIGN
	Example	\$NUM-ONE-TO-TEN’q:3d WEEK1
Cliticization	Scheme	\$NUM-ONE-TO-TEN’q:n BASE-SIGN’q:n
	Example	\$NUM-ONE-TO-TEN’q:3d WEEK1’q:3d
Affixation	Scheme	BASE-SIGN’q:n
	Example	WEEK1’q:3d

To find the sign tokens relevant for our study, we ran a query that selected all observations fitting the criteria described above and automatically annotated it according to our categorization. Furthermore, the query matched each observation with metadata on the participant producing the construction, e.g., the age group of the participant and their region. The result was 2992 observations of the signs listed above occurring in one of the three relevant constructions.

3.5. Hypotheses

The main prediction for our study is that numeral incorporation is the result of a language change process, specifically a grammaticalization process. The property of language change that we will be testing on the basis of corpus data is the progressive change in language usage throughout time. Following the apparent time hypothesis (see also Section 2.3; Bailey et al. 1991), we will test this by looking at the different age groups represented in the corpus and comparing their language use. Furthermore, we use the cline

of phrasal construction, cliticization, and affixation as a basis of looking at the progressive change in form and will be comparing the usage of each construction by age group.

The idea of this perspective on language change is that the newly developing form, in this case, the incorporated form, is used more by younger people than by older people (see Hypothesis 1).

Hypothesis 1. *The older participants are, the less likely they are to use incorporated constructions (clitics and affixes).*

We hypothesize that complex numerals might be the original source of the numeral handshake as a unit of meaning. While we cannot test this directly without diachronic data, this presumption implies that numerals exclusively or almost exclusively occur in incorporated constructions (see Hypothesis 2).

Hypothesis 2. *Complex numerals have a (very) strong tendency to be used in incorporated constructions.*

Temporals are the signs that are the most well-known for their tendency to incorporate numeral handshapes. We hypothesize them to also show a tendency to occur with incorporations, but we expect this tendency to be weaker than in complex numerals (see Hypothesis 3).

Hypothesis 3. *Temporals have a fairly strong tendency to be used in incorporated constructions.*

The signs that we summarize in the “miscellaneous” group are signs that we think are just starting to show incorporations. Thus, we expect them to still prefer use in phrasal constructions (perhaps even strongly prefer) (see Hypothesis 4).

Hypothesis 4. *Miscellaneous signs show some incorporations but overall tend to be used in phrasal constructions.*

In summary, according to sign group, we expect a decreasing amount of incorporation in this order: complex numerals; temporals; and miscellaneous. With regard to the age group, we expect more incorporations the younger people are.

3.6. Analysis

The analysis of the data was completed in R Studio using R version 4.2.1, the tidyverse package (Wickham et al. 2019), the lme4 package (Bates et al. 2015), and the car package (Fox and Weisberg 2019). A generalized linear mixed model was fitted to the data using the glmer function. The outcome variable was the categorization of the construction, a categorical variable with three levels: phrase, clitic, and affix. In order to fit the model, the variable was dummy coded using the levels 0, 1, and 2 (respectively corresponding to phrase, clitic, and affix). The outcome is discrete count data with three different levels, calling for a binomial model.

The model included two predictors: sign group and age group. As described above, the signs in question were grouped into three categories based on their semantics: complex numerals, temporals, and miscellaneous. This grouping is a categorical variable with three different levels. The second predictor is the age of the participant. As age is represented in the corpus in terms of age group, this is likewise a categorical variable with four levels: 18–30, 31–45, 46–60, and 61+.

Two random effects were added to the model: sign gloss and participant. Sign gloss was added in case some signs show particularly strong tendencies towards (non-)incorporation and participants were added in case some people have particularly strong preferences for (non-)incorporation. Sign gloss is a categorical variable with nine different levels, consist-

ing of the nine incorporating signs we selected for this study. Participant is a categorical variable with 305 levels, consisting of the participant IDs of all participants in the dataset.

The resulting model formula is **Categorization ~ 1 + Sign Group + Age Group + (1 | Sign Gloss) + (1 | Participant)**.

Originally, region was also supposed to be included in the model as a random effect. Unfortunately, there were convergence issues when more than two random effects were included. Instead, we ran a model with region as a fixed effect to test whether regionality has an effect on use of incorporation (Categorization ~ 1 + Region + (1 | Sign Gloss) + (1 | Sign Group)). The results show non-significant slope estimates for all regions except for one: Leipzig has a little more incorporation than the other regions with an estimate of 1.1862 and a *p*-value of 0.0488. A post-hoc ANOVA on the model shows that region is not a significant factor overall (*p* = 0.1451). Thus, region does not seem to have a systematic influence on the use of incorporated forms and may be excluded.

4. Results

The data collection of occurrences of the nine selected signs resulted in a total of 2992 observations. For an overview of the absolute sign frequencies per age group, see Table A2 in Appendix E. In this section, the results of the fitted generalized linear mixed model will be shown.

With regard to sign group, the model came up with the estimates shown in Table 2.

Table 2. Results showing the influence of the sign’s group.

	β	Standard Deviation	z-Value	p-Value
Intercept	7.8104	1.2794	6.105	<i>p</i> < 0.001 *** ¹⁷
Temporals	−1.8503	1.2491	−1.481	0.13853
Miscellaneous	−8.4218	1.5581	−5.405	<i>p</i> < 0.001 ***

The intercept shows the level of incorporation for signs in the complex numerals group used by 18–30 year olds. As expected, this is a rather high value, and the slope estimates for temporals and signs in the miscellaneous group show a decreasing tendency regarding incorporations. For the latter group in particular, this tendency is very strong. This indicates that the most incorporations are produced by 18–30 year olds in complex numerals. Temporals show slightly less incorporation than complex numerals as shown by the negative estimate, but the difference is not strong enough to be significant. The two signs in the miscellaneous groups show a lot less incorporation than complex numerals, and this difference is very highly significant.

The estimates for the predictor age group are shown in Table 3.

Table 3. Results showing the influence of the participants’ age group.

	β	Standard Deviation	z-Value	p-Value
Intercept	7.8104	1.2794	6.105	<i>p</i> < 0.001 ***
31–45	−1.4134	0.6745	−2.095	0.03613 *
46–60	−1.8799	0.6819	−2.757	0.00584 **
61+	−2.9070	0.6982	−4.163	<i>p</i> < 0.001 ***

As expected, all slope estimates are negative, meaning that the reference group of 18–30 year olds is the most likely to use incorporated forms. Those who are 31–45 and 46–60 years old are significantly less likely to use incorporated forms than 18–30 year olds. Participants in the 61+ age group are significantly less likely to use incorporated forms than 18–30 year olds.

A post-hoc ANOVA confirms that both factors have significant effects on the use of numeral incorporation overall (see Table 4).

Table 4. Results of a post-hoc ANOVA confirming the significance of sign group and age group.

	χ^2	Degrees of Freedom	Pr (>Chisq)
Sign Group	33.845	2	$p < 0.001$ ***
Age Group	19.135	3	$p < 0.001$ ***

The estimated marginal means for incorporation according to the model can be seen in Figure 5. These allow for a clear comparison between the different categories.

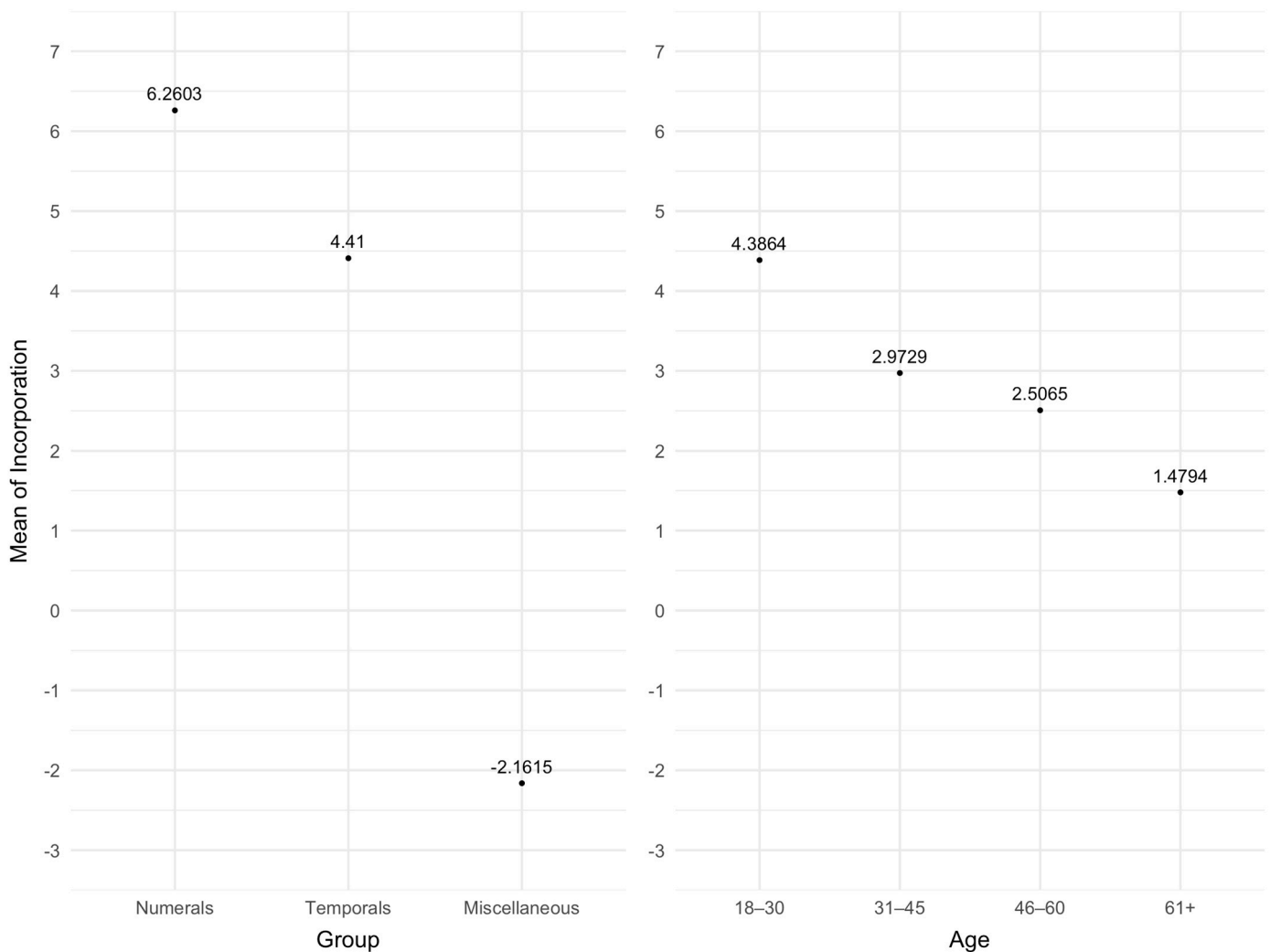


Figure 5. Means of Incorporation as calculated by the linear regression model.

The means show that the sign group has a strong effect on the use of incorporated forms with complex numerals having the highest mean (6.2603), and miscellaneous signs having by far the lowest mean (−2.1615). This overall pattern is expected, though we anticipated the difference between complex numerals and temporals to be bigger, as we suggested complex numerals as a potential origin of numeral incorporation. The miscellaneous signs match our expectations as possible future host class expansions that are just starting to be used. Concerning the age of participants, we see the expected pattern of younger people using more numeral incorporations (4.3864) than older people (1.4794).

The two middle groups fit the overall pattern, with the mean of the 46–60 group being slightly lower (2.5065) than that of the 31–45 group (2.9729). However, the means of these two groups are very close to each other, so it seems that they use numeral incorporation in similar amounts.

The relative frequencies of the different constructions per age group and per sign group can be seen in Figure 6.

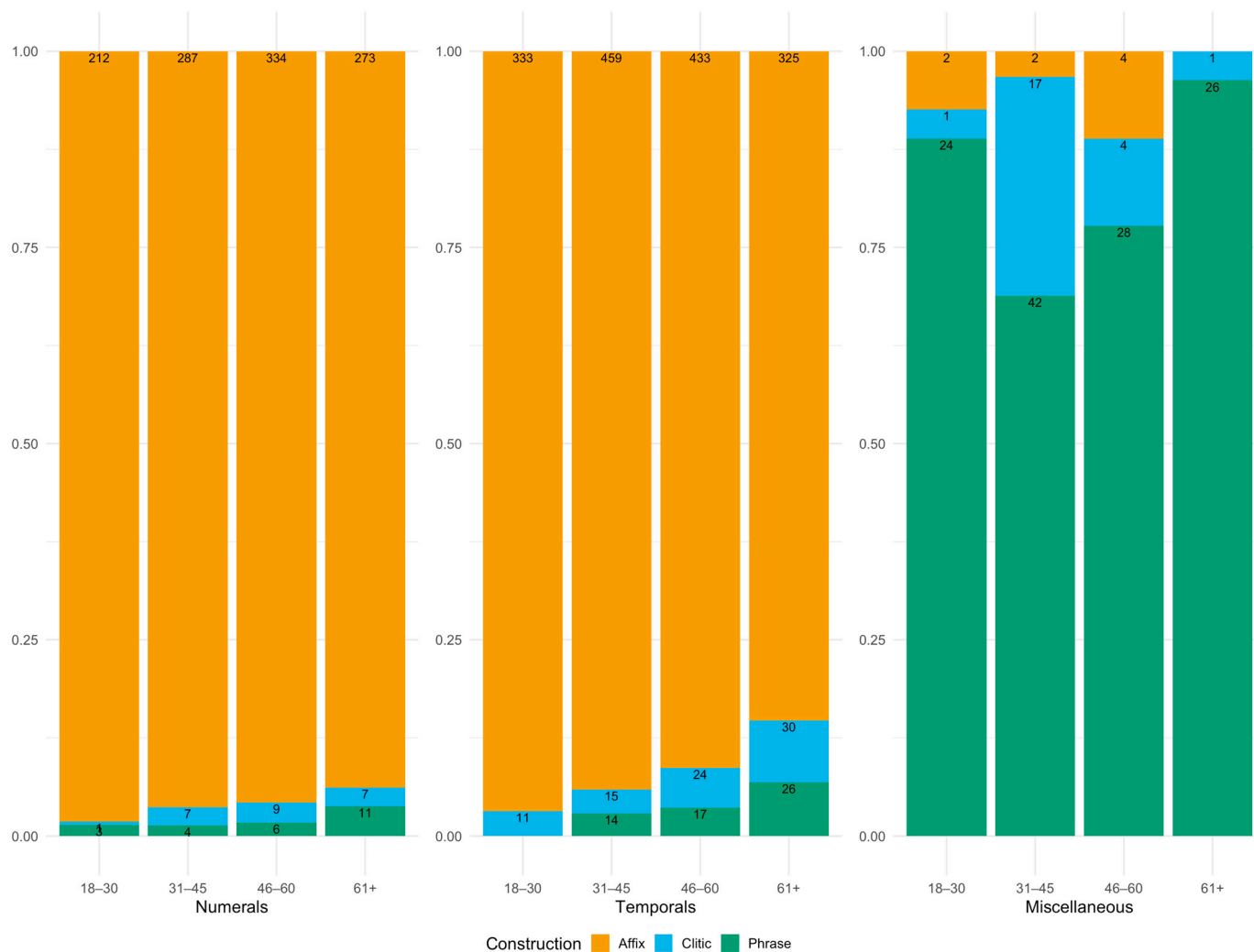


Figure 6. Construction frequencies per age group and sign group.

The visual difference between the first two groups and the third group is striking: While complex numerals and temporals are largely incorporated and only contain some phrasal constructions, the two signs in the miscellaneous group are used in phrasal constructions much more than in incorporated forms, though uses of affixes and clitics still occur.

Regarding the age groups, complex numerals and temporals again show very similar patterns: the youngest age group shows the strongest preference for incorporated forms with barely any phrasal constructions. As the participants get older, the preference for affixed constructions grows weaker (though all groups show that same preference for affixes), and the use of both clitic and phrasal constructions gradually increases. In the miscellaneous group, all age groups show a preference for phrasal constructions. This preference is the weakest in the age group including 31–45 year olds at just under 70% phrasal constructions. The oldest age group shows the strongest preference for phrases with only a single occurrence of an incorporated form, which is a cliticization. As the

youngest age group shows the second-strongest preference for phrasal constructions, there is no consistent pattern regarding the usage by different age groups in the miscellaneous group.

It is possible that individual signs show singular usage patterns. The usage of each sign per age group is visualized in Figure 7.

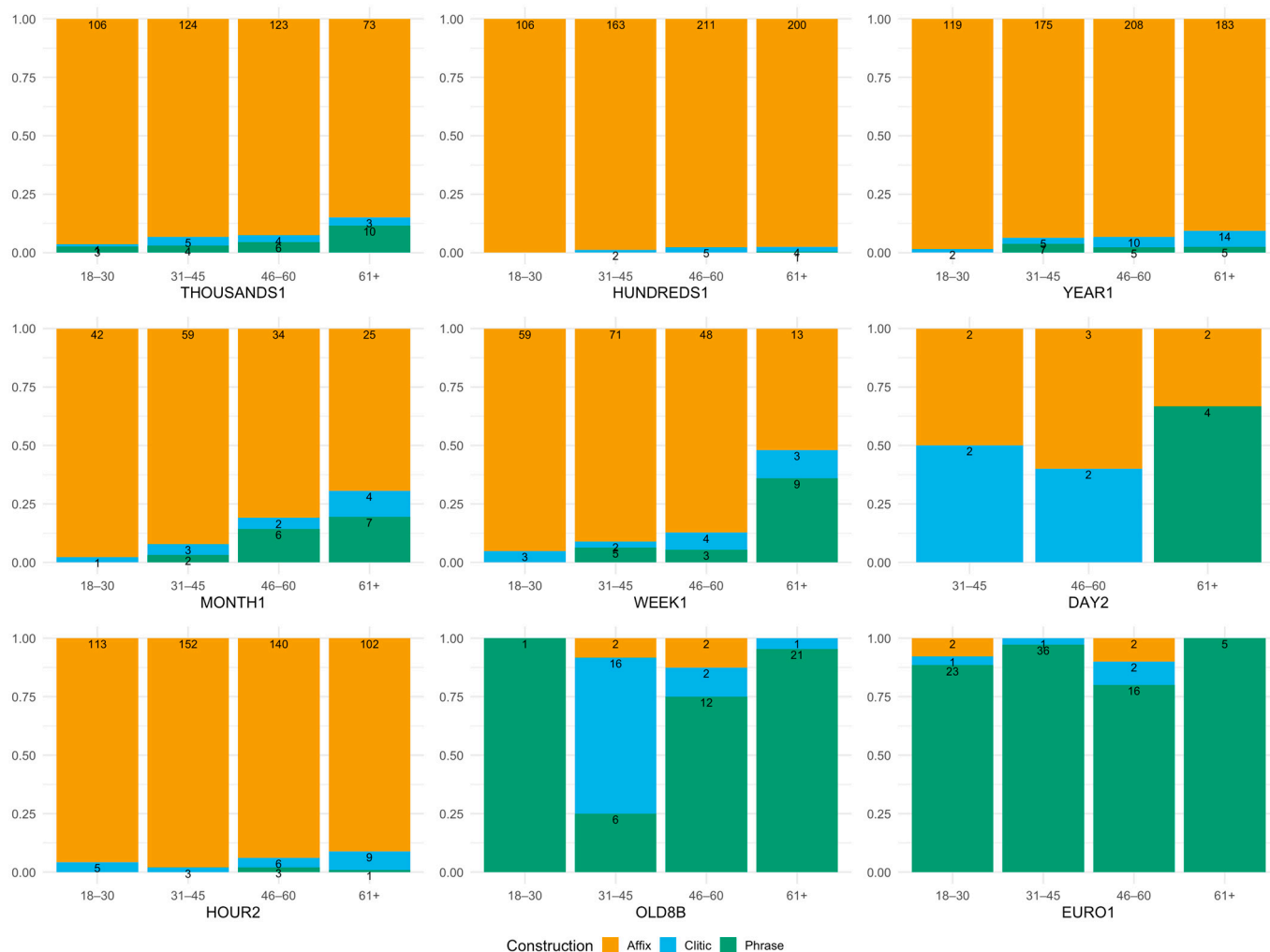


Figure 7. Construction frequencies usage per sign.

The two complex numerals, \$NUM-THOUSANDS1 and \$NUM-HUNDREDS1, behave rather differently from each other, with THOUSANDS containing quite a few phrasal constructions and HUNDREDS containing only one in the oldest age group. All temporals, except DAY2 (that is YEAR1, MONTH1, WEEK1, and HOUR2), show the expected general pattern of decreasing use of incorporations with increasing age. DAY2 does not follow this pattern beyond only having phrasal constructions in the oldest age group. It should be noted that this sign has by far the fewest occurrences overall at only 15 total tokens (for comparison: the sign with the second-fewest tokens is OLD8B at 63 tokens). The signs grouped as miscellaneous, OLD8B and EURO1, both show strong preferences for phrasal constructions in the oldest age group. Additionally, OLD8B contains exclusively phrasal constructions in the youngest age group, though this intersection consists of only one token. For EURO1, the age group of 31–45 year olds also shows a strong preference for phrasal constructions, while the youngest age group and the 46–60 age group both contain some, though few, incorporated tokens. OLD8B, when considered without the one token of the youngest age group, actually shows the same pattern as the temporals: a decreasing use of incorporations with increasing age though the temporals have a general preference towards

affixes, while OLD8B has a general preference towards phrases. Nonetheless, the overall pattern and the high number of clitics in the 31–45 age group, in particular, is consistent with the expected usage pattern.

5. Discussion

5.1. Hypotheses Evaluation

The results showed how participants of different ages use three categories of signs with regards to numeral incorporation, including three levels of incorporation: phrasal (not incorporated), clitic (partially incorporated), and affix (completely incorporated). Keep in mind that our dataset only contains quantified uses from 1–10 of the signs in question, so all of our frequency observations are made upon this basis. The results of our study will now be evaluated with regards to the hypotheses, repeated below for convenience.

Hypothesis 1. *The older participants are, the less likely they are to use incorporated constructions (clitics and affixes).*

Hypothesis 2. *Complex numerals have a (very) strong tendency to be used in incorporated constructions.*

Hypothesis 3. *Temporals have a fairly strong tendency to be used in incorporated constructions.*

Hypothesis 4. *Miscellaneous signs show some incorporations but overall tend to be used in phrasal constructions.*

The correlation between participants’ age and use of incorporations, which is described in the first hypothesis, was confirmed by the data. The oldest and the youngest age groups do indeed bookend the progression with the fewest and the most uses of incorporations, respectively. The two middle groups, likewise, fit the expected pattern with the group of 31–45 year olds using more incorporations than the group of 46–60 year olds. The graduality of this change is underlined by the fact that none of the differences between the adjacent age groups are significant (see Table 5), though the overall effect of age is highly significant (see Table 4 in Section 4). According to the apparent time approach, this correlation between age and usage frequency indicates language change. In our case, the change is from the usage of only phrasal constructions over cliticizations towards the use of affixes. Based on the use of phrasal constructions by older people, it seems that this construction was more common in an early stage of DGS, or it may have been the only possible construction to express this meaning. The gradual change throughout the middle age groups indicates that DGS progressed to include increasingly more incorporations over time. The youngest age group uses exclusively incorporated forms for temporals and very few phrasal constructions for complex numerals, indicating that in the most contemporary forms of DGS, the incorporations may be the only form remaining in use.

Table 5. Results of a paired post-hoc test on age group.

	β	Standard Error	Degrees of Freedom	z.Ratio	p.Value
(18–30)–(31–45)	1.413	0.675	Inf	2.095	0.1545
(31–45)–(46–60)	0.466	0.500	Inf	0.933	0.7873
(46–60)–61+	1.027	0.497	Inf	2.066	0.1644

The complex numerals show the same age progression, with older people using the most phrasal constructions and young people using the least. With a total of 24 phrasal constructions in the complex numerals group (0.0208%), we do indeed observe the tendency towards incorporated forms as predicted by the second hypothesis. However, the tendency

is not as strong as expected. In Section 3.3, we suggested that the complex numerals may be the origin of numeral incorporation and the grounds of a reanalysis of the numeral handshape as an affix. If this were true, we should be observing almost exclusive usage of incorporated forms, or at least a significantly lower percentage of phrasal constructions than in the temporals group. With 0.0208% phrasal constructions in the complex numerals and 0.0338% phrasal construction in the temporals, this is not the case. The model estimates presented in Table 2 in Section 4 confirm that the incorporation usage in complex numerals does not significantly differ from the usage in temporals. Thus, while we can technically confirm our hypothesis of strong tendencies towards incorporation in complex numerals, we do not come to the expected conclusion of complex numerals as the origin for numeral incorporation.

Another interesting observation when looking at the complex numerals is the difference between the individual signs. Almost all of the phrasal constructions in the complex numerals occur with \$NUM-THOUSANDS1, \$NUM-HUNDREDS1 is only once used in a phrasal construction. A closer look at this singular token reveals that this is actually not exactly the construction we were looking for. The intended meaning of the construction is the year 1940, which would usually be signed \$NUM-TEEN1'q:9 \$NUM-HUNDREDS1 \$NUM-TENS1'q:4 "nineteen-hundred-forty," which would not have appeared in our data due to the lack of a \$NUM-ONE-TO-TEN gloss. However, in the given observation, the signer signs \$NUM-ONE-TO-TEN1A'q:9 \$NUM-HUNDREDS1 \$NUM-TENS1'q:4 "nine-hundred-forty." It is questionable whether this observation should have been part of our data. However, since we did not manually check all of the 2992 observations compiled in the automatic annotation-based data selection process, such deviations in the data could not be detected and, therefore, not be excluded. Disregarding this singular phrasal usage with \$NUM-HUNDREDS1 makes the difference between the two complex numerals even more drastic, as it is then only \$NUM-THOUSANDS1 that is used in phrasal constructions at all. We cannot explain this difference on the basis of our data. On a phonological level, the two are so similar that this seems unlikely to be the cause. One possibility is that \$NUM-HUNDREDS1 is used more frequently overall (see Table A2 in Appendix E) with 692 total appearances in our dataset versus 462 for \$NUM-THOUSANDS1. If numeral incorporation is still emerging in complex numerals as well, it is plausible that it would first be established in the more frequent numerals and then spread to the less frequent ones. This, again, speaks against complex numerals as the origin of numeral incorporation.

The group of temporals, which is our main testing ground for incorporation usage, shows the expected pattern of younger people using more incorporated forms than older people. However, the overall tendency of temporals towards affixation in particular is stronger than we would have expected with 1550 affixations, 80 cliticizations, and 57 phrasal constructions in total. Even in the oldest age group, 93% of temporals have an incorporated numeral handshape (85% affixations). This indicates that the emergence process has likely started long before the oldest generation represented in our corpus and has progressed quite far. This is also supported by the fact that the younger generation not only does not prefer the phrasal construction, but in fact does not use it at all, as all 344 tokens show some kind of incorporation. The temporals' pattern also supports our claim of the cliticization construction as an in-between stage in the emergence of the numeral affix, as the use of the clitic, just like that of the phrasal construction, is the strongest in the oldest age group and the weakest in the youngest. Overall, the pattern shown here is in line with our hypothesis of numeral handshapes emerging through a language change process.

The last group of signs, termed the "miscellaneous" group, shows no clear overall pattern. However, with 21% total incorporating tokens, we can say that the signs certainly can be used in incorporating constructions. Yet the preference for phrasal constructions is very clear in all age groups. A look at the individual signs shows that EURO1 has only a few incorporations. None of the incorporations are in the oldest age group, but there is no overall pattern visible. OLD8B, in contrast, is used in a pattern more fitting to our expectations. Disregarding the youngest age group, for which there is only one token of

this sign, the expected pattern of more incorporations in the second youngest age group and the lowest number of incorporations in the oldest age group can be found once more. However, a closer examination of these tokens reveals that 13 of the 23 incorporations were produced by the same participant, and that 18 of them are from the same region (Cologne). The latter is less surprising as OLD8B is a regional sign used in the west of Germany, but the fact that over half of the tokens are from the same participant does skew our data. We will thus refrain from drawing conclusions beyond the possibility of incorporations for both signs. Hypothesis 4 is thus supported, but our claims regarding usage in different age groups cannot be verified in this group.

Besides the results regarding these hypotheses, we also observed that the phonological constraints regarding the base sign in DGS described by [Mathur and Rathmann \(2011\)](#) hold in our data for the most part but are not as strict as previously described. While all of the temporals that we considered (YEAR1, MONTH1, WEEK1, DAY2, and HOUR2) use the 1-handshape (extended index) in their unmodified form, we were also able to observe incorporations in OLD8B and EURO1, neither of which has a 1-handshape. Their handshapes are still rather unmarked, but nonetheless, these uses may indicate a shift in the phonological constraints. The 1-handshape constraint still holds for the temporals, however. For WEEK1 and DAY2, we also looked at lexical variants that do not have the 1-handshape: WEEK2 (46 tokens) and DAY1A/DAY1B (157 tokens). Neither of these show incorporating tendencies.

In conclusion, the data have shown that age groups do in fact differ in the way they use numeral incorporation, and that the resulting pattern supports a language change hypothesis. More research is needed both in regard to the usage of numeral incorporation in temporal signs to see whether it will go on to progress in a way consistent with language change, and also with regard to our miscellaneous group with signs that may incorporate more in the future. That being said, our main hypothesis was not only that numeral incorporation is the result of language change, but specifically the result of grammaticalization. In the following, we will discuss to what extent the properties of numeral incorporation are consistent with a grammaticalization analysis.

5.2. Numeral Incorporation as Grammaticalization

The phenomenon known as grammaticalization is associated with various different steps the construction in question typically goes through. Here, we will look at the parameters suggested by [Lehmann \(1985\)](#) and [Heine and Kuteva \(2007\)](#) and discuss whether they apply to numeral incorporation. We will first list the parameters and then describe the properties of numeral incorporation with respect to these parameters. Keep in mind that some of the categories that Lehmann, Heine, and Kuteva use to describe their parameters are either not transferrable to sign languages or may not have been researched in sign languages yet, so that their status is unclear.

[Lehmann's \(1985\)](#) grammaticalization parameters:

1. Attrition: loss of semantic, phonological, and morphological properties.
2. Paradigmaticization: emergence of morphological paradigms.
3. Obligatorification: the new construction becomes obligatory, alternatives disappear.
4. Condensation: decrease in the element's scope.
5. Coalescence: increase in bondedness.
6. Fixation: loss of variation in positioning.

[Heine and Kuteva's \(2007\)](#) grammaticalization parameters:

1. Reinterpretation: item is reinterpreted in a given context.
2. Extension: item can be used in new contexts.
3. Desemanticization/semantic bleaching: meaning components that are not compatible with the new usage are lost.
4. Decategorialization: loss of morphosyntactic properties.
5. Erosion: Loss of phonetic substance.

With regards to **reinterpretation**, we originally suggested complex numerals as the origin of numeral incorporation with their handshape being reinterpreted as an affix rather than just a phonological parameter. As described in Section 5.1 above, our data does not support this hypothesis and we, consequently, cannot discern a reinterpretation process in the emergence of numeral incorporation. Nonetheless, we argue that an **extension** has taken place, or rather, is currently taking place as OLD8B and EURO1 are starting to be used in incorporated constructions. Furthermore, we suggested a cline for the emergence of numeral incorporation, starting with the phrasal construction featuring two fully specified signs: the cliticized construction in which the signs begin to merge, and, finally, the affixed construction in which the two signs have become one sign. Compared to the fully specified numerals in the phrasal constructions, the numeral handshapes in the incorporated constructions show **phonological attrition/erosion** (loss of three of their four phonological parameters) and **morphological attrition** (change from a fully specified independent sign and free morpheme to an affix and bound morpheme). The base sign likewise undergoes a phonological reduction as its handshape is deleted. Due to phonological constraints on incorporation, not all numeral signs can be reduced to numeral handshapes: morphologically complex numerals such as \$NUM-TENS2A'q:4 "forty" feature a movement that is integral to the meaning of the sign and can thus not be reduced to their handshape. As a consequence of these constraints, a **closed paradigm** of numeral handshapes, ranging from 1 to 10, emerges.

Due to the simultaneous nature of the incorporation, the bond (**coalescence**) between the numeral handshape and the base sign is very strong, as they become one sign. In Example (7) (see Appendix C), we observe a very interesting construction: HOUR2B'q:3d HALF6 "three-and-a-half hours." The numeral handshape for "three" is incorporated into the sign HOUR2B. The sign HALF6, which semantically belongs to "three," and would have immediately followed the numeral sign in a phrasal construction (as we can see in several other constructions), cannot come between the numeral handshape and the base sign here, confirming their bondedness at least on a phonological level. We have found at least four observations like this (all with a sign meaning "half"), produced by four different participants. Exploring the exact nature of this construction on a morphological level is beyond the scope of this paper, but the structural separation of "three" and "half" is quite interesting and would be a great topic of study (perhaps combined with an exploration of whether the incorporating base sign can still be modified, that is, whether constructions such as BEAUTIFUL1A WEEK1'q:3 "three beautiful weeks" are possible).

Parameters that do not seem to fit for numeral incorporation at all are the ones referring to a reduction in semantic content (**semantic attrition/desemanticization**). Numerals are fairly simple in their semantics to begin with, and there does not seem to be any meaningful difference between the phrasal \$NUM-ONE-TO-TEN1A'q:3 WEEK1 and the affixed WEEK1'q:3. As we did not include a semantic analysis in our study, we cannot rule out minor semantic differences, but as it stands now, we see no shift in meaning.

The parameter of **obligatorification** may not apply yet, as there is the alternative of using the phrasal construction. However, the fact that the participants in the youngest age group are using exclusively incorporated forms for temporals may be an indicator for an ongoing obligatorification process that will be completed in the future.

The parameters of **decategorialization** and **condensation** (reduced scope) are difficult to judge for numeral incorporation, as the morphosyntactic properties of signs have not been researched enough to describe them for the phrasal construction, nor for the affixed construction. Regarding the morphological properties of incorporated forms, we can see in our data that the incorporated constructions can be modified further to mark morphological categories such as regularity ("every three weeks"). However, as this category is specific to temporal expressions, the modification does not indicate a general consideration of the incorporated construction as a single noun with the same morphosyntactic properties as other nouns. A conclusion regarding a decategorialization of the construction cannot be made.

Lastly, the parameter of **fixation**, described by [Lehmann \(1985\)](#) as the loss of variation in positioning, cannot be applied to this construction due to the simultaneity of the affix and the root. The phrasal construction tends to occur in the order of the numeral, followed by the base sign. Since the affix is a piece of simultaneous morphology, its position to the root cannot be described as preceding or following and thus cannot match that of the numeral in the phrasal construction.

In conclusion, we believe that too many essential components of a grammaticalization process are missing to claim that numeral incorporation is the result of a grammaticalization process. In particular, the lack of semantic bleaching/desemanticization on the one hand, and the fact that the complex numerals considered here do not hold up as a possible base for reanalysis on the other hand, indicate that we are not dealing with a grammaticalization process.

5.3. Alternative Approaches

Having concluded that numeral incorporation is in an ongoing process of language change but does not seem to be an instance of grammaticalization, the question arises if there might be another known language change process that might be a better fit for the description of numeral incorporation. In this section, we want to discuss some alternatives in the light of our study's results. Approaches described in this section include constructionalization, lexicalization, chunking, and compounding. Some of these may be treated as parts of a larger grammaticalization process (e.g., chunking) or as a larger language change category, under which grammaticalization can be subsumed (e.g., constructionalization). However, as all of these processes can also occur without being classified as grammaticalization, we believe them to be worthy of discussion here as alternatives.

[Traugott and Trousdale \(2013\)](#) describe **constructionalization** as the creation of a new form–meaning pair with an arbitrary association between form and meaning (p. 1). In the case of two pre-existing elements forming a new construction, they clarify that the result exhibits changes in syntax, morphology, and meaning, e.g., the meaning of “cupboard” is non-compositional as it cannot fully be derived from its components “cup” and “board.” This is similar to the “semantic bleaching” parameter in grammaticalization processes and, as already described, this does not apply to numeral incorporation. Furthermore, [Traugott and Trousdale \(2013\)](#) discuss clippings and blends as lexical constructionalizations (p. 150). The forms resulting from clippings and blends do not look entirely dissimilar to numeral incorporations, as they likewise combine some of the phonological aspects of one element with some of the phonological aspects of another element, e.g., the word “brunch” combining the “br” from “breakfast” with the “unch” from “lunch.” However, as Traugott and Trousdale point out, clippings and blends do not arise gradually (p. 150). This stands in direct contradiction to the findings of our study, as numeral incorporation has developed gradually both in terms of the temporal development (as younger age groups use more incorporation than older age groups) and in terms of the form gradually emerging (shown by the phrase—clitic—affix cline). Thus, constructionalization cannot account for all of the aspects of (the emergence of) numeral incorporation.

Lexicalization, as it has been described for sign languages so far, likewise includes a shift in meaning ([Frishberg 1975](#); [Liddell and Johnson 1986](#)). It furthermore assumes that each newly lexicalized form is stored as a single unit in the lexicon ([Liddell and Johnson 1986](#)) and, thus, excludes productive word formations. However, as we have shown, numeral incorporation is certainly productive.

Chunking describes the phenomenon of “units or word strings that are often produced together (. . .) becom[ing] units or chunks of their own right” ([Bybee 2011](#), p. 2). The new chunks are then stored and processed together. The prerequisite of the units being often produced together certainly fits numeral incorporation: the signs that allow for incorporation are ones that are commonly quantified and are consequently often produced with a numeral. However, [Bybee \(2011, p. 3\)](#) also points out that the meaning of the chunks usually becomes non-compositional with time, and that the association with the

original components fades until they are not activated alongside the chunk anymore. This is certainly not the case for numeral incorporation, as the meaning is very much compositional, and the association with the original number and base sign is not only given but also unlikely to fade (given the compositionality of the meaning).

Lastly, we want to consider the possibility of numeral incorporation being a case of **compounding**, as Meir (2012) and Jones (2013) suggest. In this scenario, our description of the numeral handshape as an affix would be incorrect and the numeral handshape would instead be considered a modifier with the base lexical sign acting as the head of the compound. Meir (2012, p. 101) describes two types of compounds in sign languages, the second of which “combines certain phonological parameters from two different sources to create a single sign,” and names numeral incorporation as one of the instances of this type of compounding, though she acknowledges that this is a rather unusual analysis (Meir 2012, p. 101). However, when describing affixes in sign languages, which are usually simultaneous, she adds that compounds and affixes are difficult to distinguish in sign languages (Meir 2012, p. 103). Her two criteria for differentiation are (i) the number of bases that a given element can combine with, as affixes should be more productive and combine with more bases than compound elements, and (ii) allomorphy, which, according to Meir (2012, p. 103), is more common in affixes. Regarding the productivity of numeral handshapes, we are uncertain of a reliable judgement. Our dataset included nine different signs, but including incorporations without an independent base sign would result in many more signs being added to the list of incorporating signs (e.g., \$NUM-CLOCK1A, \$NUM-GRADE1, and \$NUM-FROM-TO1, just to name a few). Furthermore, we have shown through the inclusion of OLD8B and EURO1 in our study that numeral incorporation has progressively been attached to more and more bases. Furthermore, the fact that we are not looking at a single element but a full paradigm of numeral handshapes from 1 to 10 does make the form seem more productive overall. Still, the numeral handshapes can certainly not compete with Meir’s (2012, p. 103) example of the affix “-ness” which can attach to about 3058 English words. Regarding the matter of allomorphs in numeral incorporation, we have shown handshape variations in multiple numeral signs, such as “one,” being articulated with either the thumb or the index finger (see also the overview in Appendix B Figure A2). As all of these handshapes can be incorporated, the case for multiple allomorphs for one number affix can be made. Thus, we stick to our analysis of the incorporated handshape as an affix.

In conclusion, none of the analyses described here can truly capture the numeral incorporation phenomenon in all of its aspects. This may be due to the lack of in-depth understanding of language change in sign languages overall, or it may be an indicator that this is a new category of language change.

6. Conclusions

Our corpus-based study reveals that there is good reason to assume that numeral incorporation is the result of a language change process, though likely not a grammaticalization process. The synchronic co-existence of three different constructions (phrasal, cliticized, and affixed) that are used in varying amounts by different age groups shows that change is still ongoing. This language change process lacks some of the essential properties of grammaticalization, though none of the other categorizations discussed here seem to be a good fit, either.

The observed tendencies towards incorporation in temporals suggest that incorporation in quantified temporal signs may become obligatory in the future, especially given the use of exclusively incorporated forms in the youngest age group. In the complex numeral \$NUM-HUNDREDS1, we already find exclusive incorporation indicating obligatoriness of the numeral handshape, confirming that an obligatorification is possible. We do not find an obligatory use of incorporated forms in \$NUM-THOUSANDS1, however. The reason for this difference could lie in the difference in overall frequency of the two signs, though this matter deserves further investigation.

With regards to the incorporating signs without a base sign such as \$NUM-CLOCK1A (described in Section 3.3), it is possible that these represent an ending point to the emergence of numeral affixes. Given the obligatorification tendencies we can see in our data, it is possible that \$NUM-CLOCK1A originally did have a base sign that disappeared after the incorporated construction became obligatory. Presumably, this would only happen in signs that have lexical variants that are used in non-quantified contexts and in quantified contexts surpassing an amount of 10 (such as CLOCK1), as the remaining use in these contexts would otherwise prevent the disappearance of the base sign. Considering the base signs included in our study, these are thus unlikely to disappear, as they are also used in non-quantified contexts. Still, this interpretation is a possible explanation for incorporated forms without a base sign.

Numeral incorporation in temporals seems to depend on several factors, including phonological and semantic constraints. However, as is common with corpus data, we see a lot of variation in use. Strict phonological constraints do not predict individual signers building forms such as EURO1'q:3d, thus violating the principle of handshape markedness. Though these forms are unexpected, conversation partners understand them, and it is imaginable that, one day, given an increased frequency of these forms, a new member has entered the class of host signs for numeral incorporation, overriding any previous phonological restrictions.

In the future, tendencies observed here, such as the incorporations in non-temporal signs and the beginning obligatorification of the incorporating construction, deserve further observation. When more contemporary data become available in the future, it can be used for comparison to our dataset, thus allowing for a diachronic comparison of language use throughout time. Our corpus-based results could, furthermore, be supplemented by an experimental approach to the subject, e.g., through an acceptability judgement task asking participants of different age groups to judge incorporation and phrasal constructions.

In conclusion, we have found that numeral incorporation seems to be the result of a still ongoing process of language change, though this claim cannot be made with certainty without diachronic data. While we do not believe this process to be an instance of grammaticalization, operating within the framework of grammaticalization has allowed us to describe many aspects of the usage of numeral incorporation in interesting and useful ways. It is our hope that future research will be able to discover more about the exact nature of the emergence of numeral incorporation and diachronic language change in sign languages overall.

Author Contributions: Conceptualization, F.O., A.M., S.W., G.L.; methodology, F.O., A.M., S.W., G.L.; formal analysis, F.O.; resources, F.O., A.M.; SQL queries, G.L., A.M.; writing—original draft preparation, F.O., A.M.; writing—review and editing, F.O., A.M., S.W., G.L.; visualization, F.O., S.W. All authors have read and agreed to the published version of the manuscript.

Funding: This publication 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 German Academies of Sciences and Humanities.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Informed consent was obtained by the DGS-Korpus project from all subjects involved in the study, as well as the signing models shown in the pictures.

Data Availability Statement: Some of the data used in this study are publicly available at <https://doi.org/10.25592/dgs.corpus-3.0>. The full dataset can be made available upon request.

Acknowledgments: We would like to thank the DGS corpus participants and the DGS-Korpus annotation team for their contributions to the dataset we used for this study. Furthermore, we would like to thank Anne Wienholz for her advice on the quantitative data analysis. We thank the attendees of our presentation at the Grammar and Corpora conference 2022 in Ghent for their comments that

helped us in the further development of this study. We would like to thank our reviewers and editors for their valuable feedback and insights.

Conflicts of Interest: The authors declare no conflict of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript; or in the decision to publish the results.

Appendix A



Figure A1. Minimal pairs for the parameters' handshape, orientation, location, and movement.

Appendix B

 Handshape: 1d "one"	 Handshape: 1 "one"	 Handshape: 2d "two"	 Handshape: 2 "two"
 Handshape: 3d "three"	 Handshape: 4 "four"	 Handshape: 4d "four"	 Handshape: 5 "five"
 Handshape: 6d "six"	 Handshape: 7d "seven"	 Handshape: 8d "eight"	 Handshape: 9 "nine"
 handshape: 9d "nine"	 handshape: 10 "ten"		

Figure A2. Numeral handshapes from 1 to 10, including variants.

Appendix C. DGS Examples with Glosses, Mouthings/Mouth Gestures, Translations, and Sources

(1)	AND2	\$NUM-TEEN1:5	CHILD2*	\$NUM-ONE-TO-TEN1A:5	DEAF1A	\$NUM-ONE-TO-TEN1A:5	
		Fünfzehn	Kinder	Fünf	Taub	Fünf	
	HEAVY1B*	TO-HEAR2	\$NUM-ONE-TO-TEN1A:5	CI1*	\$GEST-OFF1^		
		schwerhörig	fünf				
	There were 15 kids: 5 were deaf, 5 were hard of hearing, and 5 had a CI.						
	Public Corpus transcript:	https://doi.org/10.25592/dgs.corpus-3.0-text-1245390			Timecode:	00:10:25:10–00:10:32:31	

(2)	MY1	FREE1	TIME5A	TO-BELIEVE2B	\$NUM-ONE-TO-TEN1A:6d	\$NUM-TENS1:2d	MEMBER4*
		Freizeit	Glaube	Sechszwanzig	Mitglied		
	I think the leisure [bowling] club has about 26 members.						
	Public Corpus transcript:	https://doi.org/10.25592/dgs.corpus-3.0-text-1583950			Timecode:	00:09:36:40–00:09:41:07	

(3)	\$INDEX1*	ALSO1A	\$INDEX1	TO-TRY-OR-REHEARSAL2*	\$NUM-ONE-TO-TEN1A:6d	WEEK1A
			Aber	Probe	Sechs	Woche
	There was a probation period of six weeks.					
Public Corpus transcript:	https://doi.org/10.25592/dgs.corpus-3.0-text-1429964			Timecode:	00:10:22:37–00:10:24:34	

(4)	I1	TO-THINK1B	OLD8B	TO-THINK1B	BEGINNING1A*	TO-THINK1B
				alt		anfang
	\$NUM-ONE-TO-TEN1A:3d	OLD8B*	I1	TO-KNOW-OR-KNOWLEDGE2B^	NOTHING1B*	
	drei	alt	[MG]			
	I think I was three years old, but my memory doesn't go that far back.					
Public Corpus transcript:	https://doi.org/10.25592/dgs.corpus-3.0-text-1428225			Timecode:	00:00:01:49–00:00:07:16	

(5)	MY1	BOYFRIEND-GIRLFRIEND1	BEEN1	MONTH1*
	Mein	Freund	Gewesen	Zwei monat
	AUSTRALIA1*	\$INDEX1	VACATION8B	BEEN1
	Australien	Urlaub	Gewesen	
	My boyfriend spent two months on vacation in Australia.			
	Public Corpus transcript:	https://doi.org/10.25592/dgs.corpus-3.0-text-1289910		Timecode:

(6)	TO-COME1*	\$NUM-CLOCK1A:5	UNTIL1	EVENING2*	\$NUM-TEEN2B:1d
	mama	fünf uhr	bis	abend	elf uhr
	She [mom] came over around five in the afternoon and stayed until 11:00 at night.				
	Public Corpus transcript:	https://doi.org/10.25592/dgs.corpus-3.0-text-1176846		Timecode:	00:08:39:16–00:08:24:42

(7)	FREIBURG1*	LOCATION1B^	FAR1*	FROM-TO3*	HOUR2B*	HALF6	FAR1*
	freiburg		[MG]		drei stunden halb		wei{t}
	Yet it [Freiburg] is also quite far away. It took us three [and a half] hours to get there.						
	Public Corpus transcript:	https://doi.org/10.25592/dgs.corpus-3.0-text-1184756			Timecode:	00:11:28:48–00:11:33:08	

Appendix D

Table A1. List of type glosses and type entries in the Public DGS Corpus.

Glossname	DOI of Superordinate Type
WEEK1A	https://doi.org/10.25592/dgs.corpus-3.0-type-14967
WEEK1B	https://doi.org/10.25592/dgs.corpus-3.0-type-48443
WEEK1C	https://doi.org/10.25592/dgs.corpus-3.0-type-75108
WEEK2	https://doi.org/10.25592/dgs.corpus-3.0-type-16795
YEAR1A	https://doi.org/10.25592/dgs.corpus-3.0-type-15127
YEAR1B	https://doi.org/10.25592/dgs.corpus-3.0-type-15554
\$NUM-CLOCK1A	https://doi.org/10.25592/dgs.corpus-2.0-type-72542
CLOCK1	https://doi.org/10.25592/dgs.corpus-3.0-type-16979
\$NUM-ONE-TO-TEN1A	https://doi.org/10.25592/dgs.corpus-1.0-type-49944
\$NUM-ONE-TO-TEN1B	https://doi.org/10.25592/dgs.corpus-1.0-type-50050
\$NUM-ONE-TO-TEN1C	https://doi.org/10.25592/dgs.corpus-1.0-type-50052
\$NUM-ONE-TO-TEN1D	https://doi.org/10.25592/dgs.corpus-1.0-type-50054
\$NUM-TEEN1	https://doi.org/10.25592/dgs.corpus-1.0-type-49945
\$NUM-TENS1	https://doi.org/10.25592/dgs.corpus-1.0-type-49948
MONTH1	https://doi.org/10.25592/dgs.corpus-3.0-type-15420
DAY1A	https://doi.org/10.25592/dgs.corpus-3.0-type-14038
DAY1B	https://doi.org/10.25592/dgs.corpus-3.0-type-13269
DAY2	https://doi.org/10.25592/dgs.corpus-3.0-type-18466
HOUR2A	https://doi.org/10.25592/dgs.corpus-3.0-type-85194
HOUR2B	https://doi.org/10.25592/dgs.corpus-3.0-type-54108
HOUR2C	https://doi.org/10.25592/dgs.corpus-3.0-type-85822
EURO1	https://doi.org/10.25592/dgs.corpus-3.0-type-17418

Table A1. *Cont.*

Glossname	DOI of Superordinate Type
OLD5A	https://doi.org/10.25592/dgs.corpus-3.0-type-14440
OLD8B	https://doi.org/10.25592/dgs.corpus-3.0-type-14503
\$NUM–HUNDREDS1	https://doi.org/10.25592/dgs.corpus-3.0-type-49953
\$NUM–THOUSANDS1	https://doi.org/10.25592/dgs.corpus-1.0-type-49954
BEAUTIFUL1A	https://doi.org/10.25592/dgs.corpus-3.0-type-15596
PLEASE1B	https://doi.org/10.25592/dgs.corpus-3.0-type-78494
TO–COMPREHEND1	https://doi.org/10.25592/dgs.corpus-3.0-type-15301
TO–BE–SILENT1A	https://doi.org/10.25592/dgs.corpus-3.0-type-15047
TO–SAW1	https://doi.org/10.25592/dgs.corpus-3.0-type-13139
CLASS1	https://doi.org/10.25592/dgs.corpus-3.0-type-15218
MILK2C	https://doi.org/10.25592/dgs.corpus-3.0-type-17832
DIFFICULT1	https://doi.org/10.25592/dgs.corpus-3.0-type-13152
GERMAN–MARK	https://doi.org/10.25592/dgs.corpus-3.0-type-17377
\$NUM–GERMAN–MARK1	https://doi.org/10.25592/dgs.corpus-1.0-type-17919

Appendix E

Table A2. Absolute sign frequencies per age group.

	18–30	31–45	46–60	61+	Totals
THOUSANDS1	110	133	133	86	462
HUNDREDS1	106	165	216	205	692
YEAR1	121	187	223	202	733
MONTH1	43	64	42	36	185
WEEK1	62	78	55	25	220
DAY2	0	4	5	6	15
HOUR2	118	155	149	112	534
OLD8B	1	24	16	22	63
EURO1	26	37	20	5	88
Totals	587	847	859	699	2992

Notes

- ¹ Depending on the sign, it is possible to form 11 to 19 with a productive movement. However, in general, 11 and 12 have a tendency to be expressed by lexicalized suppletive forms in DGS.
- ² Pfau and Steinbach (2021) claim that in DGS only the numerals 1 to 5 can be incorporated in lexical base signs due to the two-handedness of the numerals from 6 to 10. In our corpus data, the two-handed numerals between 6 to 10 are incorporated. The movement of the lexical base sign is produced by both hands in this case.
- ³ It is common in literature on sign languages to refer to specific sign types via labels in capital letters (called glosses). See Section 3.2. for more detail on glosses and the glossing conventions adhered to in the annotation of the DGS corpus. For this introduction, we use a simplified gloss annotation.
- ⁴ Numeral and base sign share the same location in this instance, and the origin of the location in the affixed form is thus opaque. However, analyses of other incorporations show that the location matches the base sign.
- ⁵ The handshake allomorphs for “one” in ASL as well as in DGS are unmarked. It is the selected thumb or index finger as in Figure 2b.
- ⁶ <https://doi.org/10.25592/dgs.corpus-3.0>.

- ⁷ See [Schulder and Hanke \(2022\)](#) for details on the FAIR and CARE policy within the project. Research on the total of the reference corpus requires an additional license agreement for researchers outside of the DGS-Korpus project. As all authors of this paper are researchers working in the project, we had unrestricted access to the data.
- ⁸ See [Konrad et al. \(2022, p. 7\)](#) for glossing conventions and the definition of the term gloss: “A gloss is a German word that corresponds—on the subtype level—to a core meaning (keyword) of the sign.”
- ⁹ Mouthing refers to the visible mouth movement of silently articulated German words. Mouth gestures are meaningful mouth movements not related to the articulation of a word and are annotated with the label [MG], subsuming all types of mouth gestures.
- ¹⁰ Deviant in this context is used as a label for tokens whose form in actual execution of the sign differs in some aspect to the form assigned to the type as citation form in iLex. It covers aspects of adaptations due to performance and signed context, such as assimilations, as well as aspects of variation and modification that have not yet been categorized in more detail. The annotation as “deviant” only indicates some difference to the citation form.
- ¹¹ HamNoSys, short for Hamburg Notation System for Sign Languages, is a notation system for sign languages ([Hanke 2004](#)).
- ¹² For more information on the prefixes used in the glosses, see [Konrad et al. \(2022, pp. 14–16\)](#).
- ¹³ In the DGS corpus, signs of the category \$NUM end by specifying the numeral handshape. This is accomplished by adding a qualifier that includes the numbers 1–10.
- ¹⁴ The qualifier 'numinc was also included in this study. It occurs only in addition to the qualifier 'q. As the differences between 'q and 'numinc + 'q are minor and irrelevant to our study, they will not be described here.
- ¹⁵ See [Loos and Konrad \(2022, p. 20\)](#) for a list of all qualifiers, their occurrences, functions, and codes as used in detailed corpus annotation.
- ¹⁶ The sign EURO1 is a relatively new sign as the Euro was introduced to Germany as a currency in 2002. The sign is produced with the handshape of extended index and middle finger, resembling the two lines of the € symbol. One of the signs used for the previous currency, D-Mark, has no fully specified base form and incorporates numerals from 1 to 10. It is possible that EURO1 started incorporating numerals due to this use of \$NUM-GERMAN-MARK1.
- ¹⁷ The asterisks are commonly used to indicate different levels of significance. They are used as follows: *** for $p < 0.001$, ** for $0.001 \leq p < 0.01$, and * for $0.01 \leq p < 0.05$.

References

- Bailey, Guy, Tom Wikle, Jan Tillery, and Lori Sand. 1991. The Apparent Time Construct. *Language Variation and Change* 3: 241–64. [[CrossRef](#)]
- Bates, Douglas, Martin Mächler, Ben Bolker, and Steve Walker. 2015. Fitting Linear Mixed-Effects Models Using Lme4. *Journal of Statistical Software* 67: 1–48. [[CrossRef](#)]
- Brentari, Diane. 2012. 3. Phonology. In *Sign Language*. Edited by Roland Pfau, Markus Steinbach and Bencie Woll. Berlin and Boston: De Gruyter Mouton, pp. 21–54. [[CrossRef](#)]
- Bybee, Joan L. 2006. From Usage to Grammar: The Mind’s Response to Repetition. *Language* 82: 711–33. [[CrossRef](#)]
- Bybee, Joan L. 2011. Chapter 6. Usage-based Theory and Grammaticalization. In *The Oxford Handbook of Grammaticalization*, Online ed. Edited by Bernd Heine and Heiko Narrog. Oxford: Oxford University Press, pp. 69–78. [[CrossRef](#)]
- Chinchor, Nancy. 1982. *Morphological Theory and Numeral Incorporation in American Sign Language*. Ann Arbor: Brown University.
- Dachkovsky, Svetlana. 2022. Emergence of a Subordinate Construction in a Sign Language: Intonation Ploughs the Field for Morphosyntax. *Glossa: A Journal of General Linguistics* 7: 1–41. [[CrossRef](#)]
- Fox, John, and Sanford Weisberg. 2019. *An R Companion to Applied Regression*, 3rd ed. Thousand Oaks: Sage.
- Frishberg, Nancy. 1975. Arbitrariness and Iconicity: Historical Change in American Sign Language. *Language* 51: 696. [[CrossRef](#)]
- Hanke, Thomas. 2004. HamNoSys—Representing Sign Language Data in Language Resources and Language Processing Contexts. Paper presented at 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, Lisbon, Portugal, May 30; Edited by Oliver Streiter and Chiara Vettori. Paris: ELRA, pp. 1–6. Available online: <https://www.sign-lang.uni-hamburg.de/lrec/pub/04001.pdf> (accessed on 12 May 2023).
- Hanke, Thomas, and Jakob Storz. 2008. iLex—A Database Tool for Integrating Sign Language Corpus Linguistics and Sign Language Lexicography. Paper presented at the LREC2008 3rd Workshop on the Representation and Processing of Sign Languages: Construction and Exploitation of Sign Language Corpora, Marrakech, Morocco, June 1; Edited by Onno Crasborn, Eleni Efthimiou, Thomas Hanke, Ernst D. Thoutenhoofd and Inge Zwitterlood. Marrakech: European Language Resources Association (ELRA), pp. 64–67. Available online: <https://www.sign-lang.uni-hamburg.de/lrec/pub/08011.pdf> (accessed on 13 February 2023).

- Hanke, Thomas, Marc Schulder, Reiner Konrad, and Elena Jahn. 2020. Extending the Public DGS Corpus in Size and Depth. Paper presented at the LREC2020 9th Workshop on the Representation and Processing of Sign Languages: Sign Language Resources in the Service of the Language Community, Technological Challenges and Application Perspectives, Marseille, France, May 16; Paris: European Language Resources Association (ELRA), pp. 75–82. Available online: <https://aclanthology.org/2020.signlang-1.12.pdf> (accessed on 13 February 2023).
- Hanke, Thomas, Reiner Konrad, and Gabriele Langer. 2023. Chapter 7. Exploring regional variation in the DGS Corpus. In *Advances in Sign Language Linguistics*. Edited by Ella Wehrmeyer. Amsterdam and Philadelphia: John Benjamins, pp. 192–218. [CrossRef]
- Hanke, Thomas, Reiner Konrad, Gabriele Langer, Anke Müller, and Sabrina Wähl. 2017. Detecting Regional and Age Variation in a Growing Corpus of DGS. Paper presented at the Workshop Corpus-Based Approaches to Sign Language Linguistics: Into the Second Decade, Birmingham, UK, July 24.
- Heine, Bernd, and Tania Kuteva. 2007. *The Genesis of Grammar: A Reconstruction*. Oxford and New York: Oxford University Press.
- Himmelman, Nikolaus P. 2004. Lexicalisation and Grammaticalization: Opposite or Orthogonal? In *What Makes Grammaticalization? A Look from Its Fringes and Its Components*. Edited by Walter Bisang, Nikolaus P. Himmelman and Björn Wiemer. Berlin and New York: Mouton de Gruyter, pp. 21–44. [CrossRef]
- Hopper, Paul J., and Elizabeth C. Traugott. 2003. *Grammaticalization*, 2nd ed. Cambridge: Cambridge University Press. [CrossRef]
- Isard, Amy, and Reiner Konrad. 2022. MY DGS—ANNIS: ANNIS and the Public DGS Corpus. Paper presented at the LREC2022 10th Workshop on the Representation and Processing of Sign Languages: Multilingual Sign Language Resources, Marseille, France, June 25; Edited by Eleni Efthimiou, Stavroula-Evita Fotinea, Thomas Hanke, Julie A. Hochgesang, Jette Kristoffersen, Johanna Mesch and Marc Schulder. Marseille: European Language Resources Association (ELRA), pp. 73–79. Available online: <https://www.sign-lang.uni-hamburg.de/lrec/pub/22034.pdf> (accessed on 13 February 2023).
- Janzen, Terry. 2012. 34. Lexicalization and Grammaticalization. In *Sign Language. An International Handbook*. Edited by Roland Pfau, Markus Steinbach and Bencie Woll. Berlin and Boston: De Gruyter Mouton, pp. 816–40. [CrossRef]
- Jones, Vanessa L. 2013. Numeral Incorporation in American Sign Language. In *Theses and Dissertations*. Grand Forks: The University of North Dakota, vol. 1551, Available online: <https://commons.und.edu/theses/1551/> (accessed on 11 May 2023).
- Konrad, Reiner, Thomas Hanke, Gabriele Langer, Susanne König, Lutz König, Rie Nishio, and Anja Regen. 2022. *Public DGS Corpus: Annotation Conventions/Öffentliches DGS-Korpus: Annotationskonventionen. Working Paper*. Version 4.1. Hamburg: DGS-Korpus-Projekt, June 14. [CrossRef]
- Ktejik, Mish. 2013. Numeral Incorporation in Japanese Sign Language. *Sign Language Studies* 13: 186–210. [CrossRef]
- Labov, William. 1963. The Social Motivation of a Sound Change. *WORD* 19: 273–309. [CrossRef]
- Lehmann, Christian. 1985. Grammaticalization: Synchronic Variation and Diachronic Change. *Lingua e Stile* 20: 303–18.
- Liddell, Scott K. 1996. Number Incorporating Roots & Non-Incorporating Prefixes in American Sign Languages. *Sign Language Studies* 92: 201–26.
- Liddell, Scott K., and Robert E. Johnson. 1986. American Sign Language Compound Formation Processes, Lexicalization, and Phonological Remnants. *Natural Language and Linguistic Theory* 4: 445–513. [CrossRef]
- Loos, Cornelia, and Reiner Konrad. 2022. *Detailed Annotation and Qualifiers. Working Paper*. Version 1. Hamburg: DGS-Korpus-Projekt, April 4. [CrossRef]
- Mathur, Gaurav, and Christian Rathmann. 2011. Two Types of Nonconcatenative Morphology in Signed Languages. In *Deaf around the World*, Online ed. Edited by Gaurav Mathur and Donna Jo Napoli. Oxford: Oxford University Press, pp. 54–82. [CrossRef]
- Meir, Irit. 2012. 5. Word Classes and Word Formation. In *Sign Language*. Edited by Roland Pfau, Markus Steinbach and Bencie Woll. Berlin and Boston: De Gruyter Mouton, pp. 77–111. [CrossRef]
- Pfau, Roland, and Markus Steinbach. 2011. Grammaticalization in Sign Languages. In *The Oxford Handbook of Grammaticalization*, Online ed. Edited by Bernd Heine and Heiko Narrog. Oxford: Oxford University Press, pp. 683–95. [CrossRef]
- Pfau, Roland, and Markus Steinbach. 2021. Number in Sign Languages. In *The Oxford Handbook of Grammatical Number*. Edited by Patricia Cabredo Hofherr and Jenny Doetjes. Oxford: Oxford University Press, pp. 644–60. [CrossRef]
- Quer, Josep, Carlo Cecchetto, Caterina Donati, Carlo Geraci, Meltem Keleşir, Roland Pfau, and Markus Steinbach, eds. 2017. Part 2: Phonology. In *SignGram Blueprint. A Guide to Sign Language Grammar Writing*. Berlin and Boston: De Gruyter Mouton, pp. 17–68. [CrossRef]
- Sagara, Keiko, and Ulrike Zeshan. 2016. Semantic Fields in Sign Languages—A Comparative Typological Study. In *Semantic Fields in Sign Languages: Colour, Kinship and Quantification*. Edited by Ulrike Zeshan and Keiko Sagara. Berlin and Boston: De Gruyter Mouton, pp. 3–38. [CrossRef]
- Schulder, Marc, and Thomas Hanke. 2022. How to Be FAIR When You CARE: The DGS Corpus as a Case Study of Open Science Resources for Minority Languages. Paper presented at the 13th International Conference on Language Resources and Evaluation (LREC 2022), Marseille, France, June 20–25; Edited by Nicoletta Calzolari, Frédéric B chet, Philippe Blache, Khalid Choukri, Christopher Cieri, Thierry Declerck, Sara Goggi, Hitoshi Isahara, Bente Maegaard, Joseph Mariani and et al. Paris: European Language Resources Association (ELRA), pp. 164–73. Available online: <http://www.lrec-conf.org/proceedings/lrec2022/pdf/2022.lrec-1.18> (accessed on 13 February 2023).
- Schulder, Marc, Dolly Blanck, Thomas Hanke, Ilona Hofmann, Sung-Eun Hong, Olga Jeziorski, Lutz K nig, Susanne K nig, Reiner Konrad, Gabriele Langer, and et al. 2021. *Data Statement for the Public DGS Corpus. Working Paper*. Version 2. Hamburg: DGS-Korpus-Projekt, November 11. [CrossRef]

- Semushina, Nina, and Rachel I. Mayberry. 2019. Numeral Incorporation in Russian Sign Language: Phonological Constraints on Simultaneous Morphology. *Sign Language Studies* 20: 83–131. [[CrossRef](#)] [[PubMed](#)]
- Taub, Sarah F. 2001. *Language from the Body: Iconicity and Metaphor in American Sign Language*. Cambridge and New York: Cambridge University Press.
- Traugott, Elizabeth Closs, and Graeme Trousdale. 2013. *Constructionalization and Constructional Changes*. Oxford: Oxford University Press. [[CrossRef](#)]
- von Mengden, Ferdinand. 2010. Chapter I. Linguistic Numeral Systems. In *Cardinal Numerals: Old English from a Cross-Linguistic Perspective*. Berlin and New York: De Gruyter Mouton, pp. 12–71. [[CrossRef](#)]
- Wickham, Hadley, Mara Averick, Jennifer Bryan, Winston Chang, Lucy McGowan, Romain François, Garrett Grolemond, Alex Hayes, Lionel Henry, Jim Hester, and et al. 2019. Welcome to the Tidyverse. *Journal of Open Source Software* 4: 1686. [[CrossRef](#)]
- Zeshan, Ulrike, Cesar Ernesto Escobedo Delgado, Hasan Dikyuva, Sibaji Panda, and Connie de Vos. 2013. Cardinal Numerals in Rural Sign Languages: Approaching Cross-Modal Typology. *Linguistic Typology* 17: 357–96. [[CrossRef](#)]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.