The Noun Phrase in British Sign Language

Heidi Elaine Proctor

University College London

PhD Linguistics

August 2025

Declaration

I, Heidi Proctor, confirm that the work presented in this thesis is my own. Where information has been derived from other sources, I confirm that this has been indicated in the thesis.

Abstract

The noun phrase (NP) has not been extensively researched in signed languages. This thesis presents the first systematic investigation of NP structure in British Sign Language (BSL), drawing on data from the BSL Corpus and an innovative large-scale online judgement task involving 92 deaf signers. It examines the linear order of NP elements–nouns, adjectives, numerals, and demonstratives–and evaluates their distribution in relation to typological predictions, including Greenberg's Universal 20, the subset of orders generated by Cinque's (2005) derivational model, and the principle of homomorphism, whereby linear order transparently reflects semantic scope. The corpus analysis shows that most BSL NPs are short, with 76% consisting of a single element, while longer NPs exhibit a wide range of element orders. However, both the corpus and the judgement task data show a strong overall preference for sequences that resemble English word order. In the corpus data, adjectives are predominantly prenominal, but judgement task participants preferred postnominal adjectives in short NPs, suggesting the influence of language ideologies that emphasise structural divergence from English. Homomorphic NP orders were judged significantly more typical, especially when modifiers were widely spaced in scope. The results also support the predictions of Universal 20 and Cinque's model. Statistical modelling found no significant effect of signers' age of BSL acquisition, teaching experience, or comfort with English on NP order preferences. The findings suggest a nuanced picture, with BSL NP structure being shaped by both cognitive pressures and contact with English, while also being mediated by community ideologies. This research contributes to our understanding of NP typology in signed languages and offers practical implications for BSL pedagogy. The study underscores the importance of aligning teaching materials with authentic usage to support accurate and inclusive BSL education.

Impact Statement

British Sign Language (BSL) is increasingly taught in formal educational settings. As BSL becomes more widely incorporated into school curricula and adult education programmes, it is essential that teaching programmes and interpreter training accurately reflect how the language is used by deaf people in the UK. Without this alignment, learners may acquire a distorted version of BSL that diverges from natural usage. To ensure that BSL education is both linguistically valid and pedagogically effective, instructional content must be grounded in systematic analysis of authentic language data.

This thesis contributes to that goal by providing the first systematic investigation of noun phrase (NP) structure in BSL, a core component of the language's grammar. We examined how BSL signers combine nouns with optional modifiers–adjectives, numerals, and demonstratives–using both a corpus of naturalistic signing and a large-scale judgement task. Our findings show that while BSL NP structure broadly aligns with English, there is a notable divergence in short NPs consisting of a noun and an adjective. In the judgement task, participants rated [Noun Adjective] order (e.g., BUS RED) as more typical than [Adjective Noun] order (e.g., RED BUS). However, for longer NPs with additional modifiers, [Adjective Noun] order was preferred. In contrast, our corpus data showed that signers used [Adjective Noun] order 84% of the time, even in short NPs.

This discrepancy may reflect a belief within parts of the deaf community that BSL should differ structurally from English. Our research demonstrates that while postnominal adjectives are viewed as more typical in short NPs, the dominant pattern in natural signing involves placing adjectives before the noun. This insight is important for educators, including those who train interpreters, as well as curriculum developers and assessment designers. Ensuring that teaching materials reflect actual usage will help learners–whether deaf or hearing–develop a more accurate understanding of BSL grammar.

These findings also have broader implications for how BSL is conceptualised and taught. Structural similarities with English are sometimes downplayed in favour of emphasising difference, yet such overlap does not compromise BSL's status as a fully-fledged language. This research shows that BSL NP structure is shaped by multiple factors: cognitive pressures that influence how signers structure NPs, contact with English that encourages English-like sequences, and community ideologies that promote structural distinctiveness. By highlighting these influences, the study helps educators

present a more nuanced picture of BSL grammar, countering the misconception that it must diverge from English in all respects.

Situated within the Deafness, Cognition and Language Research Centre (DCAL), which plays an active role in shaping BSL policy and curriculum, this research is well positioned to shape future developments. By disseminating our findings through academic publications and engagement with educators and policymakers, we aim to support high-quality, linguistically accurate BSL instruction. Through this dissemination, we also seek to raise awareness of the urgent, ongoing issue that most deaf children grow up without full access to a natural language–and to promote efforts that address this linguistic deprivation by improving access to high-quality BSL education.

Acknowledgements

- I'm deeply grateful to my principal supervisor, Kearsy Cormier, whose expertise and thoughtful feedback shaped every stage of this research. Her consistent availability, patience and belief in my ability to complete the project–especially when I doubted it myself–made all the difference. I also thank my subsidiary supervisor, Ad Neeleman for his constructive input on syntactic theory.
- I would like to thank my examiners Johanna Mesch and David Adger for their thoughtful feedback on this thesis and for the stimulating discussion during my viva.
- Thanks to Adam Schembri for his encouragement and helpful suggestions on sociolinguistics and judgement task design, and to Trevor Johnston, whose pioneering corpus research on Auslan shaped this field and whose advice on grammatical categories informed key decisions.
- I'm profoundly thankful to Gab Hodge and Kate Rowley–my academic heroes whose generosity, insight and emotional support meant more than I can say. Velia Cardin's practical help and statistical advice were deeply appreciated, as was Neil Fox's input on judgement task design and filming. Klaus Abels's thoughtful discussion helped clarify my experimental findings.
- I also thank Matt Brown for his astute statistical advice and help with judgement task logistics. He led the way and showed me the path to completion. Thanks to Yasamin Motamedi, Marieke Schouwstra and Bodo Winter for their input on statistical methods, and to Katherine Ellis for her mentoring when I felt I'd never finish. I'm grateful to Zahra Nawab and Veronica Escobar for their friendship, support and feedback.
- Many others made this work possible. Particular thanks to the interpreters I've worked with–especially Esther, whose patience and good humour helped me improve a BSL presentation–and to the UCL staff whose behind-the-scenes support smoothed the way.
- To everyone at DCAL–those already mentioned and many others–thank you for your encouragement and generosity. You helped me begin to overcome imposter syndrome and feel I could belong in academia.
- I'm especially grateful to the deaf participants who contributed to this research–both those whose data appears in the corpus and those who took part in the judgement task. Your lived experience of BSL is at the heart of this thesis. Thanks also to those involved in corpus development and task design, including Robert Adam, Jordan Fenlon, Sannah Gulamani, and the team at Remark!–especially Marcel Hirshman and Rowan Warnock.
- To my friends–Jo, Jag, Karen M, Moley, Karen R, Dolores, Lizzie, Anna and Charlotte–thank you for your patience, encouragement and for gently steering my thoughts beyond the thesis when I needed it most. Your friendship has been a lifeline.
- To my family, thank you for backing my unconventional decision to leave a stable career and study linguistics. But then, we've never been a conventional family. Rachel and Ed, thank you for your careful proofreading and for engaging with the thesis in all its complexity.
- Finally, to Ed–thank you for your unwavering support, quiet patience and practical help. Your presence made this journey possible.

Table of Contents

List of Figures

List of Tables

Abbreviations

Adj	Adjective
ASL	American Sign Language
Auslan	Australian Sign Language
BSL	British Sign Language
CLU	Clause-like unit
DCAL	Deafness, Cognition and Language Research Centre, University College London
Dem	Demonstrative
HKSL	Hong Kong Sign Language
ISL	Israeli Sign Language
LMEM	Linear mixed-effects model
LIS	Italian Sign Language
NGT	Dutch Sign Language
NP	Noun Phrase
Num	(Cardinal) numeral
NZSL	New Zealand Sign Language
PJM	Polish Sign Language
PMI	Pointwise mutual information
SSE	Sign-supported English
TSL	Taiwanese Sign Language

Conventions

BSL Corpus Conventions

A separate annotation file is created for each BSL Corpus participant. As outlined in Cormier et al. (2017), filenames follow the convention: **RRnnGAAELT** where

- **RR** = UK region, e.g., *BM* for Birmingham, *GW* for Glasgow
- **nn** = two-digit participant identifier: unique within each region
- **G** = gender (only male or female were recorded)
- **AA** = age in years at the time of filming
- **E** = ethnicity, e.g., *W* for white, *B* for black
- **L** = language background: *D* for participants with deaf family members, *H* for those without
- **T** = task type, e.g., *N* = narrative, *C* = conversation, *I* = interview

For example, the filename **BM06F56WDC** refers to the conversation task for the sixth participant from Birmingham, a 56-year-old white woman with deaf relatives. If a signer is left-hand dominant, the filename is suffixed with **_LH**. Thus, **GW01M29WDN_LH** indicates that the first participant from Glasgow is left-handed.

Participants are also referred to using a shortened code that includes the region, participant number, and–where relevant–the task (in lowercase). For instance, the second participant from Belfast is labelled **BF02**, and their narrative data is referred to as **BF02n**.

Within BSL Corpus annotations, clause-like units (CLUs) are labelled using the full filename (not just the shortened participant code), followed by **_CLU#** and a sequential number beginning from the start of the annotation file. For example, the third CLU in the conversation annotation file for the sixth participant from Birmingham would be labelled **BM06F56WDC_CLU#003**.

Lexicalised BSL signs are also annotated within the corpus using their ID gloss–an English gloss used to uniquely represent a sign in its citation form or its morphological or phonological variants–written in small capitals. Citation forms and other English translation equivalents (*keywords*) are available in the BSL Signbank: http://bslsignbank.ucl.ac.uk. BSL Signbank is a publicly accessible resource, although researcher access is required to view ID glosses; this can be obtained by registering via the Login link on the website.

Where a single BSL sign corresponds to more than one English word, hyphens are used to join the words in the gloss (e.g., AGE-FOUR), indicating that the sign is treated as a single lexical item.

Further annotation conventions used in the BSL Corpus are detailed in Appendix A.

Noun Phrase Conventions

The following conventions are used when discussing elements of noun phrases (NPs):
Linear order within noun phrases is expressed using square brackets, such as

- [Adj Noun]

Optional elements are shown using round brackets. Thus

- [(Dem) (Num) (Adj) Noun]

includes, for example,

- [Dem Num Noun].

Linearisation of semantic scope (see Section 1.3.2.3) is expressed using nested square brackets, for example:

- [Dem [Num [Adj Noun]]]

An unordered set of NP elements is shown using curly brackets, for example

- {Num, Noun}

Angle brackets <> are used to indicate simultaneity within an NP. This includes mouthing(s) produced at the same time as manual sign(s), and separate manual signs produced simultaneously on each hand. For example,

- Adj<>Point

indicates an adjective signed with one hand and a pointing sign with the other, such as

- DEAF<>PT:PRO3SG in BL03c; see Figure 77,

Where multiple mouthings occur simultaneously with a single manual sign, they are grouped using quotation marks. For example,

- Noun<>'Noun Noun'

indicates a manual sign produced simultaneously with two mouthings, all functioning as nouns–as in

- PARENTS<>'mum dad' in BF02n (see Figure 64).

This notation is used throughout the thesis to reflect constructions where NP elements are not sequentially ordered. When each hand produces a distinct sign that forms part of the NP, both signs

are included in the analysis. This contrasts with two-handed productions of a single sign, which are treated as a single element.

Other Conventions

An asterisk (*) next to a construction indicates that this is judged to be ungrammatical.

Chapter 1 Introduction

Nouns carry much of the semantic weight in sentences (Algeo, 1995), and the noun phrase (henceforth, NP) has long been a focus of linguistic research in many spoken languages (Siewierska, 1997). While some work has explored NP structure in signed languages, British Sign Language (BSL) remains largely unexamined in this respect. This thesis addresses that gap by presenting the first systematic investigation of NP structure in BSL, drawing on both corpus data and experimental evidence.

1.1 Structure of this Thesis

This thesis investigates NP structure in BSL through a combination of theoretical analysis, typological comparison, and empirical data. Chapter 1 introduces the topic and outlines the conceptual framework, followed by a literature review that situates the study within existing research on NP structure in both spoken and signed languages. Section 1.4 presents the research questions and hypotheses that guide the investigation. Chapter 2 describes the dual methodological approach–corpus analysis and a large-scale judgement task–and explains how these methods were implemented.

Chapter 3 presents the results from both data sources, including findings on NP length and the distribution of NP modifiers. Chapter 4 discusses these results in relation to previous research and evaluates the extent to which BSL NP structure aligns with typological predictions and cognitive principles. Chapter 5 concludes the thesis by summarising the key findings and outlining implications for future research and BSL pedagogy.

1.2 Theoretical Framework

Our overall theoretical orientation is grounded in cognitive and functional linguistics (as outlined in Nuyts, 2012), which views language as emerging from general cognitive processes and our interaction with the world. Within this framework, language is conceptualized primarily as a communicative and socially embedded tool. We follow LaPolla (2003) in viewing grammatical structures as arising through the conventionalisation of recurrent patterns in discourse, which serve to guide the addressee's inferential processes in interpreting intended meaning. This perspective highlights the gradient and usage-based nature of cross-linguistic variation, offering an alternative to models that rely on discrete parameter settings.

In addition to these perspectives, our research is grounded in variationist sociolinguistics (Lucas et al., 2001), which examines how language variation and change are influenced by social factors such as signer/speaker age, region, gender or social class, providing a comprehensive understanding of language as both a cognitive and social phenomenon.

1.3 Literature Review

1.3.1 Introduction

One might assume that defining NPs is a straightforward task, yet the concept proves to be surprisingly elusive. A common starting point is the semantic view that NPs[1] refer to entities in the real world (Martínez-Insua & Pérez-Guerra, 2011). However, this assumption does not always hold. As Baker (2003, p. 97) observes, in the sentence "No letters arrived today", the noun "letters" does not refer to any specific items. On closer inspection, it becomes evident that NPs refer to our mental representations of entities, allowing us to refer to things that have never existed, such as unicorns (Rijkhoff, 2002, p. 44). Moreover, the class of nouns encompasses abstract and intangible concepts like "gender" or "ability" (Kintsch, 1972).

Despite these "fuzzy boundaries" (Thompson & Ono, 2020, p. 2), identifying NPs in a given text requires first categorising the grammatical tokens, and then extracting those that are nouns or their modifiers.

In formal (generative) linguistics, grammatical category is not treated as an inherent property of lexical items. Instead, categories such as noun or verb are viewed as structural outcomes of syntactic derivation. Central to this view is the concept of the *root*–a minimal unit of meaning that is acategorial, meaning it lacks grammatical category until it is embedded within a syntactic structure. As discussed in Lohndal (2020), both Distributed Morphology and exoskeletal approaches offer competing accounts of how categorisation occurs. Distributed Morphology proposes that roots combine with functional heads (e.g., *n* for noun, *v* for verb, *a* for adjective), which assign category during derivation. In contrast, exoskeletal approaches argue that categorisation arises from the structural position itself, without the need for overt or covert categorisers. Don (1993), working within a generative morphological framework, similarly treated grammatical category as part of the

[1] In some theoretical perspectives, the *determiner phrase* (DP) is discussed rather than the *noun phrase* (NP); see Blümel and Holler (2021) for a recent collection of papers on this issue. The terms *NP* and *DP* are equivalent for the purpose of this thesis, and *NP* will be used throughout. See Section 1.3.4.7.1 for an example of work which uses *DP* terminology when discussing signed languages.

morphosyntactic representation. He argued that even in cases of conversion–such as “walk” used as a noun versus “walk” used as a verb (Don, 1993 p. 1)–category assignment occurs through abstract affixation at the morphosyntactic level. These frameworks aim to explain the flexibility of lexical items and the cross-linguistic variation in word formation by treating categorisation as a syntactic, rather than lexical, process.

In contrast to generative approaches, which treat grammatical category as a structural outcome of syntactic derivation, cognitive-functional linguistics views categories as emergent from communicative function. Van Valin and LaPolla (1997) observed that all languages appear to have words that can be categorised as noun or verb[2], since a key function of language is to be able to predicate something (using a verb) about a referent (typically expressed as a noun). They further noted that predicate arguments are consistently realised as NPs. From this perspective, grammatical categories are not treated as fixed lexical or syntactic properties, but as emergent from communicative function. Categories arise through the interaction of semantics, pragmatics, and discourse structure, rather than being determined solely by formal syntactic mechanisms.

A third approach, distinct from both generative and cognitive-functional frameworks, is the syntactic distributional method advocated by Schachter and Shopen (2007). They proposed assigning a grammatical category to each word in a particular language based on its syntactic behaviour. They illustrated this with the English example in (1), noting that “boys” and “like” have different distributions in this language, so that (2) is ungrammatical (Schachter and Shopen, 2007 p. 2).

(1) Boys like girls

(2) *Like boys girls

Additionally, “boys” and “like” have different functions, so that, for example, “boys” can take the role of subject, whereas “like” cannot. Finally, one can use the ways in which a word can be modified to express concepts such as case, gender, number or tense. So, for example, “like” can be inflected for both number and tense, while “boys” can only be inflected for number. Therefore, Schachter and Shopen concluded that “boys” and “like” are in different grammatical categories. In contrast,

[2] There is some disagreement even over whether all languages have grammatical categories of noun and verb. Evans and Levinson (2009) cite Straits Salish, a Salishan language of the Pacific Northwest of America, as a possible counterexample. In this language, a single open class of lexical items may function as predicates, arguments or modifiers, depending on the syntactic context. However, this interpretation is contested by other authors (see Evans & Levinson, 2009, p434-435, commentary section), and no language has yet been unambiguously shown to lack this distinction entirely.

they observed that “boys” and “girls” are in the same category since they are very similar in these three respects, so that, for example, (3) is grammatical (Schachter and Shopen, 2007 p. 2).

(3) Girls like boys

Then each such category can be given a label that is useful crosslinguistically, based on the semantics of the words assigned to it: as the authors noted, the category that includes the names of most people, places and things is termed *noun*.

This approach can be effective for languages that are well-studied, especially where there is a reference grammar or large corpus available that can form the basis of assignment of grammatical categories to words as described above. However, we note that for less well-studied languages it is not clear how this approach could be used in practice to identify the grammatical categories of particular words in context.

In contrast to this syntactic approach, semantic typologists such as Haspelmath (2012) and Dryer (2018) have advocated defining grammatical categories based on meaning. Haspelmath observed that all languages appear to have many words corresponding to the semantic domains of *thing* (typically associated with nouns), *action* (verbs) and *property* (categories including adjectives). Dryer noted that adjectives may constitute a distinct grammatical category, as in English, or be integrated into the category of verbs. In such cases, they appear in relative clauses modifying nouns, as in (4)[3] from Ojibwa, an Algonquian language of the Algic language family: here, “e-gnoozi-d” is a relative clause containing “gnoozi“, a verb meaning “be tall”. This relative clause modifies the noun “nini” (meaning “man”).

(4) nini e-gnoozi-d
man rel-tall-3sg
“a tall man”
(Ojibwa, from Dryer 2018 p. 805)

Dryer (2007b) also cautioned against assigning grammatical categories in a target language based on translations into a familiar language, as this risks misrepresenting the structure of the target language. Rijkhoff (2002) further notes that in some languages, such as Samoan, grammatical category is not lexically encoded but emerges from usage in context.

[3] Abbreviations in this example are as used by Dryer: rel: relative, 3: 3rd person; sg: singular

In contrast to approaches that assume fixed grammatical categories–whether assigned through syntactic derivation, distributional behaviour, or semantic domains–this thesis does not treat grammatical category as inherent to BSL signs. Signs in BSL often appear underspecified when considered out of context, and their grammatical function is only defined within the discourse (see Section 1.3.4.5 for more detail). Consequently, we analyse each NP element based on its use in context. This context-sensitive analysis reflects a cognitive-functional orientation, in which grammatical categories are understood as arising from usage and discourse, rather than from structural derivation.

Next, we discuss past research regarding NPs; first relating to the order of elements within NPs, and then, other questions that have been addressed using large-scale language corpora. Then we provide some background about the structure of signed languages and look specifically at previous studies into NPs in this modality.

1.3.2 Order within the NP

Researchers have long been concerned with determining and accounting for word order in NPs in particular spoken languages (and a few signed languages, for details of these see Section 1.3.4.7). Before going on to discuss this research, we consider two related issues that we also encountered in our own study. Firstly, Dryer (2007a) noted that in practice most natural language production does not include all the possible elements of a NP. Indeed, Thompson & Ono (2020) observed that cross-linguistically, NPs containing more than one word are rarely found in natural conversation. Dryer pointed out that the most frequent type of NP is one containing just a noun or a pronoun. In Dryer (2007b), he noted that researchers can derive the overall order within the NP by extrapolating from observed examples involving a subset of modifiers; this then serves to describe the relative order of those elements that are present in a particular NP. So, for example, given observed data in the language under consideration with the following grammatical categories:

(5) [Noun Adj Num];

(6) [Noun Num Dem];

we can infer that the overall order within the NP in that language is:

(7) [Noun Adj Num Dem].

Secondly, in many cases, more than one order of elements is considered grammatical in a particular language. As discussed by Dryer (2007b), linguists have therefore attempted to identify a so-called

basic or *unmarked* word order, using a variety of criteria. These criteria include: (a) the most frequently occurring order; (b) the order with the least restrictive syntactic or lexical distribution; and (c) the order found in basic declarative sentences. While these criteria are sometimes stated explicitly, as in Abels (2016), they are often only implied. Additionally, without a large, representative corpus of the language, it is challenging to objectively determine the relative frequency or distribution of particular constructions. This can make it difficult to identify a basic or unmarked word order.

As discussed in Section 1.3.4.4, there is a great deal of simultaneity in both spoken and signed language production. However, despite this high level of simultaneity, research into the structure of NPs has tended to focus on linear word order. The linguistic typologist Greenberg (1963) conducted seminal work in this area, proposing 45 *Universals* based on typological data. One of these, *Universal 20*, relates to the order of the noun, adjective, numeral and demonstrative within noun phrases. We discuss Universal 20 in the next section, then go on to review research that has stemmed from this.

1.3.2.1 Greenberg's Universal 20

There are 24 logically possible ways to order the elements in a NP consisting of a noun, an adjective, a numeral and a demonstrative[4]. Universal 20 (Greenberg, 1963) relates to the order of the nominal modifiers adjective, demonstrative and (cardinal) numeral within the NP; it states that if any or all of these modifiers are prenominal, they are found in the order given in (8). (Greenberg only discusses one NP order for each language, with the implication that this is the basic unmarked NP order in the language.)

(8) [(Dem)[5] (Num) (Adj) Noun]

Universal 20 states that in contrast, if any or all of the modifiers are postnominal, they are found in one of the following two orders with (9) being much less frequent cross-linguistically than (8) and (10):

(9) [Noun (Dem) (Num) (Adj)]

(10) [Noun (Adj) (Num) (Dem)]

[4] Demonstratives are used to indicate what is being referred-to, as distinct from other entities of the same type. In English, the demonstratives are "this" and "these" (indicating entities proximal to the speaker), plus "that" and "those" (for entities distal to the speaker).

[5] Optional elements are indicated using round brackets, as defined in the Conventions section.

Greenberg's findings are grounded in typological data. He was able to investigate 30 languages, and although he acknowledged that this sample is small, his articulation of 45 linguistic universals was groundbreaking. He noted that his work assumes that all languages have identifiable grammatical categories, and stated that he used broadly semantic criteria to determine which words fall in each category, though he did not expand on his methodology for doing this. Greenberg conceded that his approach was not completely rigorous; however, we consider that this is justifiable given that his focus was on articulating a large number of diverse morphosyntactic universals. He provided some principles which may underlie his Universals: he noted that there is a *proximity hierarchy* (Greenberg, 1963, p. 24), meaning a tendency for certain elements (adjectives, in the case of NPs) to occur immediately adjacent to the head element (the noun in this instance), as compared to elements that are further down the hierarchy within the NP, namely numerals and then demonstratives. The two most common orders stated by Greenberg, (8) and (10) accord with this principle, whereas the less common order (9) does not. This idea is explored further in Sections 1.3.2.2.1 and 1.3.2.3.

Next, we discuss research that has used Greenberg's findings as a starting point and has tested his claims using a larger sample of languages. Linguists have developed theories seeking to explain the observation that when modifiers are prenominal, they are strictly ordered but when they are postnominal, there is more variation in their order. Later in this thesis, we will compare our findings from BSL with the predictions made by these theories.

1.3.2.2 Expansions Upon Greenberg's Universal 20

Greenberg's work has been built-upon by many researchers, using larger samples of the world's languages. Investigations have revealed that many different basic unmarked orders are found, not just those identified by Greenberg, and further, that these orders do not occur with equal frequency: indeed, the distribution of orders across languages is highly skewed, with some orders being observed much more frequently than others. Next, we describe four approaches to explaining the cross-linguistic distribution of orders of elements within NPs, from Cinque (2005), Abels and Neeleman (2012), Steddy and Samek-Lodovici (2011) and Medeiros (2018). Like Greenberg's, these approaches all focus on the basic or unmarked NP order in each language[6]. Cinque used a larger

[6] In contrast, Steedman (2020) took a different approach, considering all orders that are possible including those that occur as a result of contrast or focus. He proposed an analysis based on separable permutations (Bose et al., 1998) that

sample of languages than Greenberg was able to, and proposed an explanation for the exact distribution of NP orders that he observed. The other three sets of researchers offered different explanations for the same data.

1.3.2.2.1 Explanations using data in Cinque (2005)

Cinque (2005) observed that when using a larger set of languages than was possible for Greenberg, Universal 20 turns out to be both too restrictive–disallowing some attested orders–and slightly too permissive, in allowing an order–specifically, the order in (11)–which adheres to Universal 20 but is not attested in Cinque's sample of languages.

(11) [Num Noun Dem Adj]

Cinque did not list the languages included in his sample, although he provided illustrative examples for each attested order. (While this limits the extent to which the representativeness of his sample can be independently assessed, further work by Cinque addressing this issue is discussed in the following section.) In his 2005 paper, Cinque found an asymmetry within his sample of languages: only a single order ((15) below) was attested with all prenominal modifiers, whereas five orders ((18), (20), (23), (26) and (28) below) were observed with all postnominal modifiers.

Working in the generative tradition, Cinque provided a theory to explain this cross-linguistic asymmetry in NPs that contain all three types of modifier. He started with elements in the following base-generated order, which he took to be the underlying order of elements of the NP in all human languages:

(12) [Dem Num Adj Noun].

Then he allowed leftward movement of constituents that include a noun. So, for example, the order (13) can be produced by moving the constituent [Adj Noun] leftwards of the numeral.

(13) [Dem Adj Noun Num].

In contrast, the order (14) cannot be produced given these constraints on movement.

(14) [Num Dem Adj Noun].

Cinque postulated that each language uses a specific set of movements to derive its basic/unmarked

results in only two of the possible 24 NP orders being disallowed, and he found that in his sample of NP orders across and within languages, these two orders were among only three orders that were never attested. Because Steedman does not restrict his discussion to the basic order in each language, his approach cannot easily be compared with the others discussed here.

NP order. The full list of 14 orders that can be produced by this mechanism are as follows:

(15) [Dem Num Adj Noun]

(16) [Dem Num Noun Adj]

(17) [Dem Noun Num Adj]

(18) [Noun Dem Num Adj]

(19) [Adj Noun Dem Num]

(20) [Noun Adj Dem Num]

(21) [Dem Adj Noun Num]

(22) [Dem Noun Adj Num]

(23) [Noun Dem Adj Num]

(24) [Num Adj Noun Dem]

(25) [Num Noun Adj Dem]

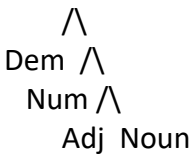
(26) [Noun Num Adj Dem]

(27) [Adj Noun Num Dem]

(28) [Noun Adj Num Dem]

Similar to Cinque's approach, Abels and Neeleman (2012) defined sets of movements and proposed that each language uses one of these to derive its NP order. They use the same data set as Cinque in terms of orders that are attested/not attested. They took the structure in Figure 1 as a starting point, and allowed each pair of sister nodes {X, Y} to be base-generated in either order; [X Y] or [Y X], so for example, the sister nodes Adj and Noun can be base-generated as either [Adj Noun] or [Noun Adj]. This produces a total of eight orders. These are the same orders that obey what Cinque terms *natural relative semantic scope*, and two of these orders are predicted by Greenberg. We discuss these eight orders further in Section 1.3.2.3.

Figure 1 A Base Structure Discussed by Abels and Neeleman (2012)



Then, in line with Cinque, they allowed leftward movement of constituents including a noun. Because of their larger set of base-generated orders, the movement operations required are simpler. Both Abels and Neeleman's, and Cinque's approaches do indeed generate all the orders that Cinque finds to be attested, and none of those that he finds are unattested.

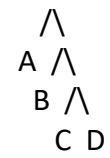
Steddy and Samek-Lodovici (2011) explained the same language data using optimality theory (Prince & Smolensky, 2004), devising a set of constraints on movement based on the distance of each element of an NP from the left-hand edge of that NP. They theorised that each language has a particular ranking of the constraints, and this produces an order within the NP that is optimal with respect to this ranking. Once again, the NP orders that are unattested in Cinque's (2005) data are exactly those which are not predicted by any possible ranking of the constraints.

Medeiros (2018) took a different approach based on theoretical computational linguistics. Rather than focusing on language production like the proposals discussed so far, he concentrated on language comprehension, modelling the human grammatical processing involved as a parser based on Knuth's (1968) stack-sorting algorithm. The parser is capable of transforming the majority of possible input orders of elements in an NP into a single order (29), which Medeiros took to be the underlying order in human cognition.

(29) [Noun Adj Num Dem].

His parser produces the order in (29) for some but not all 24 possible input orders of an NP. Those for which this does not happen share a *mid-high-low* ordering property, as follows: Given a generic tree as in Figure 2, linearisations of A-D that include a mid-high-low arrangement include B-A-C-D (B-A-C is mid-high-low). In contrast, C-D-A-B does not include a mid-high-low arrangement.

Figure 2 A Generic Tree Structure



The orders which include a mid-high-low arrangement are exactly those that Cinque found to be unattested, leading Medeiros to claim that the algorithm is a model for human comprehension: he proposed that these mid-high-low orders do not occur in natural language because the algorithm cannot sort them into the order in (29). One critique of Medeiros's proposal is that some of the orders that cannot be processed by his parser do occur in natural language in special circumstances, such as when specific intonation is used or when particular elements are focused: these would require a different parsing strategy. Aside from this, the beauty of the proposal is that a single algorithm is required to parse all the (non mid-high-low) orders of input elements. This therefore could serve as a model for how children acquire language - it provides an explanation for how they

are able to parse the input that they receive (providing it is not mid-high-low), even before they have learned the rules of the grammar to which they are being exposed. While this is an appealing theory, as Medeiros noted, one cannot simply reverse the algorithm in order to describe human language production. For this production process, language-specific parameters would be required.

Cinque (2005) also discussed the relative frequency of different orders within the NP, categorising them as found in *very many*, *many*, *few*, *very few* or *no* languages. (However, he does not quantify these terms.) His model explains this variation by defining movement operations as having varying degrees of markedness: orders which have fewer marked movements are more frequent. The other papers discussed in this section vary as to whether they consider relative frequency of orders: Abels and Neeleman (2012) did not consider this, but they did discuss it in Abels & Neeleman (under review): here, they provided a ranked list of constraints, whereby the fewer higher-ranked constraints that an NP order violates, the more frequently this order is observed among the world's languages. This analysis closely matches the frequency of NP orders observed in recent samples of the world's languages such as Dryer (2018) and Cinque (n.d.) which are discussed in Section 1.3.2.2.2. Steddy and Samek-Lodovici (2011) stated that Cinque's assertion that some of his movement rules are more marked than others can also be applied to the movements involved in their own theory, and that this will result in the same relative frequencies of word orders. They also made the point that these assertions about markedness, while they describe the observed differences in frequency, do not explain why these differences occur. Finally, Madeiros did not discuss the relative frequency of different orders. Rather, his parser simply excluded those orders which Cinque found to be unattested.

The approaches discussed so far provide elegant explanations for the dataset of languages used by Cinque (2005) from a variety of perspectives. They categorise each of the 24 possible orders of elements in a NP as either *attested* or *unattested* but of course, *unattested* actually means *unattested in this sample*: this qualification will always be necessary since it will never be possible to exhaustively investigate all languages that have ever existed or that may exist in future (Cysouw, 2010). Therefore, it is more appropriate to consider the relative frequency of different word orders within an NP rather than absolute notions of attested orders. Next, we discuss two papers that do just this, using increasingly large and representative samples of the world's (spoken) languages. Then in Section 1.3.2.3 we examine psycholinguistic approaches that avoid this sampling problem.

1.3.2.2.2 Explanations based on further data

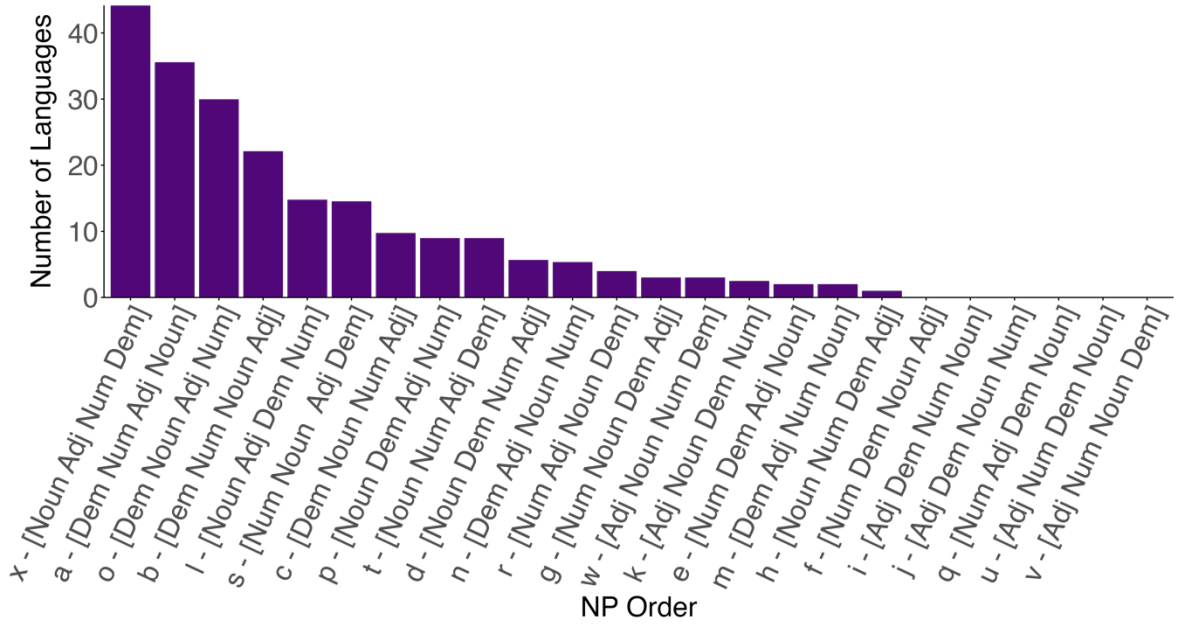
Cysouw (2010) considered the relative frequency of different orders of elements within an NP, using a larger dataset than Cinque (2005). It consisted of 276 languages and was produced by Dryer (2006). Cysouw proposed a model based on probabilistic generalisations such as (30).

(30) Noun and adjective have a strong tendency to occur immediately adjacent to each other (Cysouw, 2010, p. 282).

He then used generalised linear modelling to assign each generalisation a weighting calculated to fit Dryer's data. Cysouw observed that his model fits these observed frequencies better than does Cinque's. However, while his generalisations provide a good description of the observed NP orders, he does not provide a full explanation for why they hold true.

The most systematically sampled dataset we consider is from Dryer (2018). He used a carefully designed approach to sampling spoken languages, aiming to remove biases associated with historical and geographic relationships between languages. Figure 3 shows his adjusted frequency data based on a sample of 576 languages. (The total number of languages in this graph does not equal 576 because Dryer's bias-reducing sampling methodology that avoids over-representing any language genus or geographical area tends to reduce the number of sampled languages.)

Figure 3 Number of Languages With Each NP Order, Adjusted for Genetic Relations

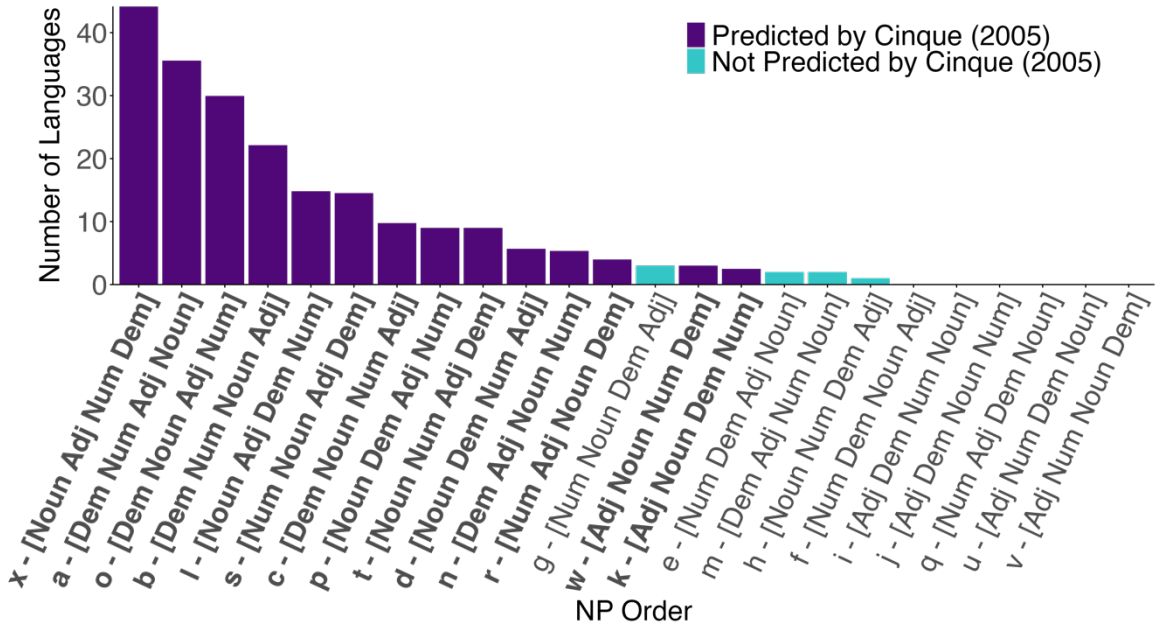


Note: Shows adjusted frequency data from Dryer (2018). Letters prefixing each NP order are those used by Cinque (2005).

Dryer (2018 p. 800) proposed five descriptive principles to account for the variation in frequency of the different orders: for example, he suggested two *iconicity principles* that together state that adjectives are generally found next to nouns (similar to Cysouw's generalisation in (30) above), with numerals further from the noun than adjectives and demonstratives further away still, when these modifiers are found on the same side of the noun. (These principles are analogous to Greenberg's proximity hierarchy discussed in Section 1.3.2.1.) The principles are motivated by the observation that semantically, adjectives tend to describe inherent properties of nouns, numerals describe more transient properties, and demonstratives do not describe properties of the noun at all. (We revisit this motivation in Section 1.3.2.3.) Another of Dryer's principles states that postnominal adjectives are preferred over prenominal adjectives. While Dryer noted that his main aim was to provide these principles as a descriptive account, he nonetheless offered tentative explanations for them where possible. One such explanation addresses the cross-linguistic tendency for adjectives to follow nouns. As noted by Culbertson et al. (2012), this preference may arise from the fact that the interpretation of many adjectives depends on the noun they modify. For example, the NP "big mouse" does not simply denote the intersection of the sets of big things and mice. Instead, the adjective "big" is interpreted relative to the typical size of a mouse, demonstrating its semantic reliance on the noun (Dryer, 2018 p. 818). In contrast, the meaning of a noun is generally independent of the adjective that accompanies it. (We return to this principle and its possible explanation in Section 1.3.2.4.)

Dryer related the work of Cinque (2005) and Abels and Neeleman (2012) to his expanded and more representative sample of the world's languages. This broader dataset altered the relative frequency of NP orders, though the overall distributional shape remained consistent. Dryer observed that Cinque's theory aligns very well with the updated data: in particular, although it permits 14 of the 24 possible orders, it disallows all the orders that Dryer also found to be unattested. However, Dryer did find examples of four of the orders that Cinque had said were unattested, albeit with low frequency: these are orders *g*, *e*, *m* and *h* in Figure 4. This is not unexpected, as if a larger sample of languages is included, the frequency of all possible orders of elements within a NP is likely to increase.

Figure 4 NP Orders That are Predicted by Cinque (2005)



Note: Shows adjusted frequency data from Dryer (2018). Letters prefixing each NP order are those used by Cinque (2005). Orders predicted by Cinque (2005) are shown in boldface.

In more recent research, Neeleman (2024) provided a critique of one of Dryer's newly-attested examples. A native speaker of the language in question, Sierra Popoluca (a Mixe-Zoque language of North America), stated that the expression quoted is actually a partitive construction meaning, for example, "two of these new sacks" rather than a straightforward NP "these two new sacks" (Neeleman, 2024, p.4). As such, it falls outside the scope of Universal 20.

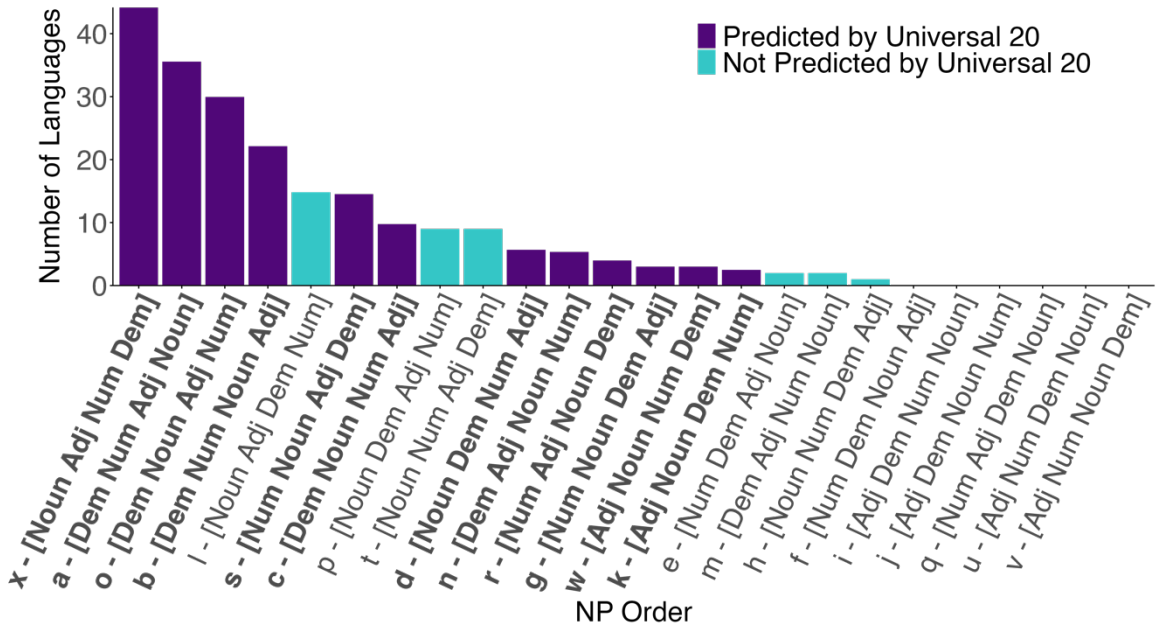
Cinque has since compiled an even larger typological dataset of around 2200 languages (Cinque, n.d.)–despite the significantly larger sample, he continued to observe that exactly the same set of 14 NP orders are attested.

There is inevitably a degree of arbitrariness in any language sample. The question of whether low- or zero-frequency NP orders are genuinely attested will likely persist as more languages are studied and existing data is examined in greater detail. Any orders that prove to be robustly attested present a challenge to theories that rule them out, such as those proposed by Cinque (2005), Abels and Neeleman (2012), Steddy and Samek-Lodovici (2011), and Medeiros (2018).

For comparison with Figure 4, Figure 5 shows the 12 NP orders permitted by Greenberg's Universal 20 on Dryer's (2018) frequency graph. Universal 20 is also successful in disallowing orders that are rarely/never attested. However, it fails to allow three orders (*l*, *p* and *t*) that have medium frequency, which Cinque's 14 predicted orders include. Greenberg's Universal 20 remains an

impressive theory, notable for its innovative methodology and approach: the fact that Greenberg was only able to consider 30 languages may explain its relatively poor fit with Dryer's data.

Figure 5 NP Orders That are Predicted by Greenberg's Universal 20



Note: Shows adjusted frequency data from Dryer (2018). Letters prefixing each NP order are those used by Cinque (2005). Orders predicted by Greenberg's Universal 20 are shown in boldface.

To complete our comparison of approaches in this section to Dryer's improved cross-linguistic sample of languages, we note that the model in Cysouw (2010) does not match the frequencies in Dryer (2018) very well. However, this is not surprising since it was designed as the best fit for a previous set of frequency data (from Dryer, 2006). It would be possible to create an updated model based on Cysouw's approach and the adjusted frequencies from Dryer (2018), which would result in a better match.

The approaches discussed so far all attempt to explain the observed asymmetry in NP orders across the world's languages: In languages where modifiers precede the noun, six logically possible orders exist when all three modifier types are present–but in practice, almost all such languages use just one of these orders (order *a* in Figure 3 - Figure 5). By contrast, languages with postnominal modifiers exhibit far greater variation, with multiple orders commonly attested. The following section introduces a theory that departs from this asymmetrical pattern: unlike theories that predict a single dominant prenominal order, it proposes a symmetrical distribution in which each NP order has a corresponding mirror-image counterpart.

1.3.2.3 Homomorphism

Having reviewed typological models that explain the asymmetrical distribution of NP orders across languages–particularly the tendency for prenominal modifiers to follow a strict order while postnominal modifiers exhibit greater variation–we now turn to a theory that offers a contrasting perspective: *homomorphism*, which proposes that NP orders reflect semantic scope transparently. This section outlines the hierarchical structure underlying this theory, presents the subset of NP orders that preserve scope relations, and reviews experimental evidence from artificial language learning and silent gesture studies that support the psychological plausibility of this structure. These findings suggest that homomorphism may reflect a general cognitive bias toward scope-based linearisation.

To illustrate the semantic scope relationships that underpin the homomorphism hypothesis, Figure 6 presents a version of the base structure in Figure 1 that is discussed by Abels and Neeleman (2012). It shows the semantic scope relationship between the elements of a NP: Rijkhoff (2002, p. 314) states this as "Modifiers tend to occur next to the part of the expression that they have in scope".

In other words, adjectives take scope over the nouns that they modify, then numerals take scope over this adjective-noun combination and finally demonstratives take scope over the whole numeral-adjective-noun combination (Culbertson & Adger, 2014).

Figure 6 A Base Structure Discussed by Abels and Neeleman (2012)

Note. Repeated from Figure 1. The colours of the branches correspond to the brackets in (31) to (38).

The eight possible orders of elements within a NP that can be produced by linearising Figure 6 are listed in (31) to (38). The nested brackets indicate the scope relations, and their colours correspond to those of the branches of the tree in Figure 6.

(31) [Dem [Num [Adj Noun]]]

(32) [Dem [Num [Noun Adj]]]

(33) [Dem [[Adj Noun] Num]]

(34) [Dem [[Noun Adj] Num]]

(35) [[Num [Adj Noun]] Dem]

(36) [[Num [Noun Adj]] Dem]

(37) [[[Adj Noun] Num] Dem]

(38) [[[Noun Adj] Num] Dem]

The other 16 logically possible orders, such as (39), cannot be produced by linearising Figure 6 and the scope relations cannot be preserved–in this case, because Num is found between Adj and Noun.

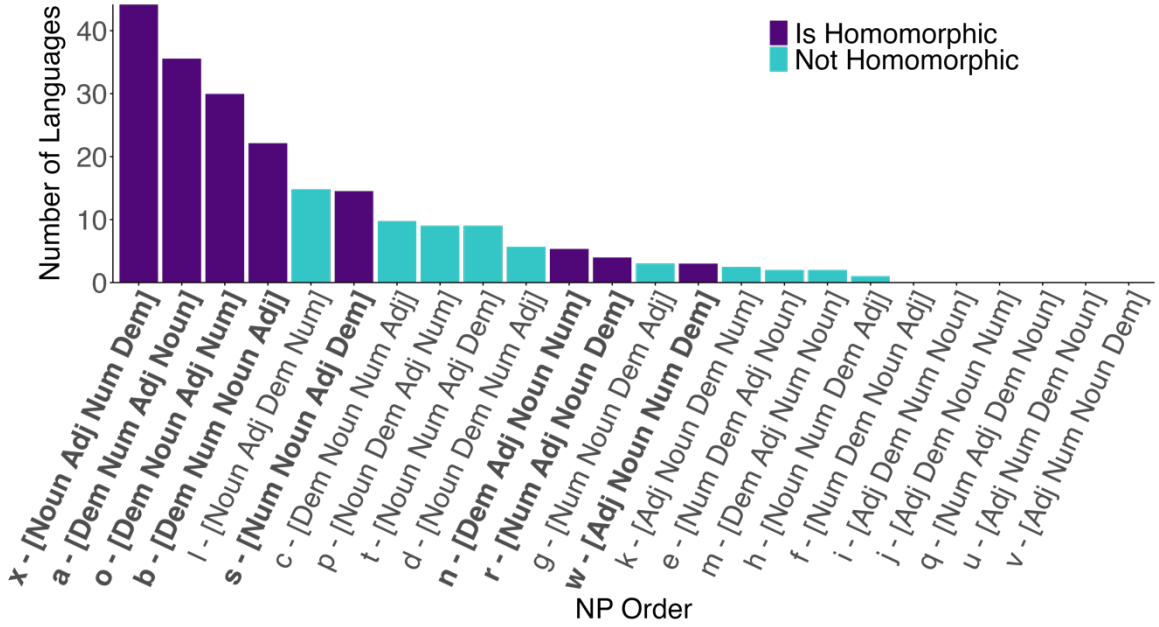
(39) [Dem Adj Num Noun].

The orders (31) to (38) feature in most of the theories and explanations about the distribution of orders within the NP that we have discussed so far. They are a restatement of Greenberg's proximity hierarchy that we introduced in Section 1.3.2.1, since the adjective is always adjacent to the noun, then the numeral is always adjacent to the noun-adjective combination, and finally the demonstrative is adjacent to the noun-adjective-numeral combination. (31) to (38) are a subset of the NP orders predicted by the theories discussed in Section 1.3.2.2: they are included in those predicted by Cinque (2005) and are the same as those base-generated by Abels and Neeleman (2012). They are discussed by Cysouw (2010) and they also obey Dryer's (2018) iconicity principles. For example, in all the orders (31) to (38), the demonstrative is further away from the noun than any other modifiers on the same side of the noun. This transparent relationship between the base structure of elements of the NP as in Figure 6 and the linear orders listed in (31) to (38) has been termed *homomorphism*[7] by Martin et al. (2020). In this thesis, we have used *homomorphic* as a shorthand for "homomorphic to the base structure in Figure 6".

In Dryer's (2018) data, the four most frequently occurring orders of elements in a NP cross-linguistically are homomorphic, as shown in Figure 7. In all, 73% of the languages in Dryer's sample are homomorphic. Therefore, it has been suggested that humans represent the underlying structure of the NP is as in Figure 6 and that then there is a transparent mapping from this to the linear orders that we observe.

[7] In previous work, this has been characterised as *isomorphism.* As discussed in Martin et al. (2020), it is better described as *homomorphism*, reserving *isomorphism* for the two orders which have a one-to-one correspondence with Figure 6. These are orders (31) and (38) where all the modifiers of the noun are found on the same side of the noun, so that the relative order of all of these elements is transparent. The remaining six orders are described as homomorphic because it is not possible to determine the relative order of all of the elements from these. For example, in (33) [Dem Adj Noun Num], the adjective and numeral are equally far from the noun, in contrast to (31) and (38), where the adjective is closer to the noun.

Figure 7 NP Orders That are Homomorphic



Note: Shows adjusted frequency data from Dryer (2018). Letters prefixing each NP order are those used by Cinque (2005). Homomorphic orders are shown in boldface. Based on Martin et al. (2019), Figure 1, bottom right[8].

In order to test this theory, Culbertson and Adger (2014) asked participants to innovate word order within NPs in a miniature artificial language and examined whether these novel word orders were homomorphic. Their artificial language consisted of NPs made up of English words but whose NP modifiers were postnominal: for example, “vase purple” for “purple vase” (Culbertson and Adger, 2014 p. 5844). They taught the language to English-speaking participants (English having pre-nominal modifiers). Each noun in the training data had only a single modifier, either an adjective, numeral or demonstrative. They then tested the participants on two-modifier utterances: a noun followed by:

(40) {Adj, Dem};

(41) {Adj, Num}; or

(42) {Dem, Num}.

Participants were asked to select the correct word order. They did indeed reliably choose constructions with postnominal modifiers that were homomorphic, rather than choosing those where the modifiers were in the same linear order as English. This result was strongest for artificial

[8] We note some differences in the NP orders’ adjusted number of languages in this figure in Martin et al. (2019) compared to Dryer’s (2018) data. For example, in Dryer (2018), order c has 9.75 languages whereas by eye, the figure appears to be about 5 in Martin et al. (2019). This means that the sequence of NP orders on the x-axis is different in our graph as compared to Martin et al. The impact is minor since the distribution of homomorphic orders is approximately the same.

language (40) (94% of responses were homomorphic) as compared to (41) and (42) (for which 70% and 75% of responses were homomorphic respectively). They note that the modifiers in (40) (adjective and demonstrative) are the furthest apart in Figure 6, and suggest that this may be why the preference for a homomorphic order for this combination is strongest. We return to this finding in our own results (Section 3.2.6.1).

Martin et al. (2019) extended Culbertson and Adger's work by showing that the preference for homomorphic orders is not limited to English speakers or postnominal modifiers: they found that Thai speakers (Thai having postnominal NP modifiers) learning a prenominal artificial language also preferred homomorphic orders. This preference was even stronger than that observed in English speakers: this aligns with the cross-linguistic typological pattern–related to that noted at the end of Section 1.3.2.2.2–where languages whose NP modifiers are prenominal have a greater tendency for these NPs to be homomorphic than do languages whose NP modifiers are postnominal.

While these two studies support the theory that people represent NPs as in Figure 6, it could also be the case that the participants transferred scope relations (rather than surface order) from their native language to the artificial language. Martin et al. (2020) addressed this concern by refining the experimental design, using novel lexical items and picture stimuli so as to remove native-language cues. Their results showed homomorphism preferences across all modifier combinations, and no evidence of metalinguistic strategies such as transfer of scope relations from participants' native language. This strengthens the case that the preference for homomorphic NP orders reflects a cognitive bias, rather than prior linguistic experience.

However, even these findings cannot entirely rule out the possibility that participants' preference for homomorphic orders reflects their prior knowledge of a homomorphic language. To test whether such preferences arise independently of linguistic experience, one would ideally recruit participants who do not know a homomorphic language. Although non-homomorphic languages do exist (around 27% of Dryer's (2018) sample of the world's languages are in this category: see Figure 7), it is very difficult to find a participant population who *only* know a non-homomorphic language, since so many homomorphic languages (including English) have a wide global reach either as a first or an additional language. Therefore, researchers have turned to a *silent gesture* experimental paradigm. This is because this approach has been shown to elicit consistent gesture orders, regardless of the basic word order in each participant's native language: these have been argued to reflect humans' underlying cognitive representation of such concepts (Goldin-Meadow et al., 2008).

In their silent gesture experiment investigating the order of the semantic roles Agent, Action and Patient, Goldin-Meadow et al. showed videos of short vignettes to native speakers of four different languages (Chinese, English, Spanish and Turkish), and asked them to describe the vignettes using gestures but no speech. Even though these languages have different basic word orders in terms of Agent, Action and Patient, the participants overwhelmingly produced the same order of gestures, regardless of the order used in their spoken language: Agent-Patient-Action.

Along the same lines, Culbertson et al. (2020) used silent gesture to explore order within the NP. Their participants' task was to describe certain images, which were designed to elicit NPs. The images consisted of either four of five (numeral) squares or triangles (noun), which were spotted or striped (adjective), and either proximal or distal to the participant (demonstrative). Participants' gesture productions were overwhelmingly homomorphic, with only 2% being non-homomorphic and a further 11% ambiguous due to repetition or the simultaneous production of more than one gesture. Assuming a transparent mapping from internal representation to linear order, this result is consistent with the hypothesis that humans' mental representations of elements of the NP are as in Figure 6. Despite the fact that the participants were English speakers, they did not produce any gesture strings in the NP order found in English ([Dem Num Adj Noun]). This suggests that the use of silent gesture does entail minimal transfer from participants' native language, although one cannot exclude the possibility that a general bias for homomorphic orders is transferred. The only way to rule this out would be to recruit participants who do not know a language whose NPs are homomorphic. As noted above, such a population is difficult to find and to study. However, Martin et al. (2024) were able to work with monolingual speakers of Kĩĩtharaka, a Bantu language spoken in rural Kenya whose NPs are not homomorphic: they are ordered as in (43).

(43) [Noun Dem Num Adj]

A modified version of the task in Culbertson and Adger (2014) was used, using pictures to teach participants a miniature artificial language consisting of invented words: there was a noun (meaning "cup"), two demonstratives (meaning "this" and "that") and two adjectives (meaning "red" and "black"). The modifiers in the artificial language were prenominal since Kĩĩtharaka has postnominal modifiers. The training data consisted of a single prenominal modifier plus the noun (a translation of "that cup", for example). At test, participants were asked to use 3 words to describe pictures that required the use of both a demonstrative and an adjective as well as the noun (so translating to "this black cup", for example). They produced significantly more homomorphic orders (like "this

black cup") than non-homomorphic orders (like "black this cup") ($p < .01$). It is possible that the Kĩîtharaka speakers were using the linear order of modifiers in their own language (demonstrative followed by adjective) in this task: however, the English-speaking participants in Culbertson and Adger (2014) produced responses that do not match the linear order in English NPs so one would expect that the Kĩîtharaka speakers would also not be using this strategy. This shared outcome across both experiments–namely, participants' consistent preference for homomorphic orders–further supports the claim that there is a general human cognitive bias towards homomorphism, even for people such as the Kĩîtharaka speakers who do not know a homomorphic language.

Culbertson et al. (2020) hypothesised that the preference for homomorphic NP orders could be explained by features of human cognition and how we relate to the world around us: the structure could be subconsciously learned, based on observations of objects in our environment. They noted that semantically, objects (nouns) are frequently associated with their properties (adjectives); for example, "wine" with the colour "red". In contrast, objects are not generally associated with numerosities (numerals) (with a few exceptions such as shoes which occur in pairs). Furthermore, objects do not have a fixed association with their location in relation to the speaker (demonstrative), since this changes depending on the context (Culbertson et al., 2020, p. 8). (This explanation aligns with Dryer's (2018) motivation of his iconicity principles discussed in Section 1.3.2.2.2.)

Culbertson et al. (2020) tested this hypothesis by operationalising the strength of association between pairs of words in a given language using the measure *pointwise mutual information* (PMI). PMI provides a measure of collocation, indicating whether a given pair of words occur together more often than would be expected given each word's frequency in the sample. For example, the pair ("red", "wine") has high PMI, since these words co-occur relatively frequently, as noted above. The authors used large corpora of 24 spoken languages that had been annotated for grammatical category, taking these datasets to be a proxy for how concepts are represented in the world. (Since grammatical category annotation was a prerequisite, they were obliged to use only spoken languages, and mostly those from *WEIRD* (Western, Educated, Industrialised, Rich, Democratic) countries. These are clearly not representative of all the world's signed and spoken languages – see Henrich et al. (2010) for a wider discussion of the drawbacks of extrapolating experimental data from WEIRD populations to all of humankind.) Culbertson et al. calculated the PMI between pairs of noun and modifier (where the modifier is either an adjective, a numeral or a demonstrative) in each corpus. As they predicted, for each language, the average PMI across all noun-adjective pairs was

higher than the average across all noun-numeral pairs, which in turn was higher than the average across all noun-demonstrative pairs. This is consistent with the structure in Figure 6, where the adjective is closest to the noun, followed by the numeral and then the demonstrative. Therefore their hypothesis that humans represent NPs as in Figure 6, as a consequence of human cognition and engagement with the world, is supported with this corpus data.

This bias towards homomorphic orders within the NP is not absolute, since examples of languages with non-homomorphic orders (such as Kîîtharaka) have been observed, albeit these are relatively low frequency as shown in Figure 7. Languages with non-homomorphic NPs may have evolved due to external influences, such as historical or cultural factors, rather than cognitive processes (cf. Martin et al., 2024). We revisit this explanation in Section 1.3.2.4.

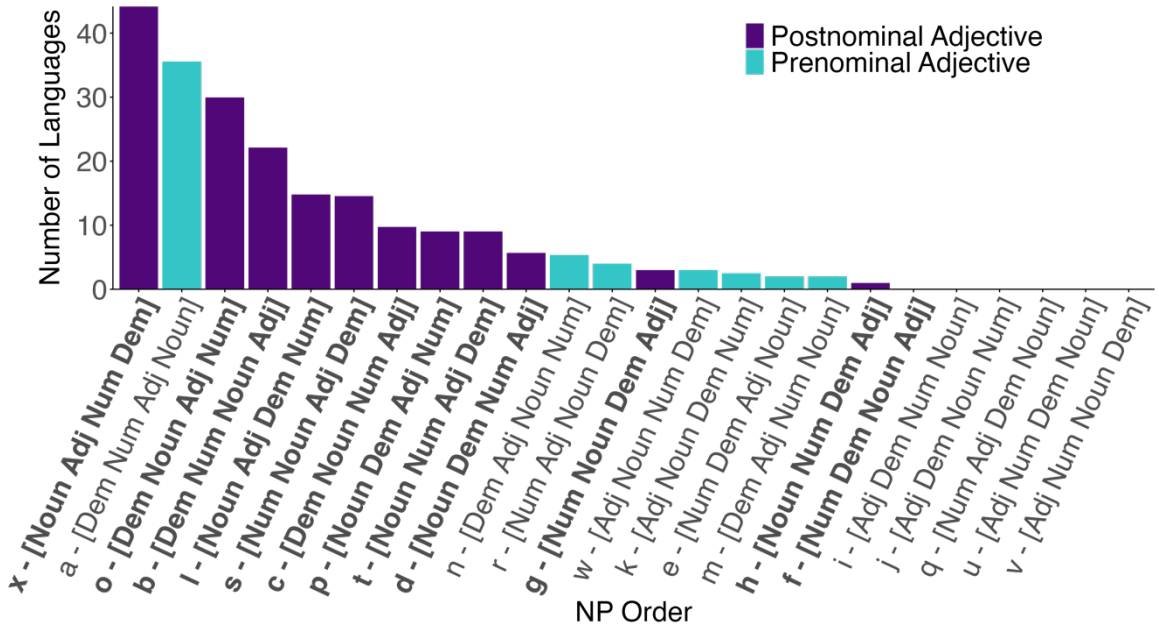
To summarise our discussion of theories about the order of elements within an NP in spoken languages: the research discussed in Sections 1.3.2.1 and 1.3.2.2 seeks to provide explanations for the observed cross-linguistic variation in the basic unmarked order in each language, specifically, that the orders are not distributed evenly among the world's languages, and while there are many different orders when all modifiers are postnominal, there is only one commonly-occurring order when all modifiers are prenominal. The theories discussed in Section 1.3.2.2.1 that seek to explain why these asymmetries are found rely on a clear-cut distinction between orders that are attested versus unattested. They face the challenge that as more data comes to light, the attested/unattested status of a given order can change, which can undermine the theory. By contrast, the research covered in Section 1.3.2.2.2 focuses on the relative frequencies of different NP orders rather than their binary attestation status, thereby offering a more stable empirical foundation that is less susceptible to revision as new data emerges. However, this tends to focus on providing principles which underlie the observed relative frequencies rather than offering fully independently motivated reasons for the observations. Lastly, we covered homomorphism. This theory is independently motivated and well supported by experimental evidence. However, its symmetrical nature means it cannot explain the observation that when modifiers are prenominal they tend to be strictly ordered but postnominal modifiers exhibit more variation in their order. This asymmetry may reflect deeper cognitive or communicative principles that challenge the explanatory power of existing linguistic frameworks. To our knowledge, there are no published accounts that fully explain this asymmetry in a way that is both independently motivated and theoretically satisfying.

One specific area of asymmetry in the distribution of orders of NP elements is the relative ordering of adjective and noun. We conclude this section with a review of the literature about the possible reasons for this.

1.3.2.4 Order of Adjective and Noun within a NP

Analysis of the data considered by Dryer (2018) that we discussed in Section 1.3.2.2.2 reveals that 75% of the languages in his sample have postnominal adjectives: this is actually slightly more than the 73% of languages that are homomorphic. This striking preponderance of postnominal adjectives is illustrated in Figure 8.

Figure 8 NP Orders That Have Prenominal Versus Postnominal Adjectives



Note: Shows adjusted frequency data from Dryer (2018). Letters prefixing each NP order are those used by Cinque (2005). Orders with postnominal adjectives are shown in boldface.

This observation about the world's spoken languages is mirrored in Culbertson et al.'s (2020) silent gesture experiment discussed in Section 1.3.2.3, where participants expressed a strong preference for creating noun phrases where the adjective occurs after the noun. As discussed by Culbertson et al. (2012), a possible explanation for the higher proportion of postnominal adjectives (both typologically across the world's languages, and in artificial language learning experiments) is Kamp and Partee's *Head Primacy Principle* which is as follows:

> In a modifier-head structure, the head is interpreted relative to the context of the whole constituent, and the modifier is interpreted relative to the local context created from the former context by the interpretation of the head. (Kamp & Partee, 1995 p.161)

This means that (as noted by Dryer, 2018 – see Section 1.3.2.2.2) for those adjectives which require their noun to be known in order to be interpretable, for example, gradable adjectives such as "tall", there is a computational cost if the adjective is presented first. This is because it has to be stored in a buffer until the related noun is perceived. In languages with postnominal adjectives, it may be that people have regularised the Head Primacy Principle, applying it to all adjectives rather than just to gradable adjectives. (No similar preference for postnominal numerals or demonstratives derives from the Head Primacy Principle since these modifiers do not rely on the meaning of their noun for their own interpretation: for example, the concept of threeness stands alone.)

In addition to processing considerations, adjective position can also reflect semantic interpretation. Although adjectives in English overwhelmingly appear in prenominal position, a small subset–typically those denoting temporary or context-dependent properties–can occur postnominally. Bolinger (1967) observed that adjectives such as "visible" may appear after the noun when interpreted as stage-level predicates, as in "the stars visible were faint", where "visible" refers to stars that happened to be visible at that moment. In contrast, "the visible stars were faint" uses "visible" as an individual-level predicate, referring to stars that are inherently bright enough to be seen. This postnominal use is rare in English and largely restricted to a small set of adjectives with context-sensitive meanings, making it a marked and exceptional construction rather than a productive pattern.

Elicitation work by Rubio-Fernandez et al. (2022) on American Sign Language (ASL) lends support to this explanation. In their study, participants were asked to use ASL to describe a target image in contrast to three foil images. The experimental design was intended to elicit a noun and an adjective in the participants' responses. They found that colour and material adjectives (such as "red" and "cotton") were more likely to be produced prenominally, as compared to gradable adjectives (such as "big") which were distributed more evenly between pre- and postnominal position. They hypothesise that this is because colour and material adjectives denote absolute properties that do not require further context for interpretation, so it can be beneficial to produce them before the noun if they are particularly salient, thereby aiding comprehension by the interlocutor. In contrast,

scalar adjectives need to be interpreted in relation to the noun that they describe so there is no communicative benefit in producing them prenominally.

The relatively high proportion of postnominal adjectives across languages may be explained by the Head Primacy Principle, which holds that adjectives relying on the noun for interpretation are more likely to appear postnominally. At the same time, there is a competing pressure for salient, absolute adjectives–such as those denoting colour or material–to precede the noun, as this can facilitate faster comprehension. We suggest that it is the interaction between these pressures along with other factors–including extralinguistic determinants such as geopolitical dominance and the relative "success at survival, propagation, and colonization" (De Lacy, 2006, p. 351) of users of languages favouring pre- or postnominal adjectives–that has shaped the cross-linguistic distribution of adjective position observed today. While these pressures help explain the overall typological pattern, it is important to note that Figure 8 shows only the basic order for each language. As discussed in Section 1.3.2.2, however, many languages also permit adjectives to appear in the alternative position–that is, prenominally in languages where the basic order is postnominal, and postnominally where the basic order is prenominal.

This concludes our discussion of previous research about the order of elements within NPs in spoken languages. Next, we introduce large scale corpora of natural language production.

1.3.3 Use of Spoken-Language Corpora to Investigate the NP

While the previous section focused on typological and experimental findings, large-scale spoken language corpora offer another valuable source of evidence about NP structure. Some of these corpora have been annotated for grammatical category, such as Zeman et al.'s (2020) collection of 150 corpora annotated using the Universal Dependencies framework, covering 90 languages. Although these provide the opportunity to analyse the lengths of NPs and the frequency of different orders of elements within them, there appears to be little previous spoken language corpus research into the order of elements within the NP. (The fact that English has a strict word order (Huddleston, 1984) may explain the lack of research into NP order using corpora of English. Instead, other corpus-based research into English has focused on examining particular examples of NP constructions; see, for example, Keizer, (2007).)

However, there has been some corpus-based investigation into the length of NPs: for example, as part of their work investigating homomorphism discussed in Section 1.3.2.3, Culbertson et al. (2020)

examined nouns with modifiers in five large corpora of English that had been annotated for grammatical category by Zeman et al. (2020). Two of these corpora included some spoken English whereas the other three consisted entirely of written English. Culbertson et al. found that in general, these English NPs were short: of all NPs that had at least one modifier (adjective, numeral or demonstrative), 83% had one modifier, 14% had two modifiers and only 3% had three or more modifiers. Similarly, as part of work focusing mainly on Polish Sign Language (PJM) that is discussed further in Section 1.3.4.7.2, Rutkowski et al. (2015) investigated the lengths of NPs in a corpus of spoken Polish (www.nkjp.uni. lodz.pl/spoken.jsp). They found that out of a total of 10545 nouns, 71% had no modifiers, 24% had one modifier, 3.8% had two modifiers and 0.2% had three modifiers[9]. They did not find any nouns with more than three modifiers. These figures for Polish are very similar for those observed by Culbertson et al. for English, as illustrated in Table 1, despite the fact that the Polish data consisted entirely of speech whereas the English data was a mix of spoken and written language.

Table 1 Comparison of the Length of NPs With a Noun and at Least One Modifier in Polish and English

Language/ Length of NP (words)	**Polish[10] (%)**	**English (%)**
2	85	83
3	13	14
4 or more	0.7	3

Note. From Rutkowski et al. (2015) and Culbertson et al. (2020).

All the work described above has focused on spoken languages. We next turn to research investigating NPs in signed languages.

1.3.4 NPs in Signed Languages

We now consider NPs in signed languages: first, we provide some background regarding sign language structure, then look at differences between signed and spoken languages regarding the production of simultaneous or sequential constructions. After that, we examine means of identifying NPs and their elements in signed languages, and then we discuss social factors which

[9] There appears to be a discrepancy in the figures in this paper because the breakdown by number of modifiers does not sum to the total number of nouns – 30 nouns are not accounted for.

[10] Because of the discrepancy in the figures in Rutkowski et al. (2015) described in the previous footnote, these percentages do not total 100.

have been found to influence how signers produce BSL. Having established this context, we review existing literature investigating NPs in signed languages.

1.3.4.1 Background to Sign Languages

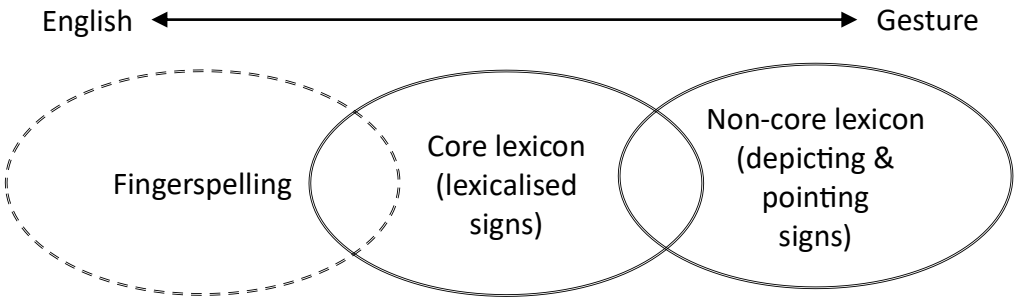
Since the 1960s, sign languages have been recognized as fully-fledged natural languages, each with its own grammatical structures and linguistic complexity (McBurney, 2012). Like spoken languages, sign languages are not generally mutually intelligible. For instance, British Sign Language (BSL) – the sign language used by deaf people in the UK and the central focus of this study – and American Sign Language (ASL) are distinct and unrelated, despite both being used in English-speaking contexts (Schembri et al., 2024). However, historical connections have shaped similarities among some sign languages. For example, as discussed in Schembri et al. (2010), BSL, Australian Sign Language (Auslan), and New Zealand Sign Language (NZSL) have a high degree of lexical overlap, and BSL and Auslan are often reported by signers to have no major grammatical differences, although this has not yet been systematically investigated. Schembri et al. note that the similarity between these languages is due to the spread of British deaf education practices and personnel to Australia and New Zealand during the nineteenth century. Despite their diversity, sign languages tend to exhibit common structural features such as the use of the space around the signer's body for grammatical purposes, a high degree of iconicity, and contact with the surrounding spoken language (Brentari, 2010; Meier, 2009). These shared characteristics support drawing meaningful parallels across sign languages and hypothesizing that similar linguistic principles may apply. As research focused specifically on British Sign Language is relatively scarce, we will, where useful and appropriate, draw on studies of other sign languages to inform the discussion and support the development of our analysis.

Having outlined the broader linguistic landscape of signed languages and situated BSL within this typological context, we now turn to a more detailed examination of the BSL lexicon.

1.3.4.2 Categorising Signs in the BSL Lexicon

A well-established analysis of the lexicon of BSL involves categorising it into two primary components: the native lexicon and the non-native lexicon. This framework, illustrated in Figure 9, is adapted from Johnston and Schembri (2007, p. 158) for Auslan. As discussed in the previous section, this adaptation is appropriate since BSL and Auslan are closely related languages.

Figure 9 Model of the BSL Lexicon



Note: Based on Johnston and Schembri (2007, p. 158).
– – – – – signifies non-native lexicon; ——— signifies native lexicon.

In this approach, the non-native lexicon (indicated with a dashed oval in the figure) is viewed as consisting of borrowings from the surrounding spoken language, English, in the form of full or partial *fingerspellings* of those words, using a manual version of the Roman alphabet (Sutton-Spence, 1999). The native lexicon (represented by solid ovals in Figure 9) is considered to be further sub-divided into the *core lexicon*, comprising all signs that could be listed in a dictionary such as BSL Signbank (Fenlon et al., 2014; see Section 2.2.1.1), plus the *non-core lexicon*, containing partly-lexical signs known as depicting signs (discussed in Section 1.3.4.2.1), plus pointing signs (covered in Section 1.3.4.2.2). However, it has been noted that in natural production, signers often incorporate gestural elements to a greater or lesser degree into a given sign, in a process dubbed *lexicalisation/de-lexicalisation* (Cormier et al., 2012). This means that in practice, there is a continuum between fully and partly lexicalised signs.

In more recent work (on ASL), Lepic (2019) takes an alternative usage-based approach, extending the concept of a continuum between fully lexicalised and partly lexicalised signs to argue that there is no sharp division between our mental representations of the lexicon and the grammar of ASL. He observes that a given instance of sign language production may contain elements that are structured (i.e., grammatical) plus elements that are holistic (i.e., retrieved from our linguistic knowledge), and that this may vary within a single communicative event, over time and among different signers. Given the structural features commonly shared across sign languages (see Section 1.3.4.1), it is likely that Lepic's analysis is also applicable to BSL. However, in this thesis we use the analysis shown in Figure 9 because the BSL Corpus is a central data source for this project, and existing corpus annotations that we rely upon use conventions that are based on this analysis, such as a prefix of FS: for fingerspellings, DS: for depicting signs and PT: for pointing signs (Cormier et al., 2017). However, we recognise the continuum proposed by Lepic, and note that this presents

ongoing challenges for assigning grammatical category to individual signs throughout the annotation process.

Having outlined the structure of the BSL lexicon and the distinction between core and non-core elements, we now turn to a closer examination of the constructions found within the non-core lexicon. These include depicting signs and pointing signs, which present challenges for traditional grammatical categorisation due to their partly lexicalised and highly iconic nature. In the following sections, we explore these constructions in more detail, beginning with depicting signs, and consider their role in the formation of noun phrases and in the broader grammar of signed languages.

1.3.4.2.1 Depicting (Classifier) Signs in Signed Languages

We briefly touched upon depicting signs (Liddell, 2003) (also called *classifier signs*) in the previous section. Here, we provide more detail plus an example. Depicting signs are not fully lexical: they include (more or less) iconic classifier handshapes such as in Figure 10. (This and other examples in this document are taken from the BSL Corpus (Schembri et al., 2013); see Appendix A for details of its annotation conventions: specifically, depicting signs are annotated with the prefix DS:. Background to the corpus is provided in Section 2.2.1.) In Figure 10, each hand's bent-finger configuration indicates a seated person. The signer is representing two people sitting opposite one another, on one side of the signing space, as part of a larger description of a train carriage. As noted by Schwager and Zeshan (2008), it is not possible to assign a single grammatical category to such signs: in Figure 10 the sign includes nominal elements (the people) plus verbal elements (the act of sitting in a particular configuration).

Figure 10 LN04c Depicting Signs Showing Two People Sitting Opposite One Another

Similarly, Rubio-Fernandez et al. (2022) observe that in ASL, the depicting sign in (44) can be interpreted as either an adjective (in English: “the sharp knife”), or a predicate (in English, “the knife is sharp”).

(44) KNIFE DS:SHARP-KNIFE (ASL, based on Rubio-Fernandez et al., 2022 p. 3)

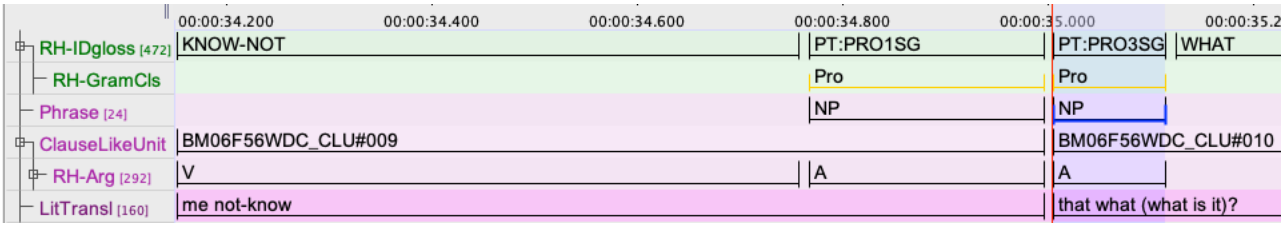
While depicting signs illustrate how signers can convey complex, often simultaneous, semantic content through partly lexicalised forms, pointing signs offer another key mechanism for reference and structure in signed discourse. Like depicting signs, pointing signs resist straightforward grammatical categorisation, yet they play a central role in identifying referents and marking locations. In the next section, we examine the various functions and forms of pointing signs, and consider how they are annotated and interpreted within the BSL Corpus.

1.3.4.2.2 Pointing Signs in Signed Languages

Pointing signs are a central feature of signed discourse, used to identify referents that are either physically present or contextually established. They can be categorised as pronominal (referring to people or objects), locative (indicating spatial locations) or determining (occurring immediately before or after a noun, with prosodic cues indicating that it forms a cohesive unit with the noun). However, these categories often overlap in practice, and considerable ambiguity can arise in interpretation (Cormier et al., 2013): for more on this, see Section 2.2.2.3.3. Appendix A.2 outlines the annotation conventions for pointing signs in the BSL Corpus. Although pointing signs form part of the non-core lexicon (see Figure 9), they can become lexicalised. One example is the BSL sign UP (https://bslsignbank.ucl.ac.uk/dictionary/words/up-2.html), which likely originated as a pointing sign but now functions as a lexical item.

Demonstrative pointing signs, translated as “this”, “that”, “these” or “those” in English, are not considered to form a distinct grammatical category in BSL. Instead, in the BSL Corpus they are categorised based on their syntactic position: as a determiner if they occur before or after a noun and form a prosodic unit with it (see Figure 33 PT:DET TIME), or as a pronominal if there is no explicit accompanying noun, as in Figure 11 (Cormier et al., 2013). (This and similar figures display relevant tiers from the BSL Corpus annotation tool ELAN, as described in Section 2.2.2.1. Only tiers containing annotations pertinent to each example are shown: In this case, tiers related to the left hand and to mouthing are not included. In each case, the caption includes a free translation of the utterance.)

Figure 11 BM06c Example of Demonstrative Point Annotated as Third Person Singular Pronoun; “I don't know what it/that is”



Pointing signs produced by the dominant hand can also be directed towards constructions that use the non-dominant hand, called *buoys* (Liddell, 2003). Buoys are held in place whilst the dominant hand continues to sign, serving to anchor the discourse. There are various types of buoy, as described in Liddell et al. (2007): the key types of buoy of relevance to NPs are the *list buoy* and the *pointer buoy*. List buoys enumerate a series of items, allowing the dominant hand to point to specific elements (e.g., the second item in the list, as in Figure 30). This point has a pronominal function (Liddell et al., 2007) and as such, forms a NP. Pointer buoys entail the non-dominant hand pointing to a location that is salient in the discourse whilst the dominant hand continues signing. This location may correspond to an NP.

Pointing signs, then, while frequently employed and functionally diverse, resist straightforward grammatical categorisation but play a central role in identifying referents and marking locations within signed discourse.

1.3.4.3 Carrier-Attribute Constructions in Signed Languages

A further structural feature of signed languages relevant to NP analysis is the carrier–attribute construction. Signed languages tend to lack a copula; instead, a carrier and an attribute are juxtaposed, as in example (45), from the Auslan Corpus. (As noted above, Auslan is historically related to BSL).

(45) FAMOUS STORY (Free translation: “The story is well-known”) (Johnston et al., 2019, p. 78)

When used as a standalone construction, this forms a verbless attributive clause where FAMOUS is an attribute outside of the NP (rather than functioning as an adjective within the NP) and STORY is a carrier noun. However, if the construction is part of a larger utterance, such as that whose translation into English is given in (46), it forms a NP consisting of the adjective FAMOUS plus the carrier noun STORY.

(46) "The well-known story called 'The hare and the tortoise'" (Johnston et al., 2019, p. 78)

Johnston (2019 p. 8) notes that the carrier can often be omitted, as in (47), again from the Auslan Corpus: the carrier is either recoverable from the context or else supplied by the conversational partner.

(47) ALIVE "(You're) alive!"

Next, we turn to the issue of simultaneity and sequentiality, and examine how these modality-specific features further affect the identification and analysis of NPs.

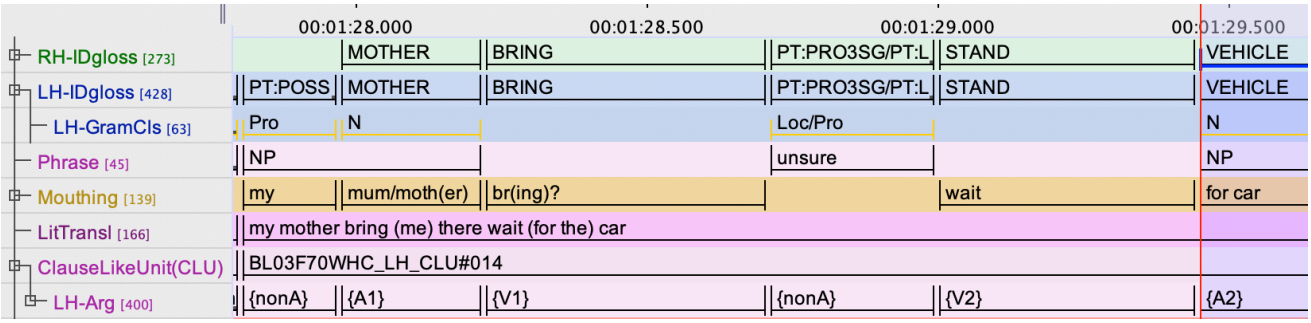
1.3.4.4 Simultaneity and Sequentiality in Signed and Spoken Languages

As a broad generalisation, in spoken languages, words are produced one after another in sequence. Vermeerbergen et al. (2007) note that although speech also includes simultaneous elements of intonation and co-speech gesture, there is more scope for simultaneity in sign language. This is because signers can use both hands plus other articulators such as the mouth and other elements of the face (including eyebrow raising/lowering and head tilts) plus the torso (including body leans), either working together to produce a single sign, or else working in parallel to produce simultaneous constructions. In terms of mouth activity, a signer may produce what is known as a *mouthing* (where they (silently or with voice) produce all or part of a word from the surrounding spoken language), or a *mouth gesture* (which does not originate from the surrounding spoken language). This mouth activity may occur in parallel with the production of the hands and other articulators. Overall, therefore, the greater simultaneity in signed languages compared to spoken languages can be viewed as a consequence of the difference in modality (Woll, 2007).

This simultaneity means that we may find constructions where multiple NP elements are produced at the same time, forming a single prosodic word. This may involve just the hands; for example, Crasborn (2011) notes that a noun with a co-occurring pointing sign has been observed in both Israeli Sign Language (ISL) and Dutch Sign Language (NGT). Or a mouthing of one or more English words may be produced alongside a manual sign, as in the highlighted portion of Figure 12, where

the signer produces the manual sign VEHICLE at the same time as mouthing “for car”. As noted by Johnston and Schembri (2007), this simultaneity can be problematic when attempting to determine a linear constituent order (and therefore, linear order within NPs) in sign languages.

Figure 12 BL03c Example of Simultaneous Mouthing and Manual Sign; “My mother brought me to stand outside as we waited for the car.”



Depicting signs can be viewed as consisting of simultaneous elements, as discussed in Section 1.3.4.2.1. Additionally, some lexical nouns can be modified to incorporate some types of adjectival information. For example, a signer can use bigger hand and arm movements to produce LARGE-BOX as opposed to simply BOX (Sutton-Spence & Woll, 1999, p. 52). (This adjective incorporation is not possible for all types of adjectives: for example, nouns cannot be modified in this way to incorporate colour.) Furthermore, certain numerals can be incorporated into signs including noun signs, so that the noun uses the handshape of the numeral to form constructions such as THREE-POUNDS (Sutton-Spence & Woll, 1999). This simultaneous morphology results in fewer standalone adjectives and numerals being produced in BSL than in English (Sutton-Spence and Woll, 1999). As with depicting signs, it is not possible to assign a single grammatical category to these constructions. We consider further the issue of determining the grammatical category of signs in the next section, and our approach to dealing with simultaneity in our data is detailed in Sections 2.2.2.4.5 and 2.2.2.5.

1.3.4.5 Identifying NPs in Signed Languages

In Section 1.3.1, we addressed general issues concerning the identification of grammatical categories. We now turn to this topic in the context of signed languages. Dryer (2007b) warned against assuming that the grammatical cate gory of a word in one (spoken) language corresponds directly to that of its translation in another–a warning equally relevant for sign languages. As discussed in Section 1.3.4.1, sign languages often share structural features that support cross-linguistic comparison. However, these similarities do not justify uncritical assumptions of grammatical equivalence. Schwager and Zeshan (2008) highlight that researchers have at times

translated a target sign into a well-documented sign language such as ASL, and then assigned it the grammatical category used in that language. This approach again risks misrepresenting the grammatical structure of the target language.

When researching the NP in PJM (see Section 1.3.4.7.2), Rutkowski et al. (2015) observed that identifying grammatical category in sign languages can be particularly challenging. We noted in Section 1.3.4.2.1 that depicting signs cannot be assigned a single grammatical category because they encapsulate a state of affairs that includes, for example, both nominal and verbal elements, as in Figure 10. The fact that in natural utterances, people produce signs that are on a continuum between fully lexical and fully non-lexical (Cormier et al., 2012; Lepic, 2019) means that one cannot necessarily categorise a sign as either (1) lexical (with a grammatical category) or (2) partly- or non-lexical (without a straightforward single grammatical category). Even signs that are plausibly lexical may exhibit ambiguity in terms of grammatical category.

In this section, we review past research into how to identify grammatical category in sign languages. Not all of this research has been transparent about whether it relates just to lexical signs or whether it includes partly- or non-lexical signs: where this is explicit or it can be inferred, we include details of this in our discussion.

Schwager and Zeshan (2008) observe that the grammatical category of a sign can often only be determined when the sign is used in context. This can also be the case for many spoken languages, as noted in Section 1.3.1. Similarly, Nadolske and Rosenstock (2007 p. 42) discuss the fact that when examining ASL, they found examples of the lexical sign DEAF with grammatical category adjective but two different syntactic functions, depending on the context: they observed it functioning as an adjective within an NP in;

(48) PT:PRO1SG MEET DEAF BOY - “I met a deaf boy”;

and as an adjectival predicate (not part of an NP) in;

(49) PT:PRO1SG DEAF - “I am deaf”.

They also observed it functioning as a noun in;

(50) DEAF CL:SIT-ALONG-RIGHT-SIDE-OF-TABLE - “The deaf people sat along the right side of the table.”

Johnston (2012) also observed that many signs in Auslan can function in more than one grammatical category, so that one cannot necessarily infer a sign’s grammatical category from its form. He reported that annotators disagreed on the grammatical category of 5-10% of all annotated signs in

the Auslan Corpus, due to varying analyses of the text. This phenomenon may highlight a key difference between signed and spoken languages, though the underlying reasons remain under-researched. We speculate that it could be attributed to the inherent ambiguity and flexibility in the structure of sign languages, and perhaps additionally, because of their minority status and because they have no written form.

Meir (2012) provides a useful overview of possible approaches to classifying all types of sign (lexical and non-lexical). She notes the difficulty of using morphological criteria such as plural markings to identify nouns because she argues that signed languages have little inflectional morphology.

Schwager and Zeshan (2008), discussing lexical signs, observe that one cannot rely on syntactic criteria for determining grammatical category, since signed languages tend to have a relatively free sign order. Instead, they propose a methodology based on semantic classes such as *entity* and *event*, since these are language-independent as they are based on general human cognition. However, they note that even when the meaning of the utterance is clear, sometimes it is still not possible to assign a grammatical category to a particular sign. They provide the following example from Indo-Pakistani Sign Language;

(51) PT:PRO1SG WORK A-LOT (adapted from Schwager & Zeshan, 2008, p. 513).

This could be interpreted as;

(52) "I work a lot" (WORK is a verb); or

(53) "My work is a lot" (WORK is a noun).

A semantic approach, based on the meaning and function of each sign in context, has also been used when annotating the Auslan Corpus for grammatical category, as noted by Johnston et al. (2015).

Despite these difficulties, work has been undertaken in several signed languages to identify patterns of morphological difference between nouns and verbs, as discussed by Ribera-Llonc et al. (2019). They found that noun signs tended to be shorter and were more likely to be accompanied by mouthing, while verb signs were more often associated with mouth gestures. They highlight early work by Supalla and Newport (1978), who proposed that in ASL, certain noun–verb pairs are distinguished by systematic movement differences: nouns are more likely than verbs to be produced using a restrained, repeated movement. Although Supalla and Newport did not describe their methodology, their findings appear to be based on contrasting citation forms of signs. Johnston

(2001) investigated whether a similar pattern occurs in Auslan (an unrelated sign language; see Section 1.3.4.1) and found evidence of such distinctions, but in a more limited set of signs. These were mostly, though not exclusively, iconic signs involving inherently reversible actions, such as DRAWER and OPEN-DRAWER/CLOSE-DRAWER. Johnston argued that examining a broader range of nouns and verbs in more naturalistic, connected signing–such as that found in corpus data–would likely yield different results.

Given that our data includes signers from a range of backgrounds, it is important to consider how social factors may affect BSL use and the structure of NPs. This is the focus of the next section.

1.3.4.6 Social Factors Affecting the Use of BSL

In this section, we consider how social factors influence BSL use and NP structure, starting with patterns of language acquisition. Only a small minority of deaf individuals–estimated at between 5% and 10% in the United Kingdom–acquire BSL from their relatives (Mitchell & Karchmer, 2004; Schembri et al., 2013). These individuals, known as *native signers*[11], are typically the children of deaf parents and are exposed to sign language from birth. Although this might suggest a stable, multigenerational transmission of BSL, we note that native signers are not directly analogous to native speakers of spoken languages. As Johnston (2022) and Fenlon and Hochgesang (2022) observe, the parents of native signers may not themselves have acquired sign language from birth, meaning that linguistic transmission is often discontinuous.

Despite their minority status, native signers have historically been the focus of much sign language research. This is often due to the desire to avoid confounds associated with late or second language (L2) acquisition. For example, Boudreault and Mayberry (2006) found that adults exposed to ASL after early childhood performed less accurately on grammaticality judgement tasks than those exposed from birth. Consequently, corpora of signed languages often prioritise native signers and *early learners* of BSL–people who do not have deaf relatives but who learned to sign by age seven–thereby excluding the majority of signers who acquire BSL later in life. The exclusion of late learners may also help explain why Schembri (2024) notes that, in almost all corpus-based research into sociolinguistic variation in BSL and Auslan that included whether participants were native signers or not, no significant difference has been found so far based on native signer status. The exception was

[11] The term *native* is problematic, because it can have connotations of language purity (Brown, 2025), and can be othering to signers with a later age of acquisition (Hou, 2024). We acknowledge this issue: see Birkeland et al. (2024) for a discussion. We follow their suggestion to retain the terminology whilst highlighting the issues surrounding it.

Stamp et al.'s (2014) study of regional variation, which found that native signers tended to use more traditional variants of signs than non-native signers. This may be due to the fact that these traditional signs were generally acquired by native signers from their older relatives, who tended to use these traditional signs. As Johnston (2022) argues, and as is now widely acknowledged in the field, including late learners of BSL is crucial for understanding how BSL is used across the wider deaf community.

In addition to variation among signers, BSL is also shaped by its contact with spoken languages. Sign language communities exist within predominantly spoken language environments, leading to extensive contact between signed and spoken languages. This contact results in lexical items from the spoken language being borrowed into the signed language, resulting in fingerspellings (Sutton-Spence, 1999; see Section 1.3.4.2) and mouthings (Proctor & Cormier, 2023). Syntactic structures may also be influenced: Bowern (2006) has documented syntactic borrowing in spoken language contact situations, and Mantovan (2017) suggests that similar processes likely occur in signed–spoken language contact. The extent of such influence may vary depending on a signer's language background, further underscoring the importance of including non-native signers in linguistic research.

Other social factors including region, gender, age, ethnicity and social class have been shown to influence how people use language (either spoken or signed) (Lucas, 2001). Previous studies have found regional variation in BSL, in the lexicon (e.g., Stamp et al., 2014) and in fingerspelling (Brown & Cormier, 2017; Sutton-Spence et al., 1990). Both regional and gender differences have been found in mouthing (Proctor & Cormier, 2023; Rentelis, 2011). The age of a signer may affect their BSL use because as Stamp (2013) notes, age is a proxy for educational experience. This is because of changes in national education policy over time: throughout most of the twentieth century, BSL was suppressed in the classrooms of residential schools for deaf children under a policy known as *oralism* (though it continued covertly in the wider school environment). Later, teaching in mainstream schools using Sign-supported English (SSE)[12] and then BSL interpretation or bilingual BSL/English was introduced (Schembri et al., 2013).

We now explore previous research into the structure of the NP in other signed languages besides BSL. As discussed in Section 1.3.4.1, sign languages share some structural features that support

[12] SSE combines signs with spoken English, using the grammar of spoken English rather than of BSL (Rowley & Cormier, 2024).

cross-linguistic comparison, which can offer valuable insights when examining BSL. At the same time, as Johnston et al. (2007) observe, comparing findings across different languages can be complicated by the use of varying methodologies, differing levels of descriptive detail, and the diverse theoretical orientations of researchers. These considerations frame our examination of how noun phrases have been analysed in individual signed languages.

1.3.4.7 The NP in Particular Sign Languages

Based on general observation, Sutton-Spence and Woll (1999) note that in general, adjectives follow the noun in BSL, although this is affected by context, including factors such as what is being focused: they state that if a signer wished to refer to hot food as opposed to cold food, they would sign HOT FOOD rather than FOOD HOT. (Sutton-Spence & Woll 1999:52). They also observe that adjectives within an NP can also precede the noun or be incorporated into the noun.

The NP has been investigated in more depth in a few signed languages to date: American Sign Language (ASL), Polish Sign Language (PJM), Italian Sign Language (LIS), Taiwanese Sign Language (TSL) and Hong Kong Sign Language (HKSL). The findings of these studies are summarised here. As noted by Meir (2012), previous sign language researchers have generally not justified the basis for their categorisation of elements into particular grammatical categories. This is the case for the research described here, with the exception of the Polish study.

1.3.4.7.1 The NP in ASL

Research into specific aspects of NPs in ASL includes the investigation of adjective position by Rubio-Fernandez et al. (2022), discussed in Section 1.3.2.4. While their study focused on adjective placement, MacLaughlin (1997) provided a more comprehensive analysis of NP[13] structure, based on the intuitions of a single deaf native signer[14]. MacLaughlin found that adjectives may appear either before or after the noun: in prenominal position, they function attributively, whereas in

[13] MacLaughlin used the terminology *determiner phrase* (DP) rather than *noun phrase* (NP). As discussed in Footnote 1, these terms are equivalent for the purpose of this thesis.

[14] Although her study involved direct collaboration with an ASL consultant, MacLaughlin is a hearing non-native signer, and her presence may have influenced her consultant's responses. In light of Schembri et al.'s (2013) discussion of filming protocols for the BSL Corpus–where only deaf signers were present during data collection to minimise the risk of contact signing, that is, signing with more influence from the surrounding spoken language–there is reason to question whether the data she elicited may have contained more English influence than if the consultant had been working with a deaf interlocutor.

postnominal position they are predicative. She also observed that numerals in ASL can occur either before or after the noun.

Similarly, Neidle, Kegl, MacLaughlin, Bahan, and Lee (2000) claim that determiners in ASL may be either prenominal or postnominal. They state that if determiners are prenominal, they have a definite interpretation whereas if they are postnominal they may be either definite or indefinite.

Neidle and Nash (2012) propose a canonical order for ASL NPs, shown in (54), although they do not provide empirical evidence for this claim. This order is homomorphic (see Section 1.3.2.3) and is one of the orders that is predicted by Greenberg's Universal 20 and by Cinque (2005).

(54) [(Dem) (Num) (Adj) Noun]

1.3.4.7.2 The NP in PJM

Rutkowski et al. (2015) investigated PJM using a subset of data from a corpus of this language consisting of 4287 tokens of connected signing taken from narrative retellings by ten deaf participants who have PJM as their first language. They were transparent in their methodology, noting that no clear morphosyntactic rules exist for identifying grammatical category within PJM and explaining that therefore their annotators used intuition about each sign in context in order to do this.

They started by identifying nouns within their corpus, then investigated modifiers of these nouns. Since they had a relatively large dataset, they were able to investigate the frequency of different lengths of NP containing a noun. They found that 887 (18%) of their sign tokens were nouns and the rest were from other grammatical categories. 682 (77%) of these nouns appeared in isolation in the NP, a further 183 (21%) had a single modifier; 20 (2%) had two modifiers and 2 (0.2%) had three modifiers.

They observed a wide variety of different arrangements of elements within NPs in PJM, the most common being NPs consisting of a noun and an adjective, and the second most common being NPs consisting of noun and a numeral. They identified that adjectives were most often postnominal: [Noun Adj] occurred in 66% of examples (*N* = 57), whereas [Adj Noun] was found 34% of the time (*N* = 30). In contrast, numerals were more often found prenominally: they found [Num Noun] order 66% of the time (*N* = 51) and [Noun Num] order in 34% of the examples (*N* = 26). They also investigated instances of a noun plus an indexical pointing sign (which they note can fulfil the grammatical function of a demonstrative or a determiner in PJM), finding that as for numerals, they

were more often prenominal: the point preceded the noun on 64% of the occasions (*N* = 12) and that the noun preceded the point 36% of the time (*N* = 8).

Although the authors examined the data through the lens of Cinque's (2005) model, only 22 constructions in the dataset contained a noun with two or three modifiers. This limited their ability to draw robust conclusions about the relative ordering of these elements. Consequently, the data does not permit a definitive assessment of whether PJM NPs conform to Greenberg's Universal 20, Cinque (2005), or to homomorphism.

Rutkowski et al. note that it could be suggested that the fact that they identified few NPs with more than one modifier means that PJM is a simpler language than spoken Polish. This claim runs counter to deaf PJM users' intuitions, so they investigated it by comparing their data to the corpus of spoken Polish conversation that we discussed in Section 1.3.3. Importantly, they did not compare PJM to a corpus of written Polish, because of the inherent differences between face-to-face sign language and the written form of spoken language. They used the same methodology as for their PJM analysis so that the results would be comparable and found a similar distribution in terms of lengths of NPs in both PJM and spoken Polish. They therefore conclude that this supports the claim that PJM's NPs are no less complex than NPs in spoken Polish. They also observed a difference in the order of noun and adjective, with spoken Polish having [Adj Noun] order in 81% of examples. This suggests that syntactic borrowing (see Bowern (2006) as discussed in Section 1.3.4.6) of the order of adjective and noun from spoken Polish to PJM has not occurred and further reinforces the understanding that PJM and spoken Polish are different languages.

1.3.4.7.3 The NP in LIS

The NP in LIS has been investigated by Mantovan (2017) using a corpus of personal narrative recordings from 162 native or near-native signers. She identified 2118 NPs within her data set that consisted of a noun plus one or more modifiers. Her corpus data tended to contain a maximum of one or two modifiers of each noun, so she extrapolated from the observed orders, as described in Section 1.3.2, to conclude that the most frequent overall order in the NP was as in (55).

(55) [(Num) Noun (Adj) (Dem)]

This contrasts with (56), which Mantovan reports as the unmarked order in earlier elicitation-based studies including Branchini (2015), Bertone (2009) and Brunelli (2011). According to her, these studies found numerals to occur postnominally: following adjectives but preceding demonstratives. In contrast, her corpus-based findings show that numerals are typically prenominal. Both (55) and (56) are homomorphic and consistent with Greenberg's Universal 20 and Cinque (2005).

(56) [Noun (Adj) (Num) (Dem)]

In spoken Italian, the dominant NP word order has demonstratives and numerals preceding the noun, with adjectives following it (Maiden & Robustelli, 2000). The comparison between LIS and spoken Italian thus presents a mixed picture. The position of demonstratives differs between the two languages, as LIS places them after the noun, while spoken Italian places them before. Regarding numeral position, Mantovan's corpus data (55) appears to match spoken Italian whereas elicited data from LIS (56) is different to spoken Italian. Adjective position, however, appears to be consistent across LIS and spoken Italian.

Mantovan also examined the effect of various linguistic and social factors on whether nominal modifiers appeared before or after their noun. The only factors that had a statistically significant effect were a) the presence of more signs between a modifier and the noun: in this case the modifier tended to be postnominal and b) signers with deaf relatives tended to produce more prenominal modifiers.

1.3.4.7.4 The NP in TSL

Zhang (2007) researched the structure of the NP in TSL. Although she did not provide a detailed account of her methodology, she noted that her findings were based on consultations with informants. According to these informants, many–but not all–possible orders of adjective, demonstrative, noun and numeral occur in TSL. Zhang concluded that all NP orders in TSL are linearisations of the base structure given in Figure 6, with the exception that demonstratives cannot occur postnominally when all three modifiers are present. She noted that the orders reported by her informants are all consistent with Greenberg's Universal 20 and can be derived using both Cinque's (2005) and Abels & Neeleman's (2012) approaches that we discuss in Section 1.3.2.2.1. Based in Zhang's analysis, we observe that since the orders that can be linearised from Figure 6 are, by definition, homomorphic to this structure, NPs reported to occur in TSL are homomorphic.

1.3.4.7.5 The NP in HKSL

NPs in HKSL have been explored by Tang and Sze (2002). They considered the relative positions of noun, numeral and a pointing sign used as a determiner, based on the intuitions of four signers. They concluded that constituent order within NPs in HKSL is as in (57)

(57) [Determiner Num Noun]

although they note that both the pointing sign and the numeral can also occur post-nominally. When the pointing sign is pre-nominal, they argue that it has a definite interpretation. When it is post-nominal, they posit that it can either be interpreted as a determiner, or as a locative point. They observed that eye gaze by the signer to a referent indicates that that referent is definite. In contrast, if the signer uses eye gaze to their addressee, they suggest that this indicates that the referent is indefinite and specific. The authors do not provide a full picture of how HKSL relates to Greenberg's Universal 20, Cinque (2005) or homomorphism since they do not discuss the position of adjectives within an NP. However, the relative position of determiner and numeral they propose does adhere to Greenberg's Universal 20 and the order of these two modifiers is homomorphic.

Having reviewed theoretical and typological accounts of NP structure, along with empirical findings from both spoken and signed languages, we now turn to the question of how such findings are obtained. The next section examines the methodological approaches used to investigate NP structure, focusing on how different data sources and elicitation techniques shape the kinds of claims that can be made.

1.3.5 Methodological Approaches

The studies reviewed so far have drawn on a range of methodological approaches, from elicited judgements to corpus analysis. In this section, we compare the strengths and limitations of different methods used to investigate signed language structure. One of the central methodological challenges in linguistics is determining how to obtain reliable evidence about language structure. Traditionally, much linguistic research, including the creation of reference grammars for specific languages, has relied heavily on native speaker/signer intuitions. In many cases, these intuitions are those of the researcher themselves, or of a small number of language consultants (see e.g., Vaux & Cooper, 1999). This approach has also been adopted in some sign language research on NPs. For instance, Zhang (2007) on TSL and Tang and Sze (2002) on HKSL both based their analyses on data elicited from a small number of informants, as discussed in Sections 1.3.4.7.4 and 1.3.4.7.5.

Similarly, for her doctoral research into ASL, MacLaughlin (1997) employed the judgements of one native signer (see Section 1.3.4.7.1). MacLaughlin observed that her choice of working with a single informant enabled her to study the grammatical system of that one individual and she remarked that if additional participants had been included, they may not have shared the same grammatical system, potentially clouding the results. However, she noted that the study of multiple informants would have enabled consideration of differences in dialect.

Neidle et al., (2000, p. 21) (whose authors include MacLaughlin) concurred with MacLaughlin's (1997) approach: although they advocated working with several individuals, they were not in favour of "random[ly] pool[ing]" the results because they felt this could mask the phenomenon being investigated. Despite the fact that MacLaughlin made it clear that her work was based on the intuitions of a single informant, this has not always been highlighted when her work is cited: see, for example, Quer et al., (2017, p. 519), which includes the following statement, with no mention of the fact that this relates only to one signer's use of ASL: "Many sign languages have also been shown to conform to [Universal 20] at varying degrees (cf. … MacLaughlin 1997 for ASL)." This has led to a tacit assumption amongst researchers that MacLaughlin's findings generalise to all ASL signers, and this is unlikely to be the case.

There are two key issues surrounding the use of native language user intuitions: (1) the fact that the consultants need to use introspection to make metalinguistic judgements and (2) the relatively small number of people consulted, as discussed in Schütze and Sprouse (2014). To address the first of these, researchers have created corpora of naturalistic language production. And to address the second, they have conducted larger-scale experiments using more participants. We discuss these methodologies in the next two sections.

1.3.5.1 Corpora

A linguistic corpus consists of a large collection of machine-readable texts, which may be in written, audio or video form, that has been collected for the purpose of linguistic analysis. Texts are chosen so as to ensure that a representative and balanced sample of the language in question is included (McEnery & Hardie, 2012). Halliday and Matthiessen (2004) note two key advantages of corpus data: Firstly, they consist of spontaneous utterances, which is important because people's actual language use can differ from their intuitions about language. And secondly, the fact that corpora are large datasets means that larger scale quantitative analysis becomes possible. Additionally, as Neidle et

al. (2000) note, corpus participants are unaware of the research questions that their contributions will address. This means that they cannot be influenced by these in the way that is possible when employing native speaker/signer intuitions.

Originally, corpora focused on written/spoken rather than signed language. The earliest corpora were not computerised: instead, the data was stored on index cards (McEnery & Hardie, 2012). Then tape recorders allowed speech recording, and the advent of video recorders enabled audio visual speech to be captured. As technology improved, it became possible to create sign language corpora, as discussed in the next section.

1.3.5.1.1 The Development of Sign Language Corpora

Historically, the development of sign language corpora has been closely tied to technological advances: the advent of video recording made it possible to capture sign language in a way that written transcription could not. However, while the annotation of spoken language corpora has increasingly benefited from automation and large-scale computational resources, this is not yet the case for sign language. Annotation of sign language data remains a predominantly manual process, which is both time-consuming and labour-intensive (Fenlon & Hochgesang, 2022). Moreover, the field still lags significantly behind spoken language research in terms of the volume of annotated material and the scale of computational power applied to annotation and analysis. This disparity presents an opportunity for innovation, as advances in annotation tools and computational methods tailored to sign language have the potential to significantly accelerate research and enrich our understanding of visual-gestural languages.

Differences in people's intuition about language as compared to their actual language use have been observed in the field of sign language syntax: For example, Oomen and Pfau's (2017) study of negation in NGT concluded that earlier observations of negation in NGT that had been based on elicited data were not supported by corpus evidence.

Corpora do have certain drawbacks. Although the aim when creating a corpus is to collect a representative sample of utterances, this can never actually be achieved since language is unbounded and any corpus is finite. Additionally, corpora can only be used to identify what constructions are found in a given language – they do not provide negative evidence (Carnie, 2013). Furthermore, corpora cannot generally be used to draw conclusions about constructions that are

rare in a given language (Den Dikken et al., 2007; Fenlon & Hochgesang, 2022). To address these issues, researchers have used native speaker/signer intuition tasks as described in the next section.

1.3.5.2 Judgement Tasks

In a judgement task, the experimenter provides a series of linguistic stimuli to each of their participants and asks them to provide a response for each. The participants may need to make a binary choice about whether the stimulus is *grammatical* or *acceptable*, or a more fine-grained response may be elicited, where the participant rates the *grammaticality*, *acceptability* or *typicality* of the stimulus on a Likert scale. While these terms are sometimes used interchangeably, they refer to distinct concepts. *Grammaticality* concerns conformity to syntactic rules; *acceptability* reflects how *natural* a sentence feels, which may be influenced by processing or context; and *typicality* relates to how expected or representative a sentence is in usage.

This experimental collection of judgements differs from the informal consultations with language users discussed in Section 1.3.5: Those consultations are typically less structured and more wide-ranging, which can elicit more nuanced responses and occasionally take the conversation in unexpected directions. Since it is not feasible to collect such detailed evidence from large numbers (perhaps scores) of language users, more focussed experimental tasks are undertaken as these are more scalable.

Historically, linguistic judgement tasks have been carried out face-to-face in a laboratory but advances in technology mean that researchers now have the option to conduct research online. As discussed by Rodd (2024), while there are many advantages to online studies such as the ability to recruit larger numbers of participants, who are more geographically spread, one disadvantage is that researchers cannot control the environment of a remote task in the way that we can in the laboratory. Rodd also notes that the hardware and internet connection available to remote participants is likely to vary widely. This means that it's important to test one's study on different device types (e.g., laptop, desktop, tablet, mobile phone) and operating systems, and on many different web browsers if the task is delivered via browser. She observes that the variety in participants' technology may introduce a confound into the experiment if researchers wish to study the variation in responses between participants because, for example, it may be that older participants' technological infrastructure is worse than that available to younger participants.

1.3.5.2.1 Instructions to Participants

In considering spoken-language judgement tasks, Schütze and Sprouse (2014) discuss what researchers are *not* looking for in terms of responses: we are not interested in whether a sentence contravenes any prescriptive grammatical rules, or whether it is comprehensible – since for example, many non-fluent language users' utterances are not grammatical but can nonetheless be understood. Neidle et al. (2000) note that this issue may be particularly prevalent for sign languages since there is such a high degree of variation in sign language production because of the diversity of the signing population, due, for example, to the large number of people with a later age of acquisition of a sign language. We are also not concerned about the truth/falsehood of a sentence or how likely it is that someone would utter it. Schütze and Sprouse also note that it is common for experimenters to suggest that participants imagine the written sentence being spoken aloud, in order to encourage the participant to consider what a native speaker would say.

Schütze and Sprouse (2014) also observe that in fact, research has shown that the instructions provided for a judgement task have relatively little effect, since experimenters cannot control how participants respond to them. They conclude that it is more fruitful for researchers to concentrate their efforts on creating their stimuli.

In terms of grammaticality judgement tasks for BSL, Fenlon (2019) conducted a small online experiment presenting 14 BSL students and six BSL teachers with a series of videos of BSL sentences and asking them "Do you think that the above order is typical of BSL?". The response options were *very typical*, *somewhat typical*, *neutral*, *somewhat atypical*, and *very atypical*.

1.3.5.2.2 Presentation of Sentences

Marty et al. (2020) investigated how different presentation formats affect task sensitivity in linguistic judgement tasks involving 100 written English sentences, recruiting around 100 participants for each combination of options. Their study included comparing the presentation of one sentence at a time (i.e., *single presentation*) with showing the sentences in theoretically-related pairs, so that each sentence in a pair forms a reference point for the other (i.e., *joint presentation*). They found that sensitivity was higher for the latter. For the joint presentation, they randomised the order in which each member of a pair was presented, to mitigate the effects of response bias, such as the strategy of always preferring the first member of a pair.

1.3.5.2.3 Response Scale

In addition to investigating the presentation of sentences, Marty et al. (2020) also considered various response scales that are used in judgement tasks. They compared a binary choice (e.g., Yes/No), a Likert scale (where the participant is presented with usually five or seven response options) and a continuous scale (with "virtually infinite" response options (p. 2)). They found that with joint presentation of sentences, either a Likert scale or a continuous scale, each with only endpoints labelled, were the most sensitive. They also noted that avoiding numerical labels on the response scale has the advantage of not involving participants in numerical reasoning – instead, spatial reasoning is used.

In summary, this section has explored various methodological approaches in linguistic research, highlighting the strengths and limitations of each. The development of linguistic corpora has provided more comprehensive and unbiased datasets, enabling large-scale quantitative analysis and reducing participant bias. However, corpora have their own limitations, such as the inability to provide negative evidence and the challenge of capturing rare constructions. Online judgement tasks offer a scalable alternative, allowing researchers to gather nuanced responses from a larger pool of participants, though they introduce variability due to differences in technology and environment. Combining these methods (see, for example, Pullum, 2017) can help balance their respective limitations, providing a more holistic approach to linguistic research. This section underscores the importance of methodological diversity when investigating language structure, especially for minority languages including signed languages.

1.4 Research Questions and Hypotheses

This project examines the overarching research question: What is the structure of the NP in BSL? We also examine the extent to which NPs in BSL are homomorphic, and whether NP order in BSL matches the cross-linguistic distributions of NP order predicted by Greenberg's Universal 20 and by Cinque (2005). Greenberg's Universal 20 and Cinque (2005) are concerned with the basic, unmarked order within the NP of each language. Therefore, we identify a dominant order in BSL NPs and compare this to these theories.

Our research questions are as follows:

RQ 1. What orders of elements within the NP occur in BSL, and at what frequency?

Previous research suggests that postnominal adjective placement is typical in BSL. Sutton-Spence

and Woll (1999) observe that BSL adjectives generally follow the noun (as noted in Section 1.3.4.7), and Dryer (2018) reports that approximately 75% of the world's spoken languages in his representative sample exhibit postnominal adjectives (see Section 1.3.2.4). These findings inform our expectations for BSL.

Hypothesis 1: Adjectives will follow the noun in our BSL data.

RQ 2. How well does the order in NPs in BSL match established predictions about the cross-linguistic distribution of NP order in different languages, such as Greenberg's Universal 20, Cinque (2005) and homomorphism (as discussed in Section 1.3.2.3)?

Hypothesis 2a: The dominant order in NPs in BSL is predicted by Greenberg's Universal 20.

Hypothesis 2b: The dominant order in NPs in BSL is predicted by Cinque (2005).

Hypothesis 2c: NPs in BSL are homomorphic.

RQ 3. How does signers' reported language background, reported experience of teaching BSL and their level of comfort with English (based on self-report as discussed in Section 2.5) affect the order of signs within their NPs?

Hypothesis 3a: Signers' reported language background will affect the degree of influence of English on their BSL. People who acquired BSL later in life will experience more influence of English than those who acquired it earlier. Therefore, they will tend to prefer NPs that are in English-like order.

Hypothesis 3b: Whether participants report having taught BSL may affect the ideologies that they hold about BSL structure. BSL teachers may be more likely to hold the ideology that since BSL and English are different languages, the order within the NP in BSL should be different to in English. Therefore, they will tend to prefer NPs that are not in English-like order.

Hypothesis 3c: Signers who have higher levels of comfort with English will experience more influence of English on their BSL. Therefore, they will tend to prefer NPs that are in English-like order.

Chapter 2 Method

2.1 Introduction

In Section 1.3.5, we discussed the advantages and disadvantages of different methodological approaches to investigating grammatical structures, including the use of corpora and large-scale judgement tasks. To operationalise our research questions, the current project combined use of the British Sign Language (BSL) Corpus (Schembri et al., 2013) and data collected using an online judgement task. We used the orders of elements within noun phrases (NPs) found in the BSL Corpus as a proxy for the orders that are found in naturally-occurring BSL as a whole. Our on-line judgement task allowed us to explore constructions that are rare or absent in our BSL Corpus data: this was particularly useful because such a large proportion (76%: see Section 3.1.1) of the NPs that we identified in the corpus contained only a single element. These cannot provide any evidence about the order of elements within NPs. This use of dual data sources is an example of *methodological pluralism* (McEnery & Hardie, 2012) which is becoming increasingly widely used across the language sciences.

All the data has been analysed using the statistical package R (R Core Team, 2024). Each of these two data sources are discussed in turn, followed by details of how each was analysed.

2.2 Use of the BSL Corpus

2.2.1 Introduction to the BSL Corpus

The BSL Corpus (Schembri et al., 2013), filmed between 2008 and 2010, is a key data source for this project. During development of the corpus, participants were chosen using non-random quota-based techniques with the aim of matching the proportions of the social factors age, gender, language background[15], region and ethnicity in the overall deaf community in the UK; for a discussion of some of these social factors, see Section 1.3.4.6. All participants in the BSL Corpus were deaf, and 95% said that they had been signing since at least the age of seven, with most of the remainder learning before age 12. Participants were selected to reflect recent changes in education policy in the UK (mentioned in Section 1.3.4.6), with approximately equal numbers of people

[15] A higher proportion of native signers (defined as in Section 1.3.4.6 as people who acquired BSL from birth from deaf relatives) were recruited than is present in the overall UK deaf population. Schembri et al. note that their aim was to maximise the number of participants who learned to sign before age seven for the reasons outlined in Section 1.3.4.6.

recruited in each of the following age brackets: 65 years and older, 51 to 64 years, 36 to 50 years and 35 years and younger.

The ethnic makeup of the corpus was 92% White and 8% non-White, closely mirroring the overall ethnic composition of the UK at the time, which was approximately 10% non-White according to the 2011 census. However, it has since become clear that the proportion of non-White individuals in the deaf community during the period of corpus data collection was likely closer to 20% (Traverse, 2019). Furthermore, as highlighted by Hodge et al. (2025), it is erroneous to assume that the signing of one individual can represent the linguistic production of all individuals within the same broad ethnic category. For instance, the signing of a particular Black deaf person cannot encapsulate the diversity of signing practices among all Black deaf individuals born in the UK, including those with African and Afro-Caribbean heritage.

Corpus participants were filmed in pairs performing a variety of linguistic tasks including recounting a personal narrative and engaging in spontaneous free conversation. For personal narratives, participants were asked ahead of time to think of a five-minute personal story–e.g., funny, sad, intriguing, etc–some participants did this while others forgot, and their narratives ended up being spontaneous. Participants' attitudes and beliefs around BSL were also explored in a series of interview questions.

We selected a subset of participants for this project with the aim of achieving a diverse mix of social factors, as shown in Table 2. To ensure representation from a relatively broad range of regions, we included people from Belfast and Glasgow as well as various English cities. In light of the earlier discussion regarding ethnicity in the corpus, and given that we were only able to analyse data from 34 participants, we did not attempt to select participants based on ethnicity. Our sample included two participants who identified as Black, while the remaining participants identified as White. Consequently, we do not analyse our results based on ethnicity.

Table 2 Participant Distribution by Region, Gender, Age and Language Background

Region	Gender		Age				Language background		Total
	Female	Male	16-35	36-50	51-64	65-88	Native	Early learner	
Belfast	2	2	2	2	0	0	3	1	4
Birmingham	3	3	0	2	2	2	4	2	6
Bristol	4	2	1	3	0	2	1	5	6
Glasgow	2	4	2	1	2	1	3	3	6
London	5	1	0	4	2	0	3	3	6
Manchester	3	3	2	0	0	4	1	5	6
Total	19	15	7	12	6	9	15	19	34

Note: No participants who learned BSL after the age of seven were included in our sample. This is because 95% of corpus participants learned BSL before age seven.

We chose files which had existing translations plus lexical and clause annotations (and, in many cases, also argument role annotations), as advised by Johnston et al., (2016): see Section 2.2.2.2 for more details. We used these existing annotations as a basis for identifying NPs. The distribution of these existing annotations is such that all the files from Birmingham, Bristol, Manchester and London are of conversation data, and all the files from Belfast and Glasgow are narratives. This means that it is not possible to avoid the confound of text type (conversation versus narrative) when comparing data from these regions. This is mitigated by the fact that there is overlap between the personal narratives and spontaneous conversations in the BSL Corpus; the personal narratives are often punctuated by interjections by the co-participant and conversely, many conversations naturally consist of some elements of personal narrative (Brown & Cormier, 2017). For example, one participant (coded BM05) held the floor throughout most of the first four minutes of his conversation with BM06, producing 378 manual tokens during this period. In contrast, BM06 only produced 26 manual tokens during this time. Additionally, as noted above, some participants forgot to prepare a personal narrative in advance, with the result that their narrative more closely resembled the spontaneity of the conversation task.

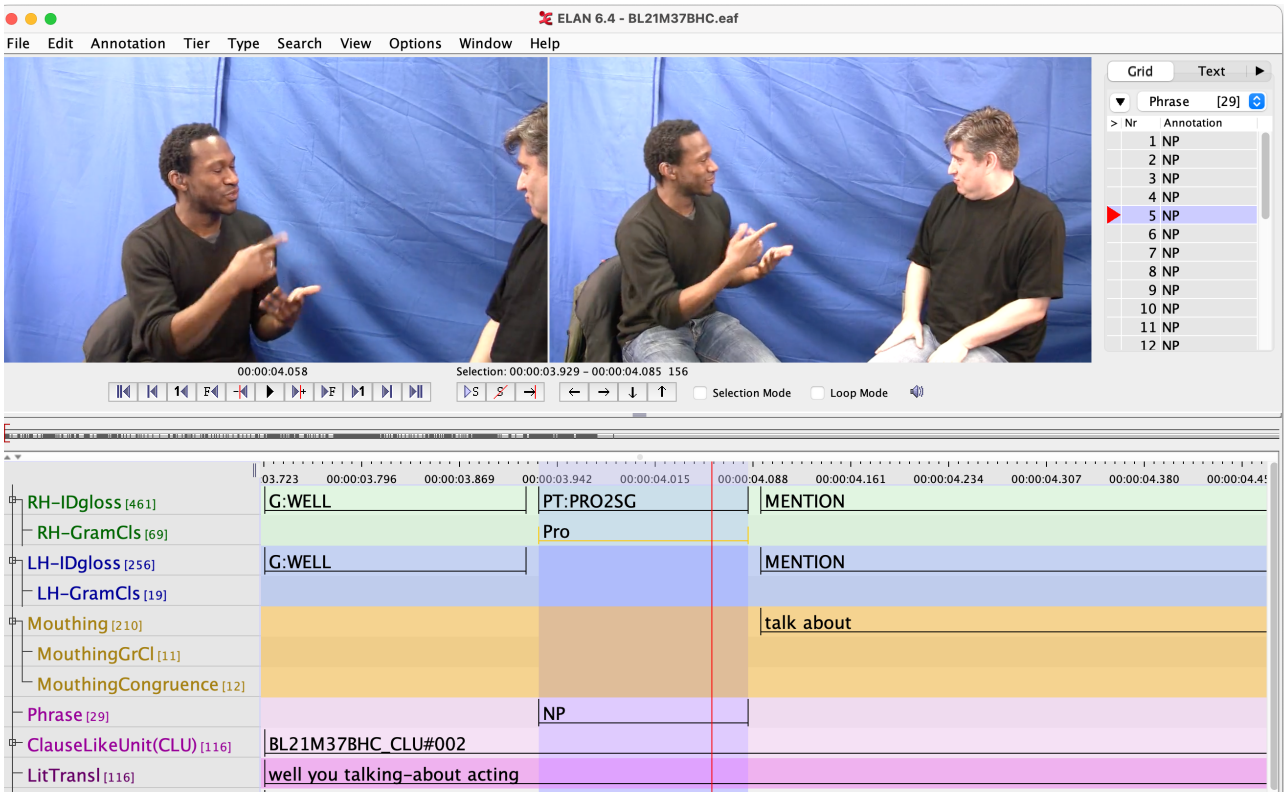
2.2.2 Coding in the BSL Corpus

2.2.2.1 Introduction

The BSL Corpus is annotated using ELAN software[16] (Crasborn & Sloetjes, 2008), which displays participant videos above a series of time-aligned annotation tiers. Each tier captures a distinct category of information, and a red vertical line indicates the current video frame, moving in synchrony with playback. These tiers thus provide temporally aligned information about what is taking place in the videos at each point in time. Annotation of the BSL Corpus is ongoing, and the present project builds on existing annotations.

There is a separate ELAN annotation file for each participant in the corpus. See Conventions section above for details of how the filenames are structured. Figure 13 shows the data for participant BL21 (on the left in the videos), who is producing a second-person singular pronominal pointing sign (meaning "you"), coded PT:PRO2SG, directed towards his partner.

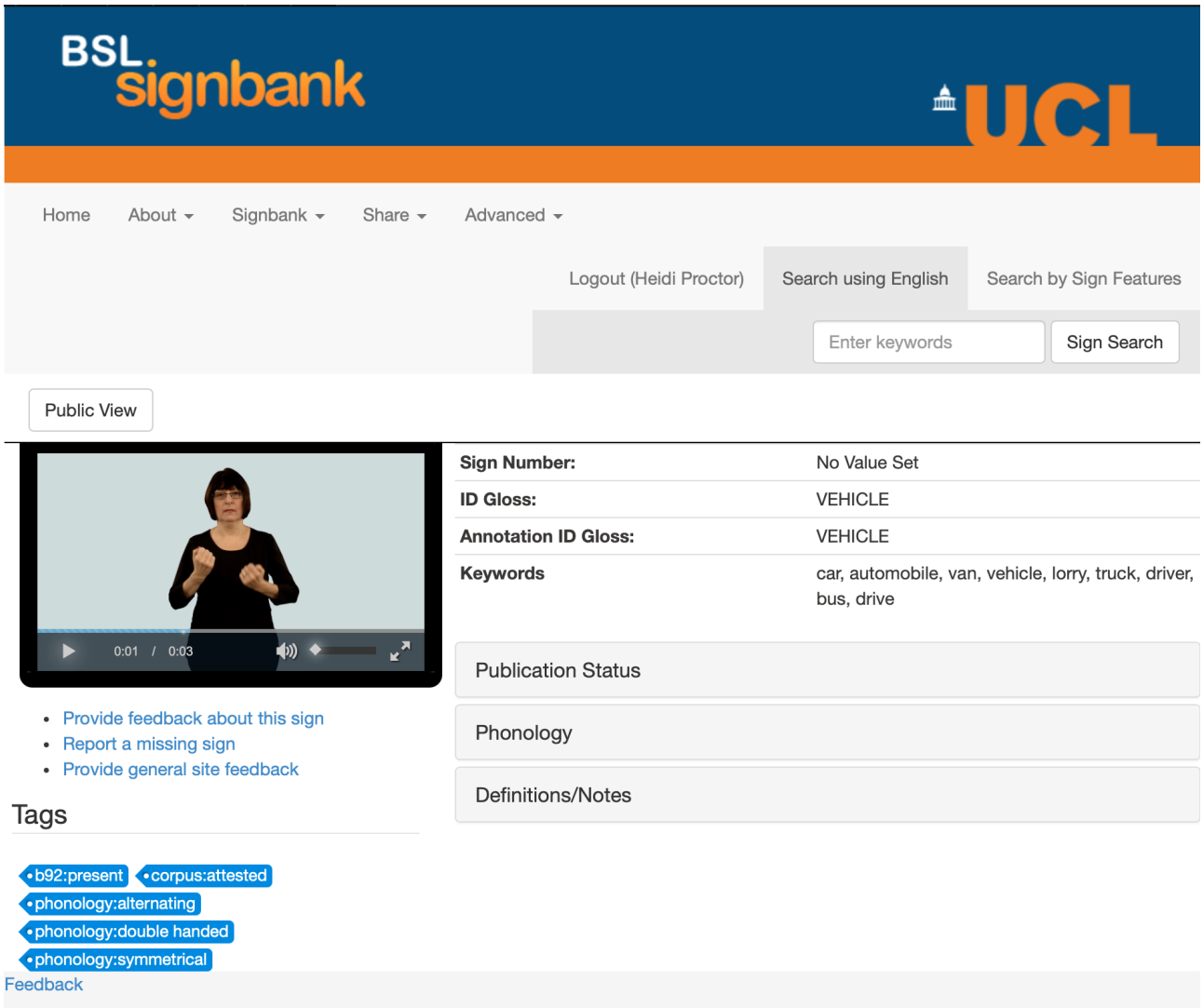
Figure 13 Screenshot of ELAN, Showing Participant Videos and Time-Aligned Tiers



[16] Downloadable from https://tla.mpi.nl/tools/tla-tools/elan/

The *RH-IDgloss* and *LH-IDgloss* tiers record the activity of the right and left hands respectively; further details are provided in Appendix A. ID gloss annotation involves categorising signs according to their lexical status. As discussed in Section 1.3.4.2, there is a continuum between lexical and non-lexical signs. Only fully lexical signs are annotated with an ID gloss, while other types–such as depicting signs (see Section 1.3.4.2.1) or pointing signs (see Sections 1.3.4.2.2 and 2.2.2.3.3)–receive alternative annotations on the *ID gloss* tiers, as described in Cormier et al. (2017). Each ID gloss corresponds to an entry in BSL Signbank (https://bslsignbank.ucl.ac.uk), a lemmatised online dictionary of BSL lexical signs. Each entry is headed by its ID gloss and includes a video of the sign plus a set of associated keywords (see Figure 14). The keywords provide possible English translations of the sign, since there is not generally a one-to-one correspondence between a BSL sign and an English word. BSL Signbank is a publicly accessible resource, although researcher access is required to view ID glosses; this can be obtained by registering via the Login link on the website.

Figure 14 Screenshot of Researcher View of BSL Signbank, Showing the Entry for the ID Gloss VEHICLE



ID gloss corpus annotations do not encode grammatical category. Instead, this information is recorded on the *RH-GramCls* and *LH-GramCls* tiers, which also correspond to the right and left hands. Some tiers use *controlled vocabularies*–predefined lists of values that annotators can select from–such as *Adj* (adjective) or *Pro* (pronoun) on the *GramCls* tiers. Appendix A provides a comprehensive list of all tiers used in this project, along with their associated controlled vocabularies.

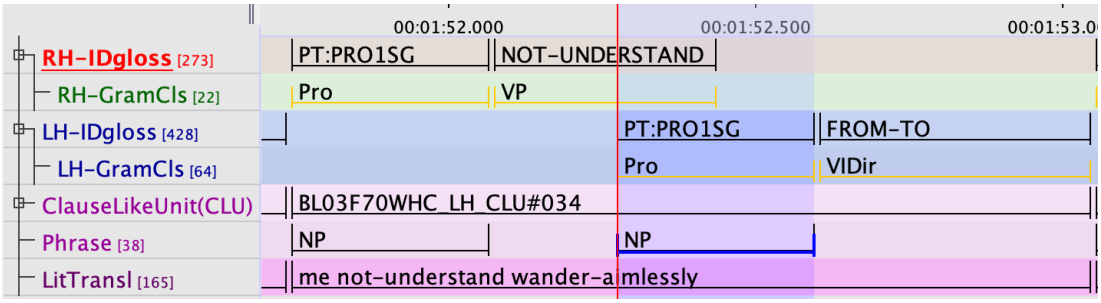
2.2.2.2 Identifying NPs

We systematically analysed the first approximately 100 sign tokens produced by the dominant hand of each participant (as indicated in the filename–see Conventions for details). We identified NPs and annotated them on a new tier called *Phrase*[17]. As part of the process, we also considered signs made with the non-dominant hand, plus mouthings. We focused on signs that have been deemed to be lexical (i.e., those that have been annotated with an ID gloss, as discussed in Section 2.2.2.1) and on pointing signs (as defined in Sections 1.3.4.2 and 1.3.4.2.2). We excluded signs whose annotation on the *ID gloss* tier indicates that they are depicting signs (i.e., where the *ID gloss* tier annotation is prefixed with *DS:*) because, as noted in Section 1.3.4.2.1, these cannot be assigned a straightforward single grammatical category. Similarly, we excluded constructed action (*CA:* prefix on *ID gloss* tier for manual elements of constructed action), which is a type of gesture where the signer enacts some aspect of the state of affairs they wish to describe (Liddell & Metzger, 1998). We did not consider manual or nonmanual modification of signs which could be considered to have an adjectival function, that are discussed in Section 1.3.4.4. We did not include *false starts* which are instances when a signer starts to produce a sign but does not complete it. Occasionally, some but not all of a signer's production at a given point in time is part of an NP – i.e., signs in other grammatical categories can be produced at the same time as elements of NPs, as in Figure 15 and Figure 16. Here, the signer's production of the verb NOT-UNDERSTAND with her right hand overlaps with the production of a pronoun (i.e., a NP) PT:PRO1SG with her left hand. When this happened, we still annotated NPs as just *NP* on the *Phrase* tier. Therefore *NP* in the *Phrase* tier should be understood as meaning that at least some of the signer's production at this point is a NP.

[17] We found a couple of instances of nested NPs: for example, in LN06c we have [TICKET ONE YEAR] (meaning "one year (season) ticket"), where the NP ONE YEAR modifies the noun TICKET, forming a larger NP [TICKET [ONE YEAR]$_{NP}$]$_{NP}$. This nested structure cannot be recorded on a single *Phrase* tier, so we annotated it as two separate NPs [TICKET]$_{NP}$ and [ONE YEAR]$_{NP}$.

Figure 15 BL03c Image Showing Simultaneous Production of NOT-UNDERSTAND with Right Hand and PT:PRO1SG with Left Hand

Figure 16 BL03c Example of verb (NOT-UNDERSTAND) overlapping with NP (PT:PRO1SG); "I didn't understand. I wandered around aimlessly"



Note. PT:PRO1SG is annotated as NP on the *Phrase* tier despite the overlapping verb.

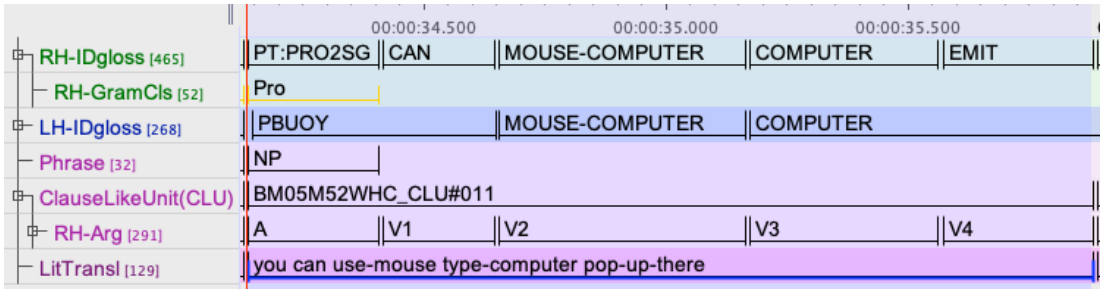
Johnston et al. (2016), in their analysis of mouth activity in Australian Sign Language (Auslan), recommend that clause annotation should precede attempts to identify the grammatical category of signs. This involves segmenting the signing stream into clause-like units (CLUs)–units that are potentially clauses (Hodge, 2013). They also advise identifying the argument role (e.g., agent, patient) of each sign within the clause, as this can inform grammatical categorisation. All of the 34 BSL Corpus files we selected for our study had already been segmented into CLUs, and 24 of these included annotations of argument structure.

We examined each sign within the context of its CLU to determine whether it formed part of an NP, following the approach used for the Auslan Corpus (Johnston et al., 2019). Argument role annotations identify predicates (annotated as *V*) and their arguments (annotated as *A*). In the BSL Corpus, only the head noun of NPs has been annotated as *A*; other elements of the NP such as adjectives, numerals and determiners have been annotated as non-arguments (*non-A*), as have other modifying clause constituents such as adverbs. Where this argument annotation was present,

we used it–alongside the observation in Section 1.3.1 that the arguments of predicates are nouns or NPs–to double-check our decisions about grammatical category.

We also considered the English words used in the literal translation of each CLU, found on the *LitTransl* tier (see Appendix A). These translations reflect aspects of BSL structure, often using multiple English words, hyphenated, to represent a single sign, as illustrated in Figure 17. Such translations can help identify whether a sign is being used nominally or verbally; for example, Figure 17 includes "use-mouse" and "type-computer" (type on a computer), suggesting that the corresponding manual signs MOUSE-COMPUTER and COMPUTER[18] have verbal interpretations in that context.

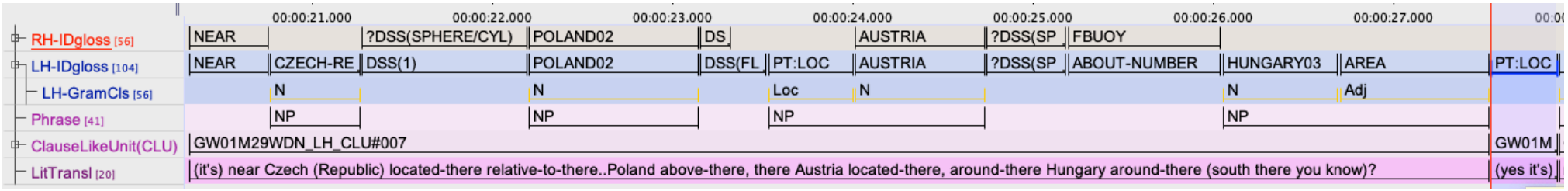
Figure 17 BM05c Example of use of Literal Translation to distinguish nouns from verbs; "You can find it on your computer".



[18] As noted in Section 2.2.2.1, ID glosses of lexical signs do not encode grammatical category; instead, this is recorded on the *GramCls* tiers. We are only annotating grammatical category of elements of noun phrases as part of this project and therefore the *GramCls* tiers are blank for these tokens MOUSE-COMPUTER and COMPUTER since these two signs are not elements of an NP. Other phrase types and grammatical categories may be annotated as part of future projects.

Figure 18 provides an example of how the segmentation of the signing stream into CLUs is useful in identifying NPs. Here, the segmentation of the highlighted PT:LOC (at the end of the extract) into its own CLU indicates that this is predicative and is not a candidate for being a NP. Had it been part of the previous CLU (CLU#007), we would have annotated it as part of the previous NP, describing the location of Hungary.

Figure 18 GW01n Example of CLU segmentation; “It's bordered by the Czech Republic, Poland, Austria and Hungary - yes it's there”



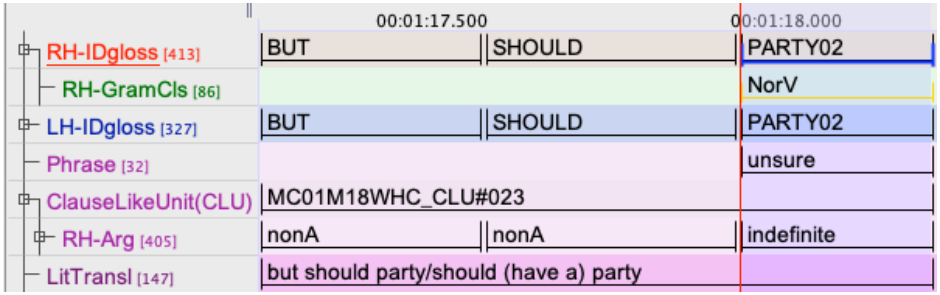
We adopted this conservative methodology rather than making any assumption about a single grammatical category that may be associated with each sign since many if not most signs can be used in more than one grammatical category, depending on context, as discussed in Section 1.3.4.5.

There are various scenarios in which there is uncertainty about how to annotate particular elements. These are discussed in the following section in relation to NPs.

2.2.2.2.1 Uncertainty in Annotating NPs

We discussed issues in identifying NPs in signed languages in Section 1.3.4.5. One such issue is that in some cases, there is more than one possible analysis of a CLU. In these instances, the CLU has previously been annotated as *indefinite* (i.e., ambiguous in terms of clausal analysis) on the right-hand or left-hand argument tier as appropriate. See Figure 19: Here the sign PARTY02 could be analysed as either a noun (an argument) or a verb (a predicate) and the two possible literal translations have been provided on the *LitTransl* tier, separated by a slash.

Figure 19 MC01c Example of Indefinite (Ambiguous) Grammatical Category, Noun or Verb; “We were supposed to have a party”



In other situations, there are portions of signing that seem to constitute a single unit, but which are either difficult to split into constituents or whose meaning is hard to establish. These have been annotated as *indeterminate*. (This contrasts with *indefinite* CLUs, whose constituents can be identified but which are ambiguous in terms of their clausal analysis.)

The identification of NPs can be difficult because of inherent ambiguity within constructions. See, for example, Figure 20. At the end of this utterance, the signer produces a pointing sign with each hand, and each point is directed to a different location, as illustrated in Figure 21. It is clear from the literal translation that one point refers to “them” (the GB Deaflympics team) and the other to “there” (Melbourne, where the Deaflympics took place). However, it is not possible to say which point refers to which concept, as reflected in the *IDgloss* tier for each sign; each has PT:PRO3SG/PT:LOC, indicating ambiguity between a pronominal and a locative point.

Figure 20 GW02n Example of inherent ambiguity; “The team had gone to Melbourne for the Deaflympics and won the bronze medal”

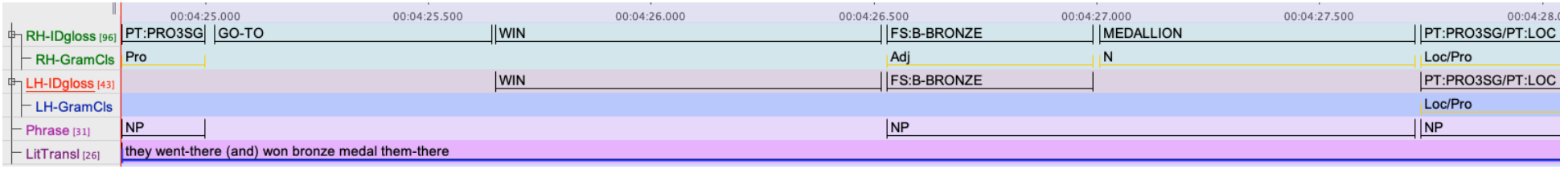
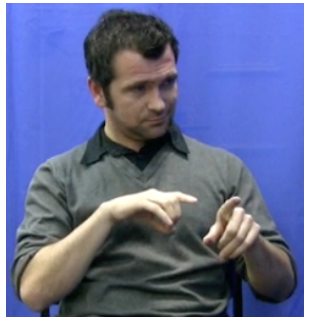
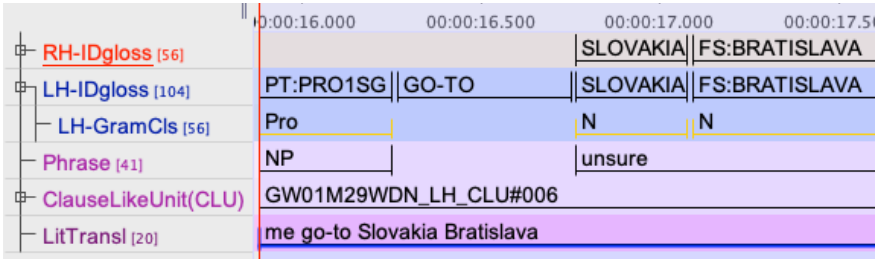


Figure 21 GW02n Image of Simultaneous Points to “them” and “there”



In other cases, more than one analysis of NP elements is possible, as in Figure 22. Here, SLOVAKIA FS:BRATISLAVA could be interpreted as either two separate NPs or one combined NP. In situations like this, and in other examples where there was any doubt about the status of elements of an NP, we annotated the *Phrase* tier as *unsure*.

Figure 22 GW01n Example of NP with more than one possible analysis; “I went to Slovakia – Bratislava”

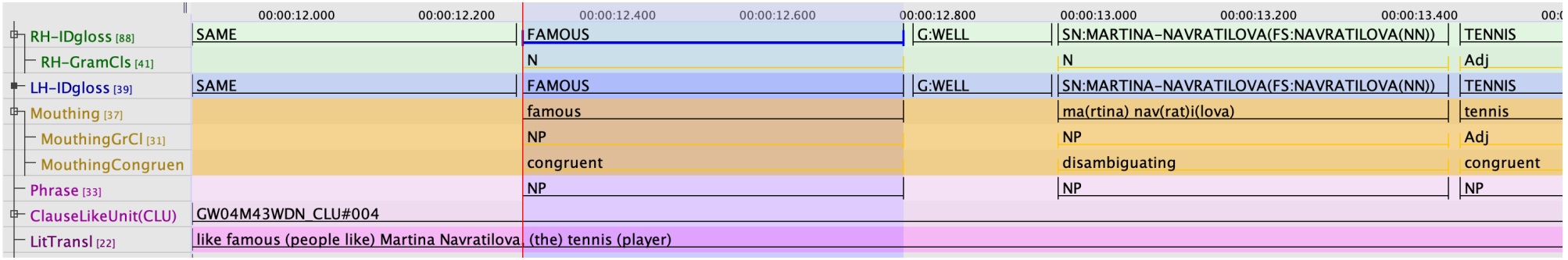


The process of corpus annotation is ongoing and iterative. In line with this, we revisited our unsure tokens as the project progressed with the aim of resolving the uncertainty in the light of experience. This conservative strategy means that only NPs where there is high confidence were annotated as such.

2.2.2.3 Assigning Grammatical Category

Having identified candidate NPs, we next assigned grammatical categories to their constituent signs. We used the *RH-GramCls* and *LH-GramCls* tiers for this. We ensured that we used the grammatical category of the sign in BSL, resisting the temptation to translate the utterance to English and use the English grammatical category. See, for example, Figure 23: the literal translation indicates that the intended meaning is something like "famous people", but the participant does not produce an overt sign for "people". The effect is that FAMOUS is being used as a noun in this context, and not as an adjective as would normally be the case in English[19].

Figure 23 GW04n FAMOUS Used as a Noun: "With celebrities like Martina Navratilova, the tennis player."

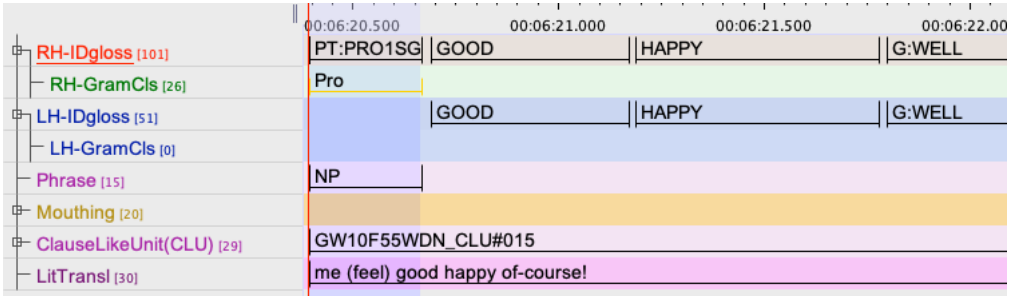


Similarly, in Figure 24, the signer does not produce an explicit verb. Such verbless attributive constructions are discussed in Section 1.3.4.3; they consist of a carrier that is a noun or pronoun (PT:PRO1SG in this case) and attributes that are predicate adjectives (GOOD and HAPPY in this example). These predicate

[19] "Famous" need not always function as an adjective in English. For example, other literal translations of this utterance include "the famous", where "famous" is functioning as a noun, or "those who are famous", where "famous" is an adjective.

adjectives are not elements of a NP. There is no copula, unlike in English; these constructions form an attributive clause which stands alone. The literal translation indicates that the verb “feel” is implicit, since it is shown in brackets[20].

Figure 24 GW10n Example of Verbless Attributive Construction: “I feel great, really happy.”



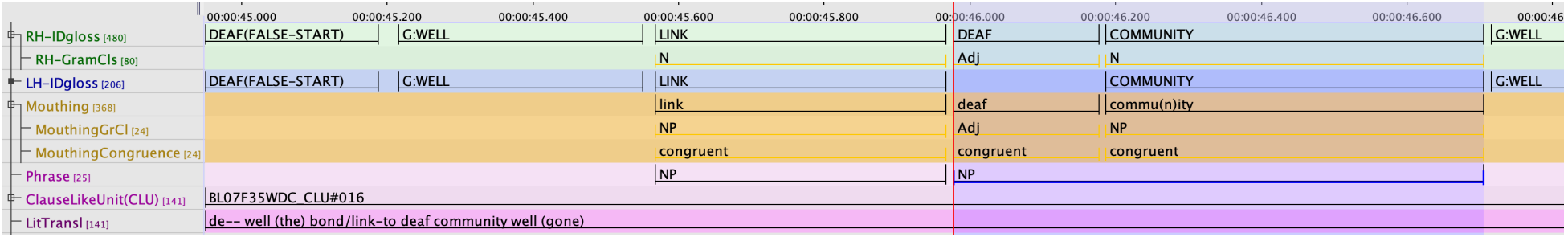
2.2.2.3.1 Discriminating Nouns and Adjectives

In many cases, it is straightforward to identify nouns and adjectives, as in Figure 20. Here we have the sign MEDALLION functioning as a noun, modified by a fingerspelled initialisation FS:B and disambiguating mouthing of the word “bronze”, here with the grammatical category of adjective.

Sometimes it can be difficult to determine whether an element is functioning as a noun or an adjective. Indeed often, more than one analysis can be possible, as noted in Section 1.3.4.5. See, for example, Figure 25, which shows one possible analysis, with COMMUNITY identified as a noun and DEAF identified its modifying adjective. Alternatively, DEAF could be analysed as another noun meaning “deaf people”. This gives us an NP meaning something like “deaf people’s community”. In such cases, following Rutkowski et al. (2015), we selected the analysis that appeared most plausible in light of the surrounding context.

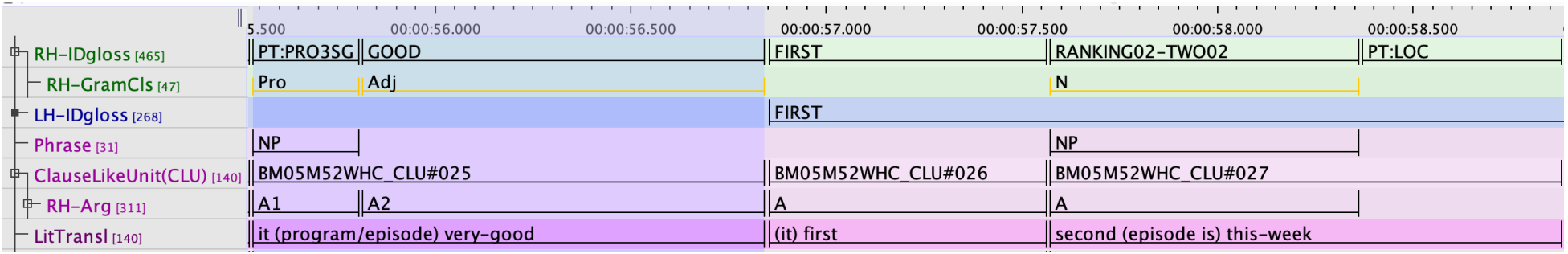
[20] Other verbs could have been selected for this literal translation, such as “am”.

Figure 25 BL07c Noun Phrase with Two Possible Grammatical Category Analyses; “But it's a shame for the close kinship we have with the deaf community.”



We consider isolated adjectives where there is nothing else present in the clause to not be elements of a NP, because we analyse them as predicate attributes whose carrier has been omitted, as discussed in Section 1.3.4.3. See, for example, Figure 26 which contains three CLUs; the second contains a single sign FIRST, meaning “(that was the) first (episode)” which is a predicate adjective and not a NP. The third CLU contains two signs, RANKING02-TWO02 PT:LOC, meaning “the second (episode is) this week”. Here RANKING02-TWO02 is functioning as a noun because it has a separate predicate, PT:LOC, meaning “this week”.

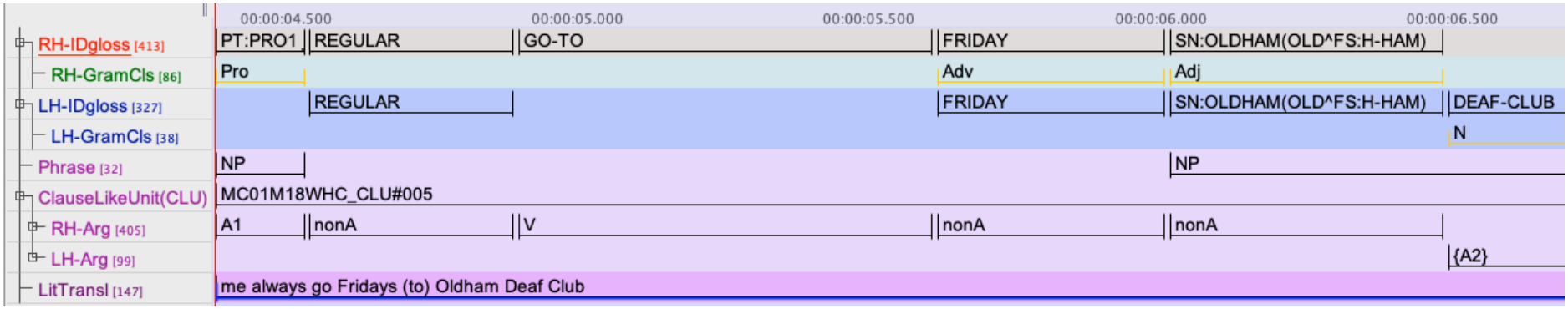
Figure 26 BM05c Predicate and non-Predicate Adjectives; “The first programme was really good. The second is this week…”



2.2.2.3.2 Discriminating Nouns from Adverbs and Verbs

Some temporal adverbs, such as “tomorrow”, can also function as nouns–for example, “tomorrow is his birthday”, as discussed in Huddleston (1984) for English. To distinguish between nouns and adverbs in our data, we relied on syntactic context and any accompanying mouthings, based on the approach taken by Cormier et al. (2012). For instance, we categorised the sign FRIDAY as an adverb in Figure 27, since it modifies the verb GO-TO and is therefore not a candidate for inclusion within a NP. Had the participant instead signed ON FRIDAY or mouthed “on Friday(s)” during the manual sign FRIDAY, we would have analysed FRIDAY as a noun.

Figure 27 MC01c Example of a Construction Including an Adverb; “I always go on Fridays to Oldham Deaf Club”



Various differences between nouns and verbs have been observed in several sign languages, as discussed in Section 1.3.4.5. It is useful to bear all the tendencies in mind during the annotation process. As observed by Johnston (2001), one should also consider the meaning of the overall utterance and syntactic information such as the order of signs, plus any non-manual features when attempting to discriminate nouns and verbs. Specifically, since mouthings have been found to accompany nouns more than verbs (as discussed in Section 1.3.4.5), the presence of mouthing can be used as an indication that the sign may be functioning as a nominal.

The same sign can function as both a noun and a verb during the course of a text; see Figure 28, where the literal translation indicates that CHANGE is functioning as a noun, with LITTLE as an adjective within the NP, modifying this noun. In contrast, Figure 29, from a few seconds later in the same turn, shows CHANGE as a verb, with LITTLE as a modifier playing an adverbial role. Here, the argument of that verb is the noun SURROUNDINGS.

Figure 28 BM22c Example of CHANGE Functioning as a Noun; "It was much the same, with a few small changes"

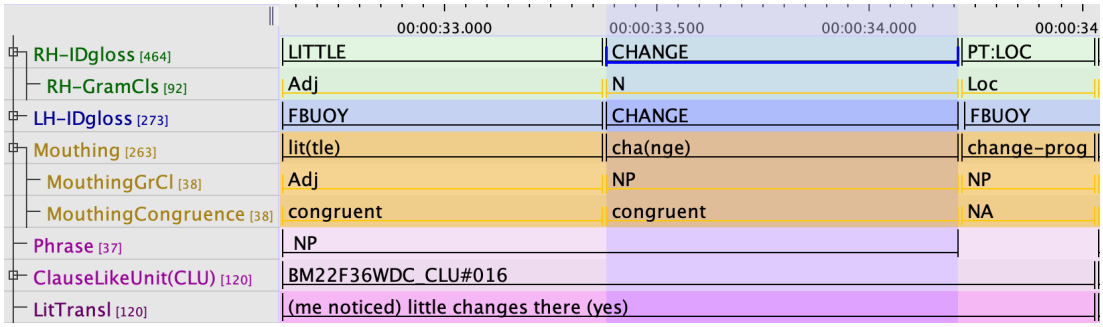
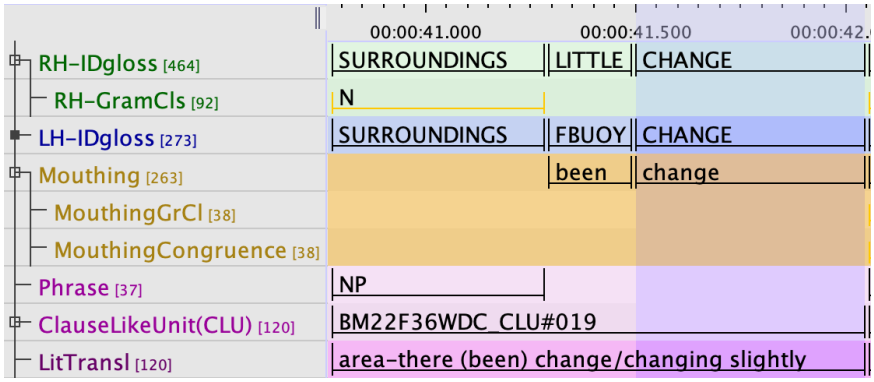


Figure 29 BM22c Example of CHANGE Functioning as a Verb; "The area … had changed a bit"



We considered prosody and meaning, plus the presence of other grammatical elements in order to distinguish between these two cases.

In some cases, it is not possible to determine whether a sign is being used as a noun or a verb. In such situations we follow Johnston et al. (2019) for Auslan in annotating these as *NorV* (noun or verb), as in the sign PARTY02 in Figure 19.

2.2.2.3.3 Pointing Signs

Unlike other annotations on the *ID gloss* tiers, the labels used to identify pointing signs in the BSL Corpus reflect the sign's grammatical function (i.e., determiner, locative or pronominal – see Section 1.3.4.2.2) and whether it is singular or plural. We therefore annotated the grammatical category tier for these signs based on these labels as listed in Table 3[21]. (As discussed in Section 2.4, after exporting our annotations from ELAN, we combine all pointing signs into a single category of *Point* for our analysis. This is because of the high degree of ambiguity re: the grammatical category of pointing signs.)

Table 3 Mapping of Label to Grammatical Category for Pointing Signs

Label	**Grammatical category**
Label is PT:DET or PT:DETPL	Det: Determiner
Label is PT:LOC or PT:LOCPL	Loc: Locative
Label contains PT:PRO	Pro: Pronoun
Label contains PT:POSS	Pro: Pronoun
Label is PT:DET/LOC	Det/Loc: Determiner/Locative
Label is PT:LOC/PRO	Loc/Pro: Locative/Pronoun
Label is PT:	Det/Loc/Pro: Determiner/Locative/Pronoun

[21] Although it is not strictly necessary to annotate both the label and the grammatical category for pointing signs, as they convey the same information, it simplifies the analysis to do so because it means that all candidates for inclusion in a NP have grammatical category tagged. We ran a cross-check using R to ensure that no errors have been made when annotating grammatical category for pointing signs.

We also find examples of signers pointing to elements of a list buoy (for background see Section 1.3.4.2.2). These are annotated as PT:LBUOY. See, for example, Figure 30 and Figure 31 where the signer is enumerating her family members. This point appears to have a pronominal function based on the semantics of the utterance and so we have annotated the grammatical category as *Pro*. This is in accordance with Liddell et al. (2007), who, as noted in Section 1.3.4.2.2, also analyse such points as pronouns.

Figure 30 LN12c Example of a point to a list buoy; “The next one (second, my cousin) is also deaf”

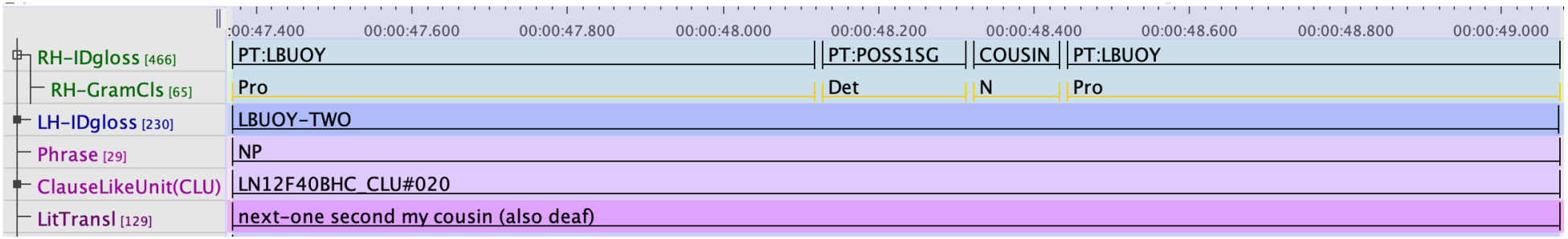
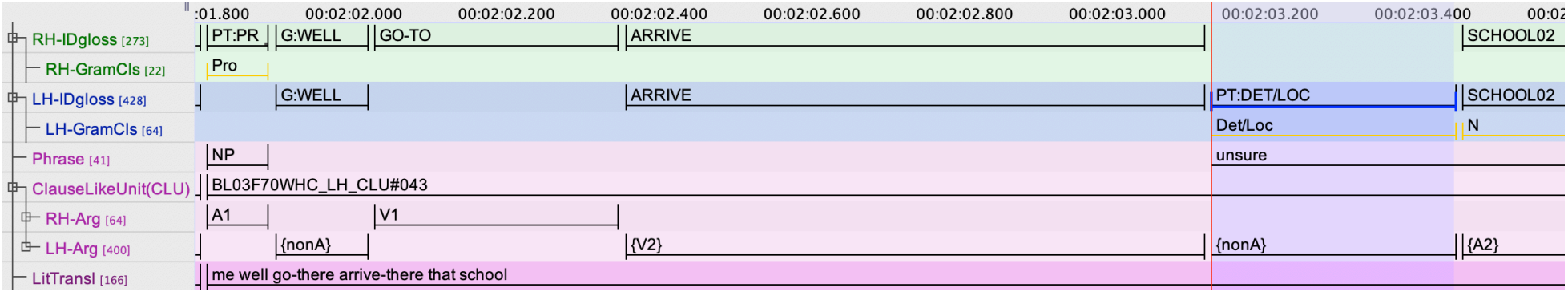


Figure 31 LN12c Image of Example of a point to a list buoy; “The next one (second, my cousin) is also deaf”

As discussed in Section 1.3.4.2.2, pointing signs can be ambiguous; for example, the highlighted sign in Figure 32 could either be a determiner point referring to the school, or a locative point referring to the location of the school. In this case, the two (or more) options are specified in the label PT:DET/LOC.

Figure 32 BL03c Example of Ambiguous Point; “I arrived at school”



Even when it is clear that a particular point is pronominal, it is not always straightforward to determine the referent of that point. See, for example, Figure 33. Here, the first PT:PRO3SG (highlighted) could refer either to “that time” or to “Sharon Hirshman”. The fact that the signer points to a different location in space here as compared to his point after Sharon’s name, as illustrated in Figure 34, suggests that they are different so that the first point refers to “that time”. This is reflected in the literal translation which starts “that time then Sharon…”, indicating that the first point means “then”. As it is adverbial, it is not annotated on the *Phrase* tier for this project.

Figure 33 GW02n Example of an Utterance with Multiple Points; “At that time Sharon Hirshman was the G.B. manager”

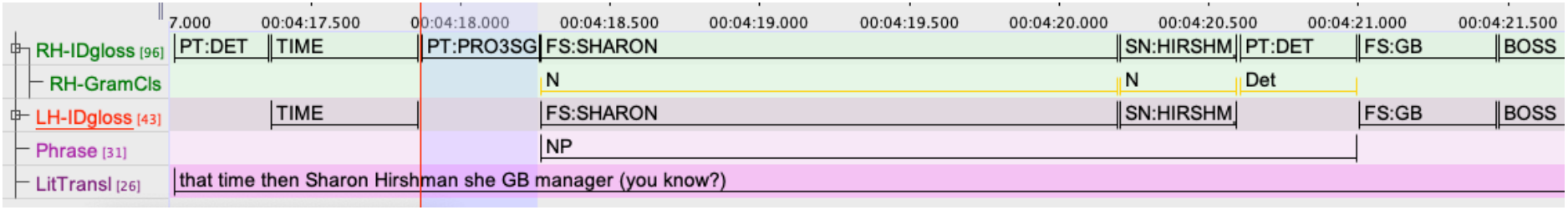
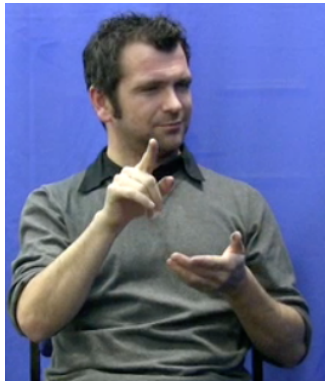
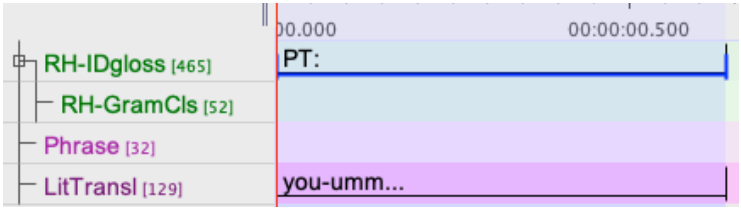


Figure 34 GW02n Images of Point first to “that time” then to “Sharon Hirshman”



Despite the fact that pointing signs can be ambiguous, in some cases it is still possible to determine whether the sign is, or is not, a candidate for inclusion within a NP. For example, many signs labelled as PT: are simply discourse markers, as in Figure 35, where the signer points to his conversational partner. In these cases, there was no need to mark the NP as *unsure* as this is clearly not a NP; instead, no annotation was made on the *Phrase* tier.

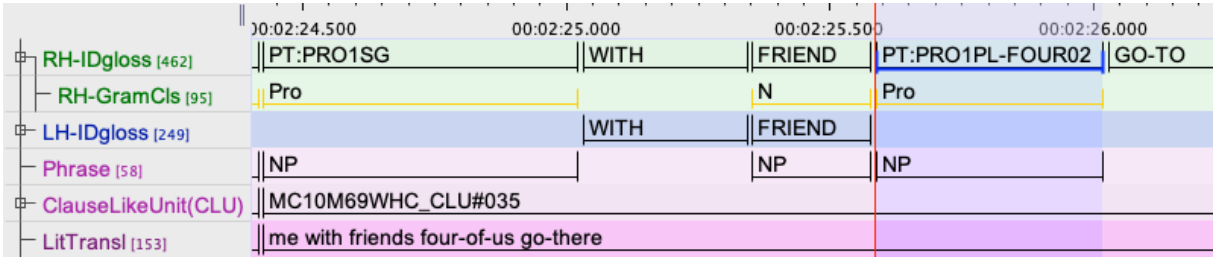
Figure 35 BM05c Example of a Point that is a Discourse Marker; “Erm?”



2.2.2.3.4 Other Grammatical Categories

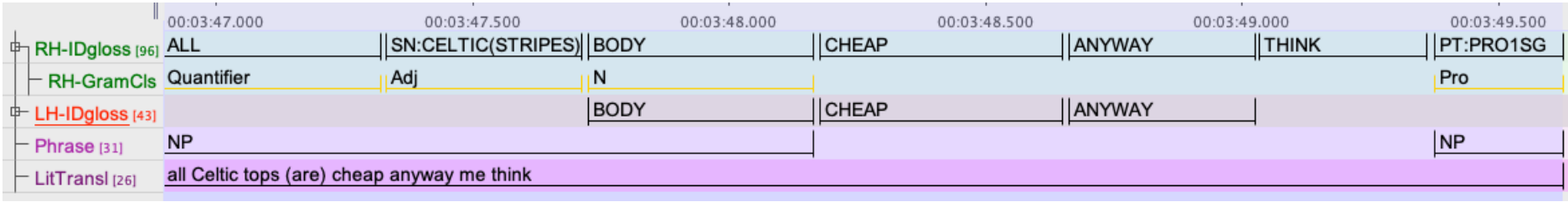
Cardinal and ordinal numerals are annotated as *Num*. The grammatical category of a sign that includes numeral incorporation is determined on the basis of its function. For example, in Figure 36, the sign PT:PRO1PL-FOUR02 (“we four”) exhibits properties of both a pronoun and a numeral. Accordingly, we assigned it the grammatical category of pronoun since its primary function is to refer to the group of friends.

Figure 36 MC10c Example of numeral incorporation functioning as a pronoun; “I went with some friends - there were four of us”



Lexical quantifiers (other than numerals) have their grammatical category annotated as *Quantifier*. These include definite universal quantifiers such as ALL, EVERY, BOTH, and indefinite existential/partitive quantifiers such as SOME, ANY. See, for example, Figure 37 which starts with the quantifier ALL.

Figure 37 GW02n Example of a Quantifier ALL; “I think all the Celtic tops are cheap anyway”

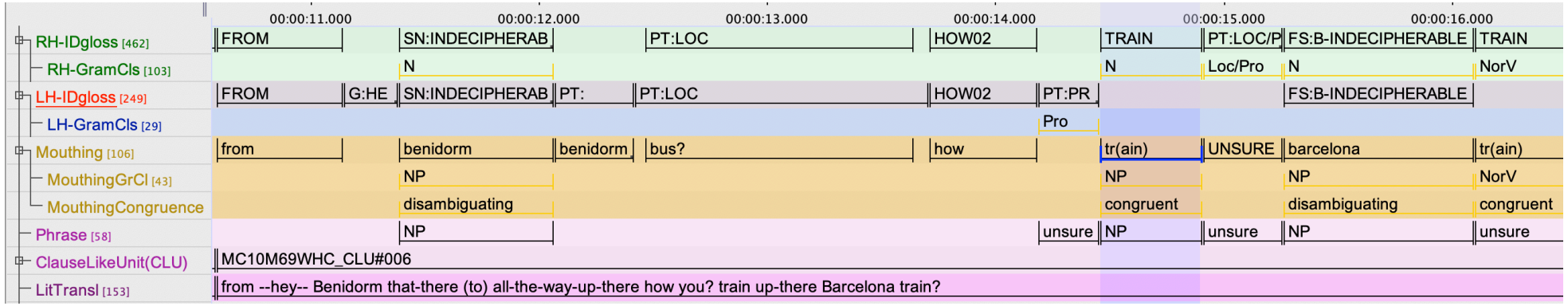


2.2.2.4 Mouthing Annotation

Important information within NPs is conveyed using mouthings, as discussed in Section 1.3.4.4 with reference to Figure 12. We annotated mouthings of full or partial English words that co-occur with elements of NPs produced on the hands, using the *Mouthing* tier. Associated tiers–*MouthingGrCl* and *MouthingCongruence*–were used to record the grammatical category of each mouthing and to indicate how its meaning relates to the target manual sign. These procedures are discussed in the following sections. A complete list of tiers used in this project is provided in Appendix A.

We annotated mouthings that co-occur with elements of NPs produced on the hands, including those marked as *Unsure* on the *Phrase* tier. No instances were found in our data where a mouthing occurred independently, without overlapping with a manual sign. For each annotated mouthing, we recorded the mouthed English word using the existing *Mouthing* tier. Mouthings are annotated in lower case (e.g., “train” rather than “TRAIN”, “i” rather than “I”). Where only part of an English word is mouthed and the intended full word is clear from context, the remainder is included in brackets. For example, in Figure 38, the signer mouths “tr–” alongside the manual sign TRAIN. (This occurs twice; the first instance is highlighted in the figure.)

Figure 38 MC10c Example of Partially-Mouthed Word “train” in “Did you go from Benidorm to Barcelona on the train?”

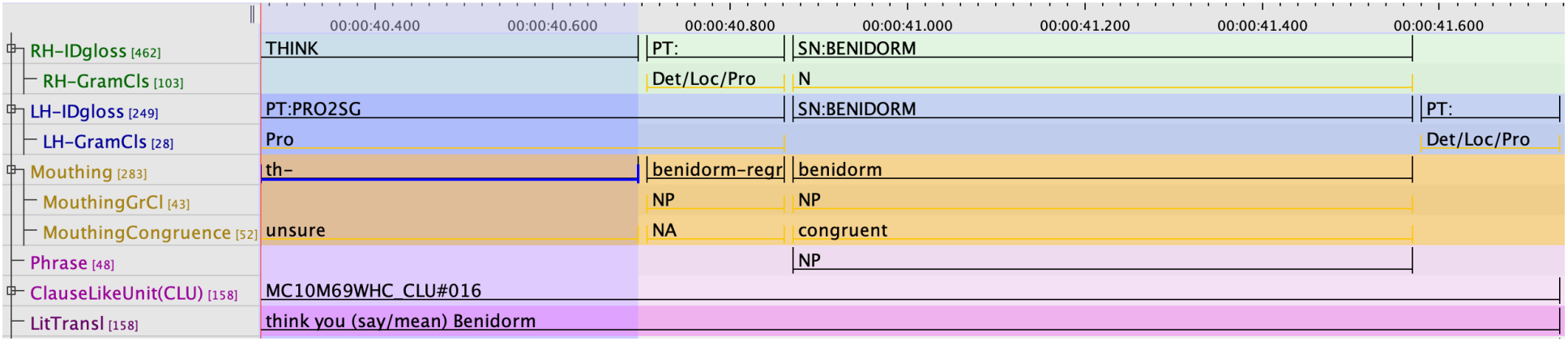


If more than one English word is mouthed alongside a single manual sign, we annotated all the words that are mouthed. We followed the above rules in terms of case and production of partial words.

2.2.2.4.1 Mouthings: Annotating Uncertainty

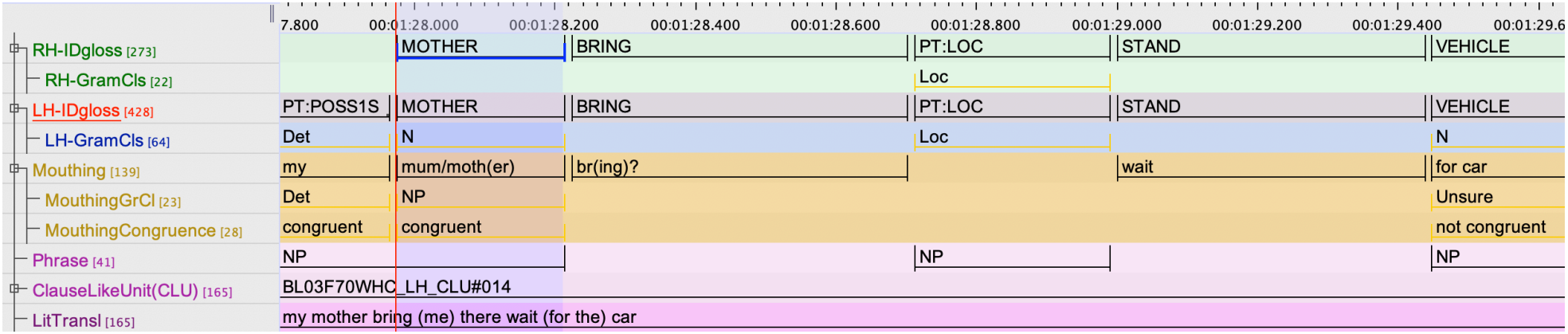
When part of an English word is produced but it is not clear what the remainder of the intended word is, we annotated the part that is visible; see Figure 39. Here, we cannot determine what the rest of the intended word is: in this case, it could be “think” or “thought”, for example.

Figure 39 MC10c Example of Ambiguous Mouthing Starting “th-“; “I think/thought you (say/mean) Benidorm.”



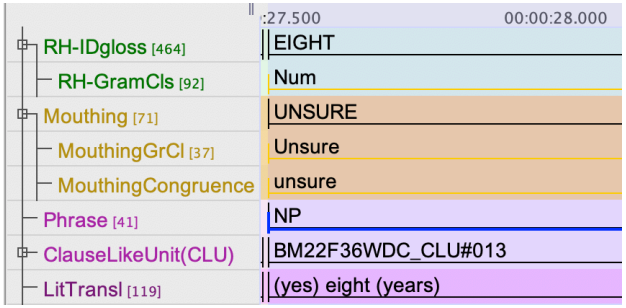
If there is more than one possibility for what the signer is mouthing, we listed the possibilities, separated by slashes, as in Figure 40, where the signer mouths either “mum” or “moth(er)”.

Figure 40 BL03c Example of Ambiguous Mouthing of “mum” or “moth(er)”; “My Mother brought me to stand outside as we waited for the car.”



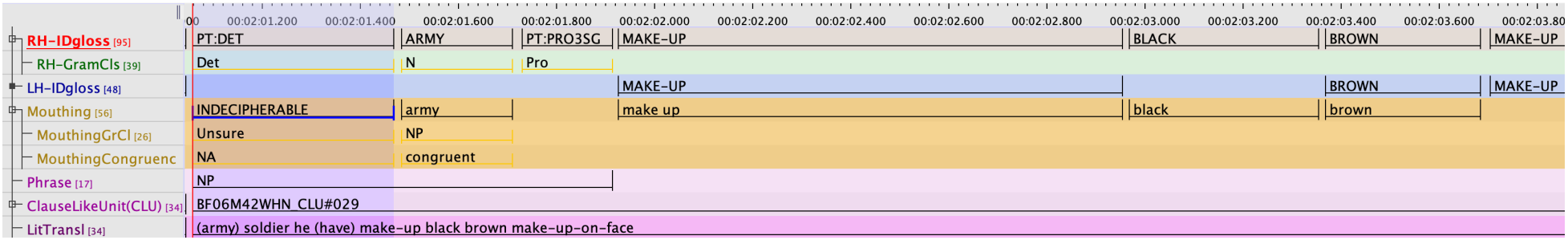
If there is visible mouth activity that might be a mouthing but it might not (e.g., it might be a mouth gesture), or, it is clearly a mouthing but the word being mouthed is unclear, we annotated *UNSURE* on the *Mouthing* tier, as in Figure 41. (We use upper case for *UNSURE* so as to distinguish this annotation from a mouthing of the actual word “unsure”.)

Figure 41 BM22c Example of a Mouthing Where the Mouthed Word is Unclear: “That was eight years later.”



If mouth activity is present and could potentially be mouthing, but cannot be identified–for example, because the signer’s mouth is at least partially obscured–we annotated *INDECIPHERABLE*, as in Figure 42. In such cases, it is not possible to determine with any confidence what the mouth activity is. This differs from *UNSURE* in that *INDECIPHERABLE* implies that no annotator would likely be able to identify the mouth activity, whereas *UNSURE* allows for the possibility that another annotator might be able to make a more confident judgement.

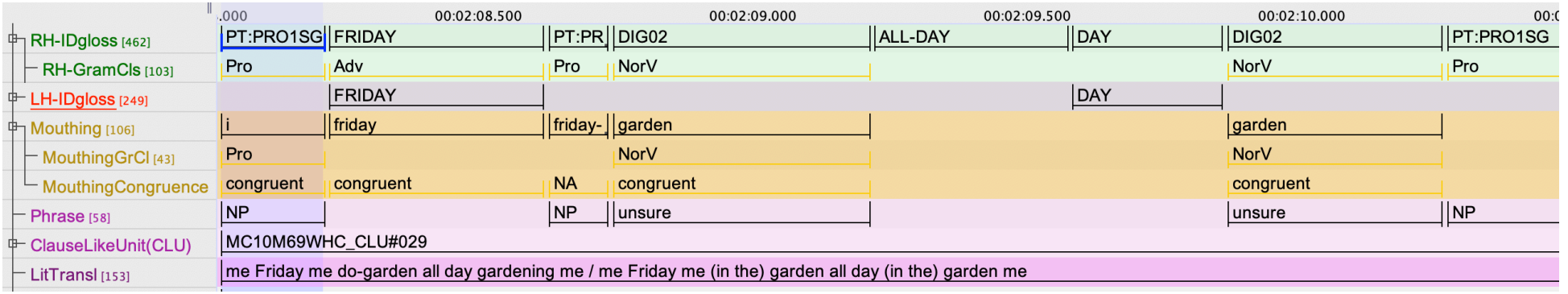
Figure 42 BF06n Example of an indecipherable mouthing: “He had camouflage cream over his face.”



2.2.2.4.2 Time Alignment of Mouthings

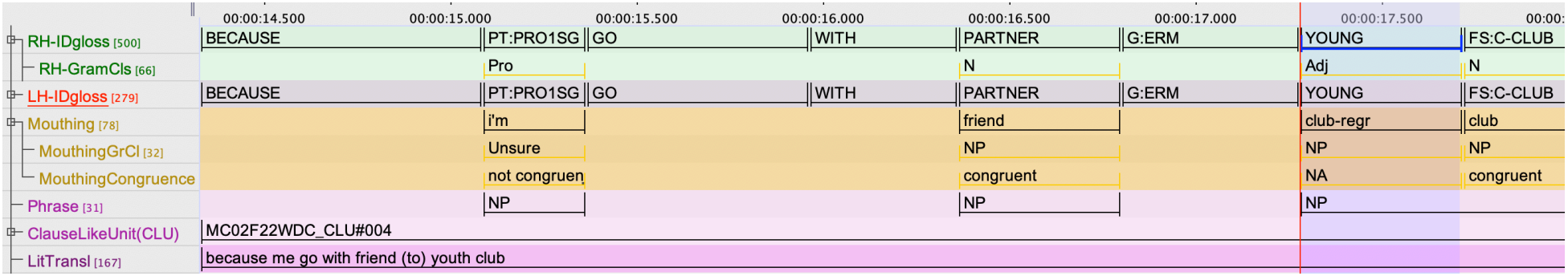
For each mouthing, a target manual sign is identified – this is the sign with which the mouthing is most strongly semantically related. Each mouthing's annotation is time aligned with this manual sign, even when the mouthing starts before or ends after the corresponding manual sign. (For an alternative approach, see Bank et al., 2011.) Figure 43 shows the signer producing a pointing sign indicating himself, and at the same time, mouthing "i" briefly. He then starts to mouth "Friday". The first mouthing "i" is semantically related to the manual sign PT:PRO1SG, so this is what we annotated here, even though the mouthing "Friday" co-occurs for longer with PT:PRO1SG. The mouthing "Friday" is annotated next, with the manual sign FRIDAY.

Figure 43 MC10c Example of Mouthing Annotation Alignment Based on Related Manual Signs "I spent all day in the garden that Friday."



When a mouthing starts before its target manual sign, and the preceding manual sign has no other accompanying mouthing, we followed Johnston et al. (2016 p. 10) in terms of the coding to be used. We annotated this as *english word-regr* on the preceding sign (with the suffix *-regr* indicating *regressive*) and *english word* on the target sign. See, for example, Figure 44 where the mouthing "club" starts whilst the manual sign YOUNG is being produced and continues whilst its target manual sign FS:C-CLUB is being articulated.

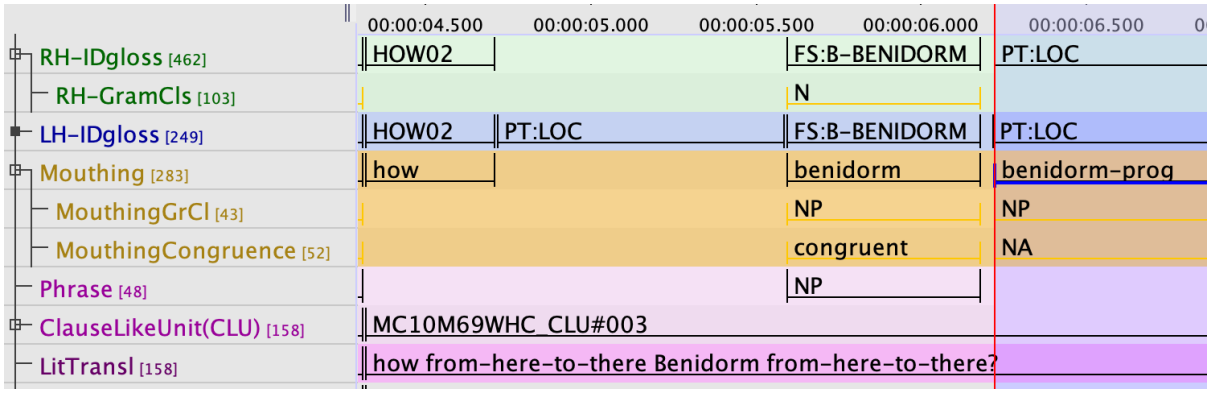
Figure 44 MC02c Example of regressive mouthing "Because I'm going to the erm... youth club with a friend"



Conversely, where a mouthing finishes after its target manual sign, and the following manual sign has no other accompanying mouthing, this is annotated as *english word* on the target sign and *english word-prog* (with the suffix *-prog* indicating *progressive*) on the following sign. For example, in Figure 45 the signer mouths "benidorm" at the same time as producing two manual signs, FS:B-BENIDORM and PT:LOC. "benidorm" has been annotated on the *mouthing* tier alongside the first manual sign, and "benidorm-prog" on the *mouthing* tier alongside the second.

When determining the length of an NP in terms of number of elements, we exclude the annotations for regressive and progressive mouthings. The main (non-regressive or -progressive) mouthing is included when calculating the NP length.

Figure 45 MC10c Example of Progressive Spreading of Mouthing "How did you get from Benidorm to Barcelona?"



Johnston et al. only annotate mouth activity on an adjacent sign when this spreads over more than 50% of the adjacent sign's duration. We take a different approach, choosing not to apply a threshold based on spread: we annotate mouthings regardless of the extent to which they overlap with adjacent signs. This approach allows us to identify which element–manual sign or mouthing–begins first, thereby enabling us to determine the linear order of elements where possible. (In cases of full simultaneity, such ordering cannot be established.)

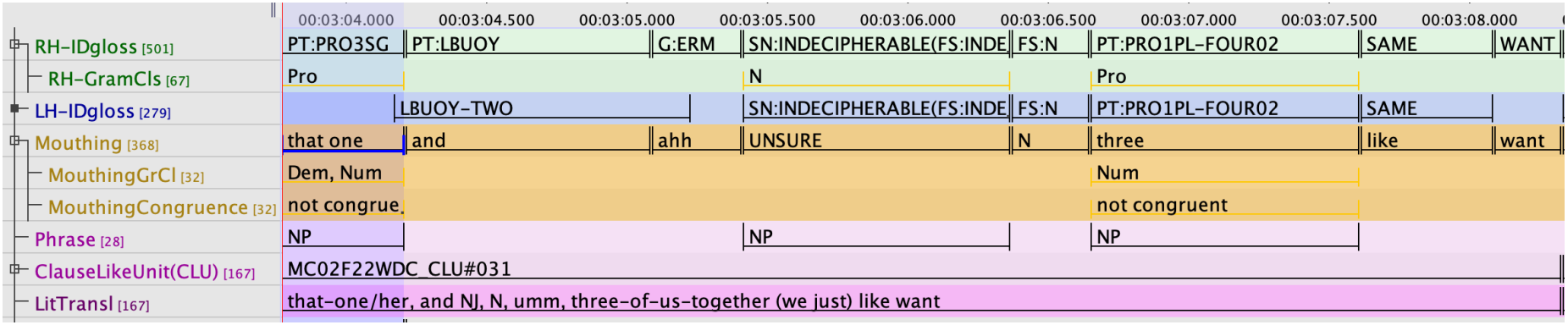
2.2.2.4.3 Grammatical Category of Mouthings

The grammatical category of each mouthed English word that is a candidate for inclusion within a NP is recorded on the existing tier *MouthingGrCl*. Nouns are annotated as *NP* (*Noun-Plain*), using the existing controlled vocabulary for this tier. (This controlled vocabulary is based on that used for manual signs, and includes other values for nouns, *Nloc* (*Noun-Locatable*) and *ND* (*Noun-Depicting*). We do not use these because the concepts *locatable/depicting* do not apply to mouthings.) If there

is more than one word that is mouthed with a given sign, each word's grammatical category is recorded here, separated by commas. Mouthings that are regressive or progressive, as in Figure 44 above, are also annotated for grammatical category. Mouthings which are annotated as *UNSURE* or *INDECIPHERABLE* are given a grammatical category of *Unsure*; see Figure 41 above for an example.

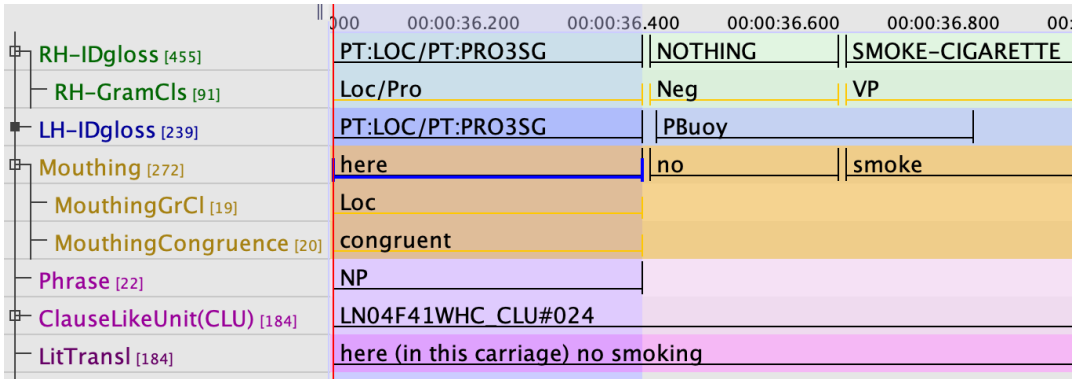
We have introduced a grammatical category of *Dem* (for demonstrative) to our controlled vocabulary for this tier – see Figure 46 for an example.

Figure 46 MC02c Example of mouthed demonstrative “that”: “That one. And, erm, NJN - the three of us kind of wanted a break from our boyfriends.”



Demonstrative as a grammatical category is not required for manual signs, as discussed in Section 1.3.4.2.2. Mouthed English words such as "there" or "here" are defined as having a grammatical category of *Loc* in the same way as the manual pointing sign PT:LOC, since they perform the same function. Figure 47 provides an example.

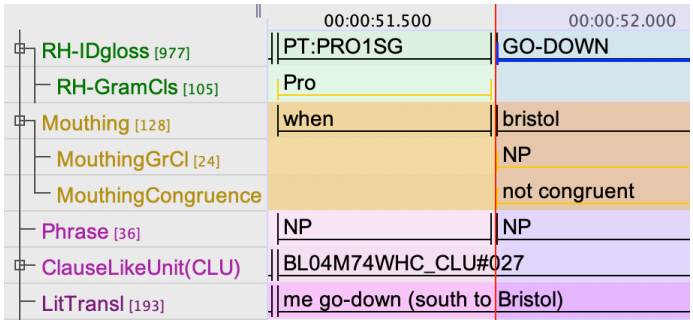
Figure 47 LN04c Example of a Mouthing of the Word "here" with Grammatical Category Loc; "Sorry I made a mistake, it was the no-smoking part of the carriage in the train."



2.2.2.4.4 Mouthing Congruence

The mouthing that co-occurs with a manual lexical sign typically corresponds to an English translation equivalent of that sign (and as such, it will be a keyword for that sign in BSL Signbank, as described in Section 2.2.2.1). However, this is not always the case. For instance, in Figure 48, the mouthing "bristol" occurs at the same time as the manual sign GO-DOWN[22] is produced. A new tier, called *MouthingCongruence* has been added to indicate whether each mouthing is congruent with the co-occurring manual sign. In this example, the mouthing is annotated as *not congruent*.

Figure 48 BL04c Example of information conveyed only on the mouth: signer mouths "bristol" alongside the manual sign GO-DOWN; "When I got to Bristol…"

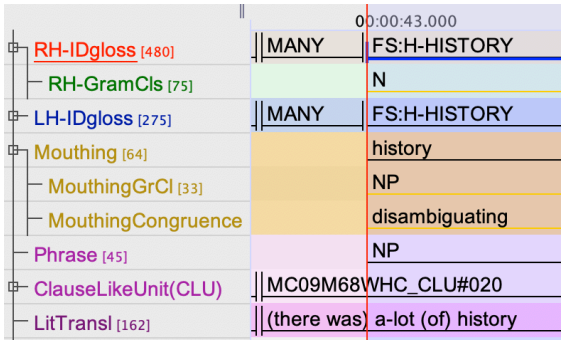


[22] Although the sign GO-DOWN can appear similar to a sign for Bristol, here a deaf native signer confirmed that the sign is GO-DOWN and not BRISTOL.

A mouthing is annotated as *congruent* to its manual sign if it is a keyword for the manual sign's ID gloss in BSL Signbank (see Section 2.2.2.1). See, for example, the mouthing of "friend" with the manual sign PARTNER in Figure 44 above. "Friend" is a keyword for this sign, so it is annotated as *congruent*.

On occasion, a mouthing can disambiguate a non-lexical manual sign, especially a fingerspelled letter or sequence of letters, or a pointing sign. For example, in Figure 49 we have the mouthing "history" alongside the fingerspelled letter -h-. This mouthing is annotated as *disambiguating*.

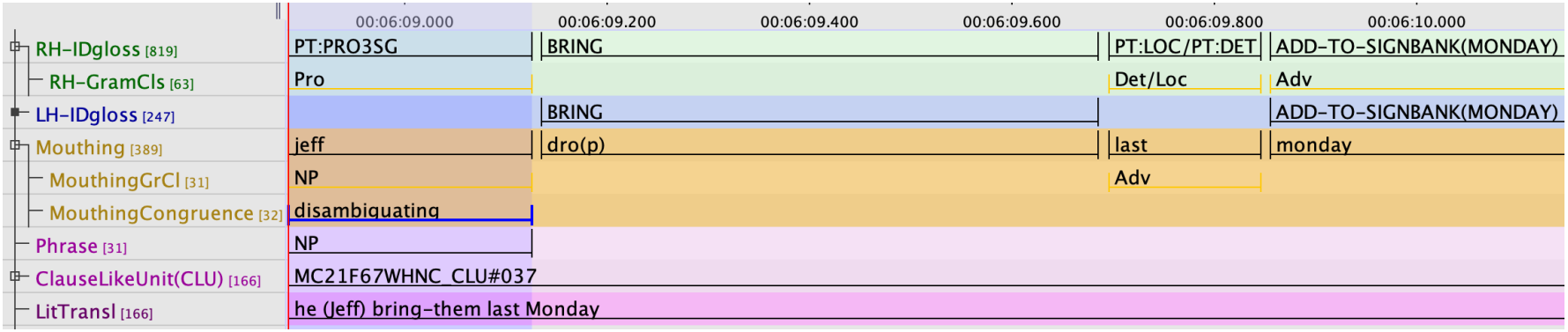
Figure 49 MC09c Example of Disambiguating Mouthing of "History"; "There's so much interesting history"



Members of the signing community may be assigned a specific personal *sign name* to identify them as individuals. Such sign names can consist of sequences of fingerspelled letters that are potentially ambiguous. Mouthings can be used to disambiguate such fingerspelled sign names–see, for example, Figure 23, which includes a disambiguation of a fingerspelling of (part of) the tennis player Martina Navratilova's name.

Mouthings can also disambiguate pointing signs, to identify who/what is being referenced. See, for example, Figure 50–here, the signer points to her right whilst mouthing “jeff”. She thereby associates the location that she indicates with the individual named Jeff.

Figure 50 MC21c Example of Mouthing of “jeff” to Disambiguate a Pointing Sign; “Jeff took them down on Monday at around lunchtime.”



When a mouthing is annotated as *UNSURE*, its mouthing congruence is annotated as *unsure*, as in Figure 41 above where it is not clear what is mouthed with the manual sign EIGHT.

Mouthing congruence is annotated as *NA* (not applicable) for *INDECIPHERABLE* mouthings. The congruence of regressive and progressive mouthings are also annotated as *NA,* as in Figure 44 above, since the concept of congruence only applies to the relationship between a mouthing and its corresponding manual sign.

Where the mouthing and the manual sign are congruent, or when the mouthing disambiguates a non-lexical manual sign, we analyse the combination as a single element within the noun phrase. Those mouthings which are annotated as *not congruent* are analysed as separate elements in the noun phrase, regardless of any simultaneity or overlap with a manual sign.

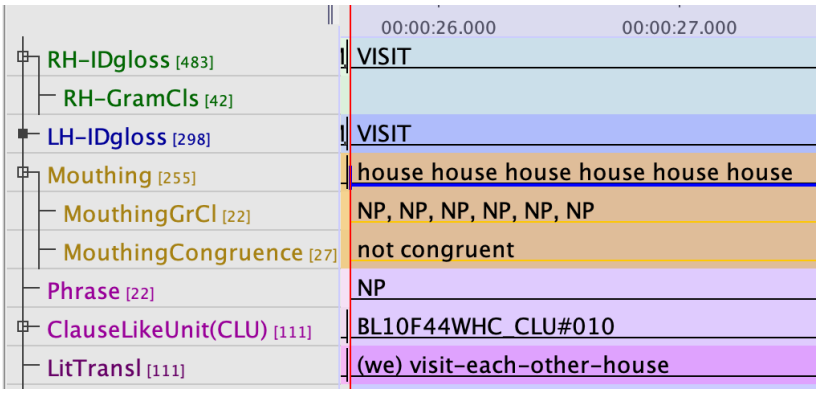
2.2.2.4.5 Multiple Mouthings with One Manual Sign

As discussed in Section 2.2.2.4.2, we align our mouthing annotations to their corresponding manual signs, rather than recording the actual start and end times of each mouthing. This means that on occasion, multiple mouthings are produced with a single manual sign (including situations when a signer mouths the same word repeatedly). We annotate such production as *not congruent,* with a mouthing grammatical category of *Unsure*. See, for example, Figure 12, where "for car" is mouthed at the same time as the manual sign VEHICLE is produced.

Such multiple mouthings with a single manual sign are split up and processed using R once the data has been exported (see Section 2.4). Only the mouthing(s) that are part of NPs are retained, and in our data, their congruence with the corresponding manual sign is updated if necessary. So, for the example in Figure 12, we dropped "for" and retained only "car". Within our exported data, the congruence of "car" with VEHICLE is set to *congruent*. We adopted this approach so that the annotations within ELAN could reflect all the signers' production, regardless of grammatical category, while allowing us to focus on NPs for this research.

Sometimes, not only are multiple mouthings produced with a single manual sign, but these mouthings constitute more than one mouthed NP. Here again, we edit our data in R to create separate NPs. See, for example, Figure 51. Here, the signer is moving the manual sign VISIT to and from her chest (referring to her own house) and different locations in space (each representing a different house). Each mouthing of "house" constitutes a separate NP so our data has been edited to reflect this.

Figure 51 BL10c Example of multiple mouthed NPs with a single manual sign; "… visiting each other at our homes..."



2.2.2.5 Simultaneous Production of Manual Signs

In addition to mouthings, we also considered cases where multiple manual signs were produced simultaneously. As discussed in Section 1.3.4.4, this includes instances of simultaneous production of different signs with each hand, as well mouthings. See for example the highlighted section of Figure 52, where the signer produces the manual sign ONE with his right hand at the same time as pointing to himself, as illustrated in Figure 53.

Figure 52 GW01n Simultaneous production of a different sign with each hand "There was one boy … who approached us straight away."

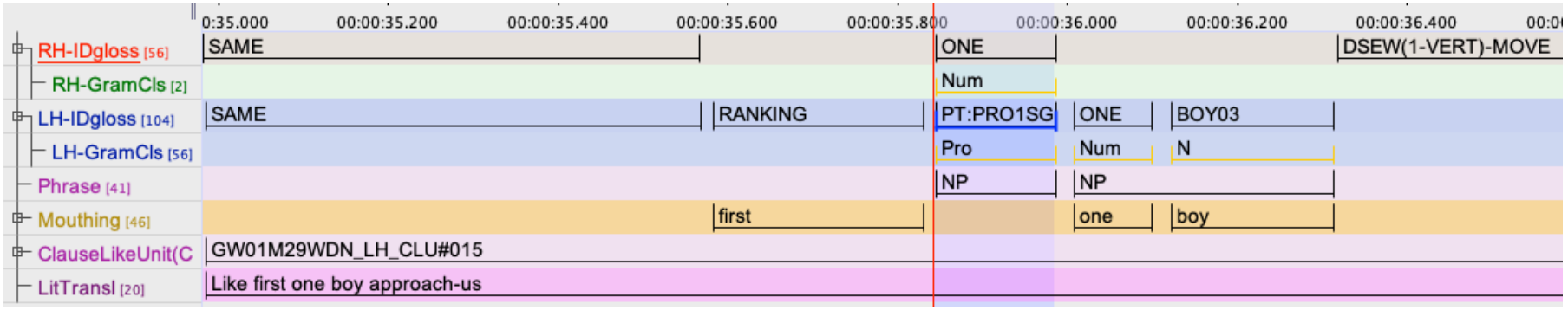
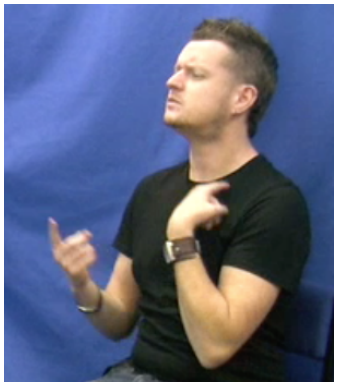


Figure 53 GW01n Image of simultaneous production of a different sign with each hand.



Where each hand is producing a separate sign as part of a NP, we annotated the grammatical category of each hand; this has the effect of including both hand's signs within the NP when the data is exported for analysis. This allows us to distinguish these from two-handed productions of a single sign, such as that shown in Figure 23. Here the grammatical category of the two-handed sign FAMOUS is only annotated on the dominant hand; in this case, this sign is only included once in our export.

Having outlined our approach to the corpus data, we now turn to a complementary method. While the BSL Corpus offers a rich source of naturalistic data, it reflects only a subset of possible NP structures and therefore cannot fully represent the range of orders that may occur in BSL. To address this, we conducted a large-scale judgement task with deaf BSL users. The next section outlines the design and implementation of this task.

2.3 Judgement Task

In this section, we describe our online judgement task. In this task, we asked deaf BSL users to watch videos of BSL sentences where the elements of the noun phrase were presented in different orders. They then rated how typical they thought that each was of the deaf BSL signing community in the UK–we used these typicality ratings as a proxy for the frequency of these orders in BSL. We chose the term "typical" rather than "natural" because, as discussed in Section 1.3.5.2, typicality refers to how expected or representative a structure is in usage, whereas naturalness can be influenced by processing or context. Our terminology follows Fenlon (2019), who used the same wording in a similar BSL judgement task (see Section 1.3.5.2.1). We collected responses from 92 participants between late July and November 2021: to our knowledge, this is the largest sample size to date for a task of this kind.

First (Section 2.3.1) we describe our task design and discuss the reasons for our approach. We list the sentences that we used and how we measured participants' responses. We include details of the demographic information that we collected about our participants. Then, in Section 2.3.2, we list our criteria for participant selection, and outline our recruitment process. In Section 2.3.3, we detail our publicity strategy, then in Section 2.3.4, we discuss how we managed the task, including monitoring the responses and we provide a breakdown of the demographic characteristics of our participants.

2.3.1 Task Design

2.3.1.1 Online task

We chose to deliver our task online using the survey platform Qualtrics (https://www.qualtrics.com/). This approach enabled us to recruit a large number of participants from all parts of the UK without incurring travel costs to/from our base in London, for either the participants or the experimenters. Although the initial design and setup of the task involved considerable investment of time and resources–particularly in piloting the task across a range of browsers and devices to ensure compatibility–it reduced the need for researcher time during data collection. We conducted the experiment during the COVID-19 pandemic, so it turned out that our choice of online paradigm was fortuitous since social distancing restrictions and other infection control measures posed additional barriers to in-person research.

One pitfall of running online experiments is that if a general link to the experiment is published, automated "bots" rather than humans may complete the task and claim the associated incentive. The chances of this were minimised in our experiment because, as discussed in Section 2.3.2, we asked potential participants to email us to express interest in our study, then we issued a unique link to the task to each participant – this was not to be shared with other people. As far as we are aware, no "bots" completed our task: our email conversations appeared to be with humans, and, as described in Section 2.5, we checked for (but did not find) patterns in the task responses that could indicate automated responses.

There are other drawbacks to online research. We discuss these in general in Section 1.3.5.2, and in the remainder of this section, we consider how they apply to our study. One issue is that our participants needed access to a computer or tablet whose internet connection was suitable for streaming our videos of BSL sentences, and whose screen was large enough to watch pairs of such

videos. This meant that mobile phones, though more ubiquitous, were not suitable for the task. We also had to test our task on a wide variety of devices and browsers rather than just the particular configuration in our lab.

The self-administered nature of our task necessitated that it be designed for completion without supervision. This requirement may have inadvertently excluded individuals with limited access to technology or those who experience difficulties using digital devices. Although data collection occurred between late July and November 2021 – after the most stringent COVID-19 lockdowns in the UK had been lifted – residual effects of the pandemic persisted. These included reduced opportunities for participants to access devices outside their homes, such as in workplaces or public libraries, potentially impacting participation.

The remote task delivery paradigm introduces additional challenges for researchers to be confident that participants are attending to the task and performing it properly. Our participants may have been subject to more distractions or interruptions than usual if they were doing the task at home as a result of COVID-19 restrictions. We mitigated this by introducing a delay before a participant could respond to each question (see Section 2.3.1.4), to encourage them to watch each sentence video rather than speeding through without providing considered answers. Additionally, we allowed participants to take breaks whenever they wished during the task, to help minimise fatigue, and we checked the completed responses for any patterns that might indicate inattention, as noted in Section 2.5. We did not identify any such patterns, so we did not exclude any participants on this basis.

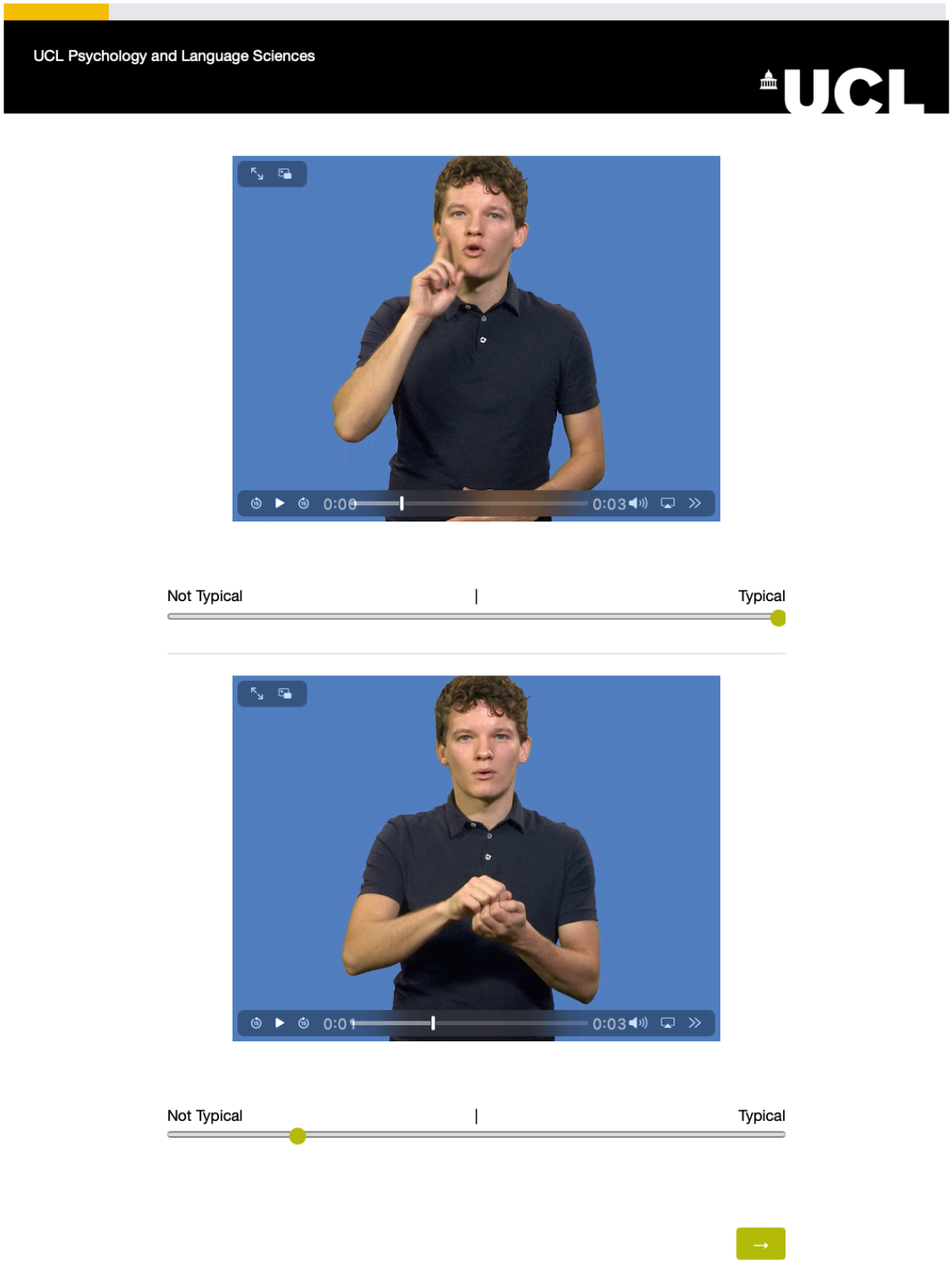
2.3.1.2 Sentence Presentation

Marty et al. (2020) found that joint presentation of sentences in theoretically-related pairs provided maximum sensitivity, as discussed in Section 1.3.5.2.2. We therefore presented our sentences in this way (see Figure 54). For example, our sentences containing [Noun Adj], and [Adj Noun] (see Table 5 below) were shown on the same screen. For the larger groups of sentences (i.e., the nouns with two or more modifiers), for each participant, we selected pairs at random from each group, so that each sentence was shown once alongside another sentence from its group. Again following Marty et al., we also displayed our sentences in random order within each pair–this meant that no particular order of sentences was favoured, serving to remove any priming effect e.g., if always presenting a certain sentence first.

The fact that sign languages have no written form means that our sentences were presented as videos. Therefore, it may be that both members of the pair are not as present in the participant's working memory as for written stimuli. We aimed to minimise this effect by choosing short sentences–our clips ranged from 3 to 5 seconds in duration–and by allowing each participant to watch each pair of sentences as often as they wished before finalising their responses to that pair. This mirrors the presentation of written sentences in an analogous task in that these could be re-read by participants.

Beneath each sentence we provided a slider for the participant to move to rate the sentence.

Figure 54 Example Screen from BSL Sentence Task showing videos of two theoretically-related sentences, with a response slider underneath each.



Note. There is an orange progress bar at the top and a green arrow button at the bottom to proceed to the next question.

2.3.1.3 Part of a Larger Task

Our NP sentences formed part of a larger task comprising 90 sentences in total. Participants were also asked to rate other types of sentences (questions, negations, and sentences with varying constituent order) as well as our NP sentences. The layout of the screens for question and negation sentence was identical to that used for our NP sentences, so we randomised the order of presentation across all these sentence types. However, a limitation of Qualtrics means that all sentences from a given group must be displayed sequentially. This meant that, for example, all six pairs of sentences with nouns with three modifiers listed in Section 2.3.1.5.3 were presented consecutively. Within each group, the order of sentence presentation was randomised for each participant.

2.3.1.4 Overall Task Structure

All the information in our task aside from the test sentences was presented in both BSL and English. The test sentences were just presented in BSL, without English translation or glossing. This was to minimise influence from English, since the research focuses on BSL structure.

Our task started with providing a standard information sheet (see Appendix B.4), listing the researchers involved in the project and their affiliations, plus ethics approval information. It also briefly described the project and listed the inclusion criteria (detailed in Section 2.3.2). It explained that participation is optional and that on completion, participants would be sent a £15 E-gift card as compensation for their time. This was followed by statutory information relating to data protection (see Appendix B.5), and then a Consent section (see Appendix B.6), in which the participant had to select a series of statements to confirm that they had seen and understood this information and that they consented to participate. They were also asked to confirm that they met the criteria for our study.

Once the participant had agreed to all the consent statements, they were shown instructions explaining that they would be shown a series of pairs of videos showing BSL sentences, and would need to move a slider to rate each sentence: Following Fenlon (2019), we asked them to judge how typical they felt the sentence was of the deaf BSL signing community. Participants were told that they could take a break at any time, and later resume where they left off–previous answers would be saved. Next, three pairs of overt practice sentences (see Appendix B.7) were presented, in the same format as the test sentences, and the participant could move each video's slider to rate the

sentence. This allowed the participant to familiarise him/herself with the task process. The responses to the practice sentences were not included in our analysis.

As noted in Section 2.3.1.2, participants were able to watch the sentence videos as many times as they wished before rating them: our task had no time limits. And as discussed in Section 2.3.1.1, we included a four-second delay before it was possible to start moving the sliders on each page, to mitigate the risk of participants selecting a response without watching the video. Participants were only permitted to move forward through the sentences – they were not able to go back to previous questions. A progress bar at the top of the screen indicated the proportion of the task completed so far.

2.3.1.5 NP sentences

We presented simple SVO sentences. The object NP in each contained a noun plus one or more modifiers–those modifiers were either a demonstrative point, an adjective or a numeral. We used demonstrative points (even though as noted in Section 1.3.4.2.2, these are not considered to form a distinct grammatical category in BSL) since demonstratives are considered in Universal 20 (see Section 1.3.2.1)–this enabled us to test predictions about Universal 20 using this data.

The modifiers appeared at different positions with respect to one another and the noun. Overall, there were 36 sentences: English glosses for these sentences are provided in Sections 2.3.1.5.1 to 2.3.1.5.3.

We worked with the model who produced our NP sentences to ensure neutral nonmanual features, neutral use of space and consistent prosody, without any pauses within each target NP. As noted in Section 1.3.4.4, there is a significant amount of simultaneity of production in spoken language, and even more in signed language. However, as a first systematic investigation of the order within NPs in BSL, this project focussed solely on the sequential order of manual signs. These were accompanied by mouthings that were congruent (see Section 2.2.2.4.4) to the manual signs so that the utterances were as naturalistic as possible. We chose not to include other non-manual features of BSL such as body leans, head tilts or brow raise/lowering, so as to avoid these non-manual features affecting participants' judgements: Our aim was to ensure that participants based their typicality ratings on the order of elements alone. We acknowledge that this made the sentences appear somewhat unnatural, and indeed several participants commented on this–see Section 4.4 for details.

2.3.1.5.1 Noun plus one modifier

All combinations of noun plus one modifier were presented after a carrier phrase (either GIRL BUY or BOY BUY), with the modifier positioned either before or after the noun. See Table 4 - Table 6.

Table 4 Judgement Task Sentences: Noun and Demonstrative Point

Code	Sentence glosses	NP structure
NP01	GIRL BUY PT VEHICLE	[Dem Noun]
NP02	GIRL BUY VEHICLE PT	[Noun Dem]

Table 5 Judgement Task Sentences: Noun and Adjective

Code	Sentence glosses	NP structure
NP03	GIRL BUY BLUE VEHICLE	[Adj Noun]
NP04	GIRL BUY VEHICLE BLUE	[Noun Adj]

Table 6 Judgement Task Sentences: Noun and Numeral

Code	Sentence glosses	NP structure
NP05	BOY BUY THREE CAKE	[Num Noun]
NP06	BOY BUY CAKE THREE	[Noun Num]

2.3.1.5.2 Noun plus two modifiers

Sentences with a noun plus two modifiers included a carrier phrase (GIRL REMEMBER or BOY EAT or GIRL BUY) followed by:

- adjective and demonstrative point,
- numeral and demonstrative point, or
- numeral and adjective.

We included all six ways to arrange the noun and its two modifiers; the sentences are given in Table 7 - Table 9:

Table 7 Judgement Task Sentences: Noun, Adjective and Demonstrative Point

Code	Sentence glosses	NP structure
NP07	GIRL REMEMBER OLD FILM PT	[Adj Noun Dem]
NP08	GIRL REMEMBER OLD PT FILM	[Adj Dem Noun]
NP09	GIRL REMEMBER FILM OLD PT	[Noun Adj Dem]
NP10	GIRL REMEMBER FILM PT OLD	[Noun Dem Adj]
NP11	GIRL REMEMBER PT FILM OLD	[Dem Noun Adj]
NP12	GIRL REMEMBER PT OLD FILM	[Dem Adj Noun]

Table 8 Judgement Task Sentences: Noun, Numeral and Demonstrative Point

Code	Sentence glosses	NP structure
NP13	BOY EAT FOUR SANDWICH PT	[Num Noun Dem]
NP14	BOY EAT FOUR PT SANDWICH	[Num Dem Noun]
NP15	BOY EAT SANDWICH FOUR PT	[Noun Num Dem]
NP16	BOY EAT SANDWICH PT FOUR	[Noun Dem Num]
NP17	BOY EAT PT SANDWICH FOUR	[Dem Noun Num]
NP18	BOY EAT PT FOUR SANDWICH	[Dem Num Noun]

Table 9 Judgement Task Sentences: Noun, Numeral and Adjective

Code	Sentence glosses	NP structure
NP19	GIRL BUY THREE BOOK OLD	[Num Noun Adj]
NP20	GIRL BUY THREE OLD BOOK	[Num Adj Noun]
NP21	GIRL BUY BOOK THREE OLD	[Noun Num Adj]
NP22	GIRL BUY BOOK OLD THREE	[Noun Adj Num]
NP23	GIRL BUY OLD BOOK THREE	[Adj Noun Num]
NP24	GIRL BUY OLD THREE BOOK	[Adj Num Noun]

2.3.1.5.3 Noun plus all three modifiers

Finally, we created sentences containing a noun plus an adjective, a numeral, and a demonstrative point following the carrier phrase BOY BRING. There are 24 possible combinations, listed in Table 10. We piloted our task including all of these 24 sentences, but early feedback indicated that this was too many similar sentences: particularly since, as noted in Section 2.3.1.3, they had to be presented sequentially as a group. Therefore, for the task proper, we decided to omit the sentences where modifiers were found on both sides of the noun: these are struck through in Table 10. These omitted items are analogous to the two-modifier sentences described in Section 2.3.1.5.2. For example, the

prenominal elements of the NP within NP30 (demonstrative then numeral) and NP36 (numeral then demonstrative) are in the same order as those in NP18 and NP14 respectively. We opted to use these shorter two-modifier sentences to assess the typicality of prenominal numeral-demonstrative order. The remaining 12 sentences still represented a relatively large set for participants to evaluate, but in the absence of further principled grounds for reduction, we retained them all.

Table 10 Judgement Task Sentences: Noun, Demonstrative Point, Adjective and Numeral

Code	Sentence glosses	NP structure
NP29	BOY BRING PT FOUR RED MUG	[Dem Num Adj noun]
~~NP30~~	~~BOY BRING PT FOUR MUG RED~~	~~[Dem Num Noun Adj]~~
NP31	BOY BRING PT RED FOUR MUG	[Dem Adj Num noun]
~~NP32~~	~~BOY BRING PT RED MUG FOUR~~	~~[Dem Adj Noun Num]~~
~~NP33~~	~~BOY BRING PT MUG FOUR RED~~	~~[Dem Noun Num Adj]~~
~~NP34~~	~~BOY BRING PT MUG RED FOUR~~	~~[Dem Noun Adj Num]~~
NP35	BOY BRING FOUR PT RED MUG	[Num Dem Adj noun]
~~NP36~~	~~BOY BRING FOUR PT MUG RED~~	~~[Num Dem Noun Adj]~~
~~NP37~~	~~BOY BRING FOUR MUG PT RED~~	~~[Num Noun Dem Adj]~~
~~NP38~~	~~BOY BRING FOUR MUG RED PT~~	~~[Num Noun Adj Dem]~~
NP39	BOY BRING FOUR RED PT MUG	[Num Adj Dem noun]
~~NP40~~	~~BOY BRING FOUR RED MUG PT~~	~~[Num Adj Noun Dem]~~
NP41	BOY BRING RED PT FOUR MUG	[Adj Dem Num noun]
~~NP42~~	~~BOY BRING RED PT MUG FOUR~~	~~[Adj Dem Noun Num]~~
NP43	BOY BRING RED FOUR PT MUG	[Adj Num Dem noun]
~~NP44~~	~~BOY BRING RED FOUR MUG PT~~	~~[Adj Num Noun Dem]~~
~~NP45~~	~~BOY BRING RED MUG PT FOUR~~	~~[Adj Noun Dem Num]~~
~~NP46~~	~~BOY BRING RED MUG FOUR PT~~	~~[Adj Noun Num Dem]~~
NP47	BOY BRING MUG PT FOUR RED	[Noun Dem Num Adj]
NP48	BOY BRING MUG PT RED FOUR	[Noun Dem Adj Num]
NP49	BOY BRING MUG FOUR PT RED	[Noun Num Dem Adj]
NP50	BOY BRING MUG FOUR RED PT	[Noun Num Adj Dem]
NP51	BOY BRING MUG RED PT FOUR	[Noun Adj Dem Num]
NP52	BOY BRING MUG RED FOUR PT	[Noun Adj Num Dem]

2.3.1.6 Response measurement

Following Marty et al. (2020) as discussed in Section 1.3.5.2.3, we chose a continuous response scale. This has the potential to reveal any gradience of judgements in what constitutes a typical sentence in BSL (as opposed to, say, a forced-choice “is this sentence typical?” where answers are

only Yes or No). Again following Marty et al., we only labelled the end points of our scale, and we used the labels *typical* and *not typical* as per Fenlon (2019). This also minimised the amount of English on the sentence pages, thereby helping the participants to focus on the BSL provided. We included a marker at the centre of our scale (see the screenshot in Figure 54), so that participants were aware of the location of the mid-point even after moving the slider.

2.3.1.7 Demographic questions after the task

After participants completed their ratings of our sentences, we collected the following demographic information from each. (See Appendix B.9 for the English version of these questions and Section 1.3.4.6 for a discussion of previous research into the effect of such social factors on the use of BSL).

1. Year of birth. We chose not to ask for participants' full date of birth so as to preserve their anonymity – this removes the possibility of being able to identify someone based on their date of birth and other demographic characteristics, which can be an issue with small populations such as the UK deaf community.

The next three questions contribute to the classification of participants as native, early or late learners of BSL, as described in Section 2.3.4.2.1 (for a definition and discussion of the concept of native signer, see Section 1.3.4.6):

2. Age of acquisition of BSL, split into: 0-7 years; 8-12 years; 13-18 years; 18 years or older.
3. Which family members are/were deaf – a check box for each of the following: Mother / primary caregiver; father / secondary caregiver; maternal grandmother; maternal grandfather; paternal grandmother; paternal grandfather.
4. How many deaf brothers and sisters who use BSL do you have, split into: Older brothers; younger brothers; older sisters; younger sisters.

After this, we asked about participants' experience of teaching BSL, and their level of comfort with English:

5. Have you ever taught BSL classes? If so, how long have you been a BSL teacher (0-5 years, 6-10 years, more than 10 years). And have you ever had training to teach BSL?
6. Do you use speech and/or lipreading as your main form of communication with hearing people in everyday situations?
7. On a scale of 1 to 7, how comfortable are you with English?

Finally, we questioned participants about their gender, where they have lived, and how they describe their ethnicity:

8. What is your gender (Female, Male, Non-binary, Other, I prefer not to say)
9. Which region of the UK do you live in now? (Categories based on Office for National Statistics, 2016)
10. Which regions of the UK have you lived in during the past ten years? Please select as many as necessary. (Categories based on Office for National Statistics, 2016)
11. What is your background? (Using the highest-level categories from the Office for National Statistics, 2013: White, Asian/Asian British, Black/African/Caribbean/Black British, Mixed/multiple ethnic groups, Other ethnic group).

We appreciate that these categories, indeed the very concept of ethnicity, is problematic (Ahmadzadeh, 2021). However, we decided to include them because we are not aware of a more suitable set of categories available at present and without this question, we would have no information about the race/ethnicity/ancestry of our participants.

2.3.1.8 Other questions at the end of the task

We gave participants the option to sign up to receive information about future studies that they might wish to participate in. If they agreed to this, we captured an email address to be used to contact them in future.

Then, we asked all participants for optional feedback at the end of the task–whether they experienced any technical problems or interruptions that could have affected their responses, and whether they had any comments about the sentences that we presented. Their responses were automatically emailed to the Study Manager as soon as the participant completed the task, and checked frequently so that they could be acted-upon if necessary, for example, if a participant reported a technical problem with Qualtrics. Indeed, we clarified our question about BSL teaching experience as a result of this feedback, as discussed in Section 2.3.4.2.5. Participants' feedback is discussed in Section 4.4.

Finally, we provided some debrief information to participants, to explain the purpose of our study. This can be found in Appendix B.11.

2.3.1.9 End of Task Processes

In addition to the emails that could be triggered by questions at the end of the task that are detailed in Section 2.3.1.8, the following emails were sent each time that a participant completed the task:

- A copy of the information sheet, data protection information and consent form (as in Section 2.3.1.4) to the participant. This was provided in English, along with a link to the same information in BSL, hosted online.
- Information about how the participant could redeem their £15 shopping voucher using the e-gift card company Tango Card (https://www.tangocard.com/survey-incentives/).

2.3.2 Participant criteria and the Recruitment Process

We aimed to recruit a large number of participants (around 90), to minimise the issues related to use of a small number of individuals' judgements that were discussed in Section 1.3.5. Our eligibility criteria were that participants should:

- Identify as deaf;
- Use BSL as their main language;
- Be born in the United Kingdom;
- Be 18 years or older.

We avoided focusing our recruitment on participants with a particular age of acquisition of BSL, such as native signers. This is because we wanted to sample the whole adult deaf BSL-using population born in the UK so as to document the full range of BSL use, as discussed in Section 1.3.4.6.

Participants were asked to email our Study Manager to express an interest in our study and in response, we asked them to reply to self-certify that they met each of the criteria. Then as part of the consent process at the start of the task itself, they were required to select checkboxes to confirm again that they met each criterion. It was not possible to progress without doing this. It is of course possible that people who did not meet our criteria may have completed the task but we believe this is unlikely: we checked for outliers in our results but did not find any, as discussed in Section 2.5.

Once a participant indicated that they met our criteria, we used Qualtrics to generate a unique link to the task and email it to them. The unique link, personal to that participant, meant that, as noted in Section 2.3.1.4, they could take a break whenever they wished, and their responses would be saved. They could then resume the task later from where they left off.

2.3.3 Publicity

In all our publicity, we directed potential participants to the project web page which contained information in both BSL and English–see Appendix B for the English text. We piloted our study extensively with colleagues who were not involved in our project, in order to arrive at our final design. At launch, we advertised a link to the project web page–since this project is based at University College London's Deafness, Cognition and Language Research Centre (DCAL), we used DCAL's Facebook and Twitter accounts, as well as emailing people who have signed up to be kept informed of DCAL's activities. Additionally, we invited everyone already signed up to DCAL's participant database who met our criteria to participate. As data collection progressed, we publicised further on social media, including emphasising that our study was open to everyone who uses BSL every day (since we received feedback that some members of the deaf community felt that research participation was too high-level for them), and also as the study progressed, targeting our messages at particular groups that were under-represented in our data. This is discussed in Section 2.3.4.2.

2.3.4 Task Management

2.3.4.1 Reminders

Periodically, we used Qualtrics to issue reminder emails to participants who had either not started the task, or who had started it but not yet completed it. We sent up to two reminders to each participant: the first was delivered when they had received their invitation at least a week previously, and the second "last chance" reminder was sent at least a week after this, giving them one further week to complete the task. In each reminder email we gave participants the opportunity to withdraw from the study if they wished.

2.3.4.2 Monitoring responses

Because we were targeting a relatively small population, we recruited a convenience sample of people that we could reach via our publicity, starting with people who follow DCAL's social media accounts and extending to their networks. Prior to the start of data collection, we agreed that we would monitor the demographic characteristics of task participants, and that we would amend our publicity and if necessary, our selection criteria, to target underrepresented groups as required.

2.3.4.2.1 Native, Early and Late Learners of BSL

As discussed in Section 1.3.4.6, a signer's language background–specifically, whether they are a native signer–may influence their BSL production. We classified participants as native signers if they were likely exposed to BSL from birth through deaf relatives. This classification was based on responses to questions in Section 2.3.1.7: participants who reported learning BSL by age seven (question 2), and having at least one older deaf family member (questions 3 and 4) from whom they are likely to have acquired the language. Participants who reported learning BSL by the age of seven but who did not have any older deaf family members were classified as *early learners*.

Those who reported learning BSL after the age of seven were designated as *late learners*. This classification was applied irrespective of whether participants reported having older deaf relatives (as was the case for 12 participants) as we did not collect information about the extent of their contact with these relatives or the degree to which those relatives used BSL. Consequently, we cannot determine the extent to which these participants may have acquired BSL through familial exposure prior to age seven. Thus, some of the late learners may actually have learned BSL much earlier.

We aimed for a balance of native, early and late learners of BSL, and this was approximately achieved. In fact, late learners were somewhat over-represented and early learners somewhat under-represented. With hindsight, the higher percentage of late learners is advantageous since this better reflects the overall demographics of the deaf population in the UK, as discussed in Section 1.3.4.6. However, given that some participants classified as late learners may have acquired BSL earlier through familial exposure, the actual distribution may be more balanced than it appears. The actual numbers of participants in each category are shown in Table 11.

Table 11 Judgement Task Participant Breakdown by Native/Early/Late Learner of BSL

	n	**Percentage**
Late learner	37	40.2
Native signer	29	31.5
Early learner	26	28.3

Note. Ordered by decreasing percentage.

2.3.4.2.2 Region

We hoped for a balance of participants from different regions of the UK. However, we were mindful that because of DCAL's location in London, many of our contacts in the deaf community live in

London and the South-East of England. This is partly because past research conducted by DCAL generally involved face-to-face participation which was more convenient for people living relatively nearby. This proved to be the case as shown in Table 12 – nearly 35% of our participants lived in London or the South-East of England. As data collection progressed, we observed a lack of participants Northern Ireland or the East Midlands, so we targeted social media publicity towards people in these regions.

Table 12 Judgement Task Participant Breakdown by Current Region

Current region	***n***	**Percentage**
London	20	21.7
South-East England	12	13
East of England	10	10.9
North-East England	7	7.6
Scotland	7	7.6
West Midlands	7	7.6
North-West England	6	6.5
Wales	5	5.4
Yorkshire and the Humber	5	5.4
South-West England	4	4.3
East Midlands	3	3.3
Northern Ireland	3	3.3
Outside of the UK	3	3.3

Note. Ordered by decreasing percentage.

2.3.4.2.3 Ethnicity

The ethnic mix of our sample broadly mirrors the ethnic breakdown of the UK population as a whole, as shown in Table 13. Asian/Asian British people and those of mixed/multiple ethnic groups were under-represented in our sample. Black/African/Caribbean/Black British people and those from other ethnic groups were over-represented. However, surveys since the 2011 Census have shown that around 20% of the UK deaf population are non-white (Wilson & Hoong Sin, 2015).

Table 13 Judgement Task Participant Breakdown by Ethnicity

	Our study		UK population[a]
Ethnicity	***n***	**Percentage**	**Percentage**
White	79	85.9	87.1
Asian / Asian British	5	5.4	7.0
Black / African / Caribbean / Black British	5	5.4	3.0
Other ethnic group	3	3.3	0.9
Mixed / multiple ethnic groups	0	0	2.0

Note. Ordered by decreasing percentage.

[a] Taken from the UK Census 2011 (Office for National Statistics, 2013).

2.3.4.2.4 Gender

The gender balance of participants in our survey is shown in Table 14. We did not attempt to balance gender recruitment since we had no reason to expect women and men to have differences in judgements.

Table 14 Judgement Task Participant Breakdown by Gender

Gender	***n***	**Percentage**
Female	62	67.4
Male	28	30.4
I prefer not to say	1	1.1
Other	1	1.1

Note. Ordered by decreasing percentage.

2.3.4.2.5 Teachers versus non-Teachers

Towards the end of the recruitment process, it became clear that we had many more respondents with experience of teaching BSL than those who did not, as shown in Table 15. Therefore latterly, we specified in our social media publicity that we were only needed participants who were non-teachers. This resulted in three more completed tasks, two from non-teachers and one from a teacher. We later further tightened this such that when potential participants contacted us, we only accepted them if they said they had not previously taught BSL. This resulted in us disallowing one person from participating.

Initially, we used the question “Have you ever taught BSL?” to determine teacher status. However, following feedback from one participant who said that they had taught BSL but only informally, we amended the question to “Have you ever taught BSL classes?” to clarify our intent. This change was made after we had collected 26 responses. Of these, 7 participants indicated that they had taught

BSL but had not received training. It is therefore possible that up to 7 of the 92 participants (7.6%) may have been mis-classified as BSL teachers if they interpreted the original question as including informal teaching.

Table 15 Judgement Task Participant Breakdown by Teacher Status

Ever taught BSL?	*n*	Percentage
Yes	65	70.7
No	27	29.3

Note. Ordered by decreasing percentage.

2.4 Analysis of Corpus Data

We exported our annotations from ELAN and used R version 4.4.0 (R Core Team, 2024) to perform various manipulations of the data including merging it with a file of participant meta-data, which contains information about various social factors such as gender, age, region and language background.

As discussed in Section 2.2.2.4.5, we cleansed our data as follows:

- We excluded from our analysis any mouthings or manual signs that were not part of an NP, but which had been exported because they co-occurred with an NP;
- We split single annotations consisting of multiple NPs into separate NPs.

Similarly, we excluded any mouthings or manual signs whose grammatical category was annotated as *Unsure*. However, we retained NPs that we had annotated as *unsure*. This label reflects uncertainty about specific aspects of the NP–such as the grammatical category of a particular sign (e.g., whether it is functioning as a noun or a verb)–rather than uncertainty about whether the unit constitutes an NP at all (see Section 2.2.2.2.1).

We also reduced all types of pointing sign (i.e., grammatical category annotations of *Pro, Det, Det/Loc, Det/Loc/Pro* and *Loc/Pro*) to a single category of *Point*; we did this because of the high degree of ambiguity as to the grammatical category of each point. This means that when discussing the order of elements in BSL and when determining in Section 3.1.3.4 whether a NP is homomorphic, we will consider nouns, adjectives and numerals plus all pointing signs taken together (the latter in lieu of demonstratives, as explained in Section 1.3.4.2.2).

We calculated the length of each NP in our data. We considered the number of elements based on BSL rather than English, so that, for example, we counted AGE-FOUR as having length 1 because it is a single sign in BSL, even though it requires two words "age four" in English. As noted in Section 2.2.2.4.2, we excluded regressive and progressive mouthings in our calculation of the length of each NP, to avoid double-counting on top of the mouthing of main, (non-regressive/progressive) mouthing.

We next turn to the analysis of our online judgment task data.

2.5 Analysis of Judgement Task Data

We exported the task results from Qualtrics and used R version 4.4.0 (R Core Team, 2024) to perform all data processing and statistical analysis. We anonymised the data by removing personally identifiable information such as names and email addresses. We then manually inspected each participant's responses for patterns indicative of inattention or automated completion[23] (e.g., providing the same rating for every sentence). No such patterns were observed, so no participants were excluded on this basis.

Qualtrics recorded the task slider responses as scores ranging from 0 to 100, with 0 corresponding to the far left of the scale (maximally atypical) and 100 to the far right (maximally typical). To mitigate individual differences in scale use, we z-transformed the scores for each sentence from each participant, as discussed in Linzen and Oseki (2018), using the formula in (58).

(58) z-score = (score – participant mean score)/participant standard deviation score

To investigate our hypothesis concerning signers' self-reported level of comfort with English (see Section 1.4), we derived a composite measure based on two of the demographic questions detailed in Section 2.3.1.7. The first variable concerned participants' reported primary mode of communication with hearing individuals in everyday situations. The second variable directly asked participants to rate their level of comfort with English on a 7-point scale (1 = very uncomfortable, 7 = very comfortable). Participants were categorised as having *higher comfort with English* if they:

- Reported using speech or lipreading as their main form of communication with hearing people; and
- Rated their comfort with English as relatively high (a score of 6 or 7 on the scale).

[23] Questions were presented to participants in random order. We would like to thank Dr Yasamin Motamedi for assistance in enabling this check by arranging the responses into the order provided to each participant.

All other participants were classified as having *lower comfort with English*. Although only the second variable directly assessed English comfort, we included the first–communication mode–as a supplementary indicator. This decision was based on the assumption that individuals who primarily use speech or lipreading in interactions with hearing people are likely to possess greater functional proficiency and confidence in English. By combining a subjective self-assessment with a behavioural indicator, we aimed to capture both perceived and enacted comfort with English, thereby enhancing the robustness of our classification.

Table 16 shows the breakdown of our participants in terms of their self-reported level of comfort with English.

Table 16 Judgement Task Participant Breakdown by Self-Reported Comfort with English

	n	**Percentage**
Lower comfort with English	65	70.7
Higher comfort with English	27	29.3

Note. Ordered by decreasing percentage.

We developed various linear models, as detailed in Sections 2.5.1–2.5.4. Linear mixed-effects models were fitted using the *lme4* 1.1.35.3 package (Bates et al., 2015), with p-values and degrees of freedom calculated using the *lmerTest* 3.1.3 package (Kuznetsova et al., 2017). For each grouping listed in these sections, we created two types of model:

1) Simple linear regression model: This model used participants' ratings of typicality for each sentence as the dependent variable. Independent variables included participants' reported age of acquisition of BSL (operationalised as native, early or late learners of BSL as discussed in Section 2.3.4.2.1), their self-reported level of comfort with English and whether they reported having taught BSL, and other independent variables as appropriate to the specific model detailed in Sections 2.5.1 – 2.5.4.
2) Linear mixed effects model (*LMEM*): This model included the same variables as the simple linear regression model, with the addition of participant number as a random intercept. For groups of six or more sentences, we also included the order of sentence presentation as another random intercept. However, for LMEMs with only two sentences, we did not include trial order as a random effect because it only has two levels (each sentence is either presented first or second in the pair): Harrison et al. (2018) suggest that models may not be accurate if factors with fewer than five levels are included as random effects. These LMEMs

with only two sentences had singularity issues, so we discarded them and report just the linear regression models.

Similarly, our LMEM for sentences NP13 to NP18 encountered singularity issues because the random effect of order of sentence presentation order was zero. Following Barr et al. (2013), we removed this variable from the model and report the resulting output.

We ordered the levels of our social factors as follows:

- Reported age of acquisition of BSL: native 0/early 1/late 2
- Have taught BSL: yes 0/no 1
- Comfort with English (based on self-report): lower 0/higher 1

For each grouping listed in Sections 2.5.1 – 2.5.4, we used ANOVA to compare the simple linear regression model to the LMEM. Where this indicated that the LMEM was a better fit to the data, we report this, and otherwise we report the simple linear regression model. For all of the models in Section 2.5, we used the raw typicality ratings, ranging from zero to one hundred, rather than the standardised z-scores, since we account for individual differences between participants in LMEMs by including participant as a random intercept.

2.5.1 Linear Models for Groups of Sentences

We constructed models to compare scores within each of our seven sentence groups, as defined in Section 2.3.1.5. For instance, we compared the six sentences whose NPs consisted of a noun, numeral and adjective (NP19 to NP24) with one-another. In each model, sentence code (e.g., NP01, NP02) served as the independent variable.

In each model, one sentence was designated the reference sentence–the sentence against which all others in the group were compared. These reference sentences are listed in Table 17. Where possible, we selected as the reference the sentence whose NP order matched the most frequent order observed in our BSL corpus data. For example, the order in NP03, [Adj Noun], was chosen over NP04, [Noun Adj] because the former order was more frequent in our corpus data:

(59) NP03 GIRL BUY BLUE VEHICLE

We were able to identify a most frequent order for our length 2 NPs and for some–though not all–of the longer NPs. In most cases, this order involved all modifiers appearing prenominally. For NPs with more than one modifier on the same side of the noun, the most frequent order also tended to

be homomorphic. To maintain consistency across models, we selected as the reference sentence the one whose NP had all modifiers in prenominal position and, where applicable–that is, for NPs with more than one modifier on the same side of the noun–followed the homomorphic order.

Table 17 Judgement Task Reference sentences

Code	Sentence glosses	NP structure
NP01	GIRL BUY PT VEHICLE	[Dem Noun]
NP03	GIRL BUY BLUE VEHICLE	[Adj Noun]
NP05	BOY BUY THREE CAKE	[Num Noun]
NP12	GIRL REMEMBER PT OLD FILM	[Dem Adj Noun]
NP18	BOY EAT PT FOUR SANDWICH	[Dem Num Noun]
NP20	GIRL BUY THREE OLD BOOK	[Num Adj Noun]
NP29	BOY BRING PT FOUR RED MUG	[Dem Num Adj Noun]

2.5.2 Linear Models to Examine Modifier Position.

We created further sets of linear models and LMEMs to examine whether there is a difference in score depending on whether a particular modifier (demonstrative point, then adjective, then numeral) is prenominal or postnominal. Here, we included the position (prenominal or postnominal) of the modifier in question as an additional fixed effect in our models. We again chose the reference group based on the structure that was most common in our BSL Corpus data: This was a prenominal modifier for all three models.

2.5.3 Linear Models to Examine Homomorphism and English-like Order.

We created both a linear model and a LMEM to investigate the effect of homomorphism on NP typicality scores, including homomorphism as a fixed effect. Similarly, we developed a linear model and LMEM to examine whether NPs in English-like order are judged to be more or less typical. Homomorphism can be determined for NPs that contain at least two different types of modifier on the same side of the noun. It cannot be determined for the remaining NPs in our judgement task. For example, the target NP in NP07 (GIRL REMEMBER OLD FILM PT) has only an adjective (OLD) before the noun (FILM) and only a demonstrative (PT) after the noun. This means that there is no relative order of modifiers either before or after the noun in this NP and homomorphism does not apply. Therefore, we included all judgement task sentences that have at least two modifiers on the same side of the noun in our homomorphism models.

We also chose to examine the effect of whether NPs are in English-like order on our typicality scores, since, as discussed in Section 3.1.3.5, our corpus work produced the robust finding that for our NPs where English-like order could be determined, 74% were in the same order as in English. Therefore, we created a further linear model and LMEM including all our judgement task sentences, with a fixed effect for whether each sentence was in the same order as in English.

Table 18 lists all our judgement task sentences, and indicates whether each is homomorphic/in English-like order. Our reference value for each factor was the value that occurred most frequently in our BSL Corpus data (see Sections 3.1.3.4 and 3.1.3.5), that is:

(a) Homomorphic: Yes

(b) English-like order: Yes

Table 18 Adherence of Judgement Task NPs to Homomorphism and English-like Order

Code	NP structure	Homomorphic?	English-like order?
NP01	[Dem Noun]	n/a	Yes
NP02	[Noun Dem]	n/a	No
NP03	[Adj Noun]	n/a	Yes
NP04	[Noun Adj]	n/a	No
NP05	[Num Noun]	n/a	Yes
NP06	[Noun Num]	n/a	No
NP07	[Adj Noun Dem]	n/a	No
NP08	[Adj Dem Noun]	No	No
NP09	[Noun Adj Dem]	Yes	No
NP10	[Noun Dem Adj]	No	No
NP11	[Dem Noun Adj]	n/a	No
NP12	[Dem Adj Noun]	Yes	Yes
NP13	[Num Noun Dem]	n/a	No
NP14	[Num Dem Noun]	No	No
NP15	[Noun Num Dem]	Yes	No
NP16	[Noun Dem Num]	No	No
NP17	[Dem Noun Num]	n/a	No
NP18	[Dem Num Noun]	Yes	Yes
NP19	[Num Noun Adj]	n/a	No
NP20	[Num Adj Noun]	Yes	Yes
NP21	[Noun Num Adj]	No	No
NP22	[Noun Adj Num]	Yes	No
NP23	[Adj Noun Num]	n/a	No
NP24	[Adj Num Noun]	No	No
NP29	[Dem Num Adj Noun]	Yes	Yes
NP31	[Dem Adj Num Noun]	No	No
NP35	[Num Dem Adj Noun]	No	No
NP39	[Num Adj Dem Noun]	No	No
NP41	[Adj Dem Num Noun]	No	No
NP43	[Adj Num Dem Noun]	No	No
NP47	[Noun Dem Num Adj]	No	No
NP48	[Noun Dem Adj Num]	No	No
NP49	[Noun Num Dem Adj]	No	No
NP50	[Noun Num Adj Dem]	No	No
NP51	[Noun Adj Dem Num]	No	No
NP52	[Noun Adj Num Dem]	Yes	No

2.5.4 Linear Models to Examine Homomorphism where Modifiers are Widely-Spaced versus Narrowly-Spaced.

In their miniature artificial language experiment discussed in Section 1.3.2.3, Culbertson and Adger (2014) found that participants were more likely to innovate homomorphic NPs when the modifiers within those NPs were widely spaced in terms of semantic scope than when the modifiers were not widely spaced. In other words, the NPs that participants selected that contained demonstratives and adjectives were more often homomorphic than those that contained either demonstratives and numerals, or numerals and adjectives.

We sought to test whether this distinction between widely- and narrowly-spaced modifiers would also be reflected in our typicality judgements. To this end, we constructed two additional linear models and LMEMs, using homomorphism as a fixed effect. These models used our sentences whose NPs have two modifiers on the same side of the noun, some of which are homomorphic, others not. One pair of models–comprising a linear model and a linear mixed-effects model–included only those sentences in which the modifiers were widely spaced in terms of semantic scope (i.e., those containing just a demonstrative and an adjective), as listed in Table 19. The other pair–again consisting of a linear model and a linear mixed-effects model–included sentences whose modifiers are adjacent in terms of semantic scope (i.e., those containing just an adjective and a numeral, or just containing a demonstrative and a numeral): see Table 20.

Table 19 Judgement Task Sentences Whose Modifiers are Widely Spaced in Terms of Semantic Scope

Code	Sentence glosses	NP structure	Homomorphic?
NP08	GIRL REMEMBER OLD PT FILM	[Adj Dem Noun]	No
NP09	GIRL REMEMBER FILM OLD PT	[Noun Adj Dem]	Yes
NP10	GIRL REMEMBER FILM PT OLD	[Noun Dem Adj]	No
NP12	GIRL REMEMBER PT OLD FILM	[Dem Adj Noun]	Yes

Table 20 Judgement Task Sentences Whose Modifiers are Not Widely Spaced in Terms of Semantic Scope

Code	Sentence glosses	NP structure	Homomorphic?
NP14	BOY EAT FOUR PT SANDWICH	[Num Dem Noun]	No
NP15	BOY EAT SANDWICH FOUR PT	[Noun Num Dem]	Yes
NP16	BOY EAT SANDWICH PT FOUR	[Noun Dem Num]	No
NP18	BOY EAT PT FOUR SANDWICH	[Dem Num Noun]	Yes
NP20	GIRL BUY THREE OLD BOOK	[Num Adj Noun]	Yes
NP21	GIRL BUY BOOK THREE OLD	[Noun Num Adj]	No
NP22	GIRL BUY BOOK OLD THREE	[Noun Adj Num]	Yes
NP24	GIRL BUY OLD THREE BOOK	[Adj Num Noun]	No

We aimed to select our reference groups based on the structures that were most common in our BSL Corpus data – see Section 3.1.3.4. Six out of the seven corpus NPs with widely-spaced modifiers (86%) were homomorphic whereas only two out of the five corpus NPs with narrowly-spaced modifiers (40%) were homomorphic. We chose to use homomorphic as the reference group in both of these pairs of models, to facilitate comparison between them.

2.5.5 Distribution of Responses

We designed our judgement task with a continuous slider, under the assumption that participants would use the entire scale. However, contrary to our expectations, participants tended to favour the extreme ends of the slider, rating most sentences as either highly typical or highly atypical. For instance, Figure 55 shows the distribution of responses to NP04 [Noun Adj], which most people rated as highly typical, with a strong peak at the upper end of the scale. The distribution of responses to NP03 [Adj Noun] was different: While many participants again rated it as highly typical, a substantial proportion judged it to be very atypical, resulting in a bimodal distribution (see Figure 56). The pattern observed for NP03 was, in fact, representative of the broader dataset. Across nearly all NP sentences, we found that although there was often a dominant trend–such as a majority rating a sentence as typical–a substantial minority consistently held the opposite view. This resulted in bimodal distributions for many items. Moreover, responses clustered strongly at the extremes of the scale, with relatively few participants selecting mid-range values.

Figure 55 Histogram of Typicality Responses to NP04 [Noun Adj]

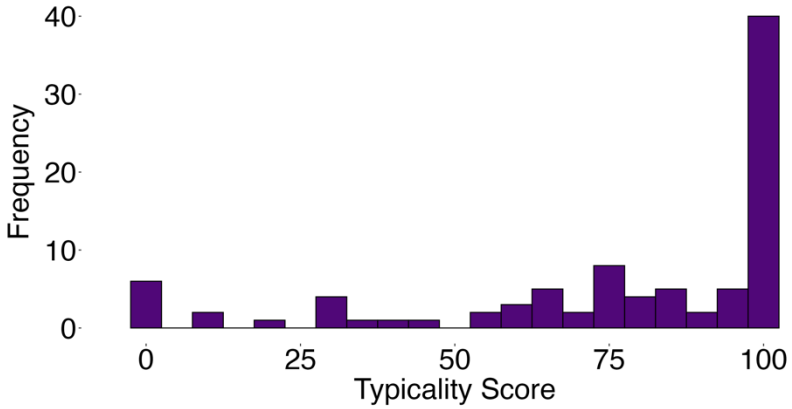
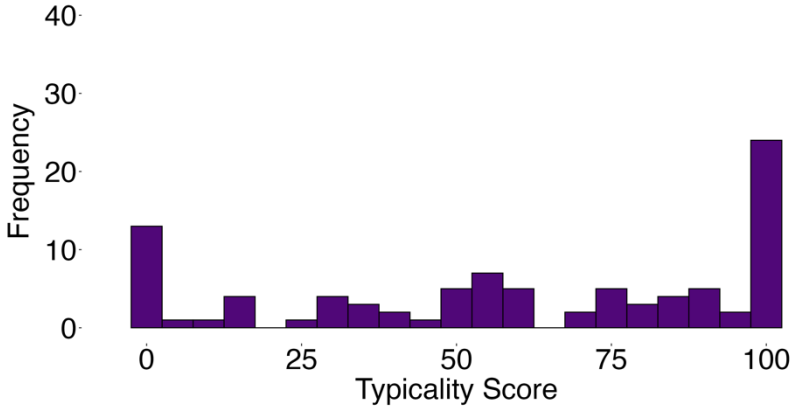


Figure 56 Histogram of Typicality Responses to NP03 [Adj Noun]



These *heavy-tailed* distributions of responses means that the assumption of normality for residuals is violated when using linear models and linear mixed-effects models. We illustrate this in Figure 57 which shows that the distribution of residuals for our linear model comparing sentence NP03 with NP04 clearly deviates from normality. Additionally, the fact that the data points in the Q-Q plot in Figure 58 do not lie on the red line of best fit also indicates that the residuals are not normally distributed. Although normality of residuals is typically required for linear models, Knief & Forstmeier (2021) provide evidence that such models without a normal distribution of residuals are still appropriate for hypothesis testing for data with this distribution, as long as there are no influential outliers. In our data, there are no such outliers, as our z-scores ranged from -2.26 to +2.12, with none below -3 or above +3.

Figure 57 Histogram of Residuals for the Linear Model Comparing NP03 and NP04

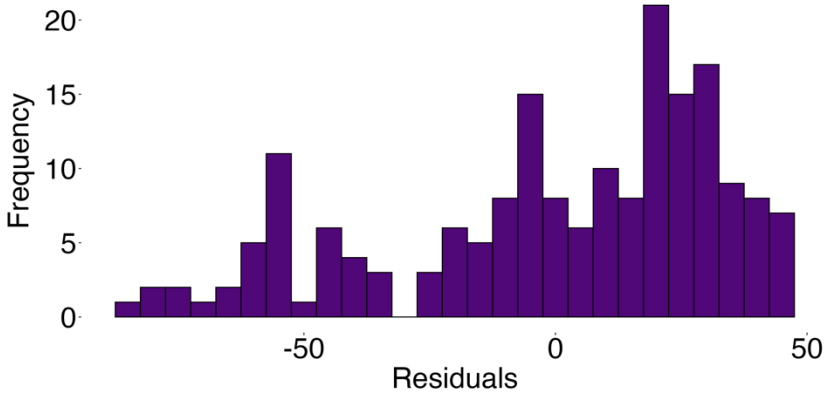
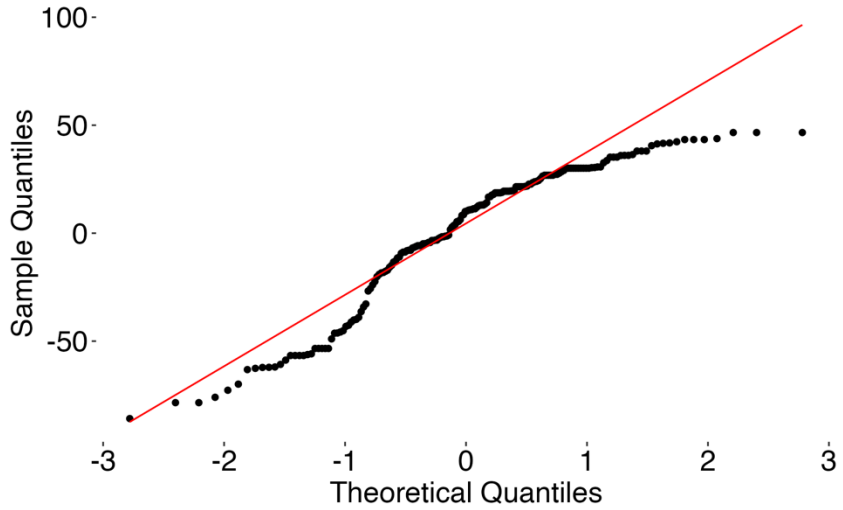


Figure 58 Q-Q Plot of Residuals for the Linear Model Comparing NP03 and NP04



2.5.6 Significance Level for Judgement Task Models

We have created many models to examine the effect of different parameters on our participants' typicality judgements of NPs in different orders. Since each of our sentences is included in more than one model, the risk of making a Type I error is increased. The Bonferroni correction is commonly used in order to mitigate this risk: however, Cabin and Mitchell (2000) note that it is not straightforward to determine exactly what degree of correction should be used for a given experimental design. And our design is particularly complex in this regard since there is partial overlap in the sentences that are included in each model–altogether, we created seven models for groups of sentences (detailed in Section 2.5.1) plus three models to investigate modifier position (Section 2.5.2), two models to examine adherence to homomorphism and English-like order respectively (Section 2.5.3), and two models to examine homomorphism in NPs which are widely- or narrowly-spaced in terms of semantic scope (Section 2.5.4), for a total of 14 models.

Each of our stimulus sentences is included in between three and six of these models, depending on the modifiers that it includes. For example, the target NP in NP03 (GIRL BUY BLUE VEHICLE) has only an adjective modifier (BLUE), so it appears in three models: the model for its sentence group (NP03 and NP04), the model investigating adjective position and the model investigating English-like order. In contrast, the target NP in NP31 (BOY BRING PT RED FOUR MUG) has a demonstrative (PT), an adjective (OLD) and a numeral (FOUR) as modifiers, and homomorphism is a meaningful concept for this sentence. It therefore appears in six models, for: (1) its sentence group (NP29 – NP52); (2) adjective position; (3) demonstrative position; (4) numeral position; (5) homomorphism; (6) English-like order.

Since each model contains a different combination of sentences, we adopted a conservative approach in terms of an adjustment for multiple models, applying the Bonferroni correction to adjust our significance cut-off (alpha) by dividing the standard $p = .05$ by 14 (for 14 models), giving an alpha of $p = .004$ for all of our models.

2.6 Combined Analysis of Corpus and Judgement Task Data

When establishing how to compare pointing signs between our corpus and judgement task data, we decided to consider together all pointing signs within NPs in the corpus, and compare these to the demonstrative pointing signs in the judgement task. This is because, as discussed in Section 2.4, we observed a great deal of ambiguity about the grammatical category of pointing signs within NPs, plus we do not consider that there is an identifiable grammatical category of demonstrative in BSL.

Together, the corpus and judgement task analyses provide complementary perspectives on NP structure in BSL. Having outlined the procedures used to collect and analyse this data, we now turn to the results. Chapter 3 presents our findings, beginning with the corpus data and followed by the judgement task outcomes.

Chapter 3 Results

3.1 Corpus Data Results

We analysed 34 files, considering the first 100 sign tokens on the dominant hand of each signer, for a total of 3400 sign tokens. This lasted between 33 seconds (participant BF01) and five minutes 23 seconds (participant BM06) for each signer, depending on how many turns they had and how long they held the floor in each turn for that part of the dialogue.

We have identified 1037 noun phrases (NPs), containing 1319 sign tokens: 39% of all 3400 sign tokens examined were in an NP. As discussed in Section 2.2.2.2.1, where there was uncertainty about the status of elements within an NP, we annotated the NP as *unsure*. There were 41 such NPs in our data, constituting 4% of the total. They are included in the NPs discussed in this chapter.

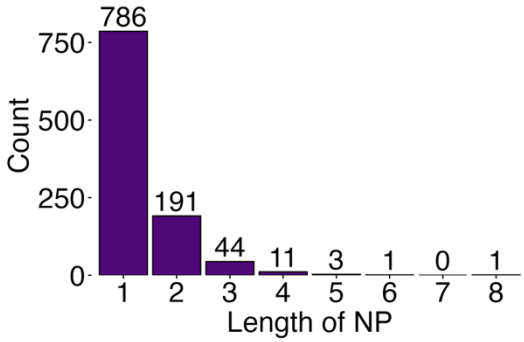
We observe a small but notable proportion of NPs in our corpus data that include simultaneous elements: Overall, 114 (11%) of our NPs include some simultaneity: much of this results from co-occurring manual signs and mouthings. As explained in the Conventions section, angle brackets <> are used to indicate simultaneity within an NP. For example, Noun<>Point indicates that the signer articulated a noun with one hand and a pointing sign with the other. Similarly, Numeral<>Numeral<>Numeral indicates that the signer produced a separate numeral with each hand, plus a third numeral as a mouthing.

In the remainder of this chapter, we will present our findings. We will cover the lengths of NPs in the British Sign Language (BSL) Corpus, the grammatical categories of elements of these NPs and then the order of these elements.

3.1.1 Lengths of NPs

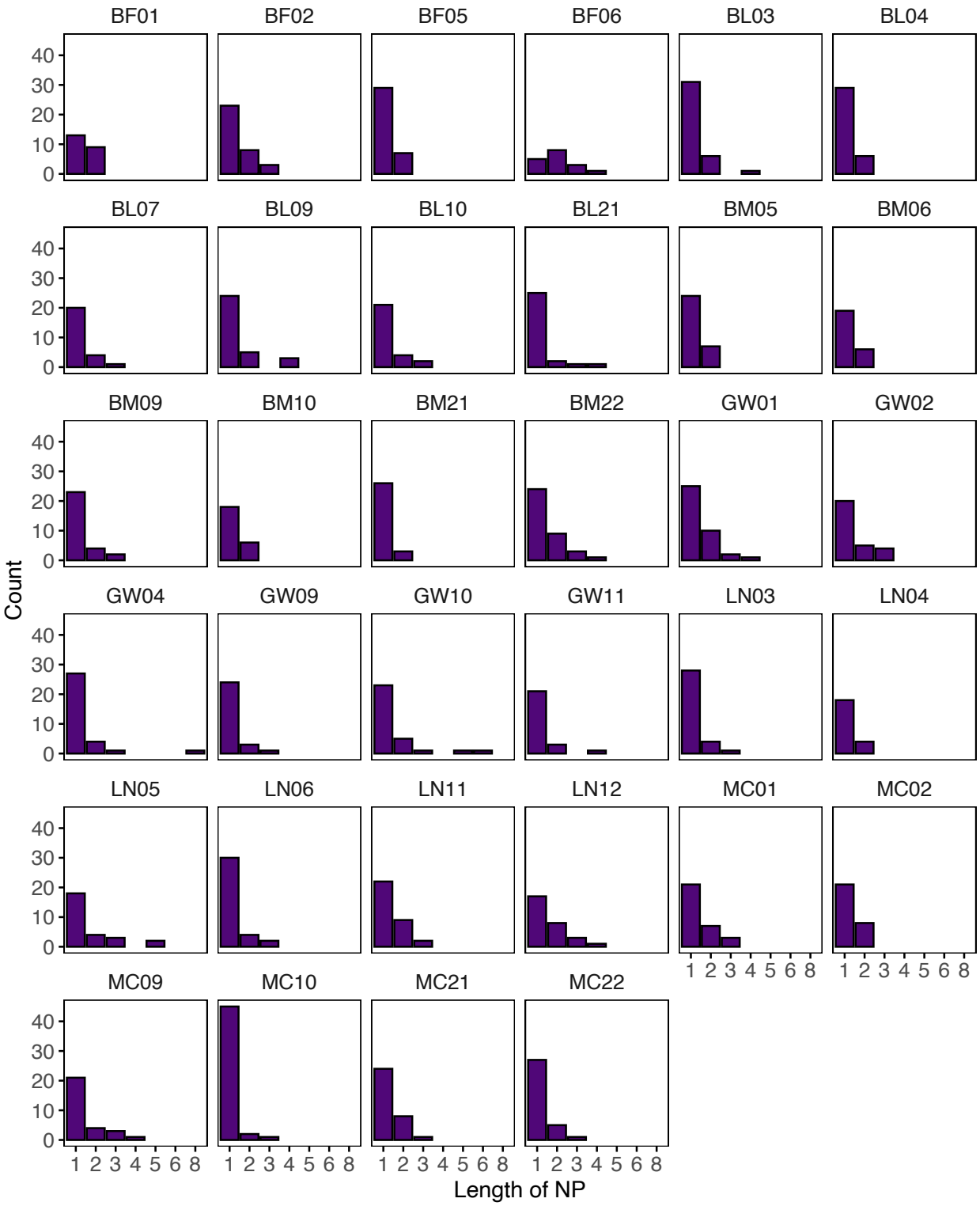
Figure 59 shows the distribution of lengths of NPs identified.

Figure 59 Frequency of NPs of Different Lengths



It is clear that in our sample from the BSL Corpus, the vast majority of NPs are very short, with 76% consisting of a single element. There is a great deal of variability between participants in the number of NPs that each produced within the 100 sign tokens that we examined, ranging from 17 to 48 NPs. Figure 60 shows the number of NPs of different lengths produced by each participant. This reveals wide variation in the distribution of lengths of NP; for example, MC10 produced 45 single-element NPs and only three longer examples. In contrast, LN05 produced 18 single-element NPs but nine longer constructions, including two containing five elements.

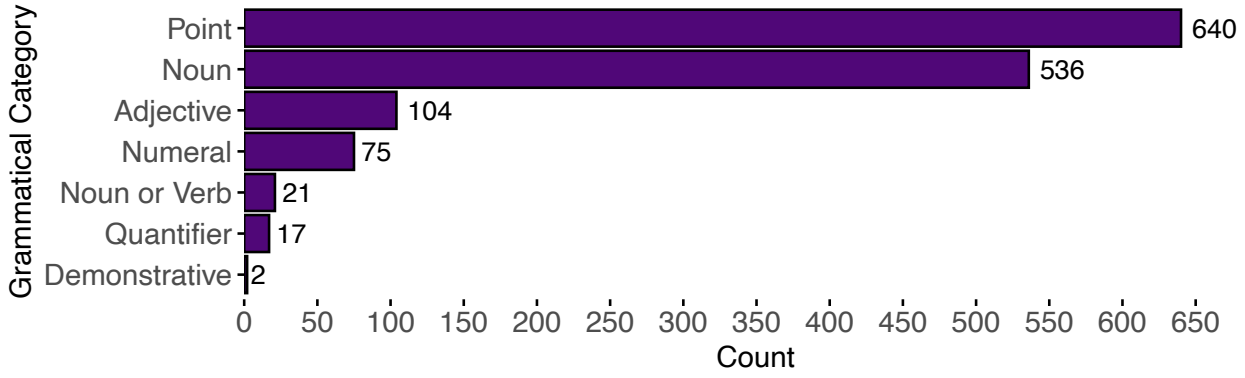
Figure 60 Frequency of NPs of Different Lengths, Split by Participant



3.1.2 Grammatical Categories of Elements of NPs

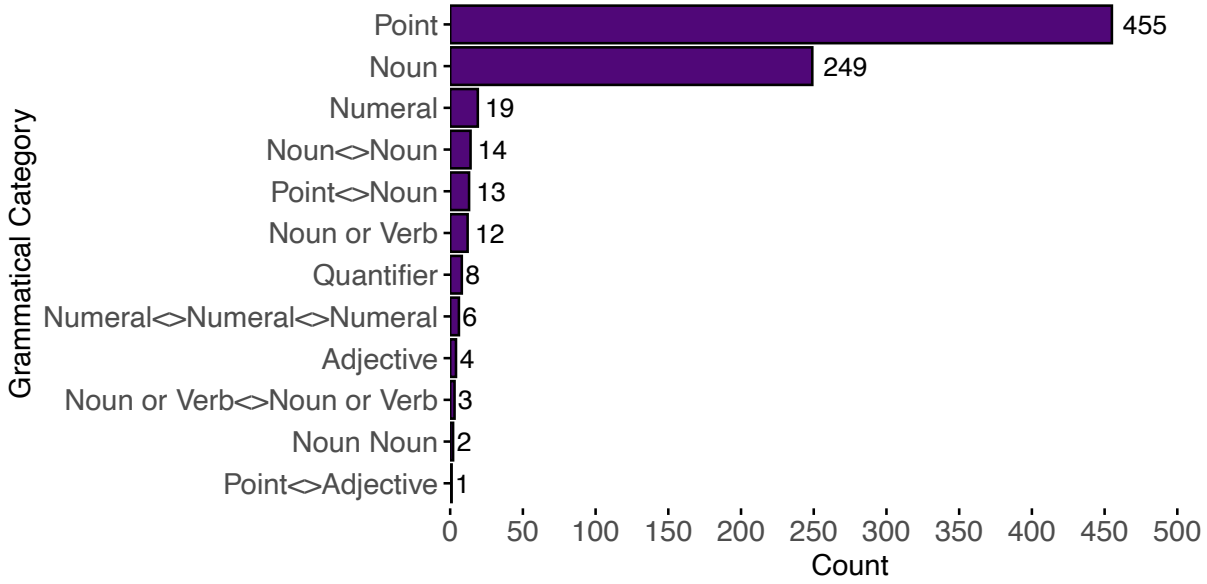
Figure 61 shows the overall distribution of grammatical categories of elements of our 1037 NPs. Around half (50.3%, *N* = 522) of our corpus NPs do not include a noun. However, many of the pointing signs identified in our data may be pronominal, and the vast majority (95.7%, *N* = 992 NPs) of our NPs do include either a noun or a pointing sign.

Figure 61 Frequency of each Grammatical Category in NPs of all lengths



As noted in Section 3.1.1, most of the NPs identified consist only of a single element. Figure 62 shows the distribution of grammatical categories of these single-element NPs.

Figure 62 Frequency of Grammatical Categories of Single-Element NPs



In the next section, we discuss longer NPs, and the order of constituents within them.

3.1.3 Order of Grammatical Categories Within NPs

This section examines the order of grammatical categories within NPs in our corpus, excluding cases with simultaneity that prevents linear analysis. We first compare prenominal and postnominal modifiers, then analyse NPs with two elements and those with three or more. Finally, we assess homomorphism in a subset of NPs that meet specific structural criteria.

In Section 1.3.4.4, we discussed simultaneous production in sign languages. When considering the order of elements within NPs, we do not include NPs that have any simultaneity within them in terms of either (1) a different sign being produced with each hand or (2) a mouthing being produced that is not congruent (see Section 2.2.2.4.4) to a co-occurring manual sign. Of our 1037 NPs, 77 had one of these types of simultaneity.

3.1.3.1 Prenominal or Postnominal Modifiers?

When calculating the proportion of prenominal or postnominal modifiers in our NPs, we excluded 68 NPs that had more than one noun, and 12 NPs that had modifiers both before and after the noun. If we consider NPs of all lengths that include one noun and at least one modifier, where the modifier(s) are either all prenominal or all postnominal, we find 108 NPs with prenominal modifiers and 15 NPs with postnominal modifiers. Of these, 44 included an adjective, with 40 (91%) appearing in prenominal position and only 4 (9%) in postnominal position.

We next discuss NPs with two elements, then those with three or more elements.

3.1.3.2 Order of Grammatical Categories in NPs with Two Elements

Figure 63 shows the distribution of grammatical categories within NPs containing two elements. Where multiple words have been mouthed at the same time as a manual sign, the mouthings are grouped together using quotation marks. For example, Noun<>'Noun Noun' indicates a manual sign produced simultaneously with two mouthings, with all three elements functioning as nouns. See Figure 64, where the manual sign PARENTS is produced at the same time as the mouthing "mum dad". Similarly, Figure 65 shows a compound manual sign MUM^FATHER produced at the same time as the mouthing "moth(er) fath(er)", giving an NP with grammatical category 'Noun Noun'<>'Noun Noun'. In both these cases, the mouthings are congruent to the manual signs, and, regardless of the fact that MUM^FATHERis a compound sign, the resulting NP has two elements. (The use of angle brackets and quotation marks to indicate simultaneity is explained in the Conventions section.)

Figure 63 Frequency of Grammatical Category Combinations for NPs with Two Elements

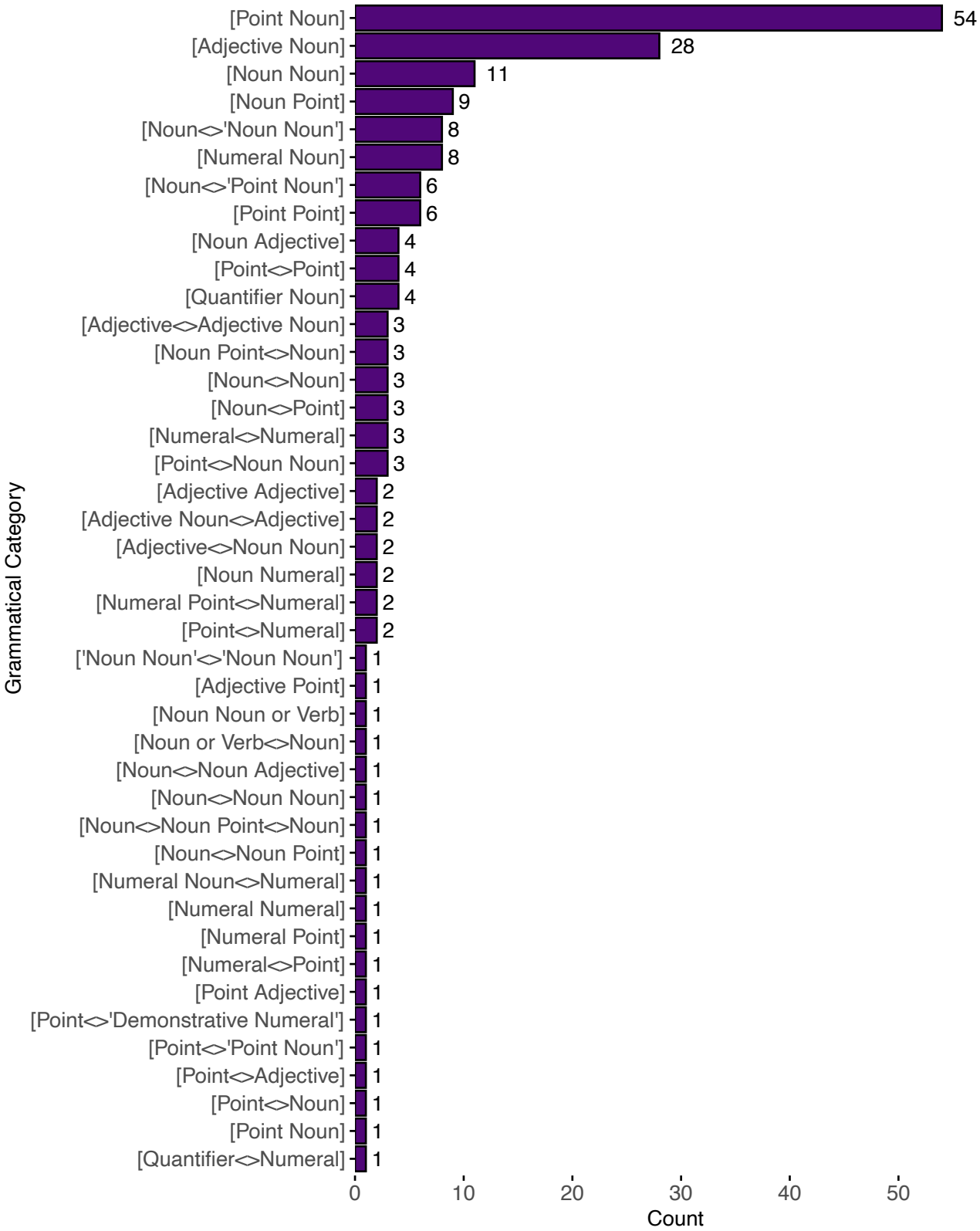


Figure 64 BF02n Example of Two Words Mouthed With One Manual Sign, Represented as [Noun<>'Noun Noun']; "So, one day when my mum and dad came over..."

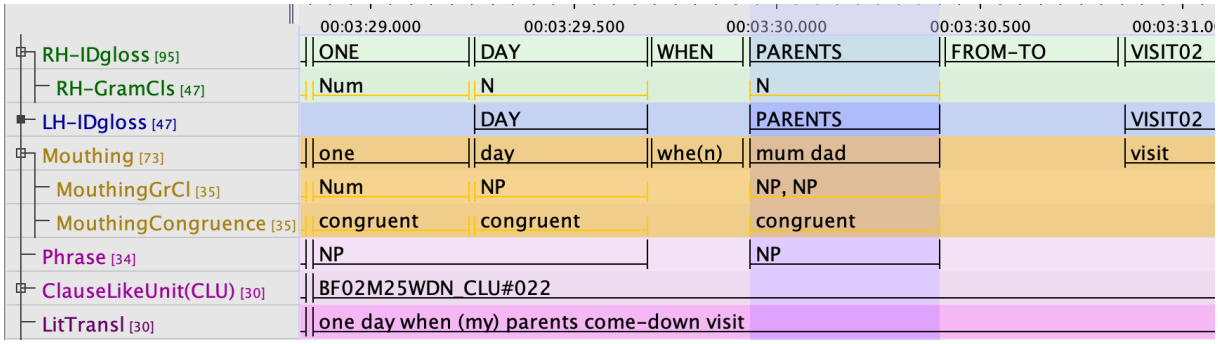
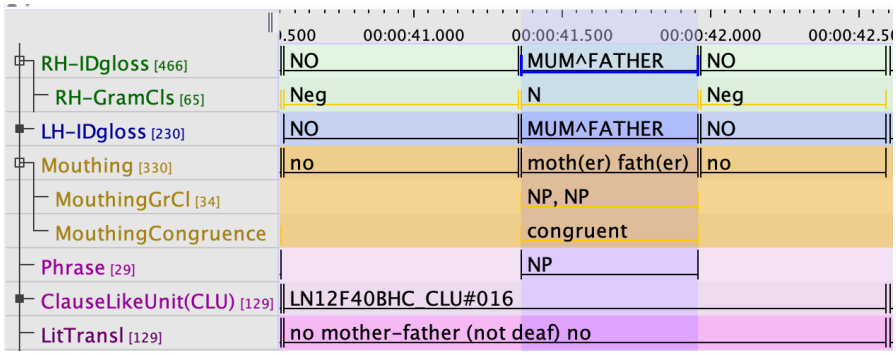


Figure 65 LN12c Example of Two Words Mouthed with a Compound Manual Sign, Represented as ['Noun Noun'<>'Noun Noun']; "No, my parents are both hearing but my sister is deaf and we have three hearing brothers."



The two most frequent sequences of two grammatical categories are [Point Noun] (as illustrated in Figure 66) and [Adj Noun] (as shown in Figure 67).

Figure 66 BL03c Example of [Point Noun] Order; "My mother was waving"

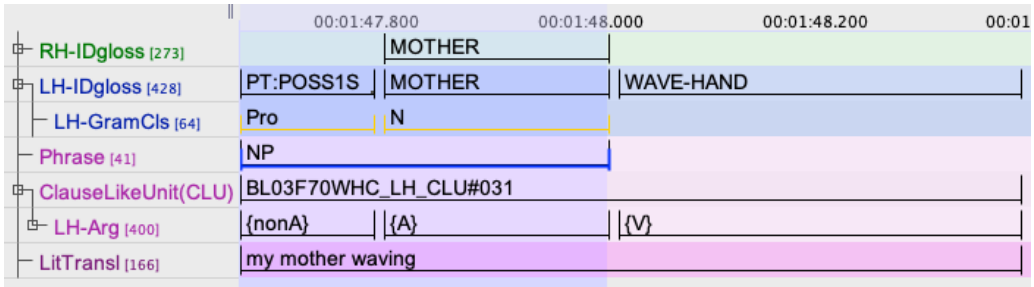
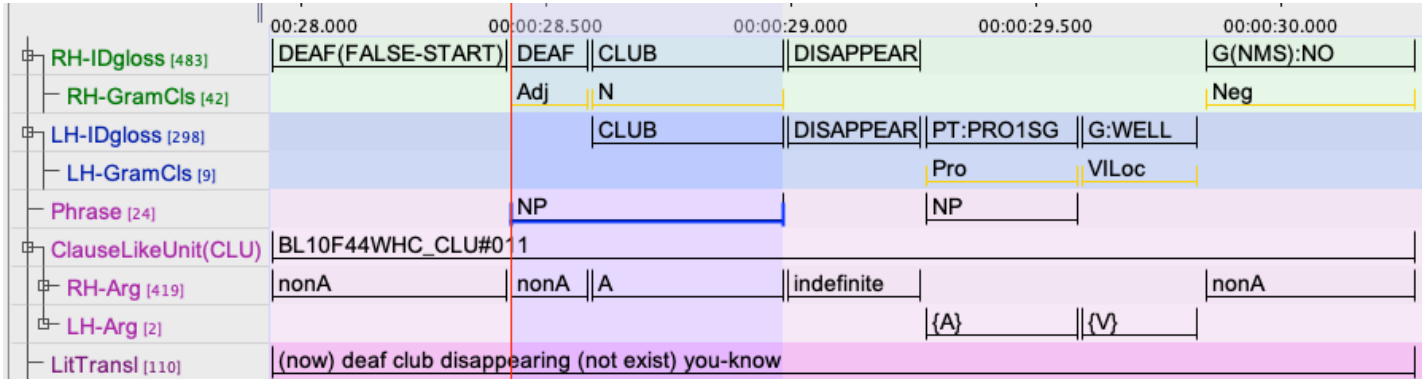


Figure 67 BL10c Example of [Adj Noun] Order; “Deaf clubs are disappearing you know”



The next two most frequent pairs of grammatical categories are [Noun Noun], illustrated in Figure 68 and [Noun Point], shown in Figure 69.

Figure 68 GW01n Example of [Noun Noun] order “Slovakia Bratislava” in “We went to Bratislava.”

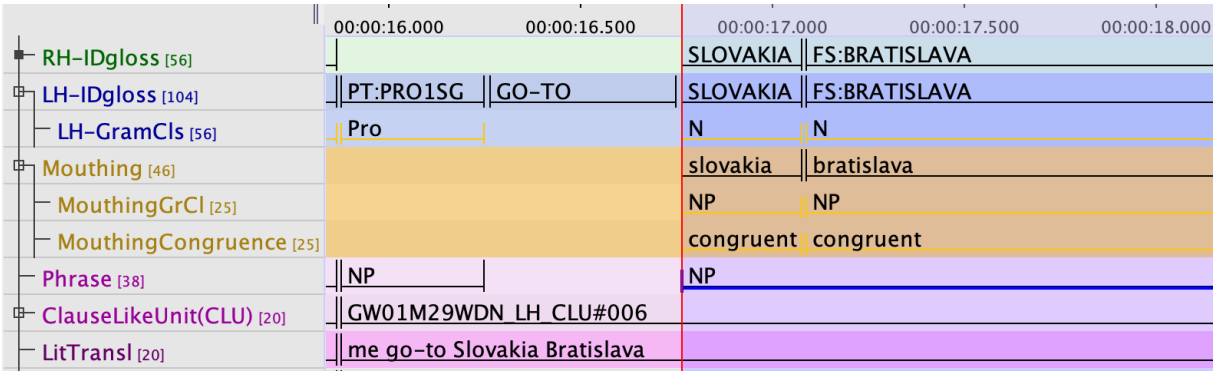
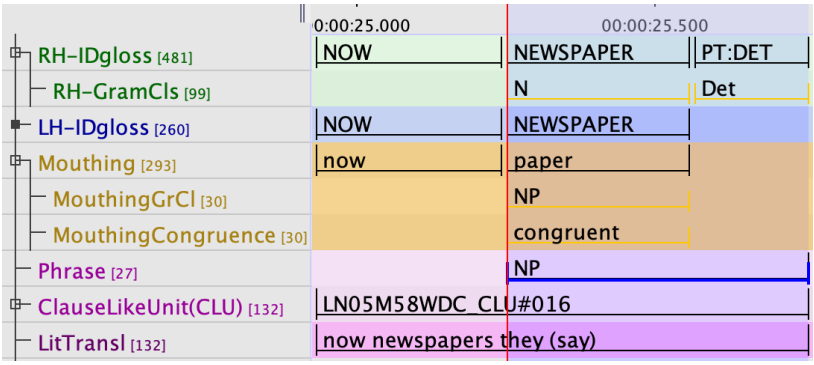


Figure 69 LN05c Example of [Noun Point] order; “Now the newspapers say…”



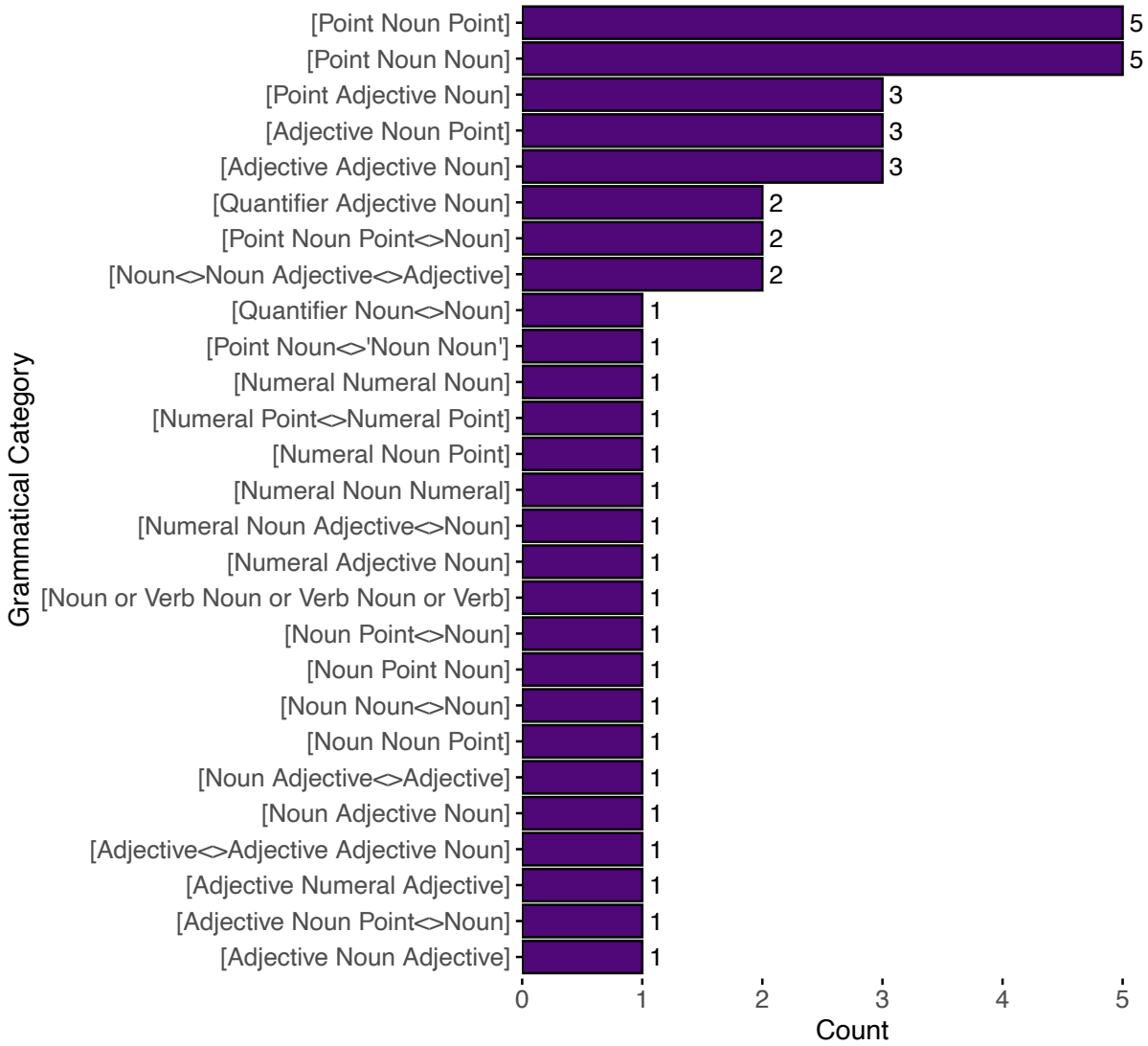
Around half (49.7%, *N* = 95) of all NPs with two elements have a prenominal modifier. A further 8% (*N* = 15) have a postnominal modifier and the remainder (42%, *N* = 81) either have no noun, or else have simultaneous elements, meaning that an order cannot be determined. So of the NPs with two elements where an order of modifier and noun can be identified, 86% have a prenominal modifier. A binomial test indicated that there are significantly more NPs with two elements that have a prenominal modifier than that have a postnominal modifier ($p < .001$).

Looking specifically at NPs that consist of a noun and an adjective, we find 28 examples of prenominal adjectives and 4 examples with a postnominal adjective: This equates to 88% prenominal. A binomial test indicated that this difference is significant ($p < .001$). So, whether we consider just adjectives or all modifiers of a noun in NPs with 2 elements, we find that the modifiers are significantly more often prenominal.

3.1.3.3 Order of Grammatical Categories in NPs with Three or More Elements

Figure 70 shows NPs that have three elements (*N* = 44), illustrating the frequency of combinations of grammatical categories in these constructions. The number of examples identified of NPs longer than 3 elements (16 in total) is too small to be able to identify patterns within the grammatical category combinations.

Figure 70 Frequency of Grammatical Category Combinations for NPs with Three Elements



We provide an example of the two most frequent orders, [Point Noun Point] in Figure 71 and [Point Noun Noun] in Figure 72.

Figure 71 BF06n Example of an NP Consisting of [Point Noun Point]: “He had camouflage cream all over his face.”

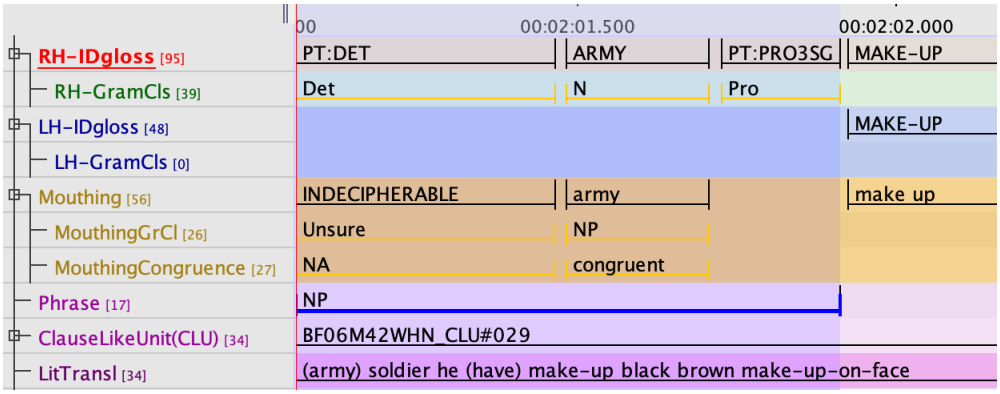
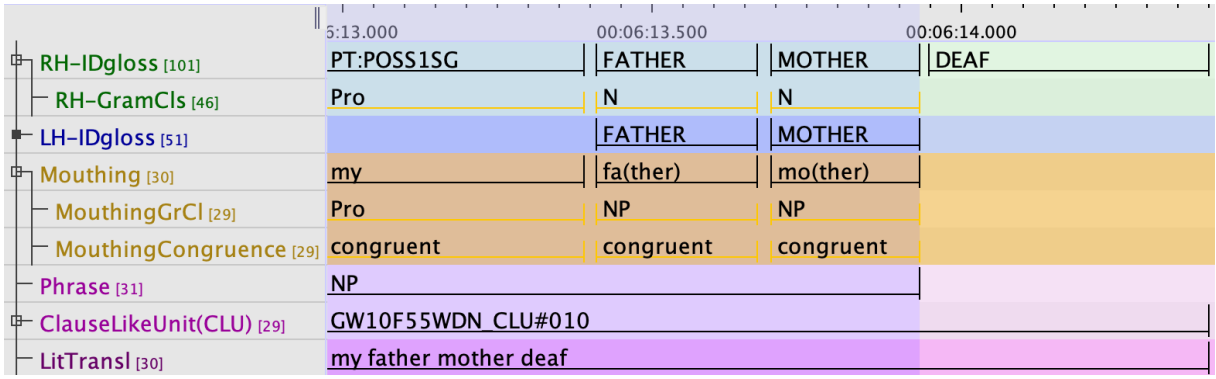


Figure 72 GW10n Example of an NP Consisting of [Point Noun Noun]: “My parents are deaf.”



In longer NPs where modifier position is identifiable, both modifiers and adjectives consistently appear in prenominal position. Of the 60 NPs in our dataset that contain three or more elements, only 13 (22%) allow us to determine whether all modifiers are prenominal or postnominal. The remaining NPs are excluded from this analysis because they either lack a noun, contain multiple nouns, have a noun that is not positioned at the start or end of the NP, or include simultaneous elements. In all 13 analysable cases, the modifiers are prenominal; none exhibit postnominal modification. Notably, 12 of these 13 NPs include adjectives among the modifiers, further supporting the observation that adjectives also favour prenominal position in BSL.

3.1.3.4 Homomorphism in Corpus NPs

We next evaluate the extent to which our corpus NPs are homomorphic. This analysis is only possible for NPs that meet certain conditions: we discuss these conditions then we provide a summary table of all the NPs where these adherences can be determined.

The conditions that must be met to determine whether an NP is homomorphic are as follows: Firstly, the NP needs to include exactly one noun. Therefore, homomorphism is not defined for the NP in Figure 73 (which has two nouns) or the NP in Figure 74 (which has no nouns).

Figure 73 BF06n Example of an NP with Two Nouns: [Point Noun Noun]: “My house windows at the back were completely destroyed.”

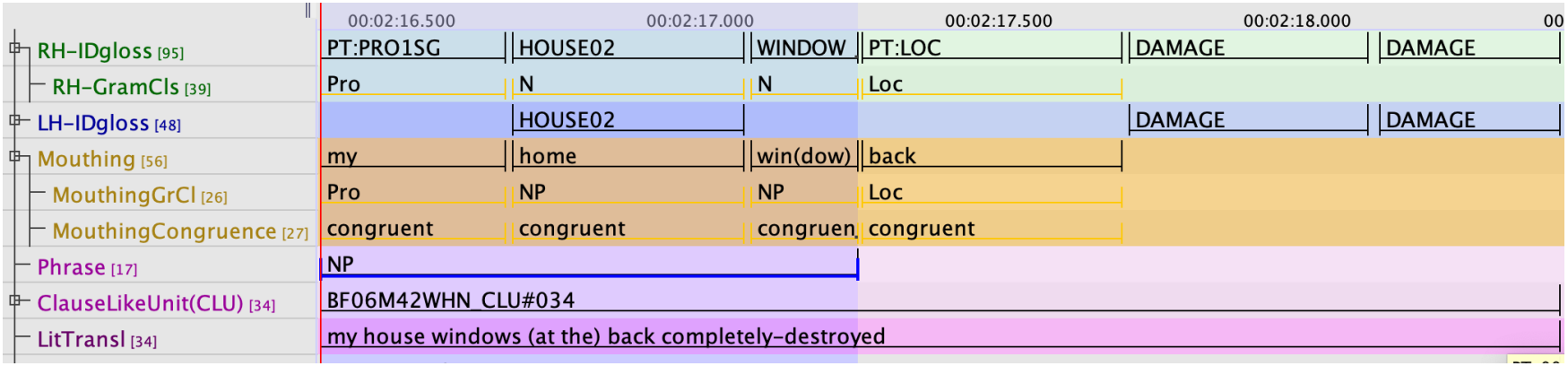
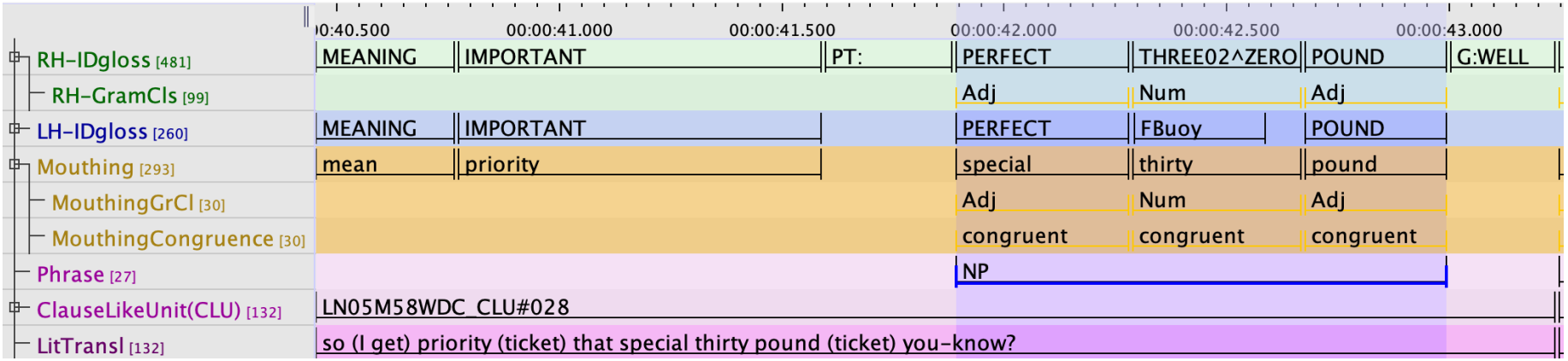


Figure 74 LN05c Example of an NP with No Nouns: [Adj Num Adj]: “So I get a priority ticket: that special thirty pound (ticket) you-know?”



Furthermore, the NP must have two or more different types of modifier (adjective, numeral or point), and these must on the same side of the noun (i.e., both types of modifier must appear prenominally or both types must appear postnominally). Homomorphism is therefore not defined for the NPs in Figure 75 and Figure 76.

Figure 75 MC22c Example of an NP Lacking Two Different Types of Modifier: [Adj Adj Noun]: "(Charlie) he wants black-brown polish."

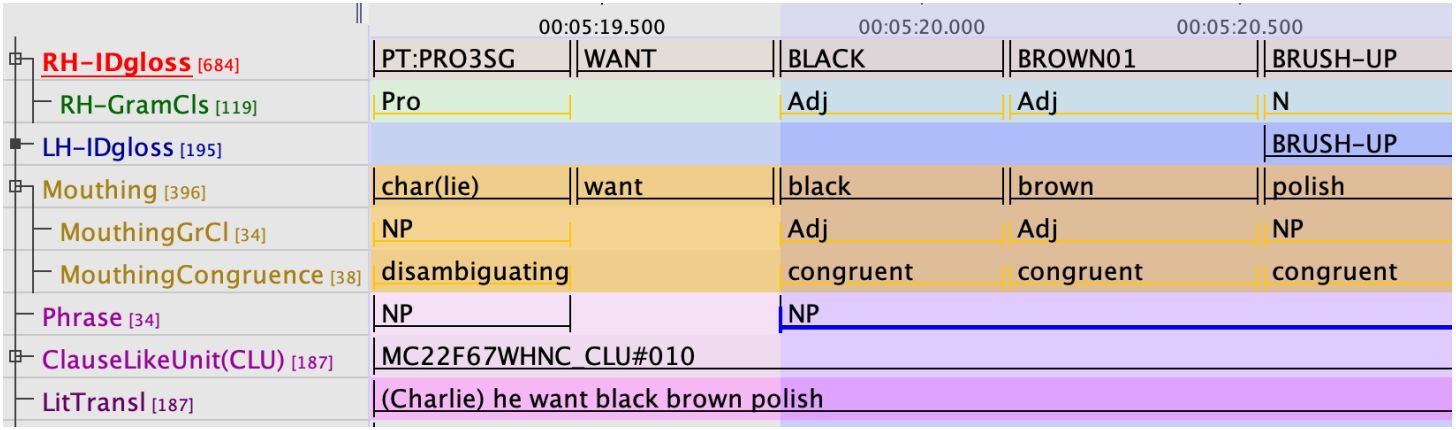
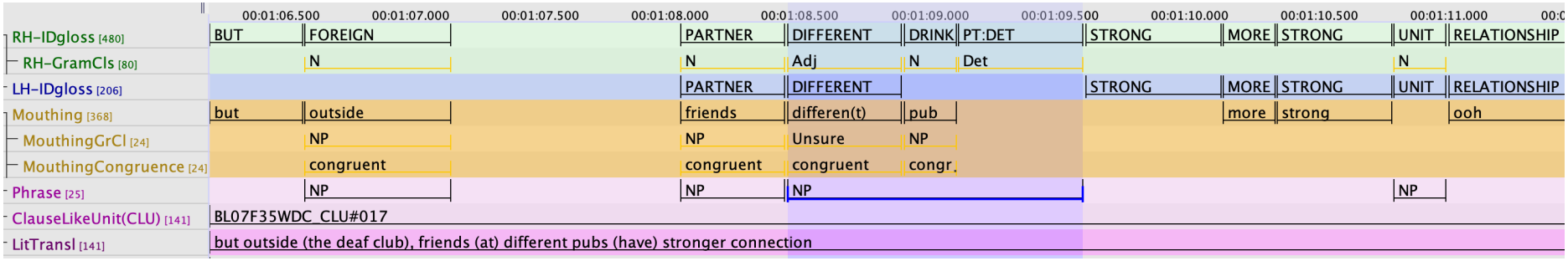
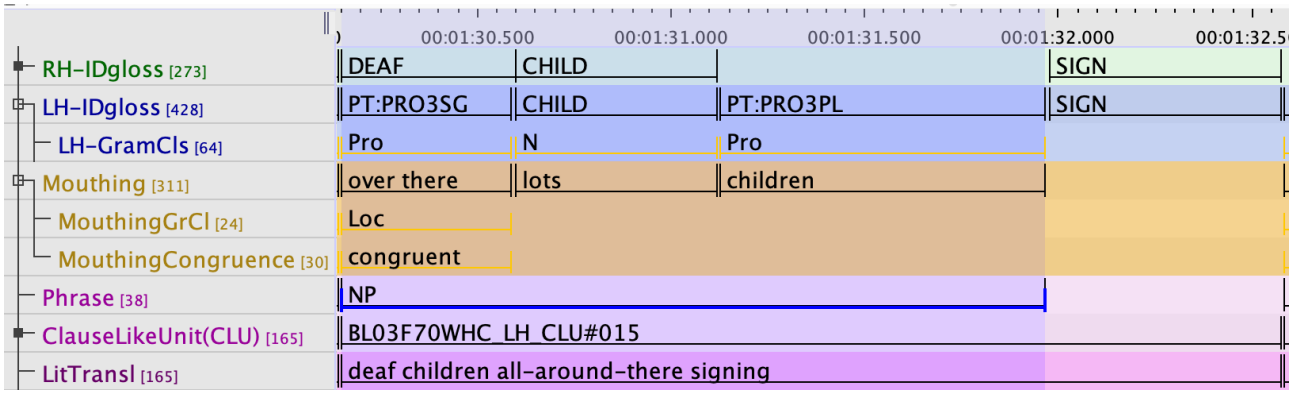


Figure 76 BL07c Example of an NP Lacking Two Modifiers on The Same Side of the Noun: [Adj Noun Point]: "but outside (the deaf club), friends at different pubs have a stronger connection."



Additionally, the NP must not include any simultaneous elements that are non-congruent, which would mean that we cannot determine a linear order of elements. So, for example, homomorphism is not defined for the NP in Figure 77 because the adjective DEAF is produced at the same time as the pronominal pointing sign PT:PRO3SG.

Figure 77 BL03c Example of an NP with Simultaneous Adjective and Point: [Adj<>Point Noun Point]: "deaf children were all around there signing."



Overall, this means that homomorphism cannot be determined for most of our corpus NPs. As a minimum, the NP must have at least 3 elements and as can be seen from Figure 59 in Section 3.1.1 above, of our 1037 NPs, 977 (94%) contain only one or two elements. Of the remaining 60 NPs, only nine (0.9% of all 1037 NPs) meet the above criteria. Glosses for these nine NPs are listed in Table 21: Seven of the nine (78%) are homomorphic: a binomial test indicated that we cannot conclude that there are more homomorphic than non-homomorphic NPs in the corpus based on this sample (p = .09). This table also indicates whether each NP where homomorphism is defined has widely-spaced modifiers in terms of semantic scope (i.e., the modifiers are {Adj, Point}) or narrowly-spaced modifiers in terms of semantic scope (i.e., the modifiers are {Adj, Num} or {Num, Point}): Five out of the six corpus NPs with widely-spaced modifiers (83%) are homomorphic whereas only two out of the three corpus NPs with narrowly-spaced modifiers (67%) are homomorphic. Fisher's exact test revealed that there is no significant association between whether an NP is homomorphic or not and whether its modifiers are widely- or narrowly-spaced (two-tailed p = 1).

Table 21 Corpus NPs and Adherence to Homomorphism

Noun phrase	Grammatical categories	Participant code	Literal translation of CLU	Homo-morphic	Modifier separation
GOOD<>good PLAY<>play(ers) PT:DETPL FIVE<>five SIX03<>six	[Adj Noun Point Num Num]	LN05	this-year West Ham **good players, them-there five (or) six** (players have been) left-the-team	No	Narrow
FOOTBALL02<>football PT:DET STADIUM<> ground FOOTBALL02	[Adj Point Noun Adj]	MC09	**(the) football grounds, football (grounds)** along-there?	No	Wide
ONE<>one HEARING<>hearing GIRL<>woman	[Num Adj Noun]	GW01	me one we bring maybe (me) think eight deaf-people altogether with **one hearing woman** too	Yes	Narrow
ONE<>one RANKING<>first OLD02<>'old old' BOY02<>boy PT:PRO3SG	[Num Adj Adj<>'Adj Adj' Noun Point]	GW10	later **one (boy) first (one) older boy**, his partner born their daughter!	Yes	Narrow
PT:PRO3SG NATIONAL<>nat(ional) BRITAIN<>british THEATRE<>theatre	[Point Adj Adj Noun]	BL21	well me know -- me (only) know **that-there Brit-- National British Theatre** there	Yes	Wide
PT:PRO1SG SCOTLAND<>scottish DEAF<>deaf UNIT	[Point Adj Adj Noun]	GW01	**we Scottish deaf group** all oral	Yes	Wide
PT:POSS1SG HOUSE<>home IMPROVE02<>(im)prove	[Point Adj Noun]	BM09	well (money) for my holiday or **my home improvements** repairs sort-out-things stuff-like-that	Yes	Wide
PT:POSS1SG<>my BEST<>best NIGHT<>night	[Point Adj Noun]	MC01	me remember **my best night**, my best night	Yes	Wide
PT:POSS1SG<>my BEST<>best NIGHT<>night	[Point Adj Noun]	MC01	me remember my best night, **my best night**	Yes	Wide

Note. Includes whether NP modifiers are narrowly or widely separated in terms of semantic scope. The parts of Literal Translation of CLU that relate to the NP are shown in boldface. Sorted by Grammatical Category.

3.1.3.5 Effect of Language Background on Order of Grammatical Categories in NPs

As per Table 2 above, 15 of our corpus participants were native signers and the remaining 19 were early learners of BSL (as defined in Section 1.3.4.6). In order to address Hypothesis 3a (restated below for convenience), we investigated whether this difference in language background affects whether they produce BSL NPs where the order of modifiers is the same as in English.

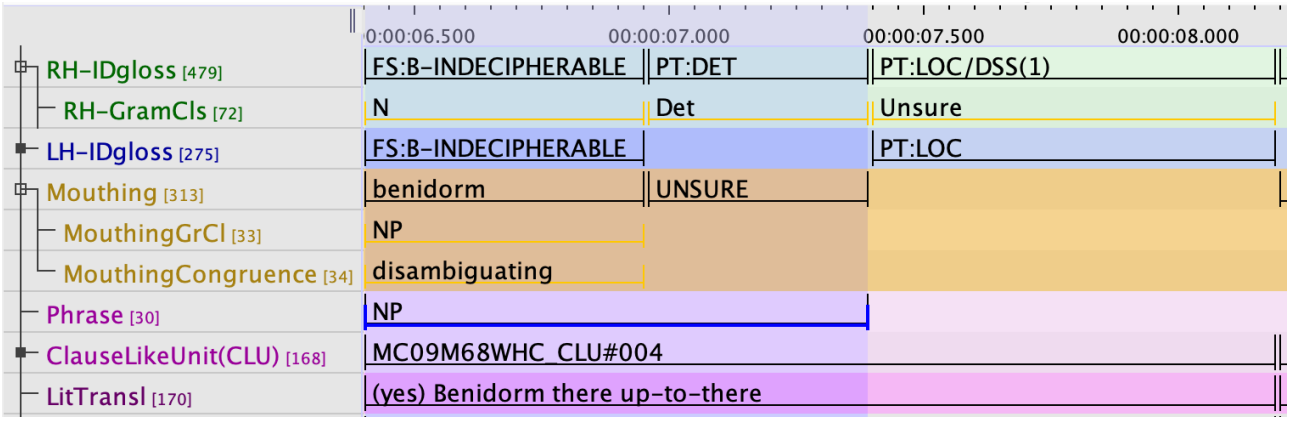
> **Hypothesis 3a:** Signers' reported language background will affect the degree of influence of English on their BSL. People who acquired BSL later in life will experience more influence of English than those who acquired it earlier. Therefore, they will tend to prefer NPs that are in English-like order.

In general, an NP is in the same order as English if it has the following structure:

- [(Dem) (Num) (Adj) Noun]

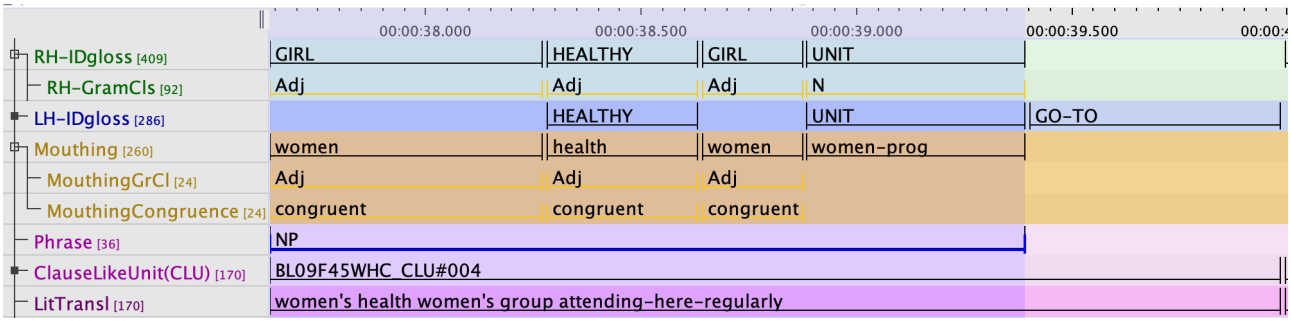
It is possible to determine whether an NP is in the same order as English when that NP contains more than one element and exactly one noun. Consequently, we cannot determine this for the NPs above in Figure 73 (containing two nouns) and Figure 74 (which has no nouns). Where an NP has simultaneous elements, sometimes it is not possible to establish a linear order of elements, meaning that we can't determine whether it is in the same order as English. This is the case for Figure 77 above (featuring a simultaneous adjective and pointing sign). However, for some NPs with simultaneous elements, we can determine an NP order: for example, when the simultaneity is because mouthing disambiguates fingerspelling, we can treat the fingerspelling and its mouthing as a single unit and then assess whether it is in the same order as English. This is the case for GW02 FS:B-BRONZE<>bronze MEDALLION (see Figure 20), where we can consider the grammatical categories to be [Adj Noun], meaning that this NP is in the same order as English. In contrast, Figure 78 provides an example of where we can determine a linear order. However, here, the order does not match English: The mouthing "benidorm" disambiguates the fingerspelled sequence, meaning that the NP has grammatical categories [Noun PT].

Figure 78 MC09c Example of NP That is Not in English-like Order: “Yes, Benidorm.”



A few NPs contained more than one adjective: for these, we confirmed on a case-by-case basis whether the order of these adjectives matched the order of adjectives in English. For example, in Figure 79, we analyse the grammatical structure as [Adj Adj Adj Noun]. This is based on the fact that the mouthing “women” is congruent with the manual sign GIRL, so the combination GIRL<>women is treated as a single adjective. Similarly, the mouthing “health” is congruent with the manual sign HEALTHY so HEALTHY<>health is likewise analysed as a single adjective. As a result, the phrase is interpreted as containing three adjectives followed by a noun. However, because GIRL<>women is repeated, the overall order does not match English adjective order.

Figure 79 BL09c Example of NP with More Than One Adjective: “It's the same for the Women's Health Group which has declined in attendance.”



In contrast, MC22 BLACK<>black BROWN01<>brown BRUSH-UP<>polish [Adj Adj Noun] (see Figure 75) contains multiple adjectives and is in English-like order.

184 out of our 1037 NPs (18%) did meet the criteria for determining English-like order. Of these, 137 NPs (74%) were the same order as English and 47 (26%) were not the same order as English. A binomial test indicated that there are significantly more NPs in English-like order in the corpus than

NPs that are not in English-like order ($p < .001$). Table 22 shows the breakdown of these NPs split by the language background of the participant that produced them.

Table 22 Number of NPs Whose Modifiers Are/Are Not in English-Like Order, Split by Language Background (Native Signer or Early Learner of BSL).

NP order	Native	Early learner
Is English-like Order	66	71
Not English-like Order	24	23

A chi-square test confirmed that there was no significant difference between the proportion of NPs produced in English order versus not in English order depending on the signer's language background: χ^2 (1, N = 184) = .12, p = .73.

In summary, our overall corpus analysis revealed that most BSL NPs are short, with a strong preference for prenominal modifiers and for English-like orders. We also found that while simultaneity occurs in a minority of cases, most NPs can be analysed in terms of linear order. In the next section, we present the results of the judgement task, which offers a complementary perspective by examining how deaf signers evaluate the typicality of different NP orders.

3.2 Judgement Task Results

We now present the findings from our judgement task. Each of the 92 participants provided typicality ratings for all the sentences listed in Section 2.3.1.5. It is important to note that our methodology does not uncover the reasons behind participants' ratings; further qualitative research, such as interviews, would be required in order to explore these motivations.

We begin by reporting results for sentences containing each type of modifier in turn (demonstrative, adjective then numeral), focussing on whether participants preferred the modifier in a prenominal or postnominal position. When only one modifier was present, preferences varied by modifier type: postnominal position was strongly typical for adjectives, whereas numerals and demonstratives showed no significant difference between pre- and postnominal positions. To examine whether these judgements changed when multiple modifiers were present, we included sentences with more than one modifier and observed a different pattern–a general dispreference for postnominal modifiers. Most multiple-modifier NPs were non-homomorphic and therefore judged less typical

overall (as discussed in Section 3.2.6), but these cases were evenly distributed across prenominal and postnominal orders, so they do not account for the dispreference for postnominal modifiers.~~,~~

We then provide results for sentences containing pairs of modifiers: for example, NP07 to NP12, which comprise the six sentences including a noun plus two modifiers (adjective and demonstrative) as listed in Section 2.3.1.5.2. This is followed by findings for the sentences that contain a noun plus all three modifiers.

Next, we provide the results of models examining the typicality of NPs based on whether or not they are homomorphic. We further analyse the homomorphism data by considering whether or not the modifiers in each NP are widely separated in terms of semantic scope. Finally, we examine typicality ratings according to whether each NP is in English-like order or not.

For each analysis, we report the results of either a linear model or a LMEM, depending on which is the better fit for our data (as discussed in Section 2.5). Where estimated typicality scores for particular sentences differ significantly from the reference sentence in their group, we report these estimates. We also include violin plots showing the z-scores for each sentence, along with the mean and standard deviation of the z-scores. Sentences or groups of sentences whose estimated ratings differ significantly from their reference sentence (at our alpha level of p = .004, as discussed in Section 2.5.6) are marked with an asterisk in our tables and graphs.

Overall, our R^2 statistics indicate that our models explain only between 1% and 29% of the variance in our data. This suggests that additional, unidentified factors are influencing participants' typicality ratings. The source of this variation remains an open question.

3.2.1 Noun and Demonstrative Point

Table 4 (repeated here as Table 23) contains our two sentences whose target noun phrases consist of a noun and a demonstrative point.

Table 23 Judgement Task Sentences: Noun and Demonstrative Point.

Code	Sentence glosses	NP structure
NP01^	GIRL BUY PT VEHICLE	[Dem Noun]
NP02	GIRL BUY VEHICLE PT	[Noun Dem]

Note: The reference sentence for our linear models is indicated by a ^ symbol.

Our linear regression model suggests that there is insufficient evidence to conclude that there is a significant difference in scores between NP01 (estimated score: 58) and NP02 (estimated score: 50) in this dataset (p = .19).

Figure 80 presents violin plots of the standardised z-scores for these sentences. On the left in purple is the distribution for NP01 and on the right in blue is the distribution for NP02. The Y-axis displays Z-scores ranging from −2 to +2. Each dot represents an individual data point (one response from one participant), while the curved outlines indicate the density of these points across the Z-score range. The mean z-score for NP01 was 0.09 and the standard deviation was 0.96. For NP02, the mean z-score was -0.09, with a standard deviation of 0.92. Both distributions are bimodal, indicating that participants were divided in their judgements: while some rated the sentences as typical, others rated them as atypical.

Figure 80 Violin Plot of Z-Scores of Typicality Ratings for Sentences Containing Just a Demonstrative.

Note. The reference sentence for our linear model is indicated by a ^ symbol.

However, if we broaden the analysis to include all sentences containing a demonstrative point, our LMEM indicates a significant dispreference (rather than preference) for postnominal demonstrative points, with an estimated score of 40, as compared to prenominal demonstrative points whose score is estimated as 45 (p = .001). The sentences are listed in Table 24 (prenominal demonstrative points) and Table 25 (postnominal demonstrative points), and the distribution of z-scores is shown in Figure 81.

As in the earlier comparison, the distributions for both prenominal and postnominal demonstratives are bimodal. This again reflects the presence of divergent views among participants: although the overall trend indicates a dispreference for postnominal demonstratives, not all such NPs were judged atypical. For example, NP07 [Adj Noun Dem] did not differ significantly from the NPs it was directly compared with (see Table 33). This pattern recurs across other NP groupings as well.

Table 24 Judgement Task Sentences that include a Prenominal Demonstrative Point.

Code	Sentence glosses	NP structure
NP01	GIRL BUY PT VEHICLE	[Dem Noun]
NP08	GIRL REMEMBER OLD PT FILM	[Adj Noun Dem]
NP11	GIRL REMEMBER PT FILM OLD	[Dem Noun Adj]
NP12	GIRL REMEMBER PT OLD FILM	[Dem Adj Noun]
NP14	BOY EAT FOUR PT SANDWICH	[Num Dem Noun]
NP17	BOY EAT PT SANDWICH FOUR	[Dem Noun Num]
NP18	BOY EAT PT FOUR SANDWICH	[Dem Num Noun]
NP29	BOY BRING PT FOUR RED MUG	[Dem Num Adj Noun]
NP31	BOY BRING PT RED FOUR MUG	[Dem Adj Num Noun]
NP35	BOY BRING FOUR PT RED MUG	[Num Dem Adj Noun]
NP39	BOY BRING FOUR RED PT MUG	[Num Adj Dem Noun]
NP41	BOY BRING RED PT FOUR MUG	[Adj Dem Num Noun]
NP43	BOY BRING RED FOUR PT MUG	[Adj Num Dem Noun]

Table 25 Judgement Task Sentences that include a Postnominal Demonstrative Point.

Code	Sentence glosses	NP structure
NP02	GIRL BUY VEHICLE PT	[Noun Dem]
NP07	GIRL REMEMBER OLD FILM PT	[Adj Noun Dem]
NP09	GIRL REMEMBER FILM OLD PT	[Noun Adj Dem]
NP10	GIRL REMEMBER FILM PT OLD	[Noun Dem Adj]
NP13	BOY EAT FOUR SANDWICH PT	[Num Noun Dem]
NP15	BOY EAT SANDWICH FOUR PT	[Noun Num Dem]
NP16	BOY EAT SANDWICH PT FOUR	[Noun Dem Num]
NP47	BOY BRING MUG PT FOUR RED	[Noun Dem Num Adj]
NP48	BOY BRING MUG PT RED FOUR	[Noun Dem Adj Num]
NP49	BOY BRING MUG FOUR PT RED	[Noun Num Dem Adj]
NP50	BOY BRING MUG FOUR RED PT	[Noun Num Adj Dem]
NP51	BOY BRING MUG RED PT FOUR	[Noun Adj Dem Num]
NP52	BOY BRING MUG RED FOUR PT	[Noun Adj Num Dem]

Figure 81 Violin Plot of Z-Scores of Typicality Ratings for All Sentences that Include a Demonstrative.

Note. The reference sentence group for our linear model is indicated by a ^ symbol and the sentence group whose score is significantly different to this are shown with an *.

3.2.2 Noun and Adjective

Table 5 (repeated here as Table 26) lists our two sentences whose noun phrase contain a noun and adjective.

Table 26 Judgement Task Sentences: Noun and Adjective.

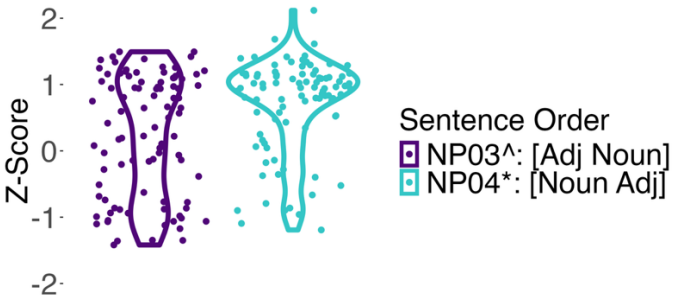
Code	Sentence glosses	NP structure
NP03^	GIRL BUY BLUE VEHICLE	[Adj Noun]
NP04*	GIRL BUY VEHICLE BLUE	[Noun Adj]

Note. The reference sentence for our linear models is indicated by a ^ symbol and sentences whose score is significantly different to this are shown with an *.

Our linear regression model reveals a significant preference for a postnominal adjective: it estimates our participants' raw score for NP03 at 62, and for NP04 significantly higher at 78 (p = .001).

Figure 82 shows the distribution of standardised z-scores for these sentences. For NP03, the mean z-score was 0.26 and the standard deviation of its z-scores was 0.91. NP04 was rated as more typical, with a mean z-score of 0.71 and a standard deviation of 0.7.

Figure 82 Violin Plot of Z-Scores of Typicality Ratings for Sentences Containing Just an Adjective.



Note. The reference sentence for our linear model is indicated by a ^ symbol and sentences whose score is significantly different to this are shown with an *.

However, if all judgement task sentences that contain an adjective are considered together, our LMEM indicates the opposite: a significant dispreference (rather than preference) for postnominal adjectives. Prenominal adjectives have an estimated score of 49, as compared to postnominal adjectives whose score is estimated as 43 ($p < .001$). The sentences are listed in Table 27 (prenominal adjectives) and Table 28 (postnominal adjectives), and the distribution of z-scores is shown in Figure 83.

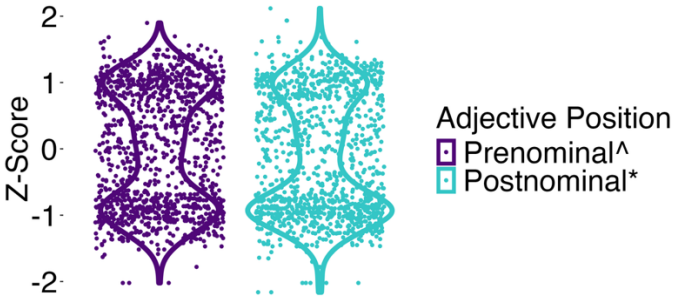
Table 27 Judgement Task Sentences that include a Prenominal Adjective.

Code	Sentence glosses	NP structure
NP03	GIRL BUY BLUE VEHICLE	[Adj Noun]
NP07	GIRL REMEMBER OLD FILM PT	[Adj Noun Dem]
NP08	GIRL REMEMBER OLD PT FILM	[Adj Dem Noun]
NP12	GIRL REMEMBER PT OLD FILM	[Dem Adj Noun]
NP20	GIRL BUY THREE OLD BOOK	[Num Adj Noun]
NP23	GIRL BUY OLD BOOK THREE	[Adj Noun Num]
NP24	GIRL BUY OLD THREE BOOK	[Adj Num Noun]
NP29	BOY BRING PT FOUR RED MUG	[Dem Num Adj Noun]
NP31	BOY BRING PT RED FOUR MUG	[Dem Adj Num Noun]
NP35	BOY BRING FOUR PT RED MUG	[Num Dem Adj Noun]
NP39	BOY BRING FOUR RED PT MUG	[Num Adj Dem Noun]
NP41	BOY BRING RED PT FOUR MUG	[Adj Dem Num Noun]
NP43	BOY BRING RED FOUR PT MUG	[Adj Num Dem Noun]

Table 28 Judgement Task Sentences that include a Postnominal Adjective.

Code	Sentence glosses	NP structure
NP04	GIRL BUY VEHICLE BLUE	[Noun Adj]
NP09	GIRL REMEMBER FILM OLD PT	[Noun Adj Dem]
NP10	GIRL REMEMBER FILM PT OLD	[Noun Dem Adj]
NP11	GIRL REMEMBER PT FILM OLD	[Dem Noun Adj]
NP19	GIRL BUY THREE BOOK OLD	[Num Noun Adj]
NP21	GIRL BUY BOOK THREE OLD	[Noun Num Adj]
NP22	GIRL BUY BOOK OLD THREE	[Noun Adj Num]
NP47	BOY BRING MUG PT FOUR RED	[Noun Dem Num Adj]
NP48	BOY BRING MUG PT RED FOUR	[Noun Dem Adj Num]
NP49	BOY BRING MUG FOUR PT RED	[Noun Num Dem Adj]
NP50	BOY BRING MUG FOUR RED PT	[Noun Num Adj Dem]
NP51	BOY BRING MUG RED PT FOUR	[Noun Adj Dem Num]
NP52	BOY BRING MUG RED FOUR PT	[Noun Adj Num Dem]

Figure 83 Violin Plot of Z-Scores of Typicality Ratings for All Sentences that Include an Adjective.



Note. The reference sentence group for our linear model is indicated by a ^ symbol and the sentence group whose score is significantly different to this are shown with an *.

3.2.3 Noun and Numeral

Table 6 (repeated here as Table 29) lists our two sentences whose noun phrase contain a noun and numeral.

Table 29 Judgement Task Sentences: Noun and Numeral

Code	Sentence glosses	NP structure
NP05^	BOY BUY THREE CAKE	[Num Noun]
NP06	BOY BUY CAKE THREE	[Noun Num]

Note. The reference sentence for our linear models is indicated by a ^ symbol.

Our linear regression model suggests that there is insufficient evidence to conclude that there is a significant difference in scores between NP05 (estimated raw score: 72) and NP06 (estimated raw score: 57) in this dataset (p = .009).

Figure 84 shows the distribution of standardised z-scores for these sentences. For NP05, the mean z-score was 0.55 and the standard deviation of its z-scores was 0.88. NP06 was rated as less typical, with a mean z-score of 0.2 and a standard deviation of 0.9.

Figure 84 Violin Plot of Z-Scores of Typicality Ratings for Sentences Containing Just a Numeral.

Note. The reference sentence for our linear model is indicated by a ^ symbol.

However, if all judgement task sentences that contain a numeral are considered together, our LMEM indicates a significant dispreference for postnominal numerals, with an estimated score of 38, as compared to prenominal numerals whose score is estimated as 44 (p < .001). The sentences are listed in Table 30 (prenominal numerals) and Table 31 (postnominal numerals), and the distribution of z-scores is shown in Figure 85.

Table 30 Judgement Task Sentences that include a Prenominal Numeral.

Code	Sentence glosses	NP structure
NP05	BOY BUY THREE CAKE	[Num Noun]
NP13	BOY EAT FOUR SANDWICH PT	[Num Noun Dem]
NP14	BOY EAT FOUR PT SANDWICH	[Num Dem Noun]
NP18	BOY EAT PT FOUR SANDWICH	[Dem Num Noun]
NP19	GIRL BUY THREE BOOK OLD	[Num Noun Adj]
NP20	GIRL BUY THREE OLD BOOK	[Num Adj Noun]
NP24	GIRL BUY OLD THREE BOOK	[Adj Num Noun]
NP29	BOY BRING PT FOUR RED MUG	[Dem Num Adj Noun]
NP31	BOY BRING PT RED FOUR MUG	[Dem Adj Num Noun]
NP35	BOY BRING FOUR PT RED MUG	[Num Dem Adj Noun]
NP39	BOY BRING FOUR RED PT MUG	[Num Adj Dem Noun]
NP41	BOY BRING RED PT FOUR MUG	[Adj Dem Num Noun]
NP43	BOY BRING RED FOUR PT MUG	[Adj Num Dem Noun]

Table 31 Judgement Task Sentences that include a Postnominal Numeral.

Code	Sentence glosses	NP structure
NP06	BOY BUY CAKE THREE	[Noun Num]
NP15	BOY EAT SANDWICH FOUR PT	[Noun Num Dem]
NP16	BOY EAT SANDWICH PT FOUR	[Noun Dem Num]
NP17	BOY EAT PT SANDWICH FOUR	[Dem Noun Num]
NP21	GIRL BUY BOOK THREE OLD	[Noun Num Adj]
NP22	GIRL BUY BOOK OLD THREE	[Noun Adj Num]
NP23	GIRL BUY OLD BOOK THREE	[Adj Noun Num]
NP47	BOY BRING MUG PT FOUR RED	[Noun Dem Num Adj]
NP48	BOY BRING MUG PT RED FOUR	[Noun Dem Adj Num]
NP49	BOY BRING MUG FOUR PT RED	[Noun Num Dem Adj]
NP50	BOY BRING MUG FOUR RED PT	[Noun Num Adj Dem]
NP51	BOY BRING MUG RED PT FOUR	[Noun Adj Dem Num]
NP52	BOY BRING MUG RED FOUR PT	[Noun Adj Num Dem]

Figure 85 Violin Plot of Z-Scores of Typicality Ratings for All Sentences that Include a Numeral

Note. The reference sentence group for our linear model is indicated by a ^ symbol and the sentence group whose score is significantly different to this are shown with an *.

3.2.4 NPs with Two Modifiers

We next consider judgement task sentences that consist of two modifiers plus a noun.

3.2.4.1 Noun, Adjective and Demonstrative Point

Table 7 (repeated here as Table 32) lists our six sentences whose noun phrase contain a noun, adjective and demonstrative point.

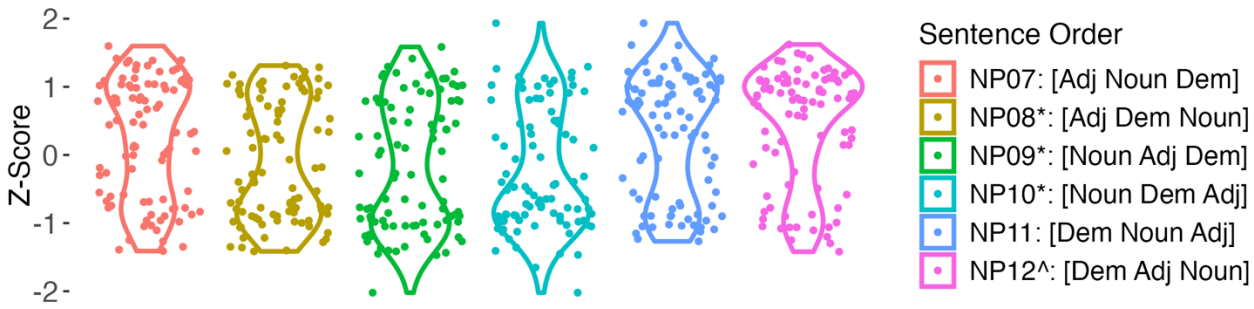
Table 32 Judgement Task Sentences: Noun, Adjective and Demonstrative Point

Code	Sentence glosses	NP structure
NP07	GIRL REMEMBER OLD FILM PT	[Adj Noun Dem]
NP08*	GIRL REMEMBER OLD PT FILM	[Adj Dem Noun]
NP09*	GIRL REMEMBER FILM OLD PT	[Noun Adj Dem]
NP10*	GIRL REMEMBER FILM PT OLD	[Noun Dem Adj]
NP11	GIRL REMEMBER PT FILM OLD	[Dem Noun Adj]
NP12^	GIRL REMEMBER PT OLD FILM	[Dem Adj Noun]

Note. The reference sentence for our linear models is indicated by a ^ symbol and sentences whose score is significantly different to this are shown with an *.

We report the results of our LMEM since this was a better fit for these sentences than our linear regression model (p = .03). The estimate of our participants' raw score for the reference sentence, NP12 is 74, and all the other sentences score less than this. The data supports a significant difference in scores for NP08, NP09 and NP10 but not for the remaining sentences in this group (p = .07 and p = .19). Figure 86 shows the distribution of z-scores for this group of sentences and Table 33 shows the sentences' mean z-score, standard deviation of z-scores plus estimated raw scores. For the significantly different sentences (NP08, NP09 and NP10) it shows the model p-value.

Figure 86 Violin Plot of Z-Scores of Typicality Ratings for Noun Phrases Consisting of Noun, Adjective and Demonstrative Point.



Note. The reference sentence for our linear model is indicated by a ^ symbol and sentences whose score is significantly different to this are shown with an *.

Table 33 Summary Statistics for Scores of Judgement Task Sentences containing Noun, Adjective and Demonstrative Point.

Code	NP07	NP08*	NP09*	NP10*	NP11	NP12^
Mean Z-Score	0.25	-0.11	-0.18	-0.25	0.31	0.46
s. d. of Z-scores	0.88	0.87	0.94	0.88	0.90	0.85
Estimated raw score	64	52	49	47	67	74
p-value	n.s.	<.001	<.001	<.001	n.s.	-

Note. The reference sentence for our linear models is indicated by a ^ symbol and sentences whose score is significantly different to this are shown with an *. *p*-value is only reported for sentences whose score is significantly different to the reference sentence.

3.2.4.2 Noun, Numeral and Demonstrative Point

Table 8 (repeated here as Table 34) lists our six sentences whose noun phrase contain a noun, numeral and demonstrative point.

Table 34 Judgement Task Sentences: Noun, Numeral and Demonstrative Point

Code	Sentence glosses	NP structure
NP13	BOY EAT FOUR SANDWICH PT	[Num Noun Dem]
NP14*	BOY EAT FOUR PT SANDWICH	[Num Dem Noun]
NP15*	BOY EAT SANDWICH FOUR PT	[Noun Num Dem]
NP16	BOY EAT SANDWICH PT FOUR	[Noun Dem Num]
NP17	BOY EAT PT SANDWICH FOUR	[Dem Noun Num]
NP18^	BOY EAT PT FOUR SANDWICH	[Dem Num Noun]

Note. The reference sentence for our linear models is indicated by a ^ symbol and sentences whose score is significantly different to this are shown with an *.

We report the results of our LMEM[24] since this was a better fit for these sentences than our linear regression model ($p < .001$). The estimate of our participants' raw score for the reference sentence

[24] This model excludes trialOrder due to singularity issues as discussed in Section 2.5.

NP18 was 43, and all the other sentences in this group score less than this. The data supports a significant difference in scores for NP14 and NP15, but not for the remaining sentences in this group (*p* ranging from .18 to .72). Figure 87 shows the distribution of z-scores for this group of sentences and Table 35 shows the sentences' mean z-score, standard deviation of z-scores plus estimated raw scores. For the significantly-different sentences (NP14 and NP15) it also shows the model *p*-value.

Figure 87 Violin Plot of Z-Scores of Typicality Ratings for Noun Phrases Consisting of Noun, Numeral and Demonstrative Point

Note. The reference sentence for our linear model is indicated by a ^ symbol and sentences whose score is significantly different to this are shown with an *.

Table 35 Summary Statistics for Scores of Judgement Task Sentences containing Noun, Numeral and Demonstrative Point

Code	NP13	NP14*	NP15*	NP16	NP17	NP18^
Mean Z-Score	-0.14	-0.50	-0.39	-0.04	-0.01	0.00
s. d. of Z-scores	0.89	0.86	0.79	0.90	0.85	0.96
Estimated raw score	36	23	26	39	41	43
p-value	n.s.	<.001	0.001	n.s.	n.s.	-

Note. The reference sentence for our linear models is indicated by a ^ symbol and sentences whose score is significantly different to this are shown with an *. *p*-value is only reported for sentences whose score is significantly different to the reference sentence.

3.2.4.3 Noun, Numeral and Adjective

Table 9 (repeated here as Table 36) lists our six sentences whose noun phrase contain a noun, numeral and adjective.

Table 36 Judgement Task Sentences: Noun, Numeral and Adjective

Code	Sentence glosses	NP structure
NP19	GIRL BUY THREE BOOK OLD	[Num Noun Adj]
NP20^	GIRL BUY THREE OLD BOOK	[Num Adj Noun]
NP21*	GIRL BUY BOOK THREE OLD	[Noun Num Adj]
NP22*	GIRL BUY BOOK OLD THREE	[Noun Adj Num]
NP23	GIRL BUY OLD BOOK THREE	[Adj Noun Num]
NP24*	GIRL BUY OLD THREE BOOK	[Adj Num Noun]

Note. The reference sentence for our linear models is indicated by a ^ symbol and sentences whose score is significantly different to this are shown with an *.

Here, our ANOVA indicates that there is no significant difference between our linear regression model and our LMEM (p = .06), so we report the simpler linear regression model. The estimate of our participants' raw score for the reference sentence NP20 was 69, and again, all other sentences score less than this. The data supports a significant difference in scores for NP21, NP22 and NP24 but not for NP19 or NP23 (p = .009 and .03 respectively). Figure 88 shows the distribution of standardised z-scores for these sentences and Table 37 shows the sentences' mean z-score, standard deviation of z-scores plus estimated raw scores. For the significantly different sentences (NP21, NP22 and NP24) it also shows the model p-value.

Figure 88 Violin Plot of Z-Scores of Typicality Ratings for Noun Phrases Consisting of Noun, Numeral and Adjective

Note. The reference sentence for our linear model is indicated by a ^ symbol and sentences whose score is significantly different to this are shown with an *.

Table 37 Summary Statistics for Scores of Judgement Task Sentences containing Noun, Numeral and Adjective

Code	NP19	NP20^	NP21*	NP22*	NP23	NP24*
Mean Z-Score	0.13	0.44	-0.35	-0.50	0.20	0.00
s. d. of Z-scores	0.92	0.89	0.85	0.79	0.86	0.92
Estimated raw score	55	69	36	30	57	50
p-value	n.s.	-	<.001	<.001	n.s.	<.001

Note. The reference sentence for our linear models is indicated by a ^ symbol and sentences whose score is significantly different to this are shown with an *. *p*-value is only reported for sentences whose score is significantly different to the reference sentence.

3.2.4.4 Summary of NPs with Two Modifiers

Across all judgement task NPs with two modifiers, we observed three consistent patterns that apply regardless of the category of modifier (adjective, demonstrative or numeral). First, NPs with both modifiers postnominal were generally dispreferred. Second, English-like order tended to be preferred. Third, NPs with one prenominal and one postnominal modifier were also rated as relatively typical. (No other judgement task sentence groups included NPs with a single prenominal modifier or a single postnominal modifier: As discussed in Section 2.3.1.5.3, we omitted three-modifier NPs that met this condition in order to keep the task manageable.)

3.2.5 Noun, Demonstrative Point, Adjective and Numeral

Table 10 (repeated here as Table 38) lists our 12 sentences whose noun phrases contain a noun, demonstrative point, adjective and numeral.

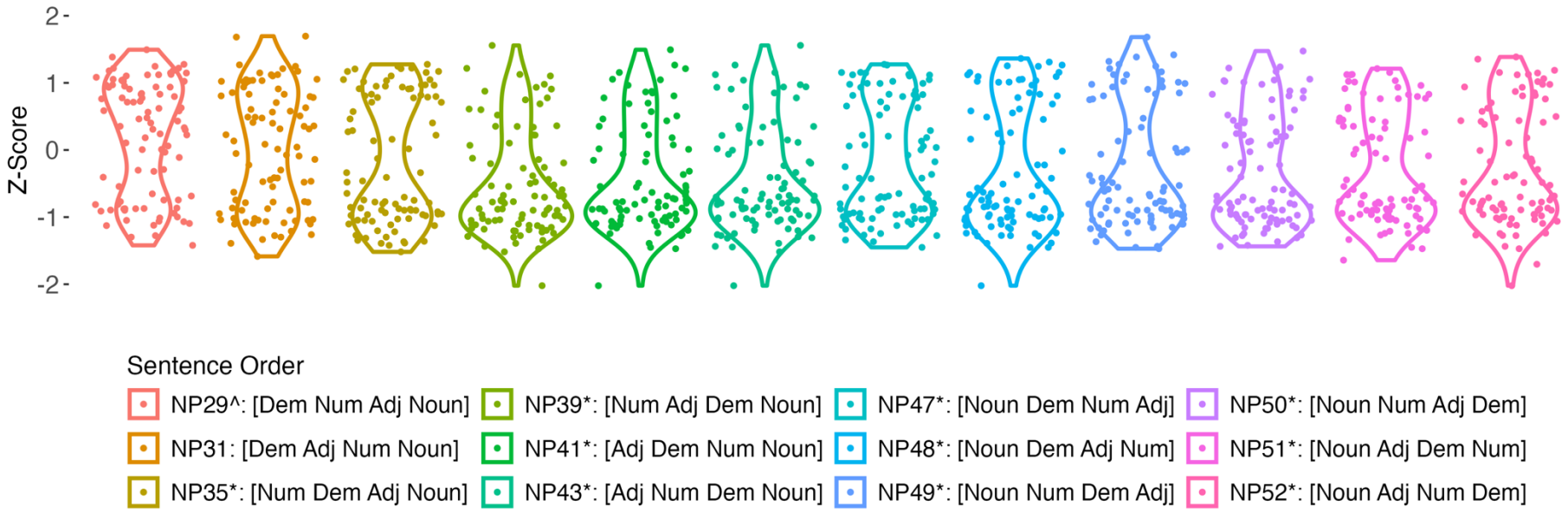
Table 38 Judgement Task Sentences: Noun plus Demonstrative Point, Adjective and Numeral

Code	Sentence glosses	NP structure
NP29^	BOY BRING PT FOUR RED MUG	[Dem Num Adj Noun]
NP31	BOY BRING PT RED FOUR MUG	[Dem Adj Num Noun]
NP35*	BOY BRING FOUR PT RED MUG	[Num Dem Adj Noun]
NP39*	BOY BRING FOUR RED PT MUG	[Num Adj Dem Noun]
NP41*	BOY BRING RED PT FOUR MUG	[Adj Dem Num Noun]
NP43*	BOY BRING RED FOUR PT MUG	[Adj Num Dem Noun]
NP47*	BOY BRING MUG PT FOUR RED	[Noun Dem Num Adj]
NP48*	BOY BRING MUG PT RED FOUR	[Noun Dem Adj Num]
NP49*	BOY BRING MUG FOUR PT RED	[Noun Num Dem Adj]
NP50*	BOY BRING MUG FOUR RED PT	[Noun Num Adj Dem]
NP51*	BOY BRING MUG RED PT FOUR	[Noun Adj Dem Num]
NP52*	BOY BRING MUG RED FOUR PT	[Noun Adj Num Dem]

Note. The reference sentence for our linear models is indicated by a ^ symbol and sentences whose score is significantly different to this are shown with an *.

We report the results of our LMEM since this was a better fit for these sentences than our linear regression model (p = .03). The estimate of our participants' raw score for the reference sentence, NP29 (which is in the same order as English), is 56 and again, all the other sentences score less than this. The data supports a significant difference in scores for all the other sentences in this group apart from NP31 (p = .12): Figure 89 provides the distribution of z-scores for these sentences and Table 39 shows the sentences' mean z-score, standard deviation of z-scores plus the estimated raw scores for each sentence. For the sentences whose raw score is significantly different to that of NP29, it also provides the model p-value.

Figure 89 Violin Plot of Z-Scores of Typicality Ratings for Noun Phrases Consisting of Noun, Demonstrative Point, Adjective and Numeral



Note. The reference sentence for our linear model is indicated by a ^ symbol and sentences whose score is significantly different to this are shown with an *.

Table 39 Summary Statistics for Scores of Judgement Task Sentences containing Noun, Demonstrative Point, Adjective and Numeral

Code	**NP29^**	**NP31**	**NP35***	**NP39***	**NP41***	**NP43***	**NP47***	**NP48***	**NP49***	**NP50***	**NP51***	**NP52***
Mean Z-Score	0.18	-0.01	-0.18	-0.53	-0.45	-0.48	-0.19	-0.33	-0.28	-0.34	-0.31	-0.31
s. d. of Z-scores	0.85	0.89	0.94	0.77	0.79	0.77	0.87	0.92	0.90	0.82	0.82	0.85
Estimated raw score	57	49	42	28	31	31	42	36	38	35	36	36
p-value	-	n.s.	.003	<.001	<.001	<.001	.002	<.001	<.001	<.001	<.001	<.001

Note. The reference sentence for our linear models is indicated by a ^ symbol and sentences whose score is significantly different to this are shown with an *. *p*-value is only reported for sentences whose score is significantly different to the reference sentence.

3.2.6 Homomorphism

We created a linear model and an LMEM that considered the effect on typicality rating of whether NPs were homomorphic or not. Table 18 in Section 2.5.3 lists the sentences included in these models along with whether or not they are homomorphic. We report the results of the LMEM since this was a better fit for our data ($p < .001$). The estimate for our reference level (homomorphic sentences) is 48, and the estimate for non-homomorphic sentences was significantly lower at 39 ($p < .001$).

Figure 90 shows that sentences that are not homomorphic were judged as relatively atypical.

Figure 90 Violin plot of Z-Scores of Typicality Ratings for Homomorphic Versus non-Homomorphic Sentences

Note. The reference sentence for our linear model is indicated by a ^ symbol and sentences whose score is significantly different to this are shown with an *.

3.2.6.1 Semantic Scope Distance Between Modifiers

In Section 1.3.2.3 we discussed Culbertson and Adger's (2014) finding that in their miniature artificial language experiment, participants more often selected homomorphic orders when the modifiers in the NP were widely spaced in terms of semantic scope. Here, we report on whether our judgement task results also accord with this theory: we compare the preference for homomorphic (as opposed to non-homomorphic) NPs between sentences where the NP modifiers were (a) widely spaced and (b) narrowly spaced in terms of semantic scope. We separated our judgement task sentences into (a) those with modifiers that are further apart in terms of the semantic scope relationship (i.e., demonstrative and adjective) and (b) those with modifiers that are adjacent in terms of semantic scope (i.e., adjective and numeral, or demonstrative and numeral).

We first discuss our judgement task sentences whose modifiers are widely separated in terms of semantic scope. They are listed in Table 40.

Table 40 Judgement Task Sentences Whose Modifiers are Widely Spaced in Terms of Semantic Scope, where Homomorphism is Defined

Code	Sentence glosses	NP structure	Homomorphic?
NP08	GIRL REMEMBER OLD PT FILM	[Adj Dem Noun]	No
NP09	GIRL REMEMBER FILM OLD PT	[Noun Adj Dem]	Yes
NP10	GIRL REMEMBER FILM PT OLD	[Noun Dem Adj]	No
NP12	GIRL REMEMBER PT OLD FILM	[Dem Adj Noun]	Yes

Our ANOVA indicates that there is no significant difference between our linear regression model and our LMEM (p = .68). We therefore report the results of the linear model since it is simpler: it estimates the score for the sentences that are homomorphic to be 62, and those that are not homomorphic to be significantly lower at 49 (p = .002). Figure 91 shows the distribution of standardised z-scores for these sentences. Participants judged non-homomorphic sentences to be relatively atypical whereas they judged homomorphic sentences to be relatively typical.

Figure 91 Violin plot of Z-Scores of Typicality Ratings for Homomorphic Versus non-Homomorphic Sentences Whose Modifiers are Widely Spaced in Terms of Semantic Scope

Note. The reference sentence group for our linear model is indicated by a ^ symbol and the sentence group whose score is significantly different to this are shown with an *.

Next, we present the results for the remaining sentences which have two modifiers: here, these modifiers are closer in terms of semantic scope. These are the sentences that contain an adjective and a numeral, or a demonstrative and a numeral. The sentences in question are listed in Table 41.

Table 41 Judgement Task Sentences whose Modifiers are Not Widely Spaced in Terms of Semantic Scope, where Homomorphism is Defined

Code	Sentence glosses	NP structure	Homomorphic?
NP14	BOY EAT FOUR PT SANDWICH	[Num Dem Noun]	No
NP15	BOY EAT SANDWICH FOUR PT	[Noun Num Dem]	Yes
NP16	BOY EAT SANDWICH PT FOUR	[Noun Dem Num]	No
NP18	BOY EAT PT FOUR SANDWICH	[Dem Num Noun]	Yes
NP20	GIRL BUY THREE OLD BOOK	[Num Adj Noun]	Yes
NP21	GIRL BUY BOOK THREE OLD	[Noun Num Adj]	No
NP22	GIRL BUY BOOK OLD THREE	[Noun Adj Num]	Yes
NP24	GIRL BUY OLD THREE BOOK	[Adj Num Noun]	No

Again, our ANOVA does not indicate a significantly better fit for our LMEM as compared to our simple linear model (p = .24), so we report the latter. There is no significant difference in the estimate of the scores for the homomorphic sentences (estimated raw score: 43) as compared to the non-homomorphic sentences (estimated raw score: 38). Figure 92 shows the distribution of z-scores for these sentences.

Figure 92 Violin Plot of Z-Scores of Typicality Ratings for Homomorphic Versus non-Homomorphic Sentences Whose Modifiers are Not Widely Spaced in Terms of Semantic Scope

Note. The reference sentence group for our linear model is indicated by a ^.

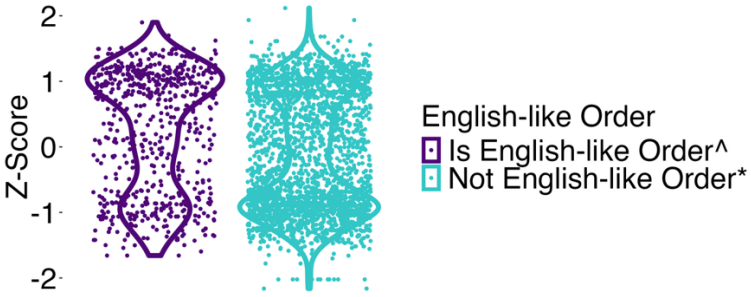
To summarise this section: our models indicate that for NPs whose modifiers are more widely separated in terms of semantic scope, homomorphic NPs are judged to be significantly more typical than non-homomorphic NPs. In contrast, for NPs whose modifiers are not widely separated in terms of scope, there is no significant difference in typicality ratings between homomorphic and non-homomorphic NPs. This accords with Culbertson and Adger's (2014) findings.

3.2.7 Distribution of Typicality Ratings by English-like Order

We created a further model with a fixed effect for whether each NP was in English-like order. Again, we report the results of our LMEM since this was a better fit for our data (p < .001). The estimate

for our reference level (sentences that are in English-like order) is 59. The estimate for sentences that are not in English-like order is significantly lower (p < .001) at 42. Figure 93 shows the distribution of typicality z-scores for sentences whose NPs are in English-like order as opposed to those that are in other orders. It is clear that sentences in English-like order were found to be highly typical. In contrast, sentences that are not in English-like order were judged as somewhat not typical.

Figure 93 Violin plots of Z-Scores of Typicality Ratings, by English-like versus not English-like Order



Note. The reference sentence for our linear model is indicated by a ^ symbol and sentences whose score is significantly different to this are shown with an *.

3.2.8 Effect of Social Factors

As discussed in Section 2.5, we included the following demographic characteristics of our participants in our LMs and LMEMs:

- Whether they were operationalised as being native, early or late learners of BSL;
- Their self-reported level of comfort with English;
- Whether they reported having taught BSL.

For all but one of our models, the data did not support a significant[25] difference in scores based on these demographic characteristics. The only significant difference emerged in the first model discussed in Section 3.2.6.1, which considers the effect of homomorphism in sentences whose modifiers are widely spaced in terms of semantic scope. (These sentences are listed in Table 40.) In this case, participants with higher self-reported comfort with English scored these sentences significantly lower in typicality (estimated score 48), compared to those who reported a lower level of comfort with English (estimated score 62, p = .003). However, there is no significant interaction between sentence homomorphism and participants' level of comfort with English regarding the typicality scores (p = .39), meaning that individuals with higher self-reported English comfort scored

[25] Significant results are defined as where p < .004, as discussed in Section 2.5.6.

this particular subset of NPs as less typical, irrespective of homomorphism. We therefore view this result as spurious.

3.2.9 Variation in Typicality Scores for a Given Sentence

We observed considerable variation in the responses of our judgement task participants. Even when the majority view of a given sentence was clear, there were generally several participants who disagreed. For example, Figure 84 above illustrates that although most people rated sentence NP05, (whose NP consists of [Num Noun]) as very typical, several scored it as atypical.

NP05 was always presented together with the same sentence, NP06 [Noun Num], as discussed in Section 2.3.1.2. However, the situation is more complex for sentences with two or three modifiers as these were paired randomly with another sentence with the same types of modifier. For example, NP24 [Adj Num Noun], was paired with one of the other sentences which contained a numeral and a noun, i.e., NP19 – NP23. The typicality score for a given sentence can be affected by its pair: for instance, NP24 tended to be rated less typical when paired with NP20 [Num Adj Noun] and more typical when paired with NP22 [Noun Adj Num]. However, this partner sentence effect should be balanced out across the whole experiment since pairings were random.

The overall variation in typicality scores for a given sentence (even after considering the sentence that it is paired with) underscores the benefit of recruiting a large sample of raters rather than relying on a small number of respondents.

This concludes our results section. In the following chapter we consider these results as they relate to our research questions, then we discuss some of the feedback that we received from our judgement task participants. Finally, we note some limitations of this project and provide some suggestions for future research.

Chapter 4 Discussion

This project was designed to investigate the structure of noun phrases (NPs) in British Sign Language (BSL) from two complementary angles: corpus analysis and judgement task experiments. The results obtained from these two methods offer distinct insights into the research questions and hypotheses formulated in Section 1.4. By combining and comparing these results, we are able to achieve a more thorough and refined understanding of the NP structure in BSL and derive more solid conclusions.

In this chapter, we will first examine how each research question and hypothesis is addressed by both sets of results. Following this, we discuss the feedback that our judgement task participants provided. We close the chapter with recommendations for future research.

4.1 RQ 1: Order and Frequency of Elements in the BSL NP

Our first research question was:

RQ 1. What orders of elements within the NP occur in BSL, and at what frequency?

Here, we consider how BSL Corpus data and our BSL judgement task results address this, first taken separately (Sections 4.1.1 and 4.1.2 respectively) and then considered together (Section 4.1.3).

4.1.1 Corpus Data

Before discussing the order of elements within NPs in BSL, we first consider the distribution of different lengths of NPs.

4.1.1.1 Lengths of NPs

As noted in Section 3.1.1, we observed that the vast majority of our NPs were short: 786 of our 1037 NPs (76%) consisted of only a single element. This aligns with previous cross-linguistic research on spoken language including Dryer (2007a), and also with Culbertson et al. (2020), who found that very few NPs had more than one modifier in English corpus data. Thompson and Ono (2020) made a similar observation in their discussion of NPs in unscripted interaction.

Comparable findings have been reported for sign languages. Mantovan (2017), for example, found that most of the nouns in her corpus of Italian Sign Language (LIS) included no more than two modifiers (see Section 1.3.4.7.3), and Rutkowski et al. (2015) reported a similar pattern based on their analysis of the Polish Sign Language (PJM) corpus (see Section 1.3.4.7.2). In their study, 77% of

the nouns identified had no modifiers within the NP–remarkably close to our figure of 76% single-element NPs. However, our methodological approach differed from that of Rutkowski et al. Whereas their analysis began by identifying nouns and then examining their modifiers, we identified all elements that constitute an NP, regardless of whether they included a noun. As a result, their approach excludes NPs consisting solely of a pointing sign–structures that we found to be very frequent: 44% (*N* = 455) of our NPs consisted of just a pointing sign. By contrast, only 24% (*N* = 249) of our BSL NPs consisted of just a noun–less than one third of the PJM figure.

Table 1 in Section 1.3.3 above provided a comparison of the lengths of NPs that contain a noun plus at least one modifier in Polish and English. We expand that here in Table 42, to add figures for PJM, from Rutkowski et al. (2015), and for BSL, from our own corpus study. We found that the NPs in our BSL data were somewhat longer compared to the other languages. This may be due to differences in methodology across the studies–such as variation in corpus design, annotation criteria (including NP inclusion thresholds)–or to differences in the types of texts included. The Polish data came from a spoken language corpus; the PJM data was based on narrative retellings; the English data came from a corpus consisting mostly of written texts, with some spoken material; and our BSL data was drawn primarily from free conversations, with some personal narratives. Further research is needed to determine whether the observed pattern is attributable to methodological factors, text type, or a genuine difference in NP length between BSL and the other languages.

Table 42 Comparison of the Percentage of NPs of Different Lengths in Polish, PJM, English and BSL

Language/ Length of NP (words/signs)	**Polish (%)**	**PJM (%)**	**English (%)**	**BSL (%)**
2	85	89	83	74
3	13	10	14	19
4 or more	0.7	1	3	7

Note. Polish and PJM data is from Rutkowski et al. (2015). English data is from Culbertson et al. (2020). BSL data is from the current corpus study. It includes NPs that include a noun and at least one modifier.

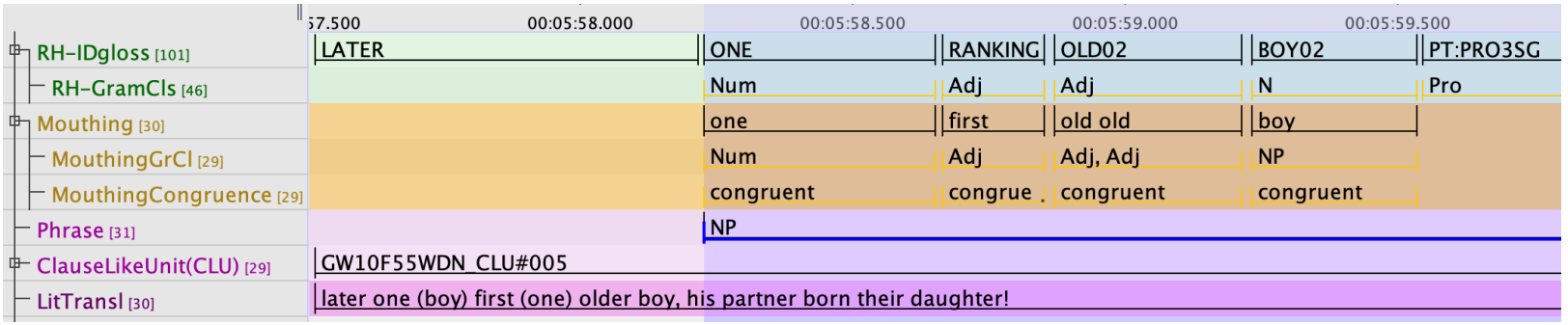
4.1.1.2 Order of Elements within NPs

4.1.1.2.1 Order of NP Elements Overall

Turning to the order of elements within our NPs: The fact that 786 of our 1037 NPs contain only a single element means that we can only discuss order within an NP in the remaining 251 NPs. In these, many orders are found, as evidenced by the large number of different orders in both two-element NPs (Figure 63, Section 3.1.3.2) and three-element NPs (Figure 70, Section 3.1.3.3). We find that for all lengths of NP, the distributions of grammatical categories are skewed, with a small number of orders occurring many times and many other orders occurring only rarely.

Regarding determining an overall order including all types of modifier, we find only a single token of an NP in our corpus data that includes all three modifiers: it is shown here as Figure 94.

Figure 94 GW10n Example of an NP Containing All Types of Modifier (Adjective, Numeral, Point): “Later, my oldest son…”



Note: CLU#005 is too long to display in a single screenshot. We present the relevant portion here; the *LitTransl* tier provides the translation of the entire CLU.

In Section 1.3.2 we noted that previous researchers have extrapolated from examples of NPs containing subsets of modifiers to determine the order for NPs containing all modifiers. However, an examination of our corpus data shows that here, multiple orders are attested, from which it is not possible to extrapolate to a single order. For example, we see from Figure 70 in Section 3.1.3.3 that if we consider our NPs with two modifiers, where we have exactly one noun and no multiple occurrences of the same type of modifier (e.g., two adjectives), we find three occurrences with a prenominal point - see (60).

(60) [Point Adj Noun]

In contrast, (61) includes a postnominal point and is also attested three times.

(61) [Adj Noun Point]

The fact that both prenominal and postnominal points are attested means that no single primary or preferred order of point, adjective and noun can be determined.

As discussed in Section 3.1.3.5, it is possible to determine whether corpus NPs are in English-like order where the NP has more than one element, has exactly one noun, and has no simultaneous elements. We found 184 such NPs and of these, 137 (74%) had the same order as in English.

4.1.1.2.2 Order of Noun and Adjective re Hypothesis 1

Crosslinguistically, postnominal adjectives are more common: Dryer (2018) reports that 75% of the world's spoken languages in his representative sample have postnominal adjectives. Sutton-Spence and Woll (1999) similarly describe BSL as generally having postnominal adjectives. These observations informed our first hypothesis: that adjectives in BSL would be postnominal. Relatedly, Culbertson et al. (2020) found that hearing non-signing participants using silent gesture showed a strong tendency to produce postnominal adjectives.

However, our corpus results did not bear out this hypothesis. As noted in Section 3.1.3.2, we found that adjectives are prenominal 88% of the time (N = 28) in NPs consisting of just a noun and an adjective. Similarly, in NPs with three or more elements including a noun and at least one adjective (Section 3.1.3.3), all 12 examples (100%) had prenominal adjectives.

It seems plausible that this discrepancy reflects a genuine pattern in the data, despite the relatively small sample size in the current study: although we identified 1037 NPs, only 44 clear-cut examples allowed us to determine adjective position. When comparing our results with Culbertson et al.'s

(2020) silent gesture findings, however, we should be cautious. Their participants were speakers of a well-studied language (English) with a consistent and stable word order, whereas BSL is under-studied and lacks well-documented stable sign orders[26].

4.1.2 Judgement Task Data

4.1.2.1 Overall Order of Elements within NPs

To address our first research question "What orders of elements within the NP occur in BSL, and at what frequency?", we used participants' typicality ratings from the judgement task as a proxy for frequency.

In the judgement task, we grouped sentences according to the set of elements they contained. For example, all six possible orders of {Adj, Num, Noun} were compared in randomised pairs–see Section 2.3.1.2. Each group included one sentence whose order matched English word order. In every group except one, the English-like sentence was rated as the most typical. This pattern was confirmed in our statistical model: NPs in English-like order were judged significantly more typical than those in non-English-like order. The exception was NP04 [Noun Adj], which was judged significantly more typical than NP03 [Adj Noun]. Sentence groups that include an adjective are considered further in Section 4.1.2.2 in relation to Hypothesis 1.

We now turn to sentence groups that contained other modifiers. First, demonstratives: as outlined in Section 3.2.1, participants showed no preference between pre- and postnominal demonstratives when the demonstrative was the only modifier. However, when combined with other modifiers, postnominal demonstratives were significantly dispreferred. A similar pattern was observed for numerals (see Section 3.2.3): no preference was expressed when the numeral was the sole modifier, but postnominal numerals were significantly dispreferred when other modifiers were present.

Next, we examine NPs with two modifiers: As discussed in Section 3.2.4.4, the strongest pattern in such NPs was a consistent dispreference for orders in which both modifiers were postnominal. Participants also tended to prefer English-like orders. Orders with one prenominal and one postnominal modifier were rated as relatively typical. When all three modifiers were present, participants again showed a consistent preference for English-like order.

[26] We thank Gabrielle Hodge for this insight.

Finally, we also considered whether combining the results for a single modifier (either adjective, demonstrative or numeral) might reveal a general preference for pre- or postnominal position. We find it very likely that the strong tendency for adjectives to occur postnominally would dominate the analysis, obscuring patterns for demonstratives and numerals, for which no significant positional preference was observed. For this reason, we did not conduct this combined analysis.

4.1.2.2 Order of Noun and Adjective re Hypothesis 1

Sentence groups containing adjectives are considered here in relation to Hypothesis 1. In Section 3.2.2, we observed that participants in our judgement task significantly preferred postnominal adjectives in NPs where the adjective is the only modifier. This finding supports Hypothesis 1, which posits that adjectives in BSL are generally postnominal. NP04 [Noun Adj] received the highest typicality rating across the whole task. As noted in the previous section, NPs with one prenominal and one postnominal modifier (numeral or demonstrative, as well as adjective) were rated as relatively typical. In contrast to the strong typicality score for NP04, the combinations involving an adjective plus other modifier(s) following the noun were consistently dispreferred.

Although the judgment task does not reveal participants' reasoning, one plausible explanation for this discrepancy is the influence of linguistic ideology. Some BSL users may subscribe to the belief that BSL and English should differ structurally as distinct languages. See, for example, this quote from the influential sign linguist Mary Brennan:

> There is little doubt that word order information has direct relevance to those engaged in teaching and learning any language. … [T]eachers of British Sign Language (BSL), for example, often claim that sign order in BSL is considerably different from word order patterning in English. (Brennan, 1994, p. 9)

Relatedly, Rowley and Cormier (2024) analysed interview responses from BSL Corpus participants regarding their perceptions of the differences between British Sign Language (BSL) and Sign-supported English (SSE), the latter of which adheres to English word order. Their findings indicate that participants frequently exaggerated the extent of syntactic divergence between BSL and SSE. For instance, some respondents claimed that wh-question words in BSL must appear sentence-finally and cannot occur at the beginning of a sentence, as they do in English. However, corpus evidence shows that initial wh-questions are in fact common in BSL (Hodge et al., 2025). This suggests that the belief in a rigid syntactic separation between BSL and English has become a widely accepted ideology within the British deaf community, and our results suggest that this ideology extends to the view that adjectives in BSL are postnominal. This perception is further perpetuated

by certain BSL teaching materials–for example, the online resources provided by Dot Sign Language (n.d.). A similar stance is evident in the Signature Level 2 BSL qualification specification, which explicitly states that candidates must "use BSL sign order, with no English structure" (Signature, 2022), without acknowledging that BSL and English word orders may align in some constructions.
It is therefore possible that participants' marked preference for a postnominal adjective in NP04 [Noun Adj] reflects an adherence to the ideology that BSL should differ structurally from English. In this case, the adjective's position may have been more noticeable, prompting participants to apply the belief that BSL favours postnominal placement. However, when the adjective appeared along with other modifiers, this strong preference disappeared. This suggests that the ideology may exert a stronger influence in simple [Noun Adj] structures. In contrast, when multiple modifiers are present, the adjective's position may be less perceptually prominent or may interact with other cognitive pressures such as homomorphism, thereby weakening the influence of ideology.

4.1.3 Synthesis of Corpus Data and Judgement Task Data

Overall, we observed that most of our corpus NPs were very short (one or two elements) and we identified a wide variety of orders of NP in our corpus data. This accords with Dryer (2007b), who found the same in many spoken languages (see Section 1.3.2) and with the previous research into signed languages that we discussed in Section 1.3.4.7. In our BSL data, we observed that the distribution of NP orders was skewed, with a few orders being very frequent and the rest occurring only rarely. This again matches other studies such as Rutkowski's investigation of PJM (see Section 1.3.4.7.2).

We found that nearly three-quarters (74%) of corpus NPs where an order of elements could be determined were in the same order as in English. In our judgement task, we found that English-like order tended to be preferred, apart from NPs containing just a noun and an adjective. These findings suggest a strong, but not overwhelming, influence of English on BSL.

It may be that the influence of English just extends to a preference for prenominal modifiers. If so, the resulting NP order in BSL [(Dem) (Num) (Adj) Noun] could be attributed to a general cognitive preference for homomorphism. Since this is the only order with all prenominal modifiers that is also homomorphic, it would emerge naturally under such pressures, regardless of the specific order in English.

Alternatively, the influence of English may be more direct, involving the transfer of specific linear orders from English into BSL. On this account, the prevalence of English-like NP order in BSL would not be due to an independent preference for homomorphism, but rather a consequence of English's dominant sociolinguistic status and the strength of its influence on BSL. In this case, the homomorphic structure of BSL NPs would be a side-effect of adopting English word order, rather than the result of a universal cognitive bias.

Our finding is similar to Neidle and Nash's (2012) observation that in American Sign Language (ASL), when all modifiers of a noun occur prenominally, their canonical order is as in (62) which matches spoken English: again, this may well be due to an influence of English as the language surrounding ASL.

(62) [(Dem) (Num) (Adj) Noun]

It is not always the case that the surrounding spoken language exerts a strong influence on the order within the NP: see for example, Rutkowski et al. (2015) (Section 1.3.4.7.2) who observed that the NP order in PJM is different to that in spoken Polish. And as discussed in Section 1.3.4.7.3, the situation regarding LIS as compared to spoken Italian is mixed: in some cases, the modifier position has been found to be the same and in other cases it differs. (The previous research into Taiwanese Sign Language (TSL) and Hong Kong Sign Language (HKSL) both found many orders within NPs to be acceptable, so that we can't draw conclusions about the degree of influence from their respective surrounding spoken languages.)

Returning to Hypothesis 1, that adjectives in BSL would be postnominal: overall, we obtained conflicting results from our corpus and judgement task data. On the one hand, contra Hypothesis 1, our corpus participants produced prenominal adjectives 91% of the time and our judgement task participants dispreferred postnominal adjectives when the target NPs included other modifiers as well as an adjective. On the other hand, supporting Hypothesis 1, when judgement task participants were asked to assess NPs where the only modifier was an adjective, they preferred this modifier to be postnominal. This one finding is in line with Halliday and Matthiessen's (2004) observation that we noted in Section 1.3.5.1, that people's intuitions about language do not necessarily reflect their actual use of that language. However, for our other results, participants' intuitions tended to match our observations of the language in use. In summary, we cannot say that Hypothesis 1 is supported.

Overall, {Noun, Adj} order in BSL tends to parallel the order in English, perhaps due to an influence of English on this aspect of BSL (cf. Bowern, 2006, as noted in Section 1.3.4.6). The contrasting finding for judgement task sentences where the only modifier was an adjective may be because of the ideology that {Noun, Adj} order in BSL should be different to English[27].

These findings highlight the importance of ensuring that BSL teaching materials reflect actual language use (see, for example, Fenlon, 2019). When teaching is based on authentic patterns of expression, learners are better equipped to communicate effectively in real-world situations. This reduces misunderstandings and enhances their ability to interact naturally with other users of the language. We hope that future research, including the current project, will inform BSL curricula such as the upcoming BSL GCSE (UK Government, 2023).

4.2 RQ 2: Comparison of NP Orders in BSL to Theoretically Predicted Orders

Our second research question and hypotheses were:

RQ 2. How well does the order in NPs in BSL match established predictions about the cross-linguistic distribution of NP order in different languages, such as Greenberg's Universal 20, Cinque (2005) and homomorphism (as discussed in Section 1.3.2.3)?

Hypothesis 2a: The dominant order in NPs in BSL is predicted by Greenberg's Universal 20.
Hypothesis 2b: The dominant order in NPs in BSL is predicted by Cinque (2005).
Hypothesis 2c: NPs in BSL are homomorphic.

Greenberg's Universal 20 and Cinque (2005) both relate to the basic unmarked order in the NP for a given language. We therefore need to identify a single order in BSL that can be compared to these theories. Based on our finding in Section 3.1.3.5 that 74% of relevant corpus NPs were in the order in (63) and on our finding in Section 3.2.6 that NPs in this order were judged as significantly more typical than those in other orders, we determine that the order shown in (63) is the dominant order in BSL. (Our data is not sufficient to determine whether this is the basic, unmarked order in BSL–we

[27] Another possible way to characterise our finding–that [Noun Adj] is preferred over [Adj Noun] but that otherwise, NP modifiers tend to be prenominal–is through the framework of Optimality Theory (Prince & Smolensky, 2004). This approach allows us to model the observed patterns as the result of competing grammatical constraints, ranked in a language-specific hierarchy. The following tendencies appear to guide NP order in BSL: (1) NPs are harmonic (i.e. their modifiers are all on the same side of the noun) (2) numerals and demonstratives are prenominal, (3) adjectives are postnominal. When multiple modifiers are present, the preference for harmonic structure (1) would take precedence over the default postnominal position of adjectives (3), resulting in all modifiers appearing prenominally. However, while this describes our data, it does not explain why it is only adjectives (and not numerals and demonstratives) that tend to be postnominal.

leave that question to future research.)

(63) [(Dem) (Num) (Adj) Noun]

4.2.1 Comparison of NP Orders in BSL to Greenberg's Universal 20

Previous research into noun phrases (NPs) in various sign languages, as discussed in Section 1.3.4.7 suggests that NPs in ASL, LIS, TSL and HKSL all adhere to Greenberg's Universal 20. It was not possible to draw conclusions about Universal 20 from the PJM data. BSL aligns with the patterns observed in these other sign languages, with the dominant order in NPs in BSL being one of the orders predicted by Greenberg's Universal. Therefore, our hypothesis is supported.

4.2.2 Comparison of NP Orders in BSL to Cinque (2005)

As discussed in Section 1.3.2.2.1, expanding beyond Greenberg (1963), Cinque (2005) found that 14 of the possible 24 orders within an NP were attested in his sample of spoken languages. We noted in Section 1.3.4.7 that the orders within the NPs of ASL, LIS and TSL as reported in various studies adhere to Cinque (2005), while there is insufficient evidence to determine whether this is the case for PJM and HKSL. Similarly, BSL exhibits one of Cinque's orders as the dominant NP order, thereby supporting our hypothesis.

4.2.3 Comparison of NP Orders in BSL to Homomorphism

Our hypothesis is as follows:
Hypothesis 2c: NPs in BSL are homomorphic.

We discussed in Section 1.3.4.7 that previous research has found NPs in ASL, LIS and TSL to be homomorphic, whereas there is not enough evidence to determine the status of NPs in PJM and HKSL. In terms of our corpus results for BSL, as discussed in Section 3.1.3.4 we observed very few NPs where homomorphism could be determined. This is because homomorphism is only defined for NPs with at least 2 different types of modifier on the same side of the noun. In fact, we found only seven tokens of homomorphic NPs and two tokens of non-homomorphic NPs (as in Table 21), so that 78% of this small sample were homomorphic: this proportion is not significantly different to what would be expected by chance. When considering NPs with widely-spaced modifiers in terms of semantic scope (i.e., with modifiers {Adj, Point}) as compared to those with narrowly-spaced modifiers (i.e., with modifiers {Adj, Num} or {Num, Point}), there was no significant difference in terms of the proportion of homomorphic sentences in each group.

In terms of our judgement task data, as discussed in Section 3.2.6, our participants judged that overall, NPs that were homomorphic were significantly more typical than non-homomorphic NPs. Our models for NPs with widely-spaced modifiers indicated a significantly more typical rating for homomorphic NPs as compared to non-homomorphic NPs. In contrast, our models for NPs with narrowly-spaced modifiers did not reveal a significant preference for homomorphism.

Our judgement task results for widely- versus narrowly-spaced modifiers align with those observed by Culbertson and Adger (2014), as discussed in Section 1.3.2.3: their participants showed a stronger preference for homomorphic NP orders when the modifiers were widely spaced in terms of semantic scope. This pattern, replicated in our study with deaf BSL users, is particularly striking given the substantial differences in experimental paradigm. Whereas Culbertson and Adger used an artificial language learning task with English-speaking participants producing novel postnominal orders, our study involved fluent BSL signers making typicality judgements about NP orders in BSL. This convergence across modalities and task types suggests that the degree of scope separation between modifiers plays a role in how readily language users map hierarchical structure onto linear order. The clearer the scope hierarchy, the more likely it is to be reflected in NP order. These findings support the idea that homomorphism reflects a general tendency to maintain a transparent relationship between semantic scope and linear order. This lends further support to the psychological plausibility of the base NP structure outlined by Abels and Neeleman (2012).

Taken together, our judgement task and corpus results suggest that our hypothesis is supported: NPs in BSL tend to be homomorphic. This tendency towards homomorphism accords with observations about previously-studied spoken languages as well as in those signed languages where sufficient data is available. Since homomorphic orders reflect the underlying semantic scope relations between NP elements, this lends weight to the theory that there is a universal abstract mental representation that influences the typology of NP word order regardless of language modality.

4.3 RQ 3: Effect of Age of Acquisition of BSL, Teaching Experience and Level of English Comfort on NP Order

Our third research question was:

RQ 3. How does signers' reported language background, reported experience of teaching BSL and their level of comfort with English (based on self-report as discussed in Section 2.5) affect the order of signs within their NPs?

Our hypotheses were as follows:

Hypothesis 3a: Signers' reported language background will affect the degree of influence of English on their BSL. People who acquired BSL later in life will experience more influence of English than those who acquired it earlier. Therefore, they will tend to prefer NPs that are in English-like order.

Hypothesis 3b: Whether participants report having taught BSL may affect the ideologies that they hold about BSL structure. BSL teachers may be more likely to hold the ideology that since BSL and English are different languages, the order within the NP in BSL should be different to in English. Therefore, they will tend to prefer NPs that are not in English-like order.

Hypothesis 3c: Signers who have higher levels of comfort with English will experience more influence of English on their BSL. Therefore, they will tend to prefer NPs that are in English-like order.

Here, we consider how our BSL Corpus data (Section 4.3.1) and our BSL judgement task results (Section 4.3.2) address this question and hypotheses. Then we synthesise both data sources in Section 4.3.3.

4.3.1 Corpus Data

Looking first at Hypothesis 3a: In terms of language background, of our 34 corpus participants, 15 were native signers and the remaining 19 were early learners of BSL. (See Table 2 for a list of all corpus participant characteristics.) We operationalise influence of English for a particular signer's BSL production as the proportion of their NPs which are in English-like order. As discussed in Section 3.1.3.5, we found no effect of language background on the order of elements within our participants' NPs: native signers and early learners of BSL produced almost the same proportion of NPs in English-like order. Therefore Hypothesis 3a is not supported. We did not select our corpus participants based on age of acquisition of BSL: this plus the fact that 95% of all corpus participants were native or early learners of BSL (see Section 2.2) resulted in the absence of any late learners of BSL amongst our cohort. The investigation of the influence of English including late learners of BSL is a possible area for future research, as noted in Section 4.5.3.

Turning to Hypothesis 3b, we do not have information about the teaching experience of most of the corpus participants that we selected. Of these 34 people, four said they were trained teachers, three

reported being teachers without having received formal training and seven more said that they were not BSL teachers. Nothing is recorded regarding teaching experience for the remaining 20 participants. Because of this, we cannot draw any conclusions about differences in NP order based on corpus participant teacher status, meaning that we cannot address Hypothesis 3b using our corpus data.

Finally, we cannot address Hypothesis 3c using BSL Corpus data because when the corpus was created, no information was collected about participants' level of comfort with English.

4.3.2 Judgement Task Data

We developed models as discussed in Section 2.5 for each of our groups of judgement task sentences, incorporating social factors (language background, comfort with English, and BSL teaching experience) as independent variables. The only significant effect of a social factor on participants' typicality scores was observed in the model analysing the four homomorphic sentences whose modifiers were widely spaced in terms of semantic scope (see Section 3.2.6.1). Specifically, participants with higher self-reported comfort with English provided significantly lower typicality scores for these sentences compared to those with lower self-reported comfort with English.
Despite adjusting our significance threshold (alpha) to account for multiple comparisons (see Section 2.5.6), we are cautious about drawing strong conclusions from this single result as it may be spurious. The scores did not depend on whether the sentences were homomorphic, which was the primary focus of this model. Consequently, this finding does not address whether signers with higher levels of comfort with English experience greater English influence on their NPs. In summary, we cannot conclude that Hypotheses 3a-c are supported.

4.3.3 Synthesis of Corpus Data and Judgement Task Data

Overall, we did not find any notable significant correlations between the demographic characteristics of either our corpus or our judgement task participants and the order of elements in their NPs. This means that we have not identified any demographic reasons for the variation in NP orders that we observed. It is possible that further studies may identify such correlations, for example, by considering a larger dataset. These and other possible areas for future research are discussed in Section 4.5.

One implication of our findings relates to Hypothesis 3b. Although we were unable to systematically test for effects of teaching experience due to incomplete metadata in the corpus, the judgement

task data did not show any significant differences in NP order preferences between participants who reported teaching BSL and those who did not. This suggests that the ideology that BSL should differ structurally from English–specifically, that adjectives should follow nouns–is not confined to those who teach BSL. Rather, it appears to be widespread across the British deaf signing community more broadly.

4.4 Judgement Task Feedback from Participants

As noted in Section 2.3.1.8, at the end of our judgement task, we asked participants whether they had any feedback about the task. Of the 92 participants, 49 people (53%) provided some feedback, though not all of it related to our NP sentences: some was about other parts of the overall task (see Section 2.3.1.3). Five people reported that they had some issues watching the videos due to temporary internet connectivity problems, but these all resolved quickly. Two other people mentioned that they did not find the typicality slider easy to use.

While many participants commented that they found the task interesting, people also often stated that they found it challenging. The most common specific negative remark was regarding the lack of non-manual features in the BSL sentences: though since the comments relate to the whole task, not just the NP part, this may be due to the fact that no mouthings were used alongside the sentences in the sentential constituent order part of the task (separate from the NP items). The NP sentences did include mouthing but other non-manual features were neutral, as noted in Section 2.3.1.5.

Four participants made explicit mention of using their intuition when responding. For example, one person stated: “I found it interesting how sentences form in my mind” and another said: “It was interesting that one can know when the sentences are not grammatically correct”. The other two participants stated that they signed the expressions themselves–in one case, “to see what felt more like BSL”. This practice of signing the expressions themselves is similar to Schütze and Sprouse’s (2014) suggestion, mentioned in Section 1.3.5.2.1, that participants imagine written sentences being spoken aloud to help them consider what a native speaker would say. In both cases, participants engage their intuitive sense of language by mentally or physically producing the expression. However, signing an expression seen in a video is more like repeating an audio sentence than reading a written one, since prosody is already present in the stimulus–unlike written stimuli, which require the reader to supply the prosody.

Fifteen people made comments about BSL grammar or vocabulary: for example, one person (who reported that they had never taught BSL) said that they felt they were being tested about the difference between BSL and Sign-supported English (SSE, defined in Section 1.3.4.6). In fact, BSL teachers were no more likely to make comments about BSL grammar or vocabulary than non-teachers: A chi-square test showed that there was no significant association between whether participants identified as teachers and whether they made a comment about BSL grammar: χ^2 (1, N = 92) = 1.74, p = .19. Similarly, no significant association was found between whether participants identified themselves as trained teachers versus teachers without training versus non-teachers and comments relating to BSL grammar: χ^2 (2, N = 92) = 1.78, p = .41.

Four participants mentioned the high degree of diversity within the BSL community. They noted that many deaf individuals are not actively involved in the community and may have limited exposure to BSL, particularly if they were educated in oralist settings. Participants highlighted differences in grammatical awareness and signing fluency, often shaped by educational background, family environment, and additional needs such as undiagnosed learning difficulties. These factors contribute to a wide range of BSL proficiency, requiring signers to adapt their communication to suit different interlocutors. Similarly, four participants noted that there is variety in how BSL is used. As a result, they found it challenging to determine how typical the BSL sentences presented in the task were of everyday usage, given the diversity of signing styles they encounter. This aligns with findings by Moriarty and Kusters (2021)[28], who observed that deaf signers interacting with deaf people from other countries often adjust their signing in ways that reflect not only linguistic sensitivity but also a moral orientation toward inclusion and mutual understanding.

Three people commented that they would have liked to see more diversity in the signers used to demonstrate the sentences: we recruited two white males because we needed models who were able to produce BSL sentences in unusual orders and with unnatural non-manual features and these individuals were the most appropriate people available. The agency that provided these signers has since increased the diversity of its staff so that it should be possible to recruit more diverse models for future projects.

Four participants noted that the test sentences could have been more realistic and/or more complex.

[28] We thank Gabrielle Hodge for drawing our attention to this research.

4.5 Recommendations for Future Research

The results of this study lay the foundations for future research into BSL NP structure. In this section, we outline several areas where future research could build on our findings, refine our methods, or address limitations in our data. These recommendations are intended to support the development of a more comprehensive understanding of NP structure in BSL and its interaction with cognitive, social and typological factors.

4.5.1 Effect of Welsh on the Structure of NPs in BSL in Wales

We discussed in Section 4.1.3 that a key finding from this project is that BSL NP order tends to match the NP order in English. We noted that the same situation has been observed for ASL and English, whereas in contrast, NP order in PJM differs from that in spoken Polish. In order to investigate the extent to which the fact that NP order in BSL tends to match English is an influence of English, one could investigate BSL NP order in signers who are surrounded by an additional spoken language as well as English, such as Welsh. NP order in Welsh differs from English in that adjectives and demonstratives are postnominal (Dryer, 2013a, 2013b), whereas Welsh and English are the same in that both Welsh and English have prenominal numerals (Dryer, 2013c). There is likely no population of BSL users who know (spoken or written) Welsh but not any (spoken or written) English, but there may be individuals who are exposed to a relatively large amount of Welsh. If NP order in the surrounding spoken language does affect NP order in BSL, then these signers' BSL NPs will tend towards the same order as Welsh, i.e., with postnominal adjectives and demonstratives, whilst retaining prenominal numerals (as found in both English and Welsh). Although such a population may not currently exist in Wales, this line of enquiry could be valuable when examining other signed languages with distinct groups of users who are exposed to spoken languages with differing NP structures.

4.5.2 Corpus Data: Numeral Incorporation and Adjective Incorporation: Consideration of Mouthings

As noted in Section 1.3.4.4, when an adjective or numeral is incorporated into a noun sign (e.g., the sign for the numeral TWO incorporated into the sign for WEEK, giving TWO-WEEKS), the two elements are produced simultaneously. Future studies could examine any mouthings produced alongside such constructions in BSL Corpus data, in order to determine a linear order. Of course, since such

mouthings will be of (parts of or whole) English words, there may be an additional influence of English here, leading signers to favour English-like order in such expressions.

4.5.3 Larger Sample Size for Corpus Data: Current or Additional Corpus

Our study identified only a small number of clear-cut NPs in the BSL Corpus where an order of elements could be established. As noted in Section 4.3.3, future research could build on the methodology of the current study by using a larger proportion of the corpus data (either more participants or more data from the same participants), potentially enabling more robust conclusions.

Future research may also benefit from including more late learners of BSL. The present study did not focus on this group because, by design, 95% of corpus participants were native or early learners of BSL. The creators of the corpus adopted this approach so as to minimise the confounding effects of language deprivation that can be more pronounced in individuals who learn BSL later in life, as discussed in Section 1.3.4.6. Specifically, the corpus was designed to maximise the number of participants who learned to sign before age seven. However, incorporating more late learners in future research would enable a more comprehensive investigation of how age of BSL acquisition influences the degree of English influence on NP structure, as discussed in Section 4.3.3. Moreover, collecting additional corpus data from late learners may enhance the real-world relevance of the findings, given that this group represents a substantial proportion of the deaf signing community in the UK.

Additionally, the BSL Corpus does not consistently include information about whether participants had experience teaching BSL. This limitation, combined with the fact that we did not consider teaching experience in our participant selection, means that–as noted in Section 4.3.1–we only have data on the teaching status of 41% (*N* = 14) of our 34 corpus participants. Consequently, we were unable to analyse this factor. Although our judgement task did not reveal any effect of teacher status, it may be valuable for future research to explore this variable using a corpus in which such metadata is systematically recorded. This would allow for testing the hypothesis that individuals who have taught BSL may be more likely to hold the ideology that NP order in BSL should differ from that in English. If they do, BSL teachers may produce a higher proportion of NPs that do not follow English-like order.

4.5.4 Include All Combinations of Length Four NPs in a Judgement Task

There are 24 possible arrangements of a noun with three modifiers (adjective, demonstrative, and numeral). The specific NPs chosen for our judgement task are detailed in Section 2.3.1.5.3. As explained there, to keep the task manageable, and because our survey platform required all NPs with three modifiers to be presented sequentially as a group, we excluded the 12 NP arrangements in which two modifiers appeared on one side of the noun and the third on the other, such as (64).

(64) NP30 BOY BRING PT FOUR MUG RED [Dem Num Noun Adj]

Instead, we used our NPs containing two of the three modifiers (NP07 – NP24) to assess the relative typicality of different orders of two modifiers on the same side of the noun. For example, (65) is equivalent to (64) in that both contain a demonstrative followed by a numeral in prenominal position.

(65) NP18 BOY EAT PT FOUR SANDWICH [Dem Num Noun]

Even with the omission of half the possible three-modifier NPs, several participants commented on the repetitive nature of the RED MUG sentences, likely because they had to be presented together as pairs in a group. It remains possible that different results would be obtained if all 24 combinations were included.

Notably, the 12 omitted NPs all have either one prenominal modifier (and two postnominal modifiers) or one postnominal modifier (and two prenominal modifiers). As noted in Section 3.2.4.4, NPs with two modifiers where one modifier was prenominal and the other was postnominal were judged as relatively typical. Inclusion of all the arrangements of NPs with three modifiers would therefore allow investigation of the typicality of these longer NPs with a single prenominal or a single postnominal modifier. We leave this to future research, which could use a different experimental design allowing these items to be interspersed with other sentences.

4.5.5 Refining Judgement Task Design

A broader suggestion for future task design would involve introducing an intermediate step between corpus analysis and large-scale testing[29]. This stage would involve using a corpus in which more–and longer–NPs have been annotated. Researchers would systematically manipulate selected corpus NPs by altering the position of their modifiers. Deaf research assistants could contribute to this process,

[29] We thank Klaus Abels for this suggestion.

helping to generate hypotheses about the typicality of these modified NPs. These hypotheses would then be tested in a broader judgement task with a larger and more diverse participant sample. This approach would enable a more systematic and theoretically grounded selection of stimuli, reducing the likelihood of including items that are ambiguous or open to multiple interpretations–an issue discussed in the following section.

4.5.6 Limitation: Participants' Interpretation of Judgement Task NP sentences

It is possible that some of our judgement task sentences may not have been interpreted as we intended by our participants. We have identified two ways in which this may have happened, and these are discussed below.

4.5.6.1 Subordinate NP

It is possible that–for any of the sentences that contain all three modifiers, and where all the modifiers are postnominal, such as

(66) NP48: BOY BRING MUG PT RED FOUR,

these may have been interpreted with a subordinate NP. In other words, these may have been interpreted as a clause plus a fragment, as in

(67) BOY BRING [MUG]$_{NP}$ [PT RED FOUR]$_{NP}$ (in English, "the boy brings [mugs]$_{NP}$, [those four red ones]$_{NP}$")

instead of the intended

(68) BOY BRING [MUG PT RED FOUR]$_{NP}$ (in English, "the boy brings [those four red mugs]$_{NP}$").

Table 43 lists all the sentences where this could have happened.

Table 43 Judgement Task Sentences that Could Have Been Interpreted with Subordinate NPs

Code	Sentence glosses	NP structure with subordinate NP
NP47	BOY BRING MUG PT FOUR RED	[Noun]$_{NP}$ [Dem Num Adj]$_{NP}$
NP48	BOY BRING MUG PT RED FOUR	[Noun]$_{NP}$ [Dem Adj Num]$_{NP}$
NP49	BOY BRING MUG FOUR PT RED	[Noun]$_{NP}$ [Num Dem Adj]$_{NP}$
NP50	BOY BRING MUG FOUR RED PT	[Noun]$_{NP}$ [Num Adj Dem]$_{NP}$
NP51	BOY BRING MUG RED PT FOUR	[Noun]$_{NP}$ [Adj Dem Num]$_{NP}$
NP52	BOY BRING MUG RED FOUR PT	[Noun]$_{NP}$ [Adj Num Dem]$_{NP}$

Johnston (2019) notes that in his study using the Australian Sign Language (Auslan) Corpus, he observed little overt marking to distinguish NPs from verbless attributive clauses or hypotactic clauses. Instead, signers appear to use the context of the utterance to determine the intended structure. It may be that other hypotactic material, such as a subordinate NP, is similarly not often overtly marked in BSL. In our experiment, we chose not to provide more context for our sentences, in order to keep the task manageable. This meant that no context was available to our participants so they would have to have use other strategies to parse the sentences. However, we did ask our signing model to produce the task sentences without any pauses within the NPs, to maximise the chances of them being interpreted as NPs.

Despite this precaution, we acknowledge the possibility that some participants may have interpreted the judgement task sentences differently from how we intended. If this occurred, it could have led them to rate these sentences as more typical. Nevertheless, all of NP47–52 were rated significantly less typical than the reference sentence for this group. This suggests that any such unintended interpretations would likely have increased the ratings, thereby making the observed differences even more compelling.

Another potential misinterpretation of some of our judgement task sentences could be viewing them as containing a verbless attributive clause. This is discussed in the following section.

4.5.6.2 Verbless Attributive Clause

Two NPs in our dataset–NP10 and NP11–may have been interpreted by participants as containing a verbless attributive clause, such as "that film is old," rather than the intended NP "that old film." This alternative interpretation is plausible because both NPs end in an adjective and follow the verb REMEMBER, which can take a verbless attributive clause as its complement.

(69) NP10: GIRL REMEMBER FILM PT OLD
Intended interpretation: GIRL REMEMBER [FILM PT OLD]$_{NP}$
Possible interpretation: GIRL REMEMBER [FILM PT]$_{NP}$ OLD

(70) NP11: GIRL REMEMBER PT FILM OLD
Intended interpretation: GIRL REMEMBER [PT FILM OLD]$_{NP}$
Possible interpretation: GIRL REMEMBER [PT FILM]$_{NP}$ OLD

No other judgement task sentences meet both criteria–ending in an adjective and following a complement-taking verb–so this ambiguity is specific to NP10 and NP11. If participants interpreted

these NPs as verbless clauses, they may have rated them as more typical than they would have if they had interpreted them as we intended.

NP10 was judged significantly less typical than its reference sentence (NP12), though the possibility of a clausal interpretation may have softened the negative rating. NP11's typicality score was not significantly different from its reference sentence, which might suggest that some participants interpreted it as a verbless clause. However, given the similarity between NP10 and NP11–they differ only in order of noun and demonstrative–it is difficult to draw firm conclusions about the role of this ambiguity in shaping participants' ratings.

Chapter 5 Conclusion

The noun phrase (NP) has been extensively studied in spoken languages and, to a lesser extent, in signed languages. However, it has never been systematically investigated in British Sign Language (BSL); this project addresses that gap. Our findings show that NPs in BSL tend to be short, and that commonly held assumptions about NP-internal word order are not always supported by naturalistic data.

Previous research has primarily focused on identifying the basic unmarked order of NP elements–noun, adjective, demonstrative, and numeral–across a wide range of spoken languages. Theoretical accounts such as Greenberg's (1963) *Universal 20* and Cinque's (2005) cartographic approach have proposed typological and formal analyses that aim to predict which NP orders are attested across languages. A key concept underlying these theories is *homomorphism*, the idea that linear order reflects semantic scope. Culbertson and Adger (2014), using an artificial language learning paradigm, found that homomorphic orders were preferred when modifiers differed widely in semantic scope (e.g., adjective and demonstrative), but not when they were closer in scope (e.g., adjective and numeral).

Traditionally, grammaticality or acceptability of NP orders has been assessed using judgements from a small number of informants. More recent work has tended to adopt one of two complementary methodologies: (1) analysis of naturalistic language corpora and (2) experimental judgement tasks with larger participant samples. In this study, we combined both approaches, using the BSL Corpus to examine naturally occurring NP structures, and an online judgement task to gather typicality ratings from a broad sample of fluent BSL users. To our knowledge, this is the first time these two methodologies have been combined in the study of a sign language, or in NP research more generally.

Our data reveals a range of NP orders in BSL. Both corpus and judgement task results suggest that the dominant order is as shown in (71), which mirrors the canonical order in English:

(71) [(Dem) (Num) (Adj) Noun]

Adjectives in BSL are generally prenominal. However, in NPs consisting only of a noun and an adjective, participants in the judgement task rated this postnominal adjective as significantly more typical. We suggest that this anomalous result may reflect language ideologies within the deaf

community–specifically, a belief that BSL should differ structurally from English, which may influence metalinguistic judgements and reinforce BSL's status as an autonomous language.

The dominant BSL NP order in (71) is consistent with both Greenberg's Universal 20 and Cinque's predictions. Moreover, BSL NPs tend to be homomorphic. We found a significant preference for homomorphic orders when modifiers were widely separated in semantic scope, but no such preference when modifiers were closer in scope. This pattern aligns with Culbertson and Adger's (2014) findings, despite the very different methodology.

Although we observed considerable variation in NP orders across both datasets, we found no significant correlations with participants' age of BSL acquisition, self-reported comfort with English, or experience teaching BSL. (Our statistical models explained only a small proportion of the observed variation, leaving its source an open question.) Notably, the absence of any difference between teachers and non-teachers suggests that the belief that BSL should structurally diverge from English–such as favouring postnominal adjectives–is not confined to educators. Rather, it appears to be a widely shared ideology within the British deaf community.

By combining corpus and experimental methods, this study takes a novel approach to investigating NP structure in a sign language. These findings provide the first systematic evidence of a preference for homomorphic NP order in a sign language. This may point to a universal cognitive representation of NP structure that shapes typological patterns across modalities.

Finally, our finding that the dominant NP order in BSL mirrors that in English challenges the assumption that BSL must differ structurally from English in all respects. This similarity may, in part, reflect the influence of language contact–given that most deaf BSL users are, to varying degrees, bilingual in BSL and in English (spoken, written, or both). Crucially, such structural overlap does not undermine BSL's status as a distinct and fully-fledged language. Rather, it highlights the complex interplay between typological universals, language contact, and community ideologies in shaping the grammar of signed languages.

Bibliography

Abels, K. (2016). The fundamental left–right asymmetry in the Germanic verb cluster. *Journal of Comparative Germanic Linguistics*, *19*(3), 179–220. https://doi.org/10.1007/s10828-016-9082-9

Abels, K., & Neeleman, A. (n.d.). *Linear Asymmetries and Incremental Parsing (under review)*.

Abels, K., & Neeleman, A. (2012). Linear Asymmetries and the LCA. *Syntax*, *15*(1), 25–74. https://doi.org/10.1111/j.1467-9612.2011.00163.x

Ahmadzadeh, Y. (2021). *Reflections on Racism in Research*. https://www.youtube.com/watch?v=t5ApTusVoBs

Algeo, J. (1995). Having a look at the expanded predicate. In *The verb in contemporary English: theory and description* (pp. 203–217). Cambridge University Press.

Baker, M. (2003). Lexical Categories: Verbs, Nouns and Adjectives. In *Cambridge Studies in Linguistics*. Cambridge University Press. https://doi.org/DOI: 10.1017/CBO9780511615047

Bank, R., Crasborn, O., & van Hout, R. (2011). Variation in mouth actions with manual signs in Sign Language of the Netherlands (NGT). *Sign Language & Linguistics*, *14*(2), 248–270. https://doi.org/10.1075/sll.14.2.02ban

Barr, D. J., Levy, R., Scheepers, C., & Tily, H. J. (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language*, *68*(3), 255–278. https://doi.org/10.1016/j.jml.2012.11.001

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models Using lme4. *Journal of Statistical Software*, *67*(1). https://doi.org/10.18637/jss.v067.i01

Bertone, C. (2009). The syntax of noun modification in Italian Sign Language (LIS). *Working Papers in Linguistics*, *19*.

Birkeland, A., Block, A., Craft, J. T., Sedarous, Y., Wang, S., Wu, A., & Namboodiripad, S. (2024). Rejecting nativeness to produce a more accurate and just Linguistics. *Language*, *100*(3), e156–e194. https://doi.org/10.1353/lan.2024.a937197

Blümel, A., & Holler, A. (2021). DP, NP, or neither? Contours of an unresolved debate. *Glossa: A Journal of General Linguistics*, *7*(1), 1–25. https://doi.org/10.16995/glossa.8326

Bolinger, D. (1967). Adjectives in English: Attribution and Predication. *Lingua*, *18*, 1–34.

Bose, P., Buss, J. F., & Lubiw, A. (1998). Pattern matching for permutations. *Information Processing Letters*, *65*, 277–283.

Boudreault, P., & Mayberry, R. I. (2006). Grammatical processing in American Sign Language: Age of first-language acquisition effects in relation to syntactic structure. *Language and Cognitive Processes*, *21*(5), 608–635. https://doi.org/10.1080/01690960500139363

Bowern, C. (2006). Syntactic Change and Syntactic Borrowing in Generative Grammar. In G. Ferarresi & M. Goldbach (Eds.), *Handbook of Historical Reconstruction*.

Branchini, Chiara. (2015). *On Relativization and Clefting: An Analysis of Italian Sign Language*. De Gruyter. https://doi.org/10.1515/9781501500008

Brennan, M. (1994). Word Order: Introducing the Issues. In M. Brennan & G. Turner (Eds.), *Word-order issues in sign language: working papers* (p. 9). University of Durham.

Brentari, D. (2010). Introduction. In D. Brentari (Ed.), *Sign Languages* (pp. 1–16). Cambridge University Press. https://doi.org/10.1017/CBO9780511712203.002

Brown, M. (2025). *Conjunction, disjunction, and scalar implicatures in British Sign Language* [University College London]. https://discovery.ucl.ac.uk/id/eprint/10209047

Brown, M., & Cormier, K. (2017). Sociolinguistic variation in the nativisation of BSL fingerspelling. *Open Linguistics*, *3*(1), 115–144. https://doi.org/10.1515/opli-2017-0007

Brunelli, M. (2011). *Antisymmetry and Sign Languages: A Comparison between NGT and LIS*. LOT Publications.

Cabin, R. J., & Mitchell, R. J. (2000). To Bonferroni or Not to Bonferroni: When and How Are the Questions. *Bulletin of the Ecological Society of America*, *81*(3), 246–248. https://www.jstor.org/stable/20168454

Carnie, A. (2013). *Syntax: A Generative Introduction* (3rd ed.). Wiley-Blackwell.

Cinque, G. (n.d.). *Dataset on the order of demonstrative, numeral, adjective, and noun across 2,200 languages*. Retrieved June 24, 2025, from https://hdl.handle.net/10278/3750226

Cinque, G. (2005). Deriving Greenberg's Universal 20 and Its Exceptions. *Linguistic Inquiry*, *36*(3), 315–332. https://doi.org/10.1162/0024389054396917

Cormier, K., Fenlon, J., Gulamani, S., & Smith, S. (2017). *BSL Corpus annotation conventions: Version 3.0*. *March*, 1–18. https://bslcorpusproject.org/wp-content/uploads/BSLCorpus_AnnotationConventions_v3.0_-March2017.pdf

Cormier, K., Quinto-Pozos, D., Sevcikova, Z., & Schembri, A. (2012). Lexicalisation and de-lexicalisation processes in sign languages: Comparing depicting constructions and viewpoint gestures. *Language and Communication*, *32*(4), 329–348. https://doi.org/10.1016/j.langcom.2012.09.004

Cormier, K., Schembri, A., & Woll, B. (2013). Pronouns and pointing in sign languages. *Lingua*, *137*, 230–247. https://doi.org/10.1016/j.lingua.2013.09.010

Crasborn, O. (2011). The Other Hand in Sign Language Phonology. *The Blackwell Companion to Phonology*, *March 2011*, 1–18. https://doi.org/10.1002/9781444335262.wbctp0010

Crasborn, O., & Sloetjes, H. (2008). Enhanced ELAN functionality for sign language corpora. *Proceedings of LREC 2008, Sixth International Conference on Language Resources and Evaluation*, 39–43.

Culbertson, J., & Adger, D. (2014). Language learners privilege structured meaning over surface frequency. *Proceedings of the National Academy of Sciences of the United States of America*, *111*(16), 5842–5847. https://doi.org/10.1073/pnas.1320525111

Culbertson, J., Schouwstra, M., & Kirby, S. (2020). From the world to word order : Deriving biases in noun phrase order from statistical properties of the world. *Language*, *96*(3). https://doi.org/10.1353/lan.2020.0045

Culbertson, J., Smolensky, P., & Legendre, G. (2012). Learning biases predict a word order universal. *Cognition*, *122*(3), 306–329. https://doi.org/10.1016/j.cognition.2011.10.017

Cysouw, M. (2010). Dealing with diversity: Towards an explanation of NP-internal word order frequencies. *Linguistic Typology*, *14*(2–3), 253–286. https://doi.org/10.1515/LITY.2010.010

De Lacy, P. V. (2006). *Markedness : reduction and preservation in phonology* [Book]. Cambridge University Press.

Den Dikken, M., Bernstein, J. B., Tortora, C., & Zanuttini, R. (2007). Data and grammar: Means and individuals. *Theoretical Linguistics*, *33*(3), 335–352. https://doi.org/10.1515/TL.2007.022

Don, J. (1993). *Morphological conversion* [PhD thesis, Utrecht University]. https://www.researchgate.net/publication/34536652

Dot Sign Language. (n.d.). *Frequently Asked Questions about British Sign Language (BSL)*. Https://Www.Dotsignlanguage.Co.Uk/Faq/.

Dryer, M. (2006). *On Cinque on Greenberg's universal 20*. Unpublished conference proceedings.

Dryer, M. (2007a). Noun Phrase Structure. In T. Shopen (Ed.), *Language Typology and Syntactic Description, Volume 3* (pp. 151–207). Cambridge University Press. https://doi.org/10.1017/CBO9780511619434.003

Dryer, M. (2007b). Word order. In T. Shopen (Ed.), *Language Typology and Syntactic Description, Volume 1* (pp. 61–131). Cambridge University Press. https://doi.org/10.1017/CBO9780511619427.002

Dryer, M. (2013a). *Order of Adjective and Noun (v2020.4)*. The World Atlas of Language Structures Online. https://doi.org/10.5281/zenodo.13950591

Dryer, M. (2013b). *Order of Demonstrative and Noun (v2020.4)*. The World Atlas of Language Structures Online. https://doi.org/10.5281/zenodo.13950591

Dryer, M. (2013c). *Order of Numeral and Noun (v2020.4)*. The World Atlas of Language Structures Online. https://doi.org/10.5281/zenodo.13950591

Dryer, M. (2018). On the order of demonstrative, numeral, adjective, and noun. *Language*, *94*(4), 798–833. https://doi.org/10.1353/lan.2018.0054

Evans, N., & Levinson, S. C. (2009). The myth of language universals: Language diversity and its importance for cognitive science. *Behavioral and Brain Sciences*, *32*(5), 429–492. https://doi.org/10.1017/S0140525X0999094X

Fenlon, J. (2019). Sign language linguistics and sign language teaching: Realigning the two fields. [Plenary presentation]. *Theoretical Issues in Sign Language Research (TISLR), Hamburg.*

Fenlon, J., Cormier, K., Rentelis, R., Schembri, A., Rowley, K., Adam, R., & Woll, B. (2014). *BSL SignBank: A lexical database of British Sign Language (First edition)*. http://bslsignbank.ucl.ac.uk

Fenlon, J., & Hochgesang, J. A. (2022a). Introduction to Signed Language Corpora. In J. Fenlon & J. A. Hochgesang (Eds.), *Signed Language Corpora* (pp. 1–17). Gallaudet University Press. https://doi.org/10.2307/j.ctv2rcnfhc.6

Fenlon, J., & Hochgesang, J. A. (Eds.). (2022b). *Signed Language Corpora*. Gallaudet University Press. https://doi.org/10.2307/j.ctv2rcnfhc

Goldin-Meadow, S., So, W. C., Özyürek, A., & Mylander, C. (2008). The natural order of events: How speakers of different languages represent events nonverbally. *Proceedings of the National Academy of Sciences of the United States of America*, *105*(27), 9163–9168. https://doi.org/10.1073/pnas.0710060105

Greenberg, J. (1963). Some universals of grammar with particular reference to the order of meaningful elements. *Universals of Language*, 73–113. https://doi.org/10.1017/CBO9781107415324.004

Halliday, M. A. K., & Matthiessen, C. M. I. M. (2004). *An Introduction to Functional Grammar*. Hodder Arnold. https://doi.org/10.4324/9780203431269

Harrison, X. A., Donaldson, L., Correa-Cano, M. E., Evans, J., Fisher, D. N., Goodwin, C. E. D., Robinson, B. S., Hodgson, D. J., & Inger, R. (2018). A brief introduction to mixed effects modelling and multi-model inference in ecology. *PeerJ*, *6*, e4794. https://doi.org/10.7717/peerj.4794

Haspelmath, M. (2012). How to compare major word-classes across the world's languages. *UCLA Working Papers in Linguistics*, *17*(3), 109–130. http://phonetics.linguistics.ucla.edu/wpl/issues/wpl17/papers/16_haspelmath.pdf

Haspelmath, M. (2020, December 5). *Acceptability judgements tell us about social norms, not about internal systems.* Diversity Linguistics Comment. https://doi.org/10.58079/nsw2

Henrich, J., Heine, S. J., & Norenzayan, A. (2010). The weirdest people in the world? *Behavioral and Brain Sciences*, *33*(2–3), 61–135. https://doi.org/10.1017/S0140525X0999152X

Hodge, G. (2013). *Patterns from a signed language corpus: Clause-like units in Auslan (Australian sign language)*.

Hodge, G., Manrique, E., Winter, B., & Cormier, K. (2025). Manual wh-signs and English wh-mouthings differentiate BSL content and polar questions. *Linguistics*. https://doi.org/10.1515/ling-2023-0238

Hou, L. (2024). *Interrogating "deaf" native signer ideologies in linguistics research.* The ROLE Collective: Rejecting Harmful Speaker/Signer Ideologies. https://www.youtube.com/watch?v=BHHWOHfGvR8

Huddleston, R. (1984). Introduction to the Grammar of English. In *Cambridge Textbooks in Linguistics*. Cambridge University Press. https://doi.org/DOI: 10.1017/CBO9781139165785

Johnston, T. (2001). Nouns and verbs in Auslan (Australian Sign Language): An open and shut case? *Journal of Deaf Studies and Deaf Education*, *6*(4), 235–257. https://doi.org/10.1093/deafed/6.4.235

Johnston, T. (2012). Lexical frequency in sign languages. *Journal of Deaf Studies and Deaf Education*, *17*(2), 163–193. https://doi.org/10.1093/deafed/enr036

Johnston, T. (2019). Clause constituents, arguments and grammatical relations in Auslan (Australian sign language): a corpus-based study. *Studies in Language*, *43*(3), 719–776. https://doi.org/10.1075/sl.18035.joh

Johnston, T. (2022). Foreword. In J. Fenlon & J. A. Hochgesang (Eds.), *Signed Language Corpora* (pp. xi–xviii). Gallaudet University Press. https://doi.org/10.2307/j.ctv2rcnfhc.5

Johnston, T., Cresdee, D., Schembri, A., & Woll, B. (2015). FINISH variation and grammaticalization in a signed language: How far down this well-trodden pathway is Auslan (Australian Sign Language)? *Language Variation and Change*, *27*, 117–155. https://doi.org/10.1017/S0954394514000209

Johnston, T., & Schembri, A. (2007). *Australian Sign Language (Auslan): An introduction to sign language linguistics*. Cambridge University Press. https://doi.org/10.1017/CBO9780511607479

Johnston, T., van Roekel, J., & Schembri, A. (2016). On the conventionalization of mouth actions in Australian Sign Language. *Language and Speech*, *59*(1), 3–42. https://doi.org/10.1177/0023830915569334

Johnston, T., Vermeerbergen, M., Schembri, A., & Leeson, L. (2007). 'Real data are messy': Considering cross-linguistic analysis of constituent ordering in Auslan, VGT, and ISL. In *Visible Variation Comparative Studies on Sign Language Structure*. De Gruyter. https://doi.org/10.1515/9783110198850.163

Johnston, T., Whynot, L., Hansford, C., Hatchard, B., Gray, M., Hodge, G., Allen, J., Shearim, G., Banna, K., Fried, D., & Beuzeville, L. De. (2019). *Auslan Corpus Annotation Guidelines - SUPERSEDED* (Issue August).

Kamp, H., & Partee, B. (1995). Prototype theory and compositionality. *Cognition*, *57*, 129–191. https://doi.org/10.1016/0010-0277(94)00659-9

Keizer, E. (2007). The English Noun Phrase. In *Studies in English Language*. Cambridge University Press. https://doi.org/10.1017/CBO9780511627699

Kintsch, W. (1972). Abstract nouns: Imagery versus lexical complexity. *Journal of Verbal Learning and Verbal Behavior*, *11*(1), 59–65. https://doi.org/10.1016/S0022-5371(72)80060-4

Knief, U., & Forstmeier, W. (2021). Violating the normality assumption may be the lesser of two evils. *Behavior Research Methods*, *53*, 2576–2590. https://doi.org/10.3758/s13428-021-01587-5/Published

Knuth, D. E. (1968). The art of computer programming / Vol.1, Fundamental algorithms / Donald E. Knuth. In *Fundamental algorithms*. Reading, Mass .

Kuznetsova, A., Brockhoff, P. B., & Christensen, R. H. B. (2017). lmerTest Package: Tests in Linear Mixed Effects Models. *Journal of Statistical Software*, *82*(13), 1–26. https://doi.org/10.18637/jss.v082.i13

LaPolla, R. J. (2003). Why languages differ: variation in the conventionalisation of constraints on inference. *Language Variation: Papers on Variation and Change in the Sinosphere and in the Indosphere in Honour of James A. Matisoff*, 113–144.

Lepic, R. (2019). A usage-based alternative to “lexicalization” in sign language linguistics. *Glossa: A Journal of General Linguistics*, *4*(1), 1–30. https://doi.org/10.5334/gjgl.840

Liddell, S. K. (2003). *Grammar, gesture, and meaning in American Sign Language*. Cambridge University Press. https://doi.org/10.1017/CBO9780511615054

Liddell, S. K., & Metzger, M. (1998). Gesture in sign language discourse. *Journal of Pragmatics*, *30*(6), 657–697. https://doi.org/10.1016/S0378-2166(98)00061-7

Liddell, S. K., Vogt-Svendsen, M., & Bergman, B. (2007). A crosslinguistic comparison of buoys. Evidence from American, Norwegian, and Swedish Sign Language. In M. Vermeerbergen, L. Leeson, & O. Crasborn (Eds.), *Simultaneity in signed languages: form and function* (pp. 187–215). John Benjamins. https://doi.org/10.1075/cilt.281.09lid

Linzen, T., & Oseki, Y. (2018). The reliability of acceptability judgments across languages. *Glossa: A Journal of General Linguistics*, *3*(1). https://doi.org/10.5334/gjgl.528

Lohndal, T. (2020). Syntactic Categorization of Roots. In *Oxford Research Encyclopedia of Linguistics*. Oxford University Press. https://doi.org/10.1093/acrefore/9780199384655.013.257

Lucas, C. (Ed.). (2001). *The sociolinguistics of sign languages*. Cambridge University Press. https://doi.org/10.1017/CBO9780511612824

Lucas, C., Bayley, R., Valli, C., Rose, M., & Wulf, A. (2001). Sociolinguistic variation. In C. Lucas (Ed.), *The Sociolinguistics of Sign Languages* (pp. 61–111). Cambridge University Press. https://doi.org/DOI: 10.1017/CBO9780511612824.006

MacLaughlin, D. (1997). *The Structure of Determiner Phrases: Evidence from American Sign Language* [Boston University]. http://louis-xiv.bu.edu/pub/asl/disserts/MacL97.pdf

Maiden, M., & Robustelli, C. (2000). *A Reference Grammar of Italian*. NTC Publishing.

Mantovan, L. (2017). *Nominal Modification in Italian Sign Language*. Walter de Gruyter Inc., Boston/Berlin and Ishara Press, Preston, UK. https://doi.org/10.1515/9781501504853

Martin, A., Adger, D., Abels, K., Kanampiu, P., & Culbertson, J. (2024). A Universal Cognitive Bias in Word Order: Evidence From Speakers Whose Language Goes Against It. *Psychological Science*, *35*(3), 304–311. https://doi.org/10.1177/09567976231222836

Martin, A., Holtz, A., Abels, K., Adger, D., & Culbertson, J. (2020). Experimental evidence for the influence of structure and meaning on linear order in the noun phrase. *Glossa: A Journal of General Linguistics*, *5*(1), 97. https://doi.org/10.5334/gjgl.1085

Martin, A., Ratitamkul, T., Abels, K., Adger, D., & Culbertson, J. (2019). Cross-linguistic evidence for cognitive universals in the noun phrase. *Linguistics Vanguard*, *5*(1), 1–13. https://doi.org/10.1515/lingvan-2018-0072

Martínez-Insua, A. E., & Pérez-Guerra, J. (2011). An open-sesame approach to English noun phrases: Defining the NP (with an introduction to the special issue). *English Language and Linguistics*, *15*(2), 201–221. https://doi.org/10.1017/S1360674311000013

Marty, P., Chemla, E., & Sprouse, J. (2020). The effect of three basic task features on the sensitivity of acceptability judgment tasks. *Glossa*, *5*(1), 72. https://doi.org/10.5334/GJGL.980

McBurney, S. (2012). 38. History of sign languages and sign language linguistics. In R. Pfau, M. Steinbach, & B. Woll (Eds.), *Sign Language: An International Handbook* (pp. 909–948). De Gruyter Mouton. https://doi.org/doi:10.1515/9783110261325.909

McEnery, T., & Hardie, A. (2012). *Corpus linguistics : method, theory and practice* (A. Hardie, Ed.). Cambridge University Press. https://doi.org/https://doi.org/10.1017/CBO9780511981395

Medeiros, D. P. (2018). ULTRA: Universal Grammar as a Universal Parser. *Frontiers in Psychology*, *9*, 1–14. https://doi.org/10.3389/fpsyg.2018.00155

Meier, R. P. (2009). Why different, why the same? Explaining effects and non-effects of modality upon linguistic structure in sign and speech. In *Modality and Structure in Signed and Spoken Languages* (pp. 1–26). Cambridge University Press. https://doi.org/10.1017/cbo9780511486777.001

Meir, I. (2012). Word classes and word formation. In *Sign Language - An international handbook* (pp. 77–111). De Gruyter Mouton. https://doi.org/10.1515/9783110261325.77

Mitchell, R. E., & Karchmer, M. A. (2004). Chasing the mythical ten percent: Parental hearing status of deaf and hard of hearing students in the United States. *Sign Language Studies*, *4*(2), 138–163. https://doi.org/10.1353/sls.2004.0005

Moriarty, E., & Kusters, A. (2021). Deaf cosmopolitanism: calibrating as a moral process. *International Journal of Multilingualism*, *18*(2), 285–302. https://doi.org/10.1080/14790718.2021.1889561

Nadolske, M. A., & Rosenstock, R. (2007). Occurrence of mouthings in American Sign Language: A preliminary study. In P. M. Perniss, R. Pfau, & M. Steinbach (Eds.), *Visible variation* (pp. 35–61). De Gruyter Mouton.

Neeleman, A. (2024). Syntax and Externalization II: Linearization. In *The Cambridge Handbook of Comparative Syntax*.

Neidle, C., Kegl, J., MacLaughlin, D., Bahan, B., & Lee, R. (2000). *The Syntax of American Sign Language*. MIT Press.

Neidle, C., & Nash, J. (2012). The noun phrase. In R. Pfau, M. Steinbach, & B. Woll (Eds.), *Sign Language - An International Handbook*. De Gruyter Mouton. https://doi.org/10.1515/9783110261325.265

Nuyts, J. (2012). Cognitive Linguistics and Functional Linguistics. In D. Geeraerts & H. Cuyckens (Eds.), *The Oxford Handbook of Cognitive Linguistics*. Oxford University Press. https://doi.org/10.1093/oxfordhb/9780199738632.013.0020

Office for National Statistics. (2013). *2011 Census: Ethnic group, local authorities in the United Kingdom*. https://webarchive.nationalarchives.gov.uk/ukgwa/20160105160709/http://www.ons.gov.uk/ons/rel/census/2011-census/key-statistics-and-quick-statistics-for-local-authorities-in-the-united-kingdom---part-1/rft-ks201uk.xls

Office for National Statistics. (2016). *Regions (Former GORs)*.

Oomen, M., & Pfau, R. (2017). Signing not (or not): A typological perspective on standard negation in Sign Language of the Netherlands. *Linguistic Typology*, *21*(1), 1–51. https://doi.org/10.1515/lingty-2017-0001

Ortega, G., Schiefner, A., & Özyürek, A. (2019). Hearing non-signers use their gestures to predict iconic form-meaning mappings at first exposure to signs. *Cognition*, *191*. https://doi.org/10.1016/j.cognition.2019.06.008

Prince, A., & Smolensky, P. (2004). *Optimality Theory: Constraint Interaction in Generative Grammar*. Blackwell Publishing Ltd. https://doi.org/10.1002/9780470756171.ch1

Proctor, H., & Cormier, K. (2023). Sociolinguistic Variation in Mouthings in British Sign Language (BSL): A Corpus-Based Study. *Language and Speech*, *66*(2), 412–441. https://doi.org/10.1177/00238309221107002

Pullum, G. K. (2017). Theory, data, and the epistemology of syntax. In M. Konopka & A. Wöllstein (Eds.), *Grammatische Variation: Empirische Zugänge und theorische Modellierung* (pp. 283–298). De Gruyter Mouton. https://doi.org/10.1515/9783110518214-016

Quer, J., Cecchetto, C., Donati, C., Geraci, C., Kelepir, M., Pfau, R., & Steinbach, M. (Eds.). (2017). *SignGram Blueprint, A Guide to Sign Language Grammar Writing*. De Gruyter Mouton. https://doi.org/10.1515/9781501511806

R Core Team. (2024). *R: A Language and Environment for Statistical Computing*. https://www.r-project.org/

Rentelis, R. (2011). *British Sign Language mouthings from the north and south of the UK*. [Unpublished master's thesis]. University of Bristol.

Ribera-Llonc, E., Espinal, M. T., & Quer, J. (2019). The noun-verb distinction in Catalan Sign Language. *Sign Language & Linguistics*, *22*(1), 1–43. https://doi.org/10.1075/sll.00027.rib

Rijkhoff, J. (2002). *The Noun Phrase*. Oxford University Press. https://doi.org/10.1093/acprof:oso/9780198237822.001.0001

Rodd, J. M. (2024). Moving experimental psychology online: How to obtain high quality data when we can't see our participants. *Journal of Memory and Language*, *134*. https://doi.org/10.1016/j.jml.2023.104472

Rowley, K., & Cormier, K. (2024). Shorter but richer versus longer with less information: Linguistic differentiation between British Sign Language and sign supported English. *Multilingua*. https://doi.org/10.1515/multi-2023-0177

Rubio-Fernandez, P., Wienholz, A., Ballard, C. M., Kirby, S., & Lieberman, A. M. (2022). Adjective position and referential efficiency in American Sign Language: Effects of adjective semantics, sign type and age of sign exposure. *Journal of Memory and Language*, *126*. https://doi.org/10.1016/j.jml.2022.104348

Rutkowski, P., Kuder, A., Czajkowska-Kisil, M., & Łacheta, J. (2015). The Structure of Nominal Constructions in Polish Sign Language (PJM): A Corpus-based. *Studies in Polish Linguistics*,

10(1), 1–15. http://www.plm.uw.edu.pl/sites/default/files2/Rutkowski et al Structure of Nominal Constructions in Polish Sign Language.pdf

Schachter, P., & Shopen, T. (2007). Parts-of-speech systems. In T. Shopen (Ed.), *Language Typology and Syntactic Description: Volume 1: Clause Structure* (2nd ed., Vol. 1, pp. 1–60). Cambridge University Press. https://doi.org/DOI: 10.1017/CBO9780511619427.001

Schembri, A. (2024, April 12). *Native signer ideologies in Auslan and BSL linguistics [Conference presentation]*. 1st ROLE Collective Symposium, University of Michigan, Ann Arbor and the University of Iowa. https://www.youtube.com/watch?v=kCqlkdL7_fk

Schembri, A., Cormier, K., Johnston, T., McKee, D., McKee, R., & Woll, B. (2010). Sociolinguistic variation in British, Australian and New Zealand Sign Languages. In *Sign Languages* (pp. 476–498). Cambridge University Press. https://doi.org/10.1017/CBO9780511712203.022

Schembri, A., Fenlon, J., Rentelis, R., Reynolds, S., & Cormier, K. (2013). Building the British Sign Language Corpus. *Language Documentation & Conservation*, *7*, 136–154.

Schembri, A., Rowley, K., & Leeson, L. (2024). Sign Languages in Britain and Ireland. In *Language in Britain and Ireland* (pp. 439–452). Cambridge University Press. https://doi.org/10.1017/9781108769617.028

Schütze, C. T., & Sprouse, J. (2014). Judgment data. In D. Sharma & R. J. Podesva (Eds.), *Research Methods in Linguistics* (pp. 27–50). Cambridge University Press. https://doi.org/10.1017/CBO9781139013734.004

Schwager, W., & Zeshan, U. (2008). Word classes in sign languages: Criteria and classifications. *Studies in Language*, *32*(3), 509–545. https://doi.org/10.1075/sl.32.3.03sch

Siewierska, A. (1997). *Constituent order in the languages of Europe: Eurotyp 20-1*. De Gruyter Mouton. https://doi.org/10.1515/9783110812206

Signature. (2022). *Qualification Specification: Level 2 Certificate in British Sign Language*. https://www.signature.org.uk/wp-content/uploads/2022/10/BSL2-2022.pdf

Stamp, R. (2013). *Sociolinguistic variation, language change and contact in the British Sign Language (BSL) lexicon (Part 1)*. [Unpublished doctoral thesis]. University College London.

Stamp, R., Schembri, A., Fenlon, J., Rentelis, R., Woll, B., & Cormier, K. (2014). Lexical variation and change in British Sign Language. *PLoS ONE*, *9*(4). https://doi.org/10.1371/journal.pone.0094053

Steddy, S., & Samek-Lodovici, V. (2011). On the Ungrammaticality of Remnant Movement in the Derivation of Greenberg's Universal 20. *Linguistic Inquiry - LINGUIST INQ*, *42*, 445–469. https://doi.org/10.1162/LING_a_00053

Steedman, M. (2020). A formal universal of natural language grammar. *Language*, *96*(3), 618–660. https://doi.org/10.1353/lan.2020.0043

Supalla, T., & Newport, E. (1978). How many seats in a chair? The derivation of nouns and verbs in American Sign Language. In P. Siple (Ed.), *Understanding Language Through Sign Language Research* (pp. 91–132).

Sutton-Spence, R. (1999). The influence of English on British Sign Language. *International Journal of Bilingualism*, *3*(4), 309–449. https://doi.org/10.1177/13670069990030040401

Sutton-Spence, R., & Woll, B. (1999). *The linguistics of British Sign Language: An introduction*. Cambridge : Cambridge University Press. https://doi.org/10.2307/417194

Sutton-Spence, R., Woll, B., & Allsop, L. (1990). Variation and recent change in fingerspelling in British Sign Language. *Language Variation and Change*, *2*(03), 313. https://doi.org/10.1017/S0954394500000399

Tang, G., & Sze, F. (2002). Nominal expressions in Hong Kong Sign Language: Does modality make a difference? In R. P. Meier, K. Cormier, & D. Quinto-Pozos (Eds.), *Modality and Structure in Signed and Spoken Languages* (pp. 296–321). Cambridge University Press. https://doi.org/10.1017/CBO9780511486777.015

Thompson, S. A., & Ono, T. (2020). Chapter 1. Introduction. In T. Ono & S. A. Thompson (Eds.), *The "Noun Phrase" across Languages: An emergent unit in interaction* (pp. 1–8). John Benjamins Publishing Company. https://doi.org/doi:10.1075/tsl.128.01tho

Traverse. (2019). *Research into the Deaf audience in the UK A review of evidence*. https://lumotv.co.uk/application/files/5316/2878/0058/Research_into_Deaf_audience_-_2019_update.pdf

UK Government. (2023, December 23). *GCSE British Sign Language (BSL)*. https://www.gov.uk/government/publications/gcse-british-sign-language-bsl

Van Valin, R. D., & LaPolla, R. J. (1997). *Syntax: Structure, Meaning, and Function*. Cambridge University Press. https://doi.org/10.1017/CBO9781139166799

Vaux, B., & Cooper, J. (1999). *Introduction to Linguistic Field Methods*. Lincom Europa. https://depts.washington.edu/uwcl/el-stec/Vaux_Cooper_1999_ch1.pdf

Vermeerbergen, M., Leeson, L., & Crasborn, O. (2007). Simultaneity in Signed Languages.: A String of Sequentially Organised Issues. In M. Vermeerbergen, L. Lorraine, & O. Crasborn (Eds.), *Simultaneity in signed languages: form and function* (pp. 1–25). John Benjamins. https://doi.org/10.1075/cilt.281

Wilson, M., & Hoong Sin, C. (2015). *Research into the Deaf audience in the UK A review of evidence Final report to the British Sign Language Broadcasting Trust*. https://www.bslzone.co.uk/application/files/4914/5320/1727/OPM_report_-_Research_into_the_Deaf_Audience_Jan_2016.pdf

Woll, B. (2007). Perspectives on linearity and simultaneity. In M. Vermeerbergen, L. Leeson, & O. Crasborn (Eds.), *Simultaneity in signed languages: form and function* (pp. 337–344). Amsterdam : John Benjamins. https://doi.org/10.1075/cilt.281.15wol

Zeman, D., Nivre, J., & Abrams, M. (2020). *Universal Dependencies 2.7*. LINDAT/CLARIAH-CZ digital library at the Institute of Formal and Applied Linguistics (ÚFAL), Faculty of Mathematics and Physics, Charles University. http://hdl.handle.net/11234/1-3424

Zhang, N. N. (2007). Universal 20 and Taiwan Sign Language. *Sign Language & Linguistics*, *10*(1), 55–81. https://doi.org/10.1075/sll.10.1.05zha

Appendix A Tiers Used for the NP Project

A.1 Introduction

Table A.1 lists all the tiers used in this project, along with the controlled vocabulary items defined for those tiers. Tiers and controlled vocabulary (*CV*) items that have been added or amended for this project are shown in italics. Table A.2 includes information about what is recorded on the *IDgloss* tiers. The *RH-GramCls, LH-GramCls* and *MouthingGrCl* tiers' controlled vocabulary items are based on those originally defined for the Auslan Corpus: more details can be found in Johnston et al. (2019).

Table A.1 Tiers Used for NP Project

Tier name	Definition
Controlled vocabulary items	
RH-IDgloss	Documents all manual activity of the right hand. See Table A.2 for details.
LH-IDgloss	As RH-IDgloss but for the left hand
RH-GramCls	Grammatical category of manual activity of the right hand.
NorV	May be a noun or a verb – not enough information to say
VP	Plain verb
VILoc	Indicating/locatable verb
VIDir	Indicating/directional verb
VD	Depicting verb
Aux	Auxiliary
N	*Noun. (Override the controlled vocabulary as we do not distinguish between plain, locatable and depicting nouns in this project)*
Det	Determiner point
Loc	Locative point
Pro	Pronominal point (or other pronominal sign such as SELF)
Loc/Pro	Point that has both a locative and a pronominal meaning
Det/Loc	Point that has both a determiner and a locative meaning

Tier name	Definition
Controlled vocabulary items	
Det/Loc/Pro	Point that is ambiguous between a determiner, a locative and a pronominal meaning
Adv	Adverb
Adj	Adjective
Num	Numeral
Quantifier	*Denotes quantity. Includes universal quantifiers such as* ALL, BOTH, *existential quantifiers such as* SOME, ANY *and distributive quantifiers such as* EACH, EVERY
Backchannel	Signal from addressee that they are attending to the discourse
DM	Discourse marker
Interact	Interactive. Expresses attitude or emotion. Not part of a larger grammatical construction. Exclamation, interjection, attention-getting
Salutation	Greeting or leaving-taking
Fragment	Incomplete or aborted utterance. Is not related to surrounding elements
Title	Title of a person or thing
Neg	Negator
Wh-ProQ	Pronominal question sign, for example, WHERE
Wh-Rel	Question sign used non-interrogatively, for example, a relative pronoun.
Conj	Conjunction.
Buoy	Handshape produced to anchor the discourse. See Table A.2.
Prep	Preposition.
Unsure	There is doubt as to the category of the sign.
LH-GramCls	As RH-GramCls but for the left hand.
Mouthing	Word (or part of a word) from the surrounding spoken language (English). See Section 2.2.2.4.
MouthingGrCl	Grammatical category of mouthing. Based on RH-GramCls and LH-GramCls tiers.
As RH-GramCls / LH-GramCls	With the following amendments:
NP	*Noun Plain. This value is used for all mouthed nouns for this project – Nloc (Noun Locatable) and ND (Noun Depicting) are not used.*

Tier name	Definition
Controlled vocabulary items	
Dem	*Demonstrative, such as "this", "that", "these", "those" – needed for mouthings although not required for manual grammatical category tiers (see Section 1.3.4.2.2).*
MouthingCongruence	
congruent	*Mouthed word is a keyword in BSL Signbank for the target manual sign*
not congruent	*Mouthed word is not a keyword in BSL Signbank for the target manual sign*
disambiguating	*Mouthed word disambiguates the target manual sign*
NA	*Corresponding mouthing is regressive or progressive, or else has been annotated as INDECIPHERABLE on the Mouthing tier*
unsure	*There is doubt about the congruence of this mouthing*
FreeTransl	Free translation of an utterance into English
LitTransl	Literal translation of a CLU into English. May not use English grammar. Gives an indication of what is explicitly conveyed and what is implicit in the utterance
Phrase	*Delimits a phrase. Currently just used for noun phrases but could be expanded in future*
NP	*Noun phrase (or, includes Noun phrase: see Section 2.2.2.2).*
unsure	*Where there is doubt.*
ClauseLikeUnit(CLU)	Delimits a single CLU (see Section 2.2.2.2). Field may contain a code relating to the participant plus a sequential number of the CLU within the file.
RH-Arg	Right-hand Argument. Records whether the manual sign produced by the right hand is a predicate, argument of a predicate or a non-argument.
V	Sign is a predicate. There is only one predicate in this CLU.
V1, V2 etc	Sign is a predicate and there are multiple predicates in this CLU. Number indicates sequential order of predicates within the CLU.
A	Sign is an argument of a predicate. There is only one argument in this CLU.

Tier name	Definition
Controlled vocabulary items	
A1, A2 etc	Sign is an argument of a predicate and there are multiple arguments in this CLU. Number indicates sequential order of predicates within the CLU.
nonA	Sign is neither a predicate nor an argument of a predicate.
Indefinite	May be either a predicate or an argument of a predicate: both analyses are possible.
Indeterminate	Sign (or series of signs) is difficult to analyse in terms of argument structure.
LH-Arg	As RH-Arg but for the left hand.

Note. New values are shown in italics.

A.2 Coding in the IDgloss Fields

The *RH-IDgloss* and *LH-IDgloss* tiers are used to record the activity of the signer's right and left hand, respectively. Table A.2 shows the labels used for different types of activity in the BSL Corpus (Cormier et al., 2017).

Table A.2 Coding of Types of Sign on the IDgloss Tiers in the BSL Corpus

Type of sign	**Corpus Label**	**More Information**
Lexical sign	ID gloss found in BSL Signbank	An ID gloss represents all of the morphological variants of a sign. It does not encode the sign's grammatical category; instead, this is recorded on the *GramCls* tiers – see Table A.1. Some ID glosses (e.g., PARTY02) have a numeric suffix; these indicate lexical variations with the same or similar meaning. Compound signs (e.g. FS:G-GRAPHIC^ART for 'graphic designer') are annotated separated by a caret.
Pointing sign	Depending on the referent of the point; • Pronominal: PT:PRO followed by a number indicating 1st, 2nd or 3rd person, then SG or PL signifying singular or plural referents; • Locative: PT:LOC; • Determining: PT:DET; • or a combination of these where there is more than one possibility, e.g., PT:LOC/PT:PRO1SG	See Section 1.3.4.2.2.

Type of sign	Corpus Label	More Information
Buoys	List buoys: LBUOY	See Section 1.3.4.2.2 for list buoys. This and other types of buoy are discussed in Liddell et al. (2007).
Depicting sign	Prefixed with DS:	See Section 1.3.4.2.1.
Fingerspelling	Prefixed with FS:	Use of a manual version of the Roman alphabet to produce words borrowed from English. See Section 1.3.4.2.
Sign name	Prefixed with SN:	Members of the signing community may be assigned a specific personal name sign, to identify them as an individual. See Section 2.2.2.4.4.
Gesture	Prefixed with G:	Bodily actions that are communicative but that lack the linguistic structure and conventionalisation of signs (Ortega et al., 2019). See Section 1.3.4.2.
Constructed action	Prefixed with G:CA:	A type of gesture, where the signer enacts some aspect of the state of affairs that they wish to describe (Liddell & Metzger, 1998). See Section 2.2.2.2.

Appendix B BSL Sentence Task: English Text – Project Web Page

In this section, we provide a copy of the English text in the web pages that formed our online judgement task.

B.1 Introduction

The BSL Sentence Task is part of a research project about the grammar of BSL. This project is led by the **UCL Deafness Cognition and Language (DCAL) Research Centre**.

Do you want to participate? You must be:

- Deaf
- Born in the United Kingdom
- 18 years or older
- Use BSL as your main language

If you decide to participate, you will complete the task online. This will take about one hour. You can stop for a break at any time.

You will see many videos of BSL sentences. For example:

[Note: 2 videos of example sentences in BSL are shown: CHEF BAKE CAKE and CHEF CAKE BAKE]

Do you think that both of these sentences are typical of BSL? Does one appear more typical than the other? We want you to watch lots of sentences and tell us what you think!

If you participate and complete the task, you will be paid £15 for your time.

Are you interested? Please email [email address[30]], "I'd like to take part in the BSL Sentence Task" and we will follow up with you from there.

B.2 Next Steps

Once we confirm that you have the right background to take part, we will send you an email with the subject **"Invitation to BSL Sentence Task"** within two working days. This will contain a link to the survey web page. **(Check your spam/junk email if you cannot find your invitation.)**

This invitation link is personal to you — please do not share it with others. If you know others who would like to participate, please send them a link to this information page instead: **https://bslcorpusproject.org/projects/bsl-syntax-project/bsl-sentence-task/**. We will then send them their own personal survey link.

It is best to do the task on a laptop or desktop computer. You can also use a tablet, such as an iPad. Please do not use a mobile phone.

Your answers will be saved after each page. You do not have to finish all the questions in one go. You can take a break at any time. If you want a short break, just leave the browser window open. If you need to stop the task and start again later, you can close the browser window or shut the

[30] Email address has been redacted in this appendix.

computer down. When you are ready to start again, click on your personal link in your email "Invitation to BSL Sentence Task" and you will be taken to where you stopped.

Once you complete the task, you will receive a voucher for £15. This can be used with many retailers, including Amazon, Argos and Marks and Spencer. You will receive an email at the end of the task explaining how to claim your voucher. You can choose which retailer or retailers to use it with.

B.3 BSL Sentence Task team

Lead Researcher: **Kearsy Cormier**

Researchers collecting data: Heidi Proctor, Gabrielle Hodge, Neil Fox, Matt Brown from UCL, and Adam Schembri from University of Birmingham

Main contact: [email address]

B.4 Information Sheet

BSL Sentence Task

Lead Researcher: Kearsy Cormier ([email address[31]]).

Researchers collecting data: Heidi Proctor, Gabrielle Hodge, Neil Fox, Matt Brown from UCL, and Adam Schembri from University of Birmingham.

Ethics approval: LING-2019-10-07 (approved by the Ethics Chair for the Linguistics Research Department).

Deafness Cognition and Language Research Centre / University College London.

1. What is the project? What will happen if I take part?

You are invited to participate in a project about the grammar of BSL. You will see many videos of a BSL signer producing two BSL sentences. You will rate how typical you think each one is of natural everyday BSL. It will take about one hour.

Once you finish, we will ask you some questions about you, e.g., your age and gender. Your answers will be recorded. They will help us understand more about BSL grammar.

2. Why am I invited? Do I have to participate?

You are invited because you are deaf, born in the UK, and use BSL as your main language. We expect that approximately 100 people will participate (adults only, no children).

It is up to you if you want to participate or not. If you decide to participate, we will ask you to read this information or watch it in BSL. We will then ask for your consent before starting the task.

You can talk about this project with other people if you want, or contact the researchers if you have any questions or if anything is unclear.

There is no penalty if you do not participate. There is no penalty if you begin the survey and later decide not to continue. You can withdraw at any time by emailing: [email address].

[31] Email address has been redacted in this appendix.

3. How will I be paid?

We will pay you when you complete the task. You will receive a £15 E-Gift Card via email. You can use this with participating retailers such as Amazon, Argos, Asda, Marks & Spencer and many others.

When you complete the task, you will receive a link via email. Click on this link to go to the E-Gift Card reward page. You can choose how much you wish to use with specific retailers. You will then receive your E-Gift Card via email.

B.5 Data Protection Privacy Notice

The data controller for this project is University College London (UCL). The UCL Data Protection Office provides oversight of UCL activities involving the processing of personal data, and can be contacted at data-protection@ucl.ac.uk.

Your personal data will be processed for the purposes outlined in this notice. For more detail, see UCL's general privacy notice for research.

Data protection legislation including the General Data Protection Regulation (GDPR) requires us to have a valid legal reason to process and use personal data about you. This is often called a 'legal basis'. GDPR requires us to clearly explain the legal basis that we rely on in order to process information about you.

The legal basis that will be used to process your personal data (your background information) is 'Public task'. This means for general benefit of the public.

The legal basis that will be used to process your special category data (sensitive background information like your deafness) is 'Research purposes'.

Your personal data will be processed so long as it is required for the research project. Your survey responses will not be stored with your name. Your responses will only be stored with your ID code to keep your personal data confidential.

B.6 Consent

If you have any questions before you start you can email [email address]. If you need to close your browser or your computer, you can restart the task by clicking on your personal link in the email 'Invitation to BSL Sentence Task'.

If you want to take part in this study, please select **ALL** of the following boxes. This means you confirm you agree. It also means you confirm that you have the background that we need.

Then **click on the arrow below to continue:**

- ☐ I am deaf.
- ☐ I was born in the United Kingdom.
- ☐ I am over 18 years old.

- [ ] I use BSL as my main language.
- [ ] I have seen the information above (in English or BSL) and I understand it.
- [ ] I understand I can email [email address] with any questions.
- [ ] I understand my consent is the legal basis for processing any personal information about me.
- [ ] I understand that all personal information will remain confidential. Any data gathered in this study will be stored securely with ID code, not my name. It will not be possible to identify me in any publications.
- [ ] I understand that my anonymised research data may be shared with, and used by, others for future research. I understand that no one will be able to identify me when these data are shared.
- [ ] I understand that I am free to withdraw from the study without penalty. If I want to withdraw, I can email if I so wish by emailing [email address].
- [ ] I consent to take part in the study.

B.7 Instructions (English Version)

Part 1

On each page, you will see a pair of videos showing BSL sentences, one video above the other.

For each video, we will ask you to use the slider underneath to indicate if you think that the sentence you see is typical BSL. Consider if the signing is typical or not typical of the deaf BSL signing community.

If you think the video is very typical for a fluent deaf BSL signer, move the slider to the far right.

If you think the video is not at all typical for a fluent deaf BSL signer, move the slider to the far left. Or if you don't understand the video, move the slider to the far left.

If you think the video is somewhat typical of BSL, move the slider to the middle.
(See BSL video above for examples.)

Sometimes there will be a drawing above the videos, like this:

When there is a drawing, the sentence in each video needs to match the drawing.

We will now show you some practice questions. Click on the arrow at the bottom of the page to move to the next page. Click here and your answer will be saved. You can only click to move forward to the next page. You cannot go back to a previous page.

Click on the arrow below to see the practice questions.

Practice Questions

[Note: Three pairs of practice questions are presented. Their order is not randomised. Their glosses are as follows:

BOY EAT CAKE

CAKE BOY EAT

[PIRATE]BR[32] LOVE CHEF

[CHEF]BR PIRATE LOVE

MAN RISE SCUBA-DIVER

RISE MAN SCUBA-DIVER]

Part 2

Practice questions finished!

There is a yellow progress bar at the top of the screen showing you how far through the task you are.

You do not have to finish all the questions in one go. You can take a break at any time. If you need to take a short break, just leave the browser window open. If you need to stop the task and start again later, you can close the browser window or shut the computer down. When you are ready to start again, click on your personal link in the email "Invitation to BSL Sentence Task" and you will be taken to where you stopped.

This task should take around one hour to complete. Once you complete the whole task, you will receive a voucher for £15. You can use your voucher with well-known retailers such as Amazon, Argos and Marks and Spencer. We will send you an email at the end of the task explaining how to claim your voucher.

Click on the arrow below to start the task.

B.8 Task Questions

[Note: Task questions are presented.]

B.9 Demographic Questions

Thank you for completing our task! Now we have a few questions about you.

First, what year were you born?

[32] BR indicates *brow raise* i.e. raised eyebrows. See Section 1.3.4.4.

How old were you when you started to use BSL?

- ○ 0-7 years
- ○ 8-12 years
- ○ 13-18 years
- ○ 18 years or older

Who in your family is (or was) deaf or hard of hearing? Please select as many as you need to:

- ☐ Mother / your primary caregiver
- ☐ Father / your secondary caregiver
- ☐ Grandmother (your mother's mother)
- ☐ Grandfather (your mother's father)
- ☐ Grandmother (your father's mother)
- ☐ Grandfather (your father's father)

Do you have any deaf brothers or sisters who use BSL?

- ○ Yes
- ○ No

[Note: The following question is displayed only if the answer to the previous question is “Yes”:]

How many deaf brothers and sisters who use BSL:

- ○ Deaf older brothers? _____
- ○ Deaf younger brothers? _____
- ○ Deaf older sisters? _____
- ○ Deaf younger sisters? _____

Have you ever taught BSL classes?

- ○ Yes
- ○ No

[Note: The following two questions are displayed only if the answer to the previous question is "Yes":]

How long have you been a BSL teacher?

- ○ 0-5 years
- ○ 6-10 years
- ○ More than 10 years

Have you ever had training to teach BSL?

- ○ Yes
- ○ No

Do you use speech and/or lipreading as your main form of communication with hearing people in everyday situations?

- ○ Yes
- ○ No

On a scale of 1 to 7, how comfortable are you with English? (1 is very uncomfortable, 7 is very comfortable).

- ○ 1 - Very Uncomfortable
- ○ 2
- ○ 3
- ○ 4
- ○ 5
- ○ 6
- ○ 7 - Very Comfortable

What is your gender?

- ○ Female
- ○ Male
- ○ Non-binary
- ○ Other
- ○ I prefer not to say

Which region of the UK do you live in now?

- ○ Scotland
- ○ North East England
- ○ Northern Ireland
- ○ North West England
- ○ Yorkshire and the Humber
- ○ East Midlands
- ○ Wales
- ○ West Midlands
- ○ East of England
- ○ South-East England
- ○ South West England
- ○ London
- ○ Outside of the UK

Which other regions of the UK have you lived in during the **past ten years**? Please select as many as necessary.

- ☐ Scotland
- ☐ North East England
- ☐ Northern Ireland
- ☐ North West England

- [ ] Yorkshire and the Humber
- [ ] East Midlands
- [ ] Wales
- [ ] West Midlands
- [ ] East of England
- [ ] South-East England
- [ ] South West England
- [ ] London
- [ ] Outside of the UK
- [ ] No other regions

What is your background?

- ○ White
- ○ Asian / Asian British
- ○ Black / African / Caribbean / Black British
- ○ Mixed / multiple ethnic groups
- ○ Other ethnic group

B.10 Other Questions

Did you enjoy this task? Are you interested in taking part in other DCAL studies? Select "Yes" below to get involved - it's quick and easy! Later, we will ask you to answer a few more questions and we'll then be able to match you with future studies you may find interesting.

- ○ Yes - please send me information about how to take part in future studies
- ○ No - please don't send me information about how to take part in future studies

[Note: The following question is displayed only if the answer to the previous question is "Yes":]

Enter the email address that we should use to contact you:

_____

Is there anything you would like to tell us about this survey? For example, did you have any technical problems or interruptions which may have affected your answers?

And do you have any other feedback about the sentences you saw?

_____

_____

_____

_____

_____

B.11 Debrief

You have now completed the task!

We can now tell you more about our study:

The aim of this project is to study the grammar of BSL, e.g., how BSL sentences are structured. We wanted to know more about BSL sign order and nonmanual features such as head shake and facial expression. We asked for ratings for a range of different sign orders and nonmanual features.

This included many structures which are rare or do not occur at all in the BSL Corpus. The ratings you provided will help us study what sign orders and nonmanuals are likely or unlikely, and whether this differs according to people's background (e.g., age, region, gender or teaching experience).

We will compare this data with information from conversations and narratives in the BSL Corpus. Together, this information will help BSL teachers and BSL learners, BSL interpreters and BSL interpreter trainers.

Please check your inbox for an email about your Tango Card - this will tell you how to claim your reward voucher.

Thank you for taking part! You can now close your browser window.