

What's in a sign?

When form features have meaning

Annika Schiefner (s4620437)

Research Master Language and Communication

Supervision:

Onno Crasborn & Mark Dingemanse

Master thesis

Course code - LET-REMA-LC1407

Nijmegen, 2019

Contents

1	Introduction	1
2	Theoretical Background	2
2.1	Iconicity	2
2.2	Sign language phonology	6
2.3	What makes a morpheme?	10
2.4	Interim summary	12
3	The development of psycholinguistic studies	13
3.1	Small factorial designs in the study of word frequency effects	14
3.2	Big data in the study of word frequency effects	16
3.3	Interim summary	19
4	The present study	20
4.1	Research Questions	21
5	Methods	22
5.1	Subjects	22
5.2	Design	23
5.3	Instruments	23
5.4	Procedure	25
5.5	Statistical analysis	26
6	Results	27
6.1	RQ1 - How do form-meaning units affect language processing in the interpretation of existing NGT signs and pseudosigns?	27
6.2	RQ2 - Do fluent users of NGT abstract to underlying form-meaning units in a perception task? (NGT signs)	30
6.3	RQ3 - Do underlying form-meaning associations provide cues towards the interpretation of novel signs?(Pseudosigns)	32
7	Discussion	34
8	Conclusion	39
9	References	40
A	Scripts and raw data	46
B	List of NGT signs and their possible translations	46

C Linguistic Background Questionnaire	49
C.1 Long version	49
C.2 Short version (WDD, Leeuwarden; Deaf club, Groningen)	55

List of Tables

1	FMUs in the present study	21
2	NGT Signs and search terms	25
3	Generalised linear mixed model for response accuracy across all items . . .	28
4	Generalised linear mixed model for response accuracy across NGT signs . .	30
5	Generalised linear mixed model for response accuracy across pseudosigns .	32
6	Examples of other form meaning associations	35

List of Figures

1	The dependency model	9
2	Relationship between subjective and count-based frequency scores	17
3	Word frequency effect in Dutch, English and French	17
4	Timeline of items in the experiment	23
5	Screenshot of an example item: DRAAIEN	24
6	Response patterns per participant by phase	29
7	Response patterns per item by FMU in NGT signs	31
8	Response patterns per item by FMU in pseudosigns	33

Abstract

Sixty years after the first studies described the structure of signs in phonological terms and thereby established sign languages as natural languages in their own right, the debate on how to best describe the structure of signs is still ongoing. The present study aims to contribute to this debate by subjecting the idea that signs are morphologically complex units that combine low-level form-meaning units in a behavioural test. Four such form-meaning units were selected for an experimental task. 25 adult native users of Sign Language of the Netherlands (NGT) were recruited to participate in a binary forced-choice task, in which they were asked to match Dutch words with signs from NGT and phonologically valid pseudosigns.

The results are in line with approaches that posit sub-lexical form-meaning units and suggest that at least two of the form-meaning units tested are valid candidates for stable associations. Participants performed better on NGT signs, implying a role for lexical processing that exceeds the mere meaning potential contributed by the FMUs and available to signers in the interpretation of novel signs.

1 Introduction

The role of iconicity in language has long been downplayed but has relatively recently seen a strong upsurge in interest. Linguists looking at a variety of languages, both spoken and signed, have contributed to a better understanding of what iconicity can contribute to models of language processing and production. One of the unresolved issues in this debate is the integration of iconic processes into theories of linguistic structure. At what level do iconic processes act and to what kinds of linguistic processing are they relevant?

The present study aims to contribute to this debate by providing evidence that iconic processes are at work at the sub-lexical level in Sign Language of the Netherlands, henceforth NGT. It studies whether native signers of NGT perceive and make use of systematic links between certain form features of NGT signs and abstract meanings that contribute to the lexical meaning of the sign in the sense of declarative schemas at a morphological level (Jackendoff & Audring, 2016). Following Van der Kooij (2002), I call these systematically occurring units form-meaning units (FMUs). Section 2 describes the theoretical background in more detail, laying the foundation for the present study.

Past studies on such FMUs in the NGT lexicon have described their presence and function based on extensive corpus analyses and observations from novel word production studies. However, no study has yet attempted to provide evidence as to their role in processing signs in a behavioural setting where signers are not explicitly instructed to access their linguistic creativity nor made aware of the presence of systematic links to guide their task performance. The present study attempts to close this gap by contributing an experimental approach tapping into linguistic processing.

This means that the present study is subject to all the pitfalls of exploratory research. Section 3 therefore illustrates how psycholinguistic studies are shaped by extensive exploration of the possibility space over time using the example of word frequency effects. While the effects of word frequency in linguistic processing are not of immediate relevance to the present study, the development of measures to accurately capture them in comprehension and production studies and across different languages clearly illustrates the process of scientific progress in a particular sub domain of linguistics. The present study shares many features with early research in word frequency effects, such as a relatively small number of stimulus items and participants, as well as a lack of understanding of the statistical properties of many aspects of the lexicon of NGT. As with written language corpora, sign language corpora are increasingly becoming available and will in future become increasingly representative of language use in different settings and across different groups of language users.

2 Theoretical Background

The present section gives an overview of the current research into the three domains that provide the theoretical backdrop for the present study. I start by defining iconicity and exploring its effects across languages and its role in linguistic theory. Subsequently, two levels of linguistic analysis, phonology and morphology are explored with a focus on the Phonological Dependency Model and the concept of FMUs, which form the basis for the experimental study reported on in Section 4.

2.1 Iconicity

Iconicity, the non-arbitrary mapping of meaning to form, has long been discounted as a potential feature of language. Instead, linguistic symbols were long seen as fully arbitrary, with words that carry meaning by convention within a community of speakers and which are composed of smaller phonemic units that do not carry meaning themselves (De Saussure, 1916; Locke, 1690). However, other philosophers of language have argued that iconicity is a second important feature of language and that speakers of all languages are sensitive to iconic relationships between the world around them and the linguistic items that refer to it (Köhler, 1929). This is certainly the case for mimetic iconicity, such as onomatopoeia in which words mimic the natural sounds they represent. Words like *cock-a-doodle-doo* (EN), *kikeriki* (DE) are subject to the phonological rules of the language in question while bearing a likeness with the sounds they stand for.

Modern linguists have learned to recognise that iconicity in language goes further than such onomatopoeic utterances. While the view that linguistic symbols are fundamentally arbitrary still has a strong hold on mainstream linguistics, many researchers have started challenging that view and describing different aspects of motivated form-meaning mappings across languages and levels of linguistic processing. On the lexical level, ideophones, words which represent all sorts of real world experiences, have been found to exist

in many languages across the world, mostly in Africa, Asia, and the Americas (Dingemanse, 2018). Additionally, iconicity has been recognised to exist at different levels of linguistic structure, including phonotactics and syntax. Vowel length is frequently used to mark durative aspects of an event, mono-syllabicity is associated with unitary events, while repeated structures tend to represent repeated events, concepts that belong together semantically or temporally are usually represented in close succession in sentence structures and narratives, etc. (Perniss, Thompson, & Vigliocco, 2010, pp. 2-3).

Most studies into spoken language iconicity investigate sensitivity to form-meaning mappings at the word level across and within a variety of languages with maximum-contrast studies, testing participants' perception on pseudowords (see for example Bremner et al., 2013; Kovic, Plunkett, & Westermann, 2010; Köhler, 1929; A. Nielsen & Rendall, 2011; 2012; Ramachandran & Hubbard, 2001; Sapir, 1929). Such studies are mostly motivated by a search for universals in human (linguistic) processing. Work on existing words, often antonym pairs, mostly comes forth from a similar motivation, showing cross-language sensitivity (Brackbill & Little, 1957; Dingemanse, Schuerman, Reinisch, Tufvesson, & Mitterer, 2016; Gebels, 1969; Imai, Kita, Nagumo, & Okada, 2008), facilitation in word learning tasks (Lockwood, Dingemanse, & Hagoort, 2016; Lockwood, Hagoort, & Dingemanse, 2016; A. Nielsen & Rendall, 2012; A. K. S. Nielsen & Rendall, 2013; Nygaard, Cook, & Namy, 2009), and processing effects (Bergen, 2004; Lockwood, Hagoort, & Dingemanse, 2016; Occelli, Esposito, Venuti, Arduino, & Zampini, 2013; Shintell, Nusbaum, & Okrent, 2006; Westbury, 2005). In such studies, participants are shown ideophones, naturally occurring words that include non-arbitrary mappings between form and meaning, from languages they are not familiar with and asked to guess the meanings of those words in forced choice paradigms.

In one such study, Dingemanse et al. (2016) show that Dutch participants are able to guess the meanings of such ideophones at above chance levels, even when foils are translations of other ideophones from the same language and semantic domain. While their participants performed well below participants in earlier studies using antonym pairs or even pseudowords designed to maximise phonetic contrast, they show that iconic mappings can be universally accessible and semantically specified. This does not negate the possibility of language specific iconic strategies, or an increased sensitivity to such mappings in speakers of languages that routinely use them as productive morphological strategies. They also show that iconicity is influenced by both segmental and supra-segmental information and that unimodal mappings, in this case mappings of natural sounds on the sound structures of words, are more transparent than those across modalities. This also implies that different modalities have different iconic affordances and may be subject to different iconic strategies. Finally, they point out that the reduced performance on the items in their experiment suggests that natural words combine both iconic and arbitrary features.

In sign language, sign level iconicity has been shown to play a role in poetry and word play, suggesting that signers are conscious of iconic links between form and meaning in their language (Sutton-Spence, 2005). In tasks stimulating less conscious processing strategies, iconicity has been found to affect similarity judgments (Vigliocco, Vinson, Woolfe, Dye, & Woll, 2005) and online-processing, even when the task did not require access to a sign's meaning (Grote & Linz, 2003; Thompson, Vinson, & Vigliocco, 2010; Vinson, Cormier, Denmark, Schembri, & Vigliocco, 2008). Iconicity has also been found to facilitate sign recognition in children (Ormel, Knoors, Hermans, & Verhoeven, 2009), early emergence of meaningful handshape distinctions (Slobin et al., 2003), early emergence of agreement marking (Casey, 2003), and sign recall in hearing children (Brown, 1980).

Orlansky and Bonvillian (1984), however, showed that the earliest signs a child learns are not more iconic than those learned later, and Meir et al. showed that iconic signs are not less prone to errors in articulation, including errors that obscure iconicity (Meier, 1982; Meier, Mauk, Cheek, & Moreland, 2008). Such findings from early childhood may be explained by the fact that iconicity relies on world knowledge. In order to understand links between forms and concepts, the signer needs to understand the basis for those links and needs to be able to abstract away from the real world concept to a representation within the linguistic and motoric constraints on sign formation. For example, in order to understand the link between the form of the NGT sign MILK and its real world referent, one must know that milk comes from cows and is obtained by milking that cow. It is not necessary to have any first-hand experience with that process, but without at least a conceptual understanding the link between form and meaning remains obscure. Additionally, linking form and meaning requires the cognitive ability to abstract away from real-world experiences to reduced representations.

Tolar, Lederberg, Gokhale, and Tomasello (2008) showed a developmental pattern in 3-year-old children which suggests a cognitive shift towards being able to explicitly recognise iconicity as a bridge between form and concept. While this may explain null effects for iconicity in very young signers, it leaves open the question why spoken language studies find earlier effects. It may well be that the type of mapping plays a role, e.g. the level of abstraction, and that spoken language correspondences tap into systematicity, a phenomenon which even very young infants are known to be able to make use of (Walker et al., 2010).

Most of the studies described above study the effects of iconicity at the word or sign level. Iconicity is often taken to be a holistic characteristic of the word or sign. This leaves open questions on which elements in signs or words carry the information that iconicity taps into and whether there may be different types of iconicity, for example in the representation of actions, shapes, or both. However, some research suggests that those different types of iconic representations may well have differential effects on learning and processing (Perniss et al., 2010, p. 10).

As early as 1977, Mandel stated that “an adequate account of (American) sign language must include the fact that the form various elements take in the language depends in part on the visual appearance of their referents” (p. 57, cited in Perniss et al., 2010). This does not mean that all features of a sign need to be iconic. However, most descriptions of iconicity in signs recognise that signs tend to combine arbitrary and iconic features. Iconicity and arbitrariness are consequently sometimes described as extremes on a continuum rather than categorically distinct notions (Perniss et al., 2010, p. 4).

In such views, signs that are so close in form to their referents that their meaning can be guessed based on the form are called transparent, while signs in which the relationship is only apparent to those who know the meaning of the sign are called translucent (Schermer, 2016). Both notions are culturally dependent, in the sense that they depict culturally recognisable features of a concept. For example, the sign for “tea” in many western sign languages can be described to depict someone holding a teabag and moving it up and down in a cup. The sign can only be iconic for people who commonly encounter tea in teabags. People from cultures that tend to brew tea from loose tea leaves would be unlikely to find the sign described above iconic.

According to Taub (2001b), iconic signs represent a selection of the salient form features of a mental image. This selection will always be subject to constraints such as salience, choice of representational strategy, and affordances of the articulators. Additionally, they are subject to language internal constraints, at the level of phonology and phonotactics, as well as morphosyntax and preferences for certain strategies in the linguistic system.

Similarly, Occhino, Anible, Wilkinson, and Morford (2017) stresses that iconicity is always construed by the individual and is therefore never an objective feature of any form, be it at a lexical or segmental level. In such an approach, the perception of a form-meaning relationship is based on an individual’s cultural and linguistic experience and cannot be truly universal. However, in my view, this is of theoretical importance only, since human perception of the world is always filtered by their perceptual systems, which are in turn developed in interaction with the sensory feedback with the world the individual encounters. Just because processes are subject to individual construction in processing, it is not, in my view, impossible to objectively measure them.

It can be both true that iconicity, just like arbitrary form-meaning links, is grounded in personal, linguistic, and cultural experience, and possible to objectively measure perceptions of iconicity across populations that are controlled for sharing or not-sharing particular aspects of those experiences or within individual language users. For example, comparing the perception of iconicity between signers of a specific sign language and non-signers from the same country implies a shared cultural background due to shared nationality and contact with the majority culture, combined with diverging linguistic experiences. As with any experimental approach, researchers will attempt to minimise the

effects of differences in personal experiences through careful experimental design and the use of appropriate statistical tests or will make these individual differences the subject of their study, trying to identify what factors lead to those individual differences.

Taking the notion that iconicity is ultimately the relationship between a construed form and concept mapping at face value integrates the perception of iconicity as a mapping of visual form to visual concept with the notion of metaphoric iconicity. If any mappings are abstractions from true experience, it is easy to include conceptual metaphor as a domain that is subject to iconic form-meaning mappings. According to Brennan (2005), such iconicity, including metaphor, is grounded in the visual world and based on metaphorically motivated gestures that are integrated into the linguistic system. All iconicity in this account is influenced by and reflective of the signing community culture.

If iconicity is the mapping of meaning to form, understanding it requires defining what we mean by form and meaning and how they are systematically connected in languages, particularly in sign languages. Therefore, the following subsections will give an account of sign language phonology and morphology.

2.2 Sign language phonology

Traditionally, phonemes are described as the smallest segments abstracted from fluent speech with potentially semantically distinctive features. They are not the smallest units of phonetic description, since their concrete realisation may depend on the phonetic context or phonotactic position, or may be subject to within- and across-speaker variation. Phonemes can be described as bundles of distinctive acoustic and/or articulatory features (cf. Phonem, *Lexikon der Sprachwissenschaft*, 2008). By transferring these concepts to the visuo-manual modality, phonemes in sign language can then be described as the smallest segments with potentially semantically distinctive function that can be abstracted from fluent signing and that consist of bundles of distinctive visual and/or articulatory features. Phonemes are thus abstract form units that can be combined to create meaningful linguistic units. Changes at this level may change the grammatical or semantic content of the resulting meaningful unit, while the units themselves are traditionally seen as not carrying meaning themselves.

Stokoe (1960) was the first to formalise such form units for a sign language, namely American Sign Language (ASL), laying the foundation for a perception of sign languages as “real” languages and as comparable to spoken languages in many regards. By employing the concept of minimal pairs, words that are distinguished by only a single form feature, to sign language, he established a rough distinction of phonological units: *tabula* (tab), *designator* (dez), and *signation* (sig). Those terms correspond to the later more common terms *location*, *handshape*, and *movement* respectively. For example, the NGT signs SUMMER and STRANGE are distinguished only by the location of the articulator (Nijen Twilhaar & Van den Bogaerde, 2016, p. 120). This formalisation not only allows

researchers to describe existing signs, but also to establish what signs are possible and impossible in a given language. Each of the units mentioned above has a fixed set of possible expressions.

Unlike Stokoe, who treats phonological units as entirely arbitrary, Friedman (1976) explicitly included iconicity in her phonological account of ASL. However, her approach has the shortcoming of being very specific to the individual item with no way of systematising signs or phonological structures. This means that there is no way to extract an inventory of phonemes for ASL from her descriptions, as no structured account of the phonological system is given. An example is the description of the ASL sign BASEBALL below:

BASEBALL: a metonymic mimic presentation of batting a ball:

A hand on A hand move outward from shoulder area, rapid form: wrists bend in neutral space, variations: speed, style of batting, hesitations, corresponding only to real world batter's action. (Friedman, 1976, p. 90)

Both systematicity and a structured account of the phonological system while accounting for iconicity at the phonological level are accounted for in the Dependency Model by Van der Kooij (2002) through linking individual phonological features to potential meanings through Semantic Implementation Rules.

According to Van der Kooij (2002), early research into sign language phonology due to its exploratory nature resulted in lists of phonetic features but provided little to no evidence to their phonological status (e.g. Brentari, 1990; 1998; Liddell & Johnson, 1989; Sandler, 1989). Following the tradition of De Saussure (1916), phonemes were seen as necessarily meaningless and their combination in morphemes as arbitrary and conventionalised. Research into classifier handshapes was the exception to this rule. Those were considered to be iconic with the interaction of phonological rules of the respective system and their iconicity described in great detail. Additionally, Brennan (1990) discovered that symbolic locations, metaphoric handshape changes, and movements can be equally meaningful and are used for sign formation both on the spot and in the coinage of new signs in the community. According to Van der Kooij (2002), classifier handshapes may be less special than is often claimed. Instead, similar mappings may exist across the whole lexicon of sign languages.

One difficulty in establishing what phonological units may be realisations of meaningful elements is that such units behave more like sound symbolism than like morphemes in many respects. Detaching the meaningful element from the unit does not result in two morphemes. Examples from spoken English are the onsets /gl/ and /tw/ which are associated with the meanings *shiny*, *sparkly* in glitter, glow, or glimmer, and with *duality* in two, twice, and twilight respectively. Additionally, those links are in themselves meaningful by convention, not necessarily due to some intrinsic properties of the forms.

Finally, research into spoken languages and especially their written form suggests a linear structure across all levels that does not easily fit the description of sign languages. Semantic motivations in signs can be componential rather than holistic despite their simultaneous realisation. For example, the classifier for round cylindrical objects in many sign languages is a C-hand, while a curved single finger represents a flat round object. The representation is thus at least two-dimensional, with the number of fingers for height and their curvedness for roundness.

Van der Kooij (2002) thus attempts to specify “formal sublexical contrastive element[s] in the underlying representation” (p. 25) based on two conditions: The feature or phonological construction needs to be *recurrent*, i.e. productive, in the lexicon and the information may not be *predictable* from the phonetic or semantic context. This precludes form elements that are obligatory in their context but only occur in a single lexical item and follows the general understanding in phonological theory that only non-redundant information is stored in the underlying phonological representation. Phonetic Implementation Rules govern the concrete phonetic realisation of those phonological specifications in a given context. An example of such a rule is the thumb position if aperture is specified, taking the value [open] or [close]: In such configurations, the thumb is opposed to the selected fingers (Van der Kooij, 2002, p. 115).

Besides articulatory affordances specified in Phonetic Implementation Rules, Van der Kooij (2002) proposes that phonological objects may be linked to semantically motivated form elements through Semantic Implementation Rules. This introduces iconicity as a property of the phonological system, rather than a holistic property of signs. Following Taub (2001a), iconicity is thus tied to individual elements in a meaning-preserving analogous mapping. This implies a fixed set of formal features that can be used to encode a concrete or abstract/metaphorical mental image. Van der Kooij (2002) calls the “representable parts of the image that are encoded by form elements [...] Meaning Aspects” (p. 30).

Figure 1 shows the phonological elements specified in the Dependency Model and indicates what values they can take. Given a specific context, those elements can be associated with specific semantic information. Additionally, their specific phonetic realisation can be specified by semantic constraints of the referent. Three of the parameters in the model (marked in italics) may change within monomorphemic signs while all other parameters are specified for the duration of the whole sign. The parameters that are subject to changes give rise to the following types of movement: path movements, i.e. changes in setting, changes in aperture, and changes in orientation.

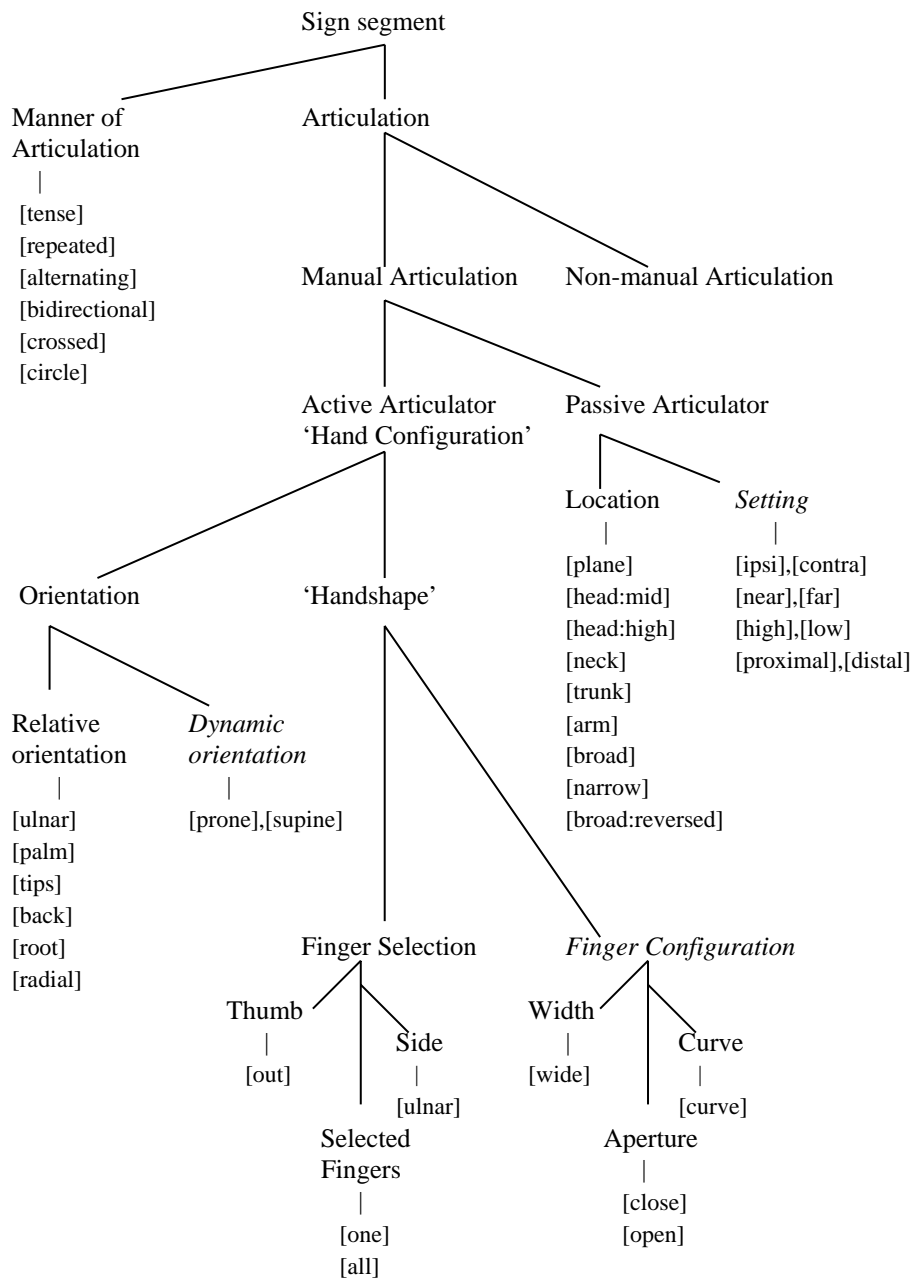


Figure 1: The dependency model with form specifications (Van der Kooij, 2002, p. 280, Fig. 5.43)

2.3 What makes a morpheme?

The Sachwörterbuch zur Sprachwissenschaft (2003) defines morphemes as the smallest, meaning-carrying units of a language, which are represented through an articulated form. Morphemes are theoretical units that are used for the scientific description of a language. A single morpheme can be realised through different elements, allomorphs, that share a meaning and appear in complementary distribution, and several morphemes can be realised in a single element. For example, { \emptyset , -e, -n, -s, -en, -er, vowel mutation, vowel mutation + -er, } are all allomorphs of the plural morpheme in German, while English *sang* combines meaning (to sing) and tense (simple past) in a single element (Morphem, Sachwörterbuch zur Sprachwissenschaft, 2003).

It follows that morphemes are defined by a systematic relationship between a specific form and meaning. Zwitserlood (2002) therefore argues that if identical structures with parallel meanings are used across different contexts, those structures should be interpreted in the same way and should therefore have the same status in both contexts. Even if those form-elements have become non-alterable in a given sign, these frozen forms still provide patterns for the formation of new, complex signs. Elements that allow for comprehensible new formations, should therefore be considered morphemic. The argument is thus ultimately based on the stability and interpretability of a given form-meaning relationship. If the relationship is stable across different contexts and types of signs, it should be considered morphemic across all those contexts and types. The morphological status should not be limited to only a subset of its uses.

If morphemes are characterised by a systematic relationship between form and meaning, they should also be productive in the formation of new signs. Language users should thus be able to form new signs and understand new signs, as long as they are produced by productive rules within the linguistics system. This is the case for NGT, where signers are able to create new signs with meaningful hand configurations and other meaningful components (Zwitserlood, 2003, p. 286). Examples of such recurring form-meaning mappings can also be found in other accounts of the structure of the sign language lexicon. For example, Brennan (2005) interprets metaphors and iconic mappings as “metaphorical morphemes” (p. 377), encoded at the sub-lexical level, such as the relative location of the two hands two each other representing hierarchical relationships or a hand-internal oscillating or fluctuating movement representing uncertainty. These morphemes are structured in metaphor sets and reoccur in sets of signs within semantic fields, essentially categorising these signs under a shared semantic component (p. 378).

Meir (2012) argues that signs may combine meaningless, i.e. phonological, as well as meaningful sub-lexical units. Depending on whether or not the sub-lexical components carry meaning, the resulting lexical items are considered to be morphologically simple or complex. The relations between forms and meanings may be iconic, though it remains unclear from Meir’s account, whether she considers arbitrary links between form and

meaning at the sub-lexical level possible. This is certainly the case in the concept of the ion-morph (Fernald & Napoli, 2000), which suggests the possibility of reoccurring form-meaning relationships in groups of words that are not motivated by transparent iconic mappings, as in the example of the shared handshape, orientation and movement in the ASL signs for *mother* and *father*, which they take to mean "parent", and the contrast between chin and forehead as locations in signs for female and male family members, respectively, across a range of kinship terms in ASL. All examples for sub-lexical form-meaning relationships provided by Meir (2012) are iconic, at least in the sense of a metaphoric mapping. Regardless of whether or not those relationships need to be iconic, the implication of such systematic links is that form overlap across signs is frequently associated with meaning overlap and that a sign's meaning may be, at least partially, componential and transparent from its building blocks.

Johnston and Ferrara (2012) formalise a continuum from partly lexicalised signs in which meaning is componential and constructed from relatively concrete and fully specified atomic units, to fully lexicalised signs in which word meaning is constructed and may be partly or completely unpredictable from its components. They argue that fully lexicalised "signs never lose their componential structure; the componential structure just becomes backgrounded to the lexical assertions" (p. 242). If required by context, signers are able to retrieve this literal composite meaning in a process that Johnston and Schembri (1999) call *de-lexification*. This dual structure is similar to the way in which spoken language multi word idioms, which can be interpreted as referring to their idiomatic meaning, or, given the appropriate context, as referring to their literal composite meaning. The likelihood of a componential reading will depend on how entrenched the idiomatic reading has become in language use (Johnston & Ferrara, 2012, p. 245). Johnston and Ferrara (2012) therefore argue that idiomatic expressions are not absent in sign languages, but are realised at the lexical level, with sub-lexical units contributing to the componential structure (p. 236). In this account, sub-lexical units are necessarily iconic in the componential structure, however, through processes of metaphoric mappings and associations that may be unknown to the present day user, the relationship between the idiomatic meaning and the componential meaning of a given fully lexicalised sign may have become completely obscured (p. 240 - 242).

Those sub-lexical units tend to be what is classified as a sign's parameters in the respective phonological model. In Meir's account, the relevant parameters are handshape, location, orientation and movement (2012). This raises the question of whether parameters of signs that can carry meaning should really be classified as morphemes. If they cannot be further decomposed phonologically, this appears to be problematic. However, single phonemes that function as morphemes are not unheard of in spoken language morphology, so according to a basic definition of morphemes as the smallest meaning-bearing

units of a language, such units would qualify to be described as phonemes on a form level and as morphemes that are instantiated by a single phoneme on a morphological level.

Jackendoff and Audring (2016) present a theory of the lexicon in which rules of grammar, both at a morphological and morphosyntactic level, are an integral part of the lexicon, stored as declarative schemas which encode relations between lexical items. Such schemas can be generative in that they are able to generate novel lexical items, and/or relational, in that they describe non-productive, prolific patterns across the lexicon. In such cases, the lexical meaning of the whole item need not necessarily be predictable from the schema. This can be illustrated easily by the way in which compounds often have lexical meanings that are related to but not completely predictable from their parts (cf. Downing, 1977). Rather than functioning as procedures creating output from input, these schemas primarily license occurrences that can be found in the utterances of language users. The lexicon thus contains words and schemas, the latter with variables instead of lexical bases. Both are interconnected in the mental lexicon through interface links, which allow for spreading of activation both via schemas and words. This represents a unified account of lexical morphology and morphosyntax, and posits a structured account of the lexicon including morphological elements and their affordances for combination (Jackendoff & Audring, 2016).

This approach is taken up by Van der Kooij and Zwitserlood (n.d.), who describe signs as morphologically complex structures that consist of meaningful form features or clusters of such form features. These structures function as morphological schemas in the sense of Jackendoff and Audring (2016) and are labelled FMUs, reflecting the two components they unite. These FMUs are part of the lexicon and can be combined to form complex lexical signs. As described above, their status as schemas implies that their relational function is their primary function, accompanied in some but not all FMUs by a generative function. The authors do not specify whether FMUs are necessarily iconic. They can certainly include different strategies of iconic mappings, including metaphoric and metonymic mappings.

The conceptualisation of FMUs as declarative schemas with a primarily relational and potentially generative function, makes important predictions for processing. Since items in the mental lexicon are connected via interface links both at the level of schemas and at the lexical-semantic level, spreading activation should be faster for such items where the two routes combine to strengthen a candidate interpretation (see Jackendoff & Audring, 2016, pp. 484-487, for examples).

2.4 Interim summary

This section has explored the notions of iconicity, phonology and morphology in sign languages. I showed that iconicity has been found to be relevant to phonological and se-

semantic processing and presented theories that aim to integrate iconic features in accounts of sign language phonology and morphology.

In the subsection on phonology in sign languages, I presented the Phonological Dependency Model, which accounts for iconic features at the phonological level and attempts to present an inventory of phonological units and formalise the phonetic and semantic constraints that account for their distribution in NGT. The inventory of phonological units as presented in Van der Kooij (2002) can be found in Figure 1.

In the subsection on morphology, I described different frameworks of morphology in sign languages, which account for sub-sign level iconicity in motivated signs. Labelled as FMUs by Van der Kooij and Zwitserlood (n.d.), these are theorised to be part of the lexicon as declarative schemas, in the sense of the framework proposed by Jackendoff and Audring (2016). This account predicts interconnectivity both via semantic or lexical links, as well as via the schemas as such.

The theories providing the background for the present study have, so far, not been extensively tested in experimental research. They have primarily been shown to be descriptively accurate in the context of linguistic analyses of various corpora of NGT. The present study thus aims to explore in how far they can be used to predict behaviour in psycholinguistic experiments. This implies that there is no prior experience as to how strong effects on processing will be or by what kind of experimental paradigm they can be accurately measured. The next section illustrates the challenges faced by such exploratory research, based on the example of word frequency effects in visual word processing studies.

3 The development of psycholinguistic studies

Psycholinguistic studies aim to establish whether theoretical models of language can be transferred to a behavioural level. In order to do so, they need to establish what predictions a theoretical model makes and how they translate to concrete behaviour. Such behaviours can subsequently be measured quantitatively or described qualitatively.

However, more often than not the translation of theoretical approaches to concrete behaviour is not entirely straightforward. While theories usually imply an effect's directionality, they rarely predict concrete effect sizes. It follows, that the task in exploring a new theory in psycholinguistic studies lies in the exploration of the possibility space of potential behaviours. What behaviours can be expected, how strong will effects be, and at what level of processing will a theory have measurable effects, are all questions that require empirical research.

The present chapter will illustrate these thoughts by reviewing the development of paradigms of word frequency effects on visual word recognition. Similarly to the development of a psycholinguistic approach to measuring the effects of FMUs in linguistic processing of NGT in this thesis, the study of factors affecting visual word recognition

initially faced the task of discovering which factors were important and how experiments needed to account for these factors. While word frequency as such is not of immediate relevance to this study, this chapter will show some parallels that will be important to keep in mind when interpreting the results from the present study and putting them into the context of various theoretical approaches.

3.1 Small factorial designs in the study of word frequency effects

The study of word frequency was already of interest in the first half of the 20th century and resulted in word lists for teachers. These lists included suggestions on age ranges at which words with certain frequencies were appropriate, as well as suggestions for the order in which words will become useful to second language learners (see for example: Thorndike, 1921; 1932; Thorndike & Lorge, 1944). These books were based on written text, extracted from the bible, children and adult literature, the United States Constitution and the Declaration of Independence, handbooks of everyday trades, and newspaper articles and personal correspondence (Thorndike, 1932, App. A). Thorndike and Lorge (1944) includes a total of 30,000 words, grouped by frequency and sorted alphabetically. Regular inflected forms, as well as rare derivations are listed under a single lexeme. Spelling variants are subsumed under the same lexeme, which is represented by standard American English spelling (p. ix-x). Word frequencies for such words with frequencies above 50 in a million words, were not specified.

1967, Kucera and Francis published basic statistics on words in American English, including automated frequency counts of about a million English words. Their counts were somewhat lower than those presented in Thorndike and Lorge (1944). The counts were based on the Brown Corpus, which was subsequently extended to include more specific information about the words through partly automated part-of-speech tagging and which remained an important corpus of American English. Unlike the Thorndike and Lorge (1944) corpus, the Brown corpus did not attempt to create a representative sample of English text as produced throughout a long period of time, but to capture present day use of American English. It was therefore restricted to English prose published in the US in 1961 (Francis & Kucera, 1964).

In 1969, Shapiro suggested that subjective frequency ratings might serve the same purpose as corpus based counts. In his study, he acquired relative frequency ratings from students from 6th and 9th grade, as well as from four adult groups with different professional and educational backgrounds. Those ratings turned out to be highly correlated the frequency counts established by Thorndike and Lorge (1944) and Kucera and Francis (1967), opening up the question of the best frequency measure.

These old studies show that two issues fundamental to the search for a good frequency measure. Ease of collection is a practical concern for anybody attempting to include word frequency in any experimental design. At the same time, different sampling techniques

imply differences in the word counts, due to differences in lexicon across different genres, as well as across time and space.

Since these first lists, word frequency has been used as one of the factors found to predict language processing in frameworks including visual word recognition and lexical decision tasks. Some other factors found to play a role in such frameworks are age of acquisition of individual items, word length, syllable count, phonological or orthographic neighbourhood density, or imageability (Balota, Yap, Hutchison, & Cortese, 2013, p. 94).

Most early studies conducted were based on small factorial designs, including 20 to 25 participants and 10 to 20 stimuli per condition. They would typically test the effects of one or two variables of interest on a certain outcome measure, often reaction times or accuracy scores. The lists of items for different conditions would be matched on other variables thought to potentially influence the results. The selection of those confounding variables was ultimately based on previous research and resulting researcher intuition. As factorial designs are unable to cope with predictors that are not categorical, or ideally binary, any scale would be reduced to a categorical variable, typically selecting items from the extremes of the scale to form the two groups (Balota et al., 2013; Brysbaert, 2018).

Word frequency would in such studies typically be included as a binary variable with high frequency (HF) and low frequency (LF) words. In a study by Ellis and Morrison (1998), LF words are defined as words that occur less than 10 times per million words and HF words as those that occur more frequently than 100 times per million words. Van Orden (1991) applies the same definition for LF words, but defines HF words as those that occur more than 55 times in a million words, based on word counts by Kucera and Francis (1967).

Such small factorial designs suffer from a number of problems. The reduction of scalar independent variables to binary categorical factors not only reduces the informative value of a study. It may also distort results through selection of items that may via other characteristics not controlled for influence response patterns. HF words, for example, are also semantically richer and more closely related to every day experiences of participants than LF words (Kuperman, 2013, p. 2). Researchers' predisposition towards the effects they expect to see in a certain framework can directly influence the selection of items. If researchers expect to see an effect of word frequency, they may select difficult, highly specialised LF words, while researchers who do not expect to see effects would be likely to select easy LF words (Brysbaert, 2018).

Additionally, small factorial designs are typically underpowered with regards to the expected effect size and are therefore not easily replicable. Brysbaert and Stevens (2018) suggest that reaction time experiments with repeated measures are replicable if they include at least 1,600 observations per condition. This translates to at least 40 participants and 40 items per condition. If this is true, experiments with approximately 20 participants and items per condition are strongly underpowered and unlikely to be replicable.

3.2 Big data in the study of word frequency effects

The solution to these issues requires sufficient amounts of data to allow for experimental designs using regression analysis, treating independent variables at the scale level appropriate to their nature. Large corpus projects combine large amounts of text with statistical information about the words included. Balota et al. (2007) presented the English lexicon project, which combines naming and lexical decision latencies from six research institutions collected from 1,260 participants for 40,481 words and non words (p. 446). Participant characteristics such as age, years of education and Shipley vocabulary age are included, as is information on their average reaction time, error rate, and percent outliers (p. 447). Item characteristics include several frequency measures, including Kucera and Francis (1967), information on orthographic neighbourhood, bigram frequency, phonological and morphological characteristics, and behavioural results from the lexical decision and naming studies (p. 449). Phonological and orthographic level information as well as behavioural results were also included for all non-words (p. 449).

Based on such data bases, it is possible to gain an overview of the actual frequency distribution of words across the lexicon. This translates to an ability to investigate the shape of the frequency curve to localise word frequency effects. Brysbaert et al. (2011) show that the frequency of words in the English lexicon is distributed across the whole range of frequencies, ranging from less than 0.1 in a million words to more than 10,000 in a million words and that subjective frequencies correlate very well with composite scores from three large corpora (see Figure 2).

Additionally, it becomes possible to show the relationship between word frequency and behavioural measures, such as lexical decision times, across the whole lexicon of a language. Keuleers, Diependaele, and Brysbaert (2010, Fig. 4) compares lexical decision reaction times for Dutch, English (Balota et al., 2007) and French (Ferrand et al., 2010) by word frequency, extracted from SUBTLEX-NL (Keuleers, Brysbaert, & New, 2010), SUBTLEX-US (Brysbaert & New, 2009) and Lexique 3.55 (Ferrand et al., 2010). Word frequencies ranged from 0.02 to almost 40,000 occurrences per million words. The graphs show that reaction times are highest for low frequency words and drop until reaching a plateau at about 1000 occurrences per million words and that this curve holds for all three languages investigated (see Figure 3).

The word frequency effect on lexical decision times is clearly strongest when comparing extremely low frequency words with words that occur 1000 times or more per million words. The long error bars on bins with very high word frequencies are due to the fact that there are very few items in those bins. Selections of words close to the 10 occurrences per million word cut-off used in earlier studies will thus lead to smaller effect sizes. Similarly, a cut-off for HF words at 55 per million words will lead to smaller effect sizes.

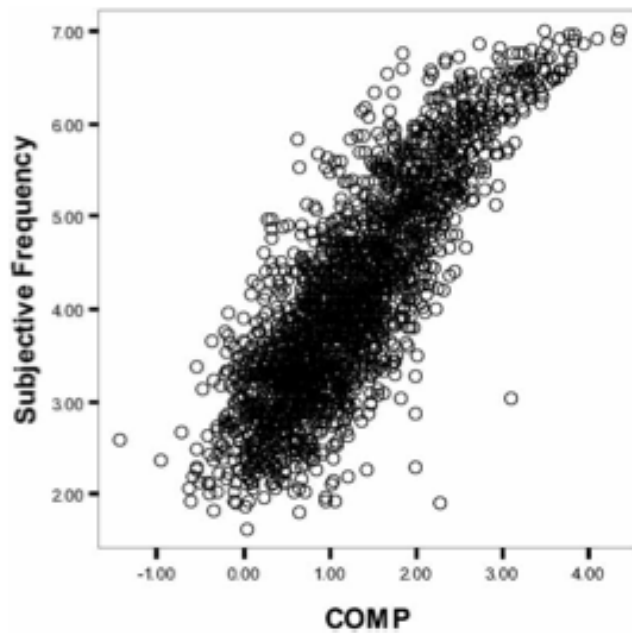


Figure 2: The relationship between the composite frequency measure, i.e. the mean of Zeno, SUBTLEX, and HAL and subjective frequency. Zeno = Zeno, Ivens, Millard, and Duvvuri (1995), SUBTLEX = Brysbaert and New (2009), HAL = Lund and Burgess (1996). The COMP frequencies are expressed in log₁₀ per million words, meaning that a value of -1 corresponds to 0.1 per million, 0 to 1 per million, and 2 to 100 per million. (Source: Brysbaert et al., 2011, Fig. 2)

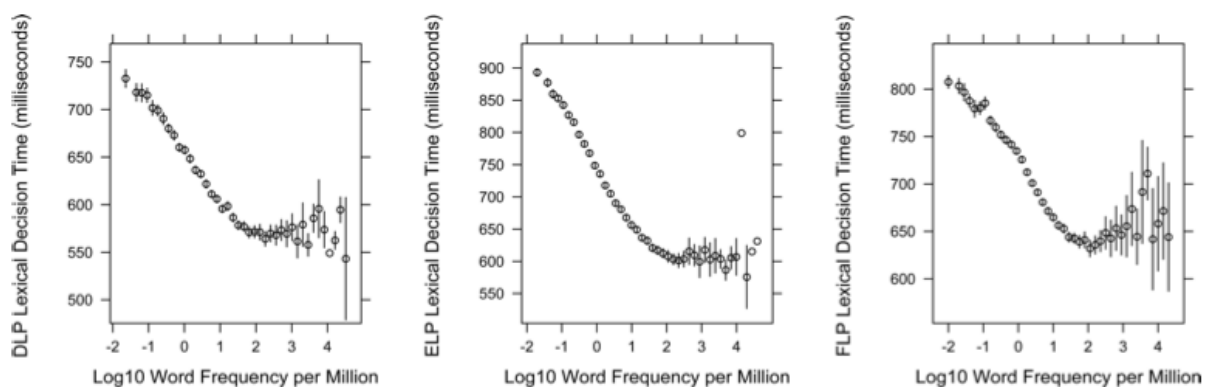


Figure 3: The relationship between reaction times in lexical decision tasks and word frequency in word stimuli from Dutch, English and French. Circles indicate the mean RT per bin of 0.15 log word-frequency; error bars indicate $2 \times SE$ (bins without error bars contained only one word). (Source: Keuleers, Diependaele, & Brysbaert, 2010, Fig. 4)

Word frequency has so far been reported as occurrences per million. The advantage of this measure over absolute word counts is its independence of the corpus size. However, it is a linear scale while the word frequency effect on lexical decision reaction times and accuracy in naming is logarithmic. The frequency effect curve in Figure 3 therefore uses \log_{10} occurrences per million to represent the effect.

This has a number of disadvantages, such as including negative values and being more difficult to interpret than a linear scale, since steps are not of equal size. For smaller corpora, with word counts close to a million items, it also suggests that the minimal possible frequency is one in a million words. This is not the case for larger corpora. As is clearly shown in Figure 3, nearly half of the frequency effect lies below the value of one in a million, necessitating the use of negative values to accurately capture the effect (Brysbaert, 2018; Van Heuven, Mandera, Keuleers, & Brysbaert, 2014).

In the SUBTLEX-UK, Van Heuven et al. (2014, pp. 8-12) therefore introduce the Zipf scale as a measure to make the word frequency effect easier to understand. The scale is logarithmic and covers a range from 1 (very low frequency) to 6 (very high frequency) or 7 for extremely frequent function words, etc. In line with a law on word frequency, developed by George Kingsley Zipf (1902–1950), it is based on \log_{10} frequency per billion words. It also includes the possibility of including a value of 0 for words not present in the corpus, by means of a Laplace transformation, which adds +1 to every absolute count, in order to avoid a logarithm of 0, which is minus infinity. This transformation implies that words not represented in a smaller corpus receive a higher Zipf value, leading to a stronger potential for bias in such data sources.

Regression analysis across large corpora allows for the investigation of factor structures, predicting certain behavioural measures. This means that the relative contribution of individual factors can be established. One such factor in the prediction of lexical decision reaction times and naming accuracy is word frequency. By comparing the contribution of different measures of word frequency to explain the variance of outcome measures, it is possible to assess their respective quality as a measurement of a theoretically important factor (Brysbaert, 2018).

Such comparisons have led to a much better understanding of word frequency effects over the years. Word frequency counts have been divided into methods that were lemma based, e.g. Kucera and Francis (1967) and Baayen, Piepenbrock, and Van Rijn (1993), and those that were based on surface word form (e.g. SUBTLEX-databases). More recent studies suggest that this does not make a difference, at least for word frequency measures based on large corpora (Baayen, Feldman, & Schreuder, 2006; Brysbaert & New, 2009).

Corpora for the computation of frequency counts draw on different textual materials. All tasks discussed so far include visual word recognition, i.e. stimuli are presented as written words. With increasing technical capacities, increasingly larger amounts of text, representing a variety of genres and language varieties were collected. Interestingly, the

most successful measure in the prediction of behavioural results in visual word recognition tasks so far have been corpora based on subtitles of movies and television series. This is noteworthy, because subtitles approximate spoken, rather than written language, and are rarely read by movie and television audiences. The word frequency effect thus seems to be stable across modalities (Brysbaert & New, 2009).

Quality of measurement seems to also depend on corpus size. Effects between low and medium frequencies were shown to be stronger in larger corpora, while there was no difference in high frequency words (Burgess & Livesay, 1998). However, this effect disappears in very large corpora with more than 16 to 30 million words (Brysbaert & New, 2009).

Finally, effects are contingent on recency of acquisition and lexicon size. Words that are newly acquired show an emerging effect and overlearning weakens the effect (Brysbaert, 2018). Subjects with a large lexicon, as established by vocabulary tests, are more likely to know low frequency words. This shifts the word frequency effect towards the lower end of the scale. This also accounts for differences between first- and second-language users disappearing after lexicon size is controlled for (Diependaele, Lemhöfer, & Brysbaert, 2013).

3.3 Interim summary

This description of the development of adequate word frequency measures to predict word processing in a variety of tasks illustrates a number of points that are relevant to the development of psycholinguistic studies in general. Small factorial designs were shown to be suitable for the exploration of effects in specific and controlled settings, provided a sufficiently high number of observations per stimulus, that can be used to discover interesting avenues for further investigation. However, an overview of the factor structure in the prediction of behavioural task results requires large amounts of data, preferably representative of the multimodal nature of language exposure.

It also shows that lexical processing is not driven by single factors, but depends on a variety of interdependent factors. Those factors are situated at different levels in the development of models. Some are characteristics of the stimuli, such as the number of syllables in a word or its frequency in a given corpus. Others are characteristics of the experiment, such as the size of the corpus from which items are drawn and the presentation mode of stimuli, e.g. auditory or visual. Finally, some are characteristics of the participants themselves, e.g. vocabulary knowledge and whether the language in question is their first language.

The surprising insight that the word frequency effect is multi-modal in nature, shows that processing effects may be driven by factors not previously considered. Certain factors not typically being included in experiments, therefore, does not necessarily mean that

those factors are less relevant. It could also indicate that nobody has yet tested the influence of a factor, because it was not theorised to be of any relevance.

For the present study, this suggests a number of important limitations. It may well be that important factors are not considered in the study presented in this thesis and that the possibility space explored for potential effects is not large enough or does not cover the full range of effects. Additionally, the small scale nature of this research recommends itself to exploratory research but does not justify strong claims about the complete structure of factors that are relevant in driving the effects in question.

4 The present study

The studies on FMUs in the NGT lexicon described above are based on an extensive corpus-analyses, as well as observations from a novel-word production study, yet it remains unclear whether those form-meaning relationships hold when signers are not explicitly asked to access their linguistic creativity and are not made aware what links may be the basis for a given task. The present study attempts to fill this gap, by contributing an experimental approach tapping into linguistic processing. This study attempts to tap into sub-conscious processing and does not involve an explanation of the underlying mechanisms to participants prior to or during testing.

If FMUs indeed function as declarative schemas, as proposed by Van der Kooij and Zwitserlood (n.d.), users of NGT should be able to use those schemas to interpret both known signs in novel contexts in which a lexical interpretation is unavailable and unknown signs. Based on the NGT data set in Global Signbank (Crasborn, Zwitserlood, Kooij, & Schüller, 2018) and a novel word elicitation task, Van der Kooij and Zwitserlood (n.d.) assembled a list of FMUs that are hypothesised to be available to fluent signers of NGT in linguistic processing. The form features in those units can be described according to the Dependency Model (Van der Kooij, 2002, see Figure 1). Every possible FMU represents an individual hypothesis within the general theoretical framework.

The form features selected for analysis in the present study are chosen to represent three distinct aspects and levels of analysis in the Dependency Model. The first is one particular type of movement, namely dynamic orientation realised as a *rotation* in the wrist. The second, relates to handshape and finger configuration, namely the presence of the feature *width*. The third is set at the level of manual articulation, looking at the coordination of the two hands, namely the presence of the feature *next-to*, where the hands are held or moved next to each other in neutral space. The last feature is not present in the dependency model as depicted in Figure 1, but has since been included in the model (Van der Kooij & Zwitserlood, unpublished toolkitpaper). All form features are associated with a specific meaning in the signs included in the study. Only the feature *width* is associated with two distinct meanings. All form and meaning features are described in Table 1.

Table 1: FMUs in the present study

Form feature	Description	Meaning	Description
Next-to	The two hands are situated next to each other	Same level	Two entities are situated at the same physical or hierarchical level
Rotation	The hand(s) change their orientation in- or outwards	Change of state	An entity changes from one state to another
Spread ¹	The selected fingers are spread	Individuation	Several separate entities of the same type
		Wideness	Abstract or physical wideness

Participants in the present study are presented with a sign including one of the FMUs, followed by a target and a foil in Dutch. The target represents a translation of a different sign which includes the same FMU as the one seen by the participant. The task is to select the word that represents a better translation of the sign, tapping into participants' sensitivity to the shared FMU in interpreting the sign.

In order to better understand the mechanism behind participants' performance, the experiment includes both existing lexical signs from NGT and phonologically possible pseudosigns. Since pseudosigns are not familiar to signers in the experiment, their interpretation can only rely on the form features present in the sign.

4.1 Research Questions

In order to test the general theory as well as the status of the individual FMUs described above, the following three research questions form the basis of this study:

RQ1: How do form-meaning units affect language processing in the interpretation of existing NGT signs and pseudosigns?

RQ2: Do fluent users of NGT abstract to underlying form-meaning units in a perception task? (NGT signs)

RQ3: Do underlying form-meaning associations provide cues towards the interpretation of novel signs? (Pseudosigns)

If participants are sensitive to the underlying FMUs and their potential contribution to the signs meaning, they would be expected to co-activate signs that share FMUs as well as

¹Note: Since pseudosigns do not have a fixed meaning, both individuation and wideness are combined as valid meaning components for pseudosigns with spread. A distinction is only made for established NGT signs.

those that have related lexical meanings and subsequently have an easier time retrieving the meanings of signs that share FMUs than those of signs that are only related via lexical semantics. However, since pseudosigns do not have a lexical meaning, co-activation of the target's translation equivalent in NGT would be expected to be weaker than for NGT signs for which both routes are available. On the other hand, FMU interpretation may only be available for unknown signs, as a creative strategy, but may be blocked for frozen lexical items, which have a fixed lexical meaning and may not allow for alternative interpretation. In that case, performance would be expected to be better on pseudosigns than NGT signs, with performance for NGT signs at chance level.

While the theoretical background makes no explicit predictions about the relative importance of the selected FMUs, results may reflect differences in salience or systematicity. These may be identical across NGT and pseudosigns, or differ. The latter would be indicative of differences in the mechanisms at work in the re-interpretation of existing signs in a context where the lexical meaning is unavailable and the interpretation of novel signs that have not been encountered previously.

The following sections give a detailed description of the methods used to answer these questions and the results of the experiment. Subsequently, the results are described and discussed in relation to the theoretical background outlined in the previous sections.

5 Methods

5.1 Subjects

For the present study, deaf adults were recruited, who reported using NGT as a regular means of communication and considered themselves native users of the language and who had no diagnosed language disorders. Three participants withdrew from the study. Their data are not included in this study.

Participants ($N = 25$) were between 18 and 54 years old (mean = 31.84). 12 participants were male, 13 female. 3 participants were left-, 22 right-handed. Participants reported a range of levels of education, ranging from HAVO to university level education (1 LEAO, 1 HAVO, 8 MBO, 11 HBO, 4 WO). About half of the participants reported having one or more deaf family members ($N = 13$). Participants were asked to self-assess their NGT and Dutch skills, in production and perception, both spoken and written for Dutch. Many participants chose not to give a score for Dutch in the auditory modality and only assessed their written production and reading comprehension. They rated their NGT skills between 7.5 and 10, with a mean of 9.27 on a ten-point scale. Dutch scores were somewhat lower ranging from 4 to 8.5 with a mean of 7.15.

All participants reported knowing other languages as well. Most common were English ($N = 23$) and International Sign ($N = 14$). Some participants also reported knowing specific other sign languages, in addition to, or instead of IS. All participants reported knowing Dutch, either as their first or as one of their other languages.

5.2 Design

The present study follows a within-subject design with the independent variables *type of sign* (NGT, pseudosign) and *FMU* (spread – individuation/wideness, next-to – same level, rotation – change of state) and the dependent variable *response-option selected* (Expected, non-expected). Item and subject were included as random effects in the statistical model.

5.3 Instruments

Main task The main task was designed and created by the author of this thesis, inspired by the methodology of tests of sensitivity to iconicity in ideophones (Dingemanse et al., 2016; Lockwood, Hagoort, & Dingemanse, 2016). It consisted of two parts, both of which followed the same setup but included different types of signs as stimuli. In the first part, participants were presented with existing NGT signs, familiar to fluent signers of NGT. The second part was composed of pseudosigns, based on the set of signs used in Ormel et al. (2016), available as non-signs in NGT data set (Ormel & Giezen, 2018) on Global Signbank (Crasborn et al., 2018).

Each test item consisted of a video of a sign or pseudosign, followed by a forced-choice selection with two options. Participants were asked to select the option that represented the better translation of the target sign (see Figure 4). The experiment was programmed and presented in OpenSesame, a Python-based, open-source and cross-platform programme (Mathôt, Schreij, & Theeuwes, 2012).



Figure 4: Timeline of items in the experiment

Stimulus selection All signs and pseudosigns were selected to include exactly one out of four FMUs under study. The translation options were selected randomly from a list

of all signs and their possible translations, as listed on Signbank and added by a native signer of NGT. For the target response, a random translation of a random sign sharing the FMU with the target sign was chosen, while the foil was a randomly chosen translation of a randomly chosen sign that does not share the FMU. For example, if the target sign was VERANDEREN (engl. *to change*, FMU = rotation), the target response could be *omdraaien* (engl. *to turn*, FMU = rotation) and the foil could be *parasol* (engl. *parasol*), which does not include the FMU *rotation* but one of the other FMUs, in this case *wideness*. In the case of signs including *individuation* and *wideness*, translations of the respective other category were excluded as foils to avoid confusion due to the form-overlap.

The target for existing signs was selected to not represent a valid translation of the stimulus sign. This was checked by automatically comparing the randomly selected option to the translations provided for the stimulus sign. It should be noted that this method may lead to occasional items where participants got a target that was a possible translation, which was not in the list. Each participant was presented with a different set of options in a different order, based on randomised lists. This should reduce the likelihood of item and position effects across participants (see Appendix A for the list generation scripts).

Existing NGT signs were retrieved from the NGT data set in Global Signbank (Crasborn et al., 2018), based on a search for phonological features potentially relevant to the FMUs (see Table 2). From this list of videos, those with translations that suggested the appropriate meaning-components were identified by the author and the lists were subsequently checked by fluent signers of NGT. The meaning components were *same level* for next-to ($N = 10$), *change of state* for rotation ($N = 10$), and *individuation* ($N = 10$) and *wideness* ($N = 9$) for spread. This resulted in a final list of 39 NGT signs (see Appendix B for both versions). Most videos were re-filmed to keep the signer and background constant across all videos. The background was set to a light grey and all signs were presented by a native signer wearing a black pullover or blouse (see Figure 5 for a screenshot).



Figure 5: Screenshot of an example item: DRAAIEN

For the pseudosigns, all items from Ormel and Giezen (2018) were categorised according to their form features. All signs that contained only one of the three form-components *next-to*, *rotation*, or *spread* were categorised accordingly. Subsequently, pseudosigns that

Table 2: NGT Signs and search terms

Form feature	Meaning component	Category	Search terms
Next-to	Same level	Handedness	2s
		Relation between articulators	Next-to
Rotation	Change of state	Orientation change	Pronation
			Pronation/Supination
			Rotation
Spread	Individuation/ Wideness	Handshape	Supination
			5
			4
			V
			W

could easily be changed to include only one of the form features were selected and categorised with the appropriate change. Finally, all selected signs were re-filmed in front of a green-screen to keep the signer and background constant across all videos. As for the NGT signs, the background was set to a light grey and all signs were presented by a native signer wearing a black pullover or blouse. The final selection of 56 signs included 17 items with *next-to*, 19 with *rotation*, and 20 with *spread*.

Linguistic Background Questionnaire The linguistic background of all participants was established using an adjusted version of the Language Background Questionnaire developed by Gullberg and Indefrey (2003). While the questionnaire contains very detailed questions about the participants’ background, only age, gender, hearing status, the highest level of education, and the combined self-assessment of NGT skills (production + perception) were used in the context of this study. The questionnaire was administered as a paper and pencil test at the end of the test session. For one participant, the self-rating of NGT skills is missing in the questionnaire. For participants tested outside Radboud University, a shortened version was administered to keep the duration of the experiments within a reasonable time frame (see Appendix C).

5.4 Procedure

The first 4 participants were tested on campus using a laptop running Debian Linux, with OpenSesame (Mathôt et al., 2012). The experiment was the first in a series of experiments from the same research group, which will not be discussed further in this thesis. Prior to participation, participants gave informed consent to participate, and

for the proposed analysis and storage of information. Participants received a financial compensation for their participation and were reimbursed for travelling expenses.

10 participants were tested at a festival at the occasion of the World Deaf Day in Leeuwarden, NL, using the same software and laptop. Participants here only filled in a strongly shortened version of the linguistic background questionnaire and were only tested on the perception task and a language test, which will not be reported upon here. As with those tested on campus, participants gave informed consent to participate, and for the proposed analysis and storage of information prior to participation. Participants at the festival were compensated for their participation with vouchers.

11 participants were tested at the deaf club in Groningen, using a different laptop with the same Linux distribution and experiment software installed. Like participants in Leeuwarden, participants in Groningen only filled in the shortened version of the linguistic background questionnaire. Some participants only participated in the perception task and a language test, while some participants also completed the other experiments mentioned above. Prior to participation, participants gave informed consent to participate, and for the proposed analysis and storage of information. Participants were compensated for their participation with vouchers according to the duration of their participation.

The experiment was presented in OpenSesame. Participants were shown instructions in NGT, followed by a shortened version in written Dutch. Subsequently, participants saw 39 NGT signs and 56 pseudosigns, with the order of each block randomised within block. Each sign was presented once followed by two options from which participants had to pick the one that was perceived as a better translation of the sign to Dutch.

Additionally, participants gave information on their linguistic background. The order of tests was changed between subjects in order to allow for multiple participants being tested at the same time, but participants always started by reading the test instructions, giving informed consent and filling in the linguistic background check.

After finishing all parts of the study, participants received a written debriefing for the study, which explained the research questions and hypotheses in Dutch. They were also invited to ask questions, which I tried to answer in sign or written Dutch. I also asked participants what they thought of the experiment. Participants regularly communicated that they did not understand the point of the task after completing the study, but were able to understand it in the debriefing procedure. This suggests that participants did not develop any conscious strategies that could influence the results.

5.5 Statistical analysis

The independent variables were experimental condition (NGT, pseudosign) and FMU with three levels for the overall analysis (next-to, rotation, spread) and four levels for the analysis within NGT signs (next-to, rotation, individuation, wideness). The dependent variable was participants' response accuracy in selecting the target response. All statistical

analyses were conducted with R 3.5.1 (R Core Team, 2018) in Rstudio 1.1.456 (Team, 2016).

To examine overall response patterns across items and participants, the response data was analysed using general mixed-effects regression with a logistic link (GLMER, binomial, lme4 version 1.1-18-1, Bates, Mächler, Bolker, & Walker, 2015). The best fitting model was established through backward selection. Within phase (NGT, pseudo), the same analysis was run to assess response patterns across FMUs. In addition to the two independent variables implemented as fixed effects, participant and item were included as random intercepts in the initial model. However, as with the fixed effects, the random effect structure was reduced if this significantly improved overall model fit (see Appendix A for a link to the scripts and raw data).

6 Results

The initial analyses served to establish that participants did not show any response bias. A single-sample t-test confirmed that participants did show any right- or left-bias in their response patterns ($t(363) = 0, p = 1$). This suggests that participants were not systematically picking responses on either side of the screen at above chance levels.

6.1 RQ1 - How do form-meaning units affect language processing in the interpretation of existing NGT signs and pseudosigns?

The model including Phase (NGT and pseudosigns) and FMU_3 (next-to, rotation, spread) as a fixed main effects and Gloss and Subject as random intercepts proved to have the best fit with the data. The relevant statistics and corresponding coefficients are reported in Table 3. Including the interaction term of Phase and FMU_3 resulted in a model with a marginally better fit. However, none of the interaction terms was significant, so the model was discarded for the more parsimonious option. The model with FMU including all five possible options (next-to, rotation, spread, wideness, and individuation) had a slightly worse fit than the one reported above. This makes sense, since spread and wideness/individuation are distributed across phases and can thus not contribute towards the explanation of variance in the respective other phase.

Participants performed above average for both FMU and Pseudosigns across FMUs (NGT: $t(103) = 13.897, p < .001$, 95% conf. interval [0.677, 0.735]; Pseudo: $t(77) = 7.447, p < .001$, 95% conf. interval [0.579, 0.637]). Figure 6 shows the distribution of expected responses per item by phase across all items and participants. It illustrates the overall effects of phase on the proportion of expected options selected.

NGT signs ($M = 0.709, SD = 0.434$) yielded the expected response at higher rates than pseudosigns ($M = 0.604, SD = 0.489, \beta = 0.563, SE = 0.116, p < .001$). Across conditions, signs including next-to yielded the highest accuracy scores ($M = 0.732, SD = 0.443$), followed by rotation ($M = 0.654, SD = 0.476$). This difference ($\beta = -0.367$,

Table 3: Estimates, standard error, and z-values of the main effects of the generalised linear mixed model assessing what factors significantly predict participants' response patterns

Fixed effects	Proportion of expected responses			
	β	SE	z	$Pr(> z)$
Intercept	1.4070	0.1380	10.193	< .001 ***
Pseudo (ref: NGT)	-0.5630	0.1164	-4.836	< .001 ***
Rotation (ref: next-to)	-0.3667	0.1478	-2.481	0.0131 *
Spread (")	-0.7667	0.1392	-5.507	< .001***
Random effects	Variance	SE		
Gloss (Intercept)	0.11287	0.3360		
Subject (Intercept)	0.02602	0.1613		

$Acc \sim Phase + FMU_3 + (1|Subject) + (1|Gloss)$

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05

$SE = 0.148$) was significant with $p = .013$. Lowest accuracy scores were achieved on spread items ($M = 0.581$, $SD = 0.494$), which was not significantly different from chance ($t(25) = 1.44$, $p = .162$), but significantly different from next-to ($\beta = -0.767$, $SE = 0.140$, $p < .001$). Notably, spread is the category with the smallest semantic overlap across items within the stimulus set, due to the combination of a form-unit with two distinct meaning-elements. Average proportions of participants choosing the expected response were significantly higher than chance for the other FMUs (p 's < .001).

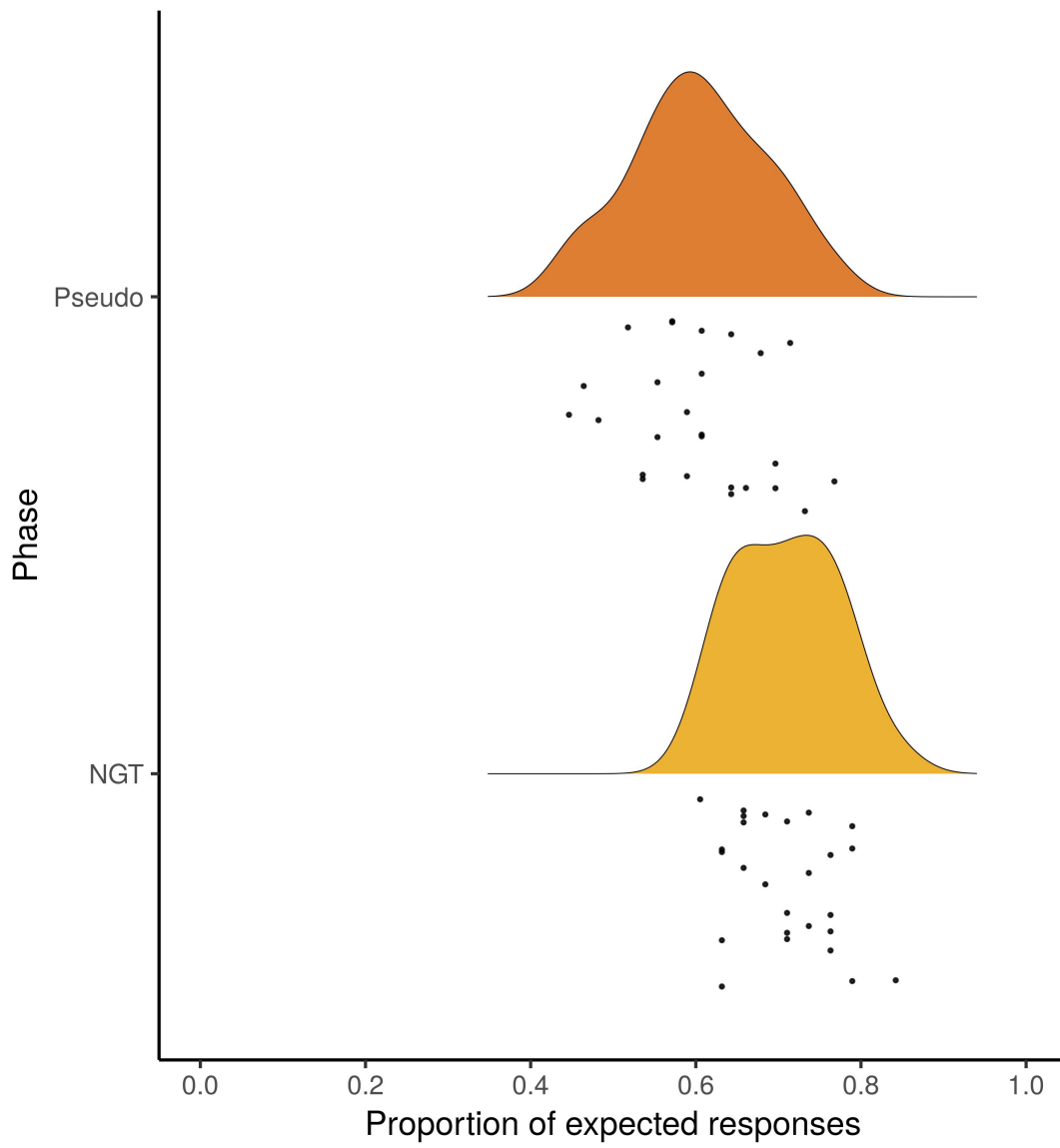


Figure 6: Response patterns per participant by phase

6.2 RQ2 - Do fluent users of NGT abstract to underlying form-meaning units in a perception task? (NGT signs)

A model for accuracy scores including FMU (next-to, rotation, wideness, individuation) as fixed factor and Subject and Gloss as random intercepts failed to converge with $\max|\text{grad}| = 0.00237122$. The model excluding Subject from the random effects structure proved to be a marginally better fit for the data. The relevant statistics and corresponding coefficients for this simpler model are reported in Table 4. Both items containing next-to ($M = 0.811$, $SD = 0.150$, $\beta = 0.936$, $SE = 0.298$, $p = .002$) and rotation ($M = 0.742$, $SD = 0.130$, $\beta = 0.584$, $SE = 0.291$, $p = .045$) yielded significantly higher proportions of expected responses than items including individuation ($M = 0.623$, $SD = 0.142$), while there was no difference between individuation and wideness ($M = 0.645$, $SD = 0.106$, $\beta = 0.114$, $SE = 0.302$, $p = .707$, see Figure 7). Note that in the NGT condition, individuation and wideness refer to two distinct FMUs, sharing the same form-element but linking it to different meaning-components. Expected responses were thus semantically similarly close to the target concept, as in the other FMU conditions, but the form-overlap was significantly larger than with other FMUs.

Table 4: Estimates, standard error, and z-values of the main effects of the generalised linear mixed model assessing what factors significantly predict participants' response patterns within NGT signs

Fixed effects	Proportion of expected responses			
	β	SE	z	$Pr(> z)$
Intercept	0.5471	0.2004	2.731	0.00632 **
Next-to(ref: individuation)	0.9360	0.2981	3.140	0.00169 **
Rotation(")	0.5836	0.2915	2.002	0.04528 *
Wideness(")	0.1139	0.3025	0.376	0.70667
<hr/>				
Random effects	Variance		SE	
Gloss (Intercept)	0.2219		0.4711	

$$Acc \sim FMU + (1|Gloss)$$

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05

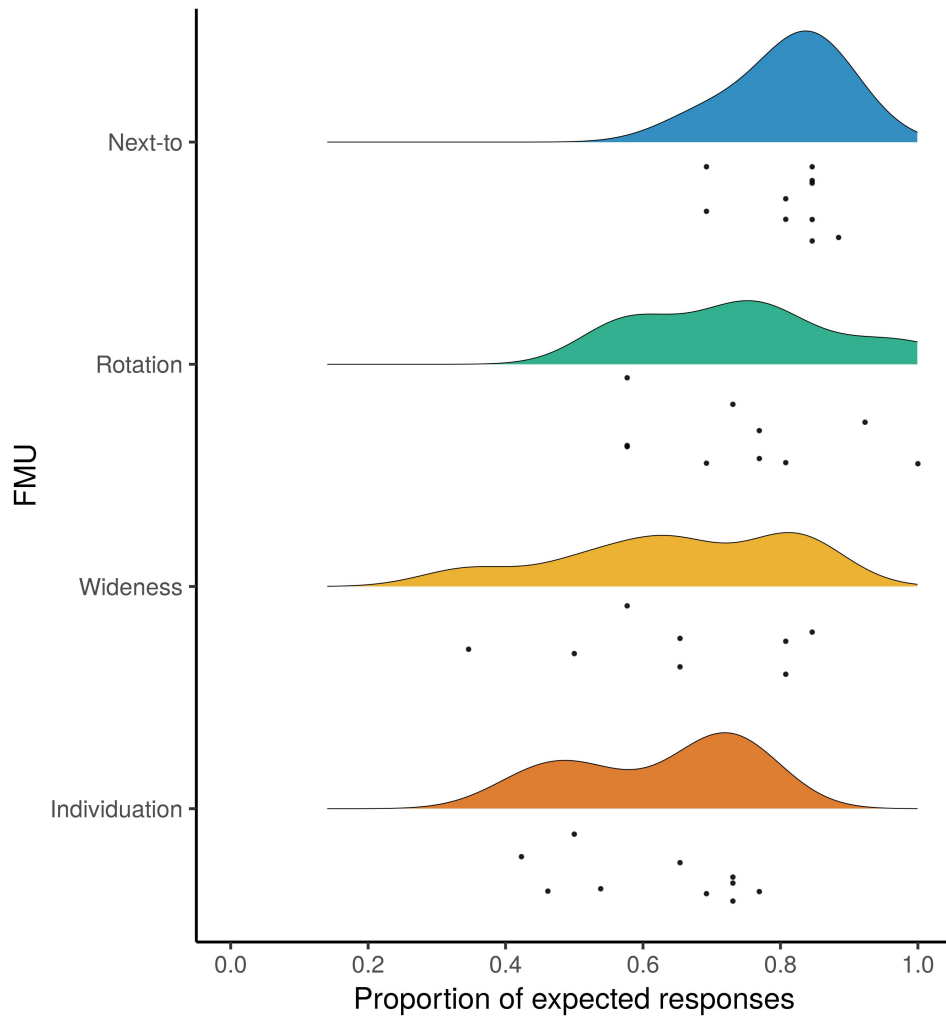


Figure 7: Response patterns per item by FMU in NGT signs

6.3 RQ3 - Do underlying form-meaning associations provide cues towards the interpretation of novel signs?(Pseudosigns)

The model including and FMU (next-to, rotation, spread) as a fixed main effects and Gloss as random intercept proved to have the best fit with the data. The relevant statistics and corresponding coefficients are reported in Table 5. Next-to items had the highest average proportion of expected responses within pseudosigns ($M = 0.685$, 0.104). Rotation items yielded significantly fewer expected responses ($M = 0.607$, $SD = 0.121$, $\beta = -0.358$, $SE = 0.163$, $p = .028$), as did pseudosigns including spread ($M = 0.533$, $SD = 0.116$, $\beta = -0.618$, $SE = 0.160$, $p < .001$; see Figure 8). As mentioned above, the proportion of expected responses on spread items was not significantly different from chance. For pseudosigns, wideness and individuation were combined into a single FMU "spread", since there is no meaning-element available in the structure of pseudosigns. These items were thus phonologically equally homogeneous as the other FMU conditions, but response options were dispersed over a wider semantic range.

Table 5: Estimates, standard error, and z-values of the main effects of the generalised linear mixed model assessing what factors significantly predict participants' response patterns within pseudosigns

		Proportion of expected responses		
Fixed effects	β	SE	z	$Pr(> z)$
Intercept	0.7833	0.1289	6.078	< .001 ***
Rotation (ref: next-to)	-0.3578	0.1632	-2.193	0.028300 *
Spread (")	-0.6185	0.1605	-3.854	< .001 ***
Random effects	Variance	SE		
Gloss (Intercept)	0.05843	0.2417		
Subject (Intercept)	0.05010	0.2238		

$$Acc \sim FMU + (1|Gloss) + (1|Subject)$$

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

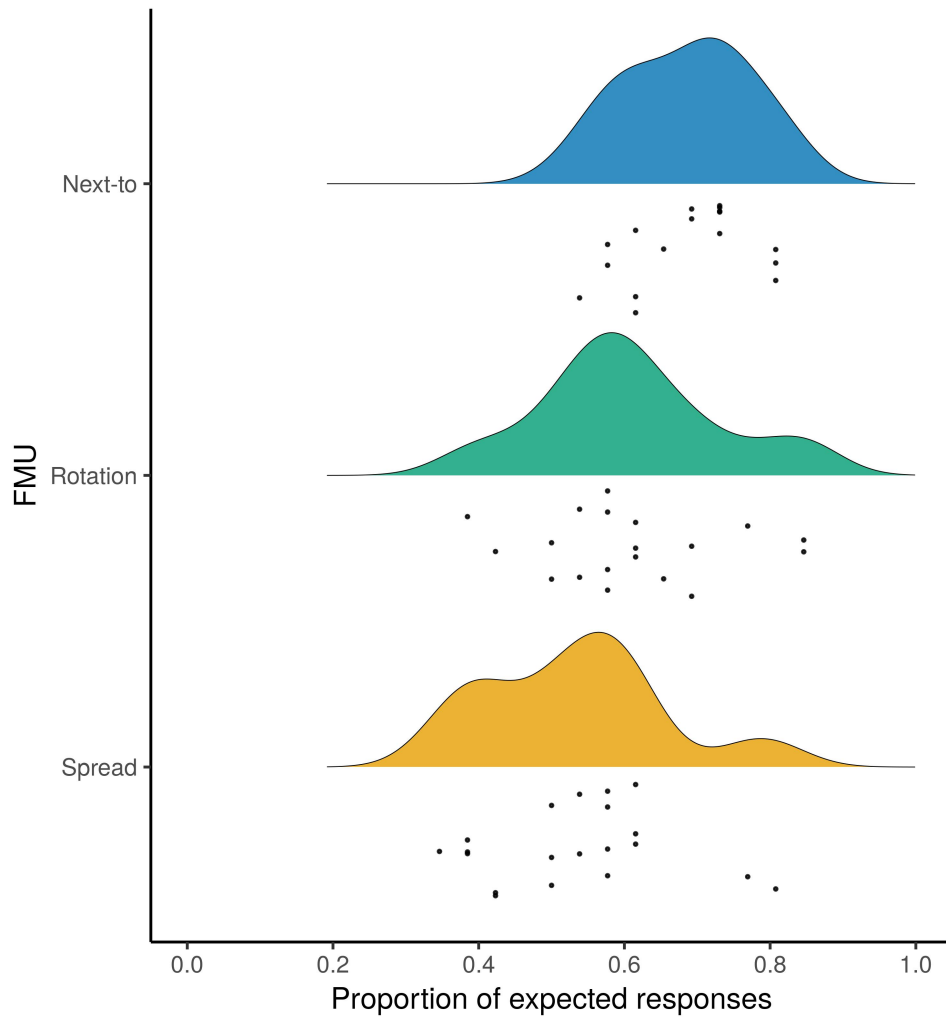


Figure 8: Response patterns per item by FMU in pseudosigns

7 Discussion

Research question 1 focussed on how FMUs and type of sign relate to response patterns across items and subjects. The present study suggests that signers are sensitive to the presence and systematicity of FMUs (FMUs) across items in the lexicon of NGT, as well as in the interpretation of unknown (pseudo-)signs. The effect was stronger in NGT than in pseudosigns, with participants choosing the expected over the non-expected response more frequently in NGT signs. Within both NGT and pseudosigns, the effect was strongest for signs including the FMU *next-to*, followed by signs with *rotation*. The smallest effect, not significantly different from chance in pseudosigns, was for signs including the feature *spread*. In NGT signs, this was separated into items with the form feature *spread* associated with *wideness* and *individuation* respectively. The effect did not differ in size between those, but was significantly different from chance.

Notably, *next-to* and *rotation* items share the property that both form and meaning differ between signs across but not within categories within the present stimulus set, that is all signs within a category share the same FMU and that FMU is not present in any signs from the other category. This means that the FMUs are very systematic and stable, at least for the selection of signs in this study. The category of *spread* items includes signs sharing the same form feature, but with a more heterogeneous distribution of meaning options for pseudosigns. Items in the categories of *wideness* and *individuation* in NGT signs, however, share a form feature: *spread*, but differ in their meaning-components, which are as homogeneous within category as for *next-to* and *rotation*.

These differences between FMUs and between pseudosigns and existing NGT signs make it possible to differentiate effects that are purely driven by form-overlap from the added similarity in meaning potentials present in signs that share an FMU. According to the FMU theory, these shared meaning-potentials should be available in the interpretation of signs, at least if the lexical interpretation is (temporarily) unavailable or does not fit contextual constraints.

Pseudosigns including *spread*, mostly relied on form-overlap, since the meaning-potential contributed a less structured set of options than for the other FMUs. Therefore, the form-meaning link may be argued to be weaker than for the other FMUs, at least within the possibility space of this experiment. Unlike pseudosigns with this FMU, NGT signs including *wideness* or *individuation* yielded responses significantly above chance. However, the responses were not significantly different from those for *spread*-pseudosigns and the other FMUs also yielded lower accuracy scores in pseudosigns than lexical signs of NGT. It is therefore unclear, whether the presence of an effect in NGT signs is due to the presence of more systematic form-meaning associations, or an underlying difference in the processing of NGT signs and pseudosigns.

This may suggest that there is a continuum of FMUs with high salience through those with low salience, to form features that are not systematically connected to a specific

meaning. Factors influencing the position of any individual FMU on that continuum may include the systematicity of the link, that is both how frequently the form occurs with a particular meaning rather than no meaning or even a different meaning, and how frequently a sub-lexical meaning component is instantiated by a particular form, the frequency of the FMU in the lexicon, its visual salience, driven by visual factors such as movement, size, and proximity to the face, and other factors that have not yet been explored (see Table 6 for examples of other functions associated with the form-features in this study). It is clear from the results discussed above, that the association of *spread* with “wideness” and “individuation” is the weakest candidate for an FMU status in the tested set.

Table 6: Examples of other form meaning associations in the NGT data set in Global Signbank (Crasborn et al., 2018)

Form feature	Example signs	Functional description
Next-to	BOWL	the two hands are next to each other because they shape part of the same, symmetrical entity
	ROLL	the combination with the alternating movement blocks an interpretation as being on the same level
Rotation	BOWL	shaping movement, including the pronation of the wrist as the hand follows the curved shape of a bowl from the brim to the bottom
	OLYMPIC	alternating rotation of the hands signalling the presence of multiple interlocked rings
Spread	POLITICS-A	arbitrary hand configuration, no obvious meaning
	WEBSITE	initialised sign, tracing the shape of a screen with W-hands

The question of why performance on pseudosigns was lower overall is an interesting one, as it provides insights into fundamental differences between pseudosigns and NGT signs. Pseudosigns are not part of the linguistic system of NGT signers. While they share certain form features with existing NGT signs, they are not interconnected with other signs at the semantic level by their word meaning. Existing NGT signs, on the other hand, are accessible via a combination of their form features, potential form meaning units, and their word meanings. That is, they co-activate signs that include the same form features, particularly those that share a form meaning unit, and signs that have a similar lexical meaning but may have completely different form elements. The semantic potential was only accessible via FMUs but not through lexical co-activation in pseudosigns. Therefore, pseudosigns had a reduced potential to co-activate lexical items with the same FMU, compared to NGT signs. This is in line with the proposal that FMUs

function as declarative schemas with both a relational and a generative function (Van der Kooij & Zwitserlood, n.d.).

Comparing the results to studies on spoken language iconicity, it is crucial to remember that neither pseudosigns nor NGT signs were optimised for maximum iconicity in the present study. In such studies, the potential for iconicity in pseudowords usually exceeded the potential in naturally occurring ideophones, since the latter included both iconic and arbitrary elements. This leads to weaker results for natural ideophones than pseudowords (Dingemanse et al., 2016). In the present study, however, the level of potential iconicity was approximately the same across both types of items, since each item only contained exactly one FMU of interest to the study. Items could, however, contain other FMUs and the resulting complexity of individual items was not controlled for. It is therefore possible, that the meaning-potential of items in one group was more complex than in the other groups.

The present findings are thus compatible with the view of sub-lexical FMUs, which initially have a meaning-potential, which can contribute to the lexical meaning of the sign. Fluent signers thus seem both to be sensitive to the contribution of FMUs to the lexical meanings of NGT signs and to be capable of making use of those FMUs when encountering novel (pseudo-)signs. However, the FMU candidates included in this study differed in their potential to drive participants' responses.

Another interpretation is also available to account for these results. Participants may simply have been sensible to the form overlap between signs in the experiment and to the semantic overlap at the lexical level for the NGT signs. This interpretation cannot be completely refuted based on the present experiment and is also compatible with the finding that NGT signs yielded more accurate responses than pseudosigns, though certain aspects of the experimental design speak against it. Participants saw signs within each condition in a random order, that is they were not grouped by FMU. This made it more difficult for them to notice the form overlap between certain items and no participant reported noticing patterns. Additionally, response options were presented as written Dutch words, rather than NGT signs, in order to discourage participants from comparing signs rather than semantic concepts.

In order to give a final response to this question, two conditions would need to be fulfilled. Firstly, the results from the present study would need to be compared against a baseline of items which have a systematic form-overlap, without the added meaning potential. This would allow for a separation of the individual contribution of form and meaning at the sub-lexical level. Secondly, the frequency with which the form-meaning pairings in this study actually occur in the NGT lexicon and the token frequency of those types needs to be established. This would be a resource-intensive process, since it would require comparing the frequency of the forms occurring in the lexicon to instances where the forms occur connected to a meaning-potential in the sense of an FMU. Additionally,

as we have seen for *spread*, form features can occur with several meanings, so it would be necessary to establish what the proportion of a particular FMU is in relation to FMUs with the same form or meaning component and to instances where the form is related to no meaning at all, or the meaning component is present in the lexical meaning with no formational feature to instantiate it.

As stated in the section on the development of psycholinguistic studies, small scale designs like the one in this thesis, are suitable for paving the way for larger, more systematic studies. In order to do this, clear theoretical predictions can help drive quantitative approaches that can systematically falsify aspects of theoretical models remain crucial. Additionally, future experimental studies will strongly benefit from a better understanding of the phenomena under study as gained from corpus-studies.

Finally, a fundamental issue illustrated by this study is the question of the nature of FMUs. While many researchers agree that there are sub-lexical components that can carry meaning and that such FMUs may be considered morphemic, the search for the smallest unit is still an open debate. To a certain extent, this is true for both the phonological and the morphological level, however, in the present study, the issue became extremely clear for the morphological level. Can single phonological units carry meaning and thereby acquire a morphological function?

While this is certainly not unheard of in spoken languages, for example in /s/ as an instantiation of the English plural morpheme, the semantic content proclaimed for FMUs is of lexical rather than grammatical nature. It is therefore much closer to concepts of semantic roots or lexical bases. Additionally, spoken language words are usually not considered to be morphologically complex to the extent that they combine multiple lexical morphemes with high levels of abstraction that each contribute their meaning potential to form a word meaning (Sprachtypologie, Lexikon der Sprachwissenschaft, 2008). We would generally call such words compounds and there are strong limitations on their complexity in many spoken languages (Downing, 1977).

Of course, the fact that European languages tend to not have extremely complex words does not make such items unavailable typologically. However, even across languages outside of Europe, lexical complexity tends to lead to words that express more complex meanings, typically expressed by phrases or sentences in European languages, rather than the equivalent of FMUs, which have meaning components that are very abstract meanings contributing to the word meaning of relatively simple lexical items.

If, however, we interpret signs to function as lexical idioms, as suggested by Johnston and Ferrara (2012), we are able to accommodate for more abstract but fully specified sub-lexical meaning components. In the present study, the context for the NGT signs, i.e. the absence of a correct translation would then have forced a componential reading, using the contribution of the shared components as a bridge to other lexical interpretations. The idiomatic, lexical meaning would then be backgrounded for this specific task. For

pseudosigns, on the other hand, no idiomatic meaning is available. They would therefore function as non-lexicalised items, whose components allow for a componential reading within the linguistic system of NGT. However, this account would not explain why NGT signs yielded more expected responses than pseudosigns on average.

The concept of morphological schemas explicitly propose a mechanism driving this difference, by suggesting two routes of activation, one based on schemas, the other on word meanings (Jackendoff & Audring, 2016). The activation based on schemas can be interpreted as activation via the componential meaning in Johnston and Ferrara (2012), while the activation via word meanings is situated at the level of idiomatic meanings. Both accounts are difficult to falsify, as a lack of effects can always be attributed to reduced salience of a certain activation path. This is particularly visible in Johnston and Ferrara's (2012) assertion that "a suitably robust context is needed in order to force a literal interpretation" (p. 245) and that "it may be very difficult to construct such a context of re-interpretation for some extremely common fully-lexical signs" (p. 245).

Following those accounts, the results of the present study would thus suggest that there can indeed be activation via both the componential or schematic structure of the signs and the idiomatic or lexical meaning. For pseudosigns, the latter is not available, leading to a reduced potential for co-activation between items. Additionally, it would seem that the generative function of the two FMUs including *spread* is less salient than for *rotation* and *next-to*, making a componential reading less available.

It follows that both NGT and pseudosigns in this study are indeed complex structures, consisting of sub-lexical components that contribute to their meaning structures. In order to interpret the signs in the present study, participants needed to de-lexicalise the NGT signs and interpret their component blend-structure in order to access the underlying FMUs' declarative function. *Rotation* with the meaning of "change of state" and *next-to* as referring to an egalitarian relationship, a literal or metaphorical "same level", would thus function as easily accessible morphological components, while the componential contribution of *spread* with the meanings "wideness" and "individuation" is less accessible or salient.

Any theory on language structure ultimately needs to be compatible with language processing and acquisition. It therefore needs to be parsimonious to a certain degree, so as to remain psychologically plausible. Positing FMUs at a level that is too small and too abstract, therefore, does not advance our understanding any more than claiming that there is no such thing as a sub-lexical structure at all. The search for the true smallest form and meaning components should therefore be informed both by theory and experimental evidence.

8 Conclusion

The present study explored signers' sensitivity to sub-lexical form-meaning units (FMUs) in lexical signs of NGT and phonologically valid pseudosigns. It showed that participants selected target responses that were related to the stimuli by a shared FMU in the translation equivalent more frequently than unrelated foils for both NGT- and pseudosigns. The effects were stronger for lexical signs, suggesting that the presence of a meaning component significantly improved participants' performance. This is in line with accounts that posit sub-lexical morphological units relevant to sign language structure and processing.

Across both conditions, items with the form feature *next-to* associated with the meaning of "being at the same (hierarchical) level" yielded the highest proportion of target responses, followed by items including *rotation* associated with a "change of state". Closest to chance was the feature *spread*, associated with "wideness" and "individuation" respectively. This may be due to a lower salience of the feature itself. However, another possibility is that the association of the form feature with two meaning components resulted in a weaker and thus less reliable form-meaning link, which was thus less exploited by participants.

The study results suggest that signs are indeed complex in that they combine a componential structure with an idiomatic reading. More studies are needed to better understand the respective contribution of phonological and semantic overlap at the word level, in comparison to the FMUs as morphological units.

9 References

- Baayen, R. H., Feldman, L. B., & Schreuder, R. (2006). Morphological influences on the recognition of monosyllabic monomorphemic words. *Journal of Memory & Language*, *55*(2), 290–313. doi: 10.1016/j.jml.2006.03.008
- Baayen, R. H., Piepenbrock, R., & Van Rijn, H. (1993). *The CELEX Lexical Database [CD-ROM]*. Philadelphia: Linguistic Data Consortium.
- Balota, D. A., Yap, M. J., Hutchison, K. A., & Cortese, M. J. (2013). Megastudies: What do millions (or so) of trials tell us about lexical processing? In J. S. Adelman (Ed.), *Visual Word Recognition Volume 1: Models and methods, orthography and phonology* (pp. 90–115). New York, NY: Psychology Press.
- Balota, D. A., Yap, M. J., Hutchison, K. A., Cortese, M. J., Kessler, B., Loftis, B., ... Treiman, R. (2007). The English Lexicon Project. *Behavior Research Methods*, *39*(3), 445–459.
- Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting Linear Mixed-Effects Models Using {lme4}. *Journal of Statistical Software*, *67*(1), 1–48. Retrieved 2018-11-01TZ, from <http://www.jstatsoft.org/v67/i01/> doi: 10.18637/jss.v067.i01
- Bergen, B. K. (2004). The Psychological Reality of Phonaesthemes. *Language*, *80*(2), 290–311. doi: 10.1353/lan.2004.0056
- Brackbill, Y., & Little, K. B. (1957). Factors determining the guessing of meanings of foreign words. *The Journal of Abnormal and Social Psychology*, *54*(3), 312–318. doi: 10.1037/h0042411
- Bremner, A. J., Caparos, S., Davidoff, J., de Fockert, J., Linnell, K. J., & Spence, C. (2013). “Bouba” and “Kiki” in Namibia? A remote culture make similar shape–sound matches, but different shape–taste matches to Westerners. *Cognition*, *126*(2), 165–172. Retrieved 2018-09-25TZ, from <https://www.sciencedirect.com/science/article/pii/S0010027712002077?via%3Dihub> doi: 10.1016/j.cognition.2012.09.007
- Brennan, M. (1990). *Word formation in British Sign Language* (PhD Thesis). University of Stockholm, Stockholm.
- Brennan, M. (2005). Conjoining Word and Image in British Sign Language (BSL): An Exploration of Metaphorical Signs in BSL. *Sign Language Studies*, *5*(3), 360–382. Retrieved from http://muse.jhu.edu/content/crossref/journals/sign_language_studies/v005/5.3brennan.html doi: 10.1353/sls.2005.0007
- Brentari, D. (1990). *Theoretical foundations of American Sign Language phonology* (PhD Thesis). University of Chicago, Chicago.
- Brentari, D. (1998). *A prosodic model of sign language phonology*. Cambridge, MA: MIT Press.
- Brown, R. (1980). Why are signed languages easier to learn than spoken languages? In W. C. Stokoe (Ed.), *Proceedings of the First National Symposium on Sign Language Research and Teaching* (pp. 9–24). Washington, DC: National Association of the Deaf.
- Brysbaert, M. (2018). *Linguistics, psycholinguistics and big data: The way to the future* [Plenary lecture]. Poznan, PL. Retrieved from https://www.youtube.com/watch?v=X6_u-uljdgI
- Brysbaert, M., Buchmeier, M., Conrad, M., Jacobs, A. M., Bölte, J., & Böhl, A. (2011). The word frequency effect: a review of recent developments and implications for the choice of frequency estimates in German. *Experimental Psychology*, *58*(5), 412–424. doi: 10.1027/1618-3169/a000123

- Brysbaert, M., & New, B. (2009). Moving beyond Kucera and Francis: A critical evaluation of current word frequency norms and the introduction of a new and improved word frequency measure for American English. *Behavior Research Methods*, *41*(4), 977–990. doi: 10.3758/BRM.41.4.977
- Brysbaert, M., & Stevens, M. (2018). Power Analysis and Effect Size in Mixed Effects Models: A Tutorial. *Journal of Cognition*, *1*(9), 1 – 20. doi: 10.5334/joc.10
- Burgess, C., & Livesay, K. (1998). The effect of corpus size in predicting reaction time in a basic word recognition task: Moving on from Kucera and Francis. *Behavior Research Methods, Instruments, & Computers*, *30*(2), 272–277. doi: 10.3758/BF03200655
- Casey, S. K. (2003). "Agreement" in gestures and signed languages: The use of directionality to indicate referents involved in actions." (Doctoral dissertation). University of California, San Diego, CA.
- Crasborn, O., Zwitserlood, I., Kooij, E. V. D., & Schüller, A. (2018). Global SignBank manual. doi: 10.13140/rg.2.2.36603.82729
- De Saussure, F. (1916). *Cours de linguistique générale*. Lausanne: Payot.
- Diependaele, K., Lemhöfer, K., & Brysbaert, M. (2013). The Word Frequency Effect in First and Second Language Word Recognition: A Lexical Entrenchment Account. *Quarterly Journal of Experimental Psychology*, *66*(5), 843–863. doi: 10.1080/17470218.2012.720994
- Dingemanse, M. (2018). Redrawing the margins of language: Lessons from research on ideophones. *Glossa: a journal of general linguistics*, *3*(1). doi: 10.5334/gjgl.444
- Dingemanse, M., Schuerman, W., Reinisch, E., Tufvesson, S., & Mitterer, H. (2016). What sound symbolism can and cannot do: Testing the iconicity of ideophones from five languages. *Language*, *92*(2), e117–e133. doi: 10.1353/lan.2016.0034
- Downing, P. (1977). On the Creation and Use of English Compound Nouns. *Language*, *53*(4), 810. Retrieved from <http://dx.doi.org/10.2307/412913> doi: 10.2307/412913
- Ellis, A. W., & Morrison, C. M. (1998). Real age-of-acquisition effects in lexical retrieval. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *24*(2), 515–523. doi: 10.1037/0278-7393.24.2.515
- Fernald, T. B., & Napoli, D. J. (2000). Exploitation of morphological possibilities in signed languages: Comparison of American Sign Language with English. *Sign Language & Linguistics*, *3*(1), 3–58. Retrieved from <http://www.jbe-platform.com/content/journals/10.1075/sll.3.1.03fer> doi: 10.1075/sll.3.1.03fer
- Ferrand, L., New, B., Brysbaert, M., Keuleers, E., Bonin, P., Méot, A., ... Pallier, C. (2010). The French Lexicon Project: lexical decision data for 38,840 French words and 38,840 pseudowords. *Behavior Research Methods*, *42*(2), 488–496. doi: 10.3758/BRM.42.2.488
- Francis, W. N., & Kucera, H. (1964). *Brown Corpus Manual* (Tech. Rep.). Providence, Rhode Island: Brown University. Retrieved from <http://clu.uni.no/icame/manuals/BROWN/INDEX.HTM>
- Friedman, L. A. (1976). *Phonology of a Soundless Language: Phonological Structure of the American sign Language* (Unpublished doctoral dissertation). University of California, Berkeley.
- Gebels, G. (1969). An investigation of phonetic symbolism in different cultures. *Journal of Verbal Learning and Verbal Behavior*, *8*(2), 310–312. doi: 10.1016/S0022-5371(69)80083-6

- Grote, K., & Linz, E. (2003). The influence of Sign Language Iconicity on Semantic Conceptualization. In W. G. Müller & O. Fischer (Eds.), *From Sign to Signing* (Vol. 3). Amsterdam/Philadelphia: John Benjamins.
- Gullberg, M., & Indefrey, P. (2003). Language background questionnaire. *Developed in the Dynamics of Multilingual Processing. Max Planck Institute for Psycholinguistics*. Retrieved from <https://www.mpi.nl/departments/other-research/research-projects/the-dynamics-of-multilingual-processing/tools>
- Imai, M., Kita, S., Nagumo, M., & Okada, H. (2008). Sound symbolism facilitates early verb learning. *Cognition*, *109*(1), 54–65. doi: 10.1016/j.cognition.2008.07.015
- Jackendoff, R., & Audring, J. (2016). Morphological schemas: Theoretical and psycholinguistic issues. *The Mental Lexicon*, *11*(3), 467–493. Retrieved from <http://www.jbe-platform.com/content/journals/10.1075/ml.11.3.06jac> doi: 10.1075/ml.11.3.06jac
- Johnston, T., & Ferrara, L. (2012). Lexicalization in Signed Languages: When is an Idiom not an Idiom? In *Selected Papers from UK-CLA Meetings* (Vol. 1, pp. 229–248).
- Johnston, T., & Schembri, A. (1999). On defining lexeme in a signed language. *Sign Language and Linguistics*, *2*, 115–185.
- Keuleers, E., Brysbaert, M., & New, B. (2010). SUBTLEX-NL: a new frequency measure for Dutch words based on film subtitles. *Behavior Research Methods*, *42*(3), 643 – 650. doi: 10.3758/BRM.42.3.643
- Keuleers, E., Diependaele, K., & Brysbaert, M. (2010). Practice Effects in Large-Scale Visual Word Recognition Studies: A Lexical Decision Study on 14,000 Dutch Mono- and Disyllabic Words and Nonwords. *Frontiers in Psychology*, *1*(174), 1 – 15. doi: 10.3389/fpsyg.2010.00174
- Kovic, V., Plunkett, K., & Westermann, G. (2010). The shape of words in the brain. *Cognition*, *114*(1), 19–28. doi: 10.1016/j.cognition.2009.08.016
- Kucera, H., & Francis, W. N. (1967). *Computational analysis of present-day American English*. Providence: Brown University Press.
- Kuperman, V. (2013). Accentuate the positive: semantic access in English compounds. *Frontiers in Psychology*, *4*(203). doi: 10.3389/fpsyg.2013.00203
- Köhler, W. (1929). *Gestalt Psychology*. New York, NY: Liveright.
- Liddell, S. K., & Johnson, R. E. (1989). American Sign Language. *Sign Language Studies*, *64*, 195–278. doi: 10.1353/sls.1989.0027
- Locke, J. (1690). *An Essay Concerning Human Understanding*. Dorset Court.
- Lockwood, G., Dingemanse, M., & Hagoort, P. (2016). Sound-symbolism boosts novel word learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *42*(8), 1274–1281. doi: 10.1037/xlm0000235
- Lockwood, G., Hagoort, P., & Dingemanse, M. (2016). How iconicity helps people learn new words: Neural correlates and individual differences in sound-symbolic bootstrapping. *Collabra*, *2*(1), 1–15. doi: <http://dx.doi.org/10.1525/collabra.42>
- Lund, K., & Burgess, C. (1996). Producing high-dimensional semantic spaces from lexical co-occurrence. *Behavior Research Methods, Instruments, & Computers*, *28*(2), 203 – 208. doi: 10.3758/BF03204766
- Mandel, M. (1977). Iconic devices in American sign language. In L. A. Friedman (Ed.), *On the other hand* (pp. 57–100). New York, NY: Academic Press.
- Mathôt, S., Schreij, D., & Theeuwes, J. (2012). OpenSesame: An open-source, graphical experiment builder for the social sciences. *Behavior Research Methods*, *44*(2), 314–324. doi: 10.3758/s13428-011-0168-7

- Meier, R. P. (1982). *Icons, Analogues, and Morphemes: the Acquisition of Verb Agreement in American Sign Language* (Unpublished doctoral dissertation). University of California, San Diego, CA.
- Meier, R. P., Mauk, C. E., Cheek, A., & Moreland, C. J. (2008). The Form of Children's Early Signs: Iconic or Motoric Determinants? *Language Learning and Development*, 4(1), 63–98. doi: 10.1080/15475440701377618
- Meir, I. (2012). Word classes and word formation. In R. Pfau, M. Steinbach, & B. Woll (Eds.), *Handbook on sign Language Linguistics* (pp. 365–387). Berlin: Mouton de Gruyter.
- Morphem. (2003). In D. Homberger (Ed.), *Sachwörterbuch zur Sprachwissenschaft* (pp. 343–345). Stuttgart, DE: Reclam.
- Nielsen, A., & Rendall, D. (2011). The sound of round: Evaluating the sound-symbolic role of consonants in the classic Takete-Maluma phenomenon. *Canadian Journal of Experimental Psychology/Revue canadienne de psychologie expérimentale*, 65(2), 115–124. Retrieved 2018-09-25TZ, from <http://doi.apa.org/getdoi.cfm?doi=10.1037/a0022268> doi: 10.1037/a0022268
- Nielsen, A., & Rendall, D. (2012). The source and magnitude of sound-symbolic biases in processing artificial word material and their implications for language learning and transmission. *Language and Cognition*, 4(02), 115–125. Retrieved 2018-09-25TZ, from http://www.journals.cambridge.org/abstract_S1866980800000843 doi: 10.1515/langcog-2012-0007
- Nielsen, A. K. S., & Rendall, D. (2013). Parsing the role of consonants versus vowels in the classic Takete-Maluma phenomenon. *Canadian Journal of Experimental Psychology/Revue canadienne de psychologie expérimentale*, 67(2), 153–163. Retrieved 2018-09-25TZ, from <http://doi.apa.org/getdoi.cfm?doi=10.1037/a0030553> doi: 10.1037/a0030553
- Nijen Twilhaar, J., & Van den Bogaerde, B. (2016). *Concise lexicon for sign linguistics*. Amsterdam ; Philadelphia: John Benjamins Publishing Company.
- Nygaard, L. C., Cook, A. E., & Namy, L. L. (2009). Sound to meaning correspondences facilitate word learning. *Cognition*, 112(1), 181–186. doi: 10.1016/j.cognition.2009.04.001
- Ocelli, V., Esposito, G., Venuti, P., Arduino, G. M., & Zampini, M. (2013). The Takete—Maluma Phenomenon in Autism Spectrum Disorders. *Perception*, 42(2), 233–241. Retrieved 2018-09-25TZ, from <http://journals.sagepub.com/doi/10.1068/p7357> doi: 10.1068/p7357
- Occhino, C., Anible, B., Wilkinson, E., & Morford, J. P. (2017). Iconicity is in the eye of the beholder: How language experience affects perceived iconicity. *Gesture*, 16(1), 100–126. Retrieved from <http://www.jbe-platform.com/content/journals/10.1075/gest.16.1.04occ> doi: 10.1075/gest.16.1.04occ
- Orlansky, M. D., & Bonvillian, J. D. (1984). The Role of Iconicity in Early Sign Language Acquisition. *Journal of Speech and Hearing Disorders*, 49(3), 287. doi: 10.1044/jshd.4903.287
- Ormel, E., & Giezen, M. (2018). *Non-signs in NGT: phonologically correct forms that could be signs, but are not (yet) used in the language*. Retrieved from <https://signbank.science.ru.nl/datasets/detail/11>
- Ormel, E., Giezen, M., Van Zuilen, M., Snijders, T., Smoll, L., Zijlstra, D., & Schiller, N. (2016). *Effects of iconicity on sign language processing - An ERP study* [Poster presentation]. Melbourne, AUS.

- Ormel, E., Knoors, H., Hermans, D., & Verhoeven, L. (2009). The Role of Sign Phonology and Iconicity During Sign Processing: The case of Deaf Children. *Journal of Deaf Studies and Deaf Education*, 14(4), 436–448. doi: 10.1093/deafed/enp021
- Perniss, P., Thompson, R. L., & Vigliocco, G. (2010). Iconicity as a general property of language: evidence from spoken and signed languages. *Frontiers in Psychology*, 1(227), 1664–1678. doi: 10.3389/fpsyg.2010.00227
- Phonem. (2008). In H. Bußmann, C. Gerstner-Link, & H. Lauffer (Eds.), *Lexikon der Sprachwissenschaft: mit 14 Tabellen* (4., durchges. und bibliogr. erg. Aufl ed., pp. 522 – 523). Stuttgart: Kröner. (OCLC: 281181768)
- R Core Team. (2018). *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing. Retrieved from <https://www.r-project.org/>
- Ramachandran, V., & Hubbard, E. (2001). Synaesthesia – A window into perception, thought and language. *Journal of Consciousness Studies*, 8(12), 3–34. Retrieved from <https://www.ingentaconnect.com/content/imp/jcs/2001/00000008/00000012/1244>
- Sandler, W. (1989). *Phonological representation of the sign: linearity and nonlinearity in American Sign Language*. Dordrecht: Foris.
- Sapir, E. (1929). A study in phonetic symbolism. *Journal of Experimental Psychology*, 12(3), 225–239. doi: 10.1037/h0070931
- Schermer, T. (2016). Lexicon. In *The Linguistics of Sign Languages: An introduction* (pp. 173 – 195). Amsterdam/Philadelphia: John Benjamins Publishing Company.
- Shapiro, B. J. (1969). The Subjective Estimation of Relative Word Frequency. *Journal of verbal learning and verbal behavior*, 8, 248–251. doi: 10.1016/S0022-5371(69)80070-8
- Shintel, H., Nusbaum, H. C., & Okrent, A. (2006). Analog acoustic expression in speech communication. *Journal of Memory and Language*, 55(2), 167–177. doi: 10.1016/j.jml.2006.03.002
- Slobin, D. I., Hoiting, N., Kuntze, M., Lindert, R., Weinberg, A., Pyers, J., . . . Thumann, H. (2003). A cognitive/functional perspective on the acquisition of "classifiers". In K. Emmorey (Ed.), *Perspectives on Classifier Constructions in Sign Languages* (pp. 271–296). Mahwah, NJ: Erlbaum.
- Sprachtypologie. (2008). In H. Bußmann, C. Gerstner-Link, & H. Lauffer (Eds.), *Lexikon der Sprachwissenschaft: mit 14 Tabellen* (4., durchges. und bibliogr. erg. Aufl ed., pp. 664 – 666). Stuttgart: Kröner. (OCLC: 281181768)
- Stokoe, W. C. (1960). Sign language structure: An outline of the visual communication system of the American deaf. *Studies in Linguistics Occasional Papers*, 8.
- Sutton-Spence, R. (2005). *Analysing Sign Language Poetry*. London, UK: Palgrave Macmillan UK. doi: 10.1057/9780230513907
- Taub, S. F. (2001a). The Analogue-Building Model of Linguistic Iconicity. In S. F. Taub (Ed.), *Language from the body: Iconicity and Metaphor in American Sign Language*. Cambridge, UK: Cambridge University Press. doi: 10.1017/CBO9780511509629.004
- Taub, S. F. (2001b). *Language from the Body: Iconicity and Metaphor in American Sign Language*. Cambridge, UK: Cambridge University Press.
- Team, R. (2016). *RStudio: Integrated Development for R*. Boston, MA: RStudio, Inc. Retrieved from <https://www.rstudio.com>

- Thompson, R. L., Vinson, D. P., & Vigliocco, G. (2010). The link between form and meaning in British Sign Language: Effects of iconicity for phonological decisions. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *36*(4), 1017–1027. doi: 10.1037/a0019339
- Thorndike, E. L. (1921). *The teacher's word book*. New York, NY: Teachers College, Columbia University.
- Thorndike, E. L. (1932). *A Teacher's Word Book of the Twenty Thousand Words Found Most Frequently and Widely in General Reading*. New York, NY: Columbia University Press.
- Thorndike, E. L., & Lorge, I. A. (1944). *A teacher's word book of 30,000 words*. New York, NY: Columbia University Press.
- Tolar, T. D., Lederberg, A. R., Gokhale, S., & Tomasello, M. (2008). The Development of the Ability to Recognize the Meaning of Iconic Signs. *Journal of Deaf Studies and Deaf Education*, *13*(2), 225–240. doi: 10.1093/deafed/enm045
- Van der Kooij, E. (2002). *Phonological Categories in Sign Language of the Netherlands: The Role of Phonetic Implementation and Iconicity*. Utrecht, NL: LOT.
- Van der Kooij, E., & Zwitserlood, I. (n.d.). *Strategies for new word formation in NGT: a case for simultaneous morphology* [unpublished manuscript]. Nijmegen.
- Van Heuven, W. J. B., Mandera, P., Keuleers, E., & Brysbaert, M. (2014). SUBTLEX-UK: A new and improved word frequency database for British English. *The Quarterly Journal of Experimental Psychology*, *67*(6), 1176–1190. doi: <https://doi.org/10.1080/17470218.2013.850521>
- Van Orden, G. C. (1991). Phonologic mediation is fundamental to reading. In D. Besner & G. W. Humphreys (Eds.), (Basic processes in reading: Visual word recognition ed., pp. 77 – 103). Hillsdale, NJ: Lawrence Erlbaum Associates Inc.
- Vigliocco, G., Vinson, D. P., Woolfe, T., Dye, M. W., & Woll, B. (2005). Language and imagery: effects of language modality. *Proceedings of the Royal Society B: Biological Sciences*, *272*(1574), 1859–1863. doi: 10.1098/rspb.2005.3169
- Vinson, D. P., Cormier, K., Denmark, T., Schembri, A., & Vigliocco, G. (2008). The British Sign Language (BSL) norms for age of acquisition, familiarity, and iconicity. *Behavior Research Methods*, *40*(4), 1079–1087. doi: 10.3758/BRM.40.4.1079
- Walker, P., Bremner, J. G., Mason, U., Spring, J., Mattock, K., Slater, A., & Johnson, S. P. (2010). Preverbal Infants' Sensitivity to Synaesthetic Cross-Modality Correspondences. *Psychological Science*, *21*(1), 21–25. doi: 10.1177/0956797609354734
- Westbury, C. (2005). Implicit sound symbolism in lexical access: Evidence from an interference task. *Brain and Language*, *93*(1), 10–19. doi: 10.1016/j.bandl.2004.07.006
- Zeno, S. M., Ivens, S. H., Millard, R. T., & Duvvuri, R. (1995). *The educator's word frequency guide*. Brewster, NY: Touchstone Applied Science.
- Zwitserlood, I. (2002). The Complex Structure of 'Simple' Signs in NGT. In J. Koppen, E. Thrift, & E. Van der Torre (Eds.), (pp. 232 – 246).
- Zwitserlood, I. (2003). *Classifying hand configurations in Nederlandse Gebarentaal (Sign Language of the Netherlands)* (PhD Thesis). Universiteit Utrecht, Utrecht, NL.

A Scripts and raw data

All scripts for list generation and analysis, as well as the raw data can be found on OSF: https://osf.io/7enu4/?view_only=8dae458cc6574842990398c8a8907f90.

B List of NGT signs and their possible translations

GLOSS	FMU	Translations	Acc
DANSEN-A.mov	IND	dansen, ballet, dans	0.50
ENZOVOORTS-C.mov	IND	enzovoorts	0.73
KIJKEN-A.mov	IND	kijken, observeren, gadeslaan, iets bestuderen, bekijken, toezien, toezicht, zien, beschouwen	0.42
REGEN-A.mov	IND	motregen, stortregen, bui, regen, miezerig	0.73
REKENEN-A.mov	IND	getal, rekenen, aantal, tellen	0.77
RIJ.mov	IND	in de rij staan, rij van mensen, rij	0.48
ROOSTER.mov	IND	tabel, rooster, schema, overzicht, Excel	0.46
TELLEN.mov	IND	tellen, paar, sommige, sommigen, aantal, aantallen	0.65
TWEEEN-A.mov	IND	wij tweeen, jullie tweeen, wij, jullie, zij, samen, elkaar, allebei, tweeling, tweeen, ons	0.54
VOORSTELLEN-A.mov	IND	dromen, voorstelling, droom, fantasie, fantaseren, verzinnen, zich iets voorstellen	0.69
COMMUNICATIE.mov	NEX	communicatie, cursus	0.85
ENE-KANT-ANDERE-KANT.mov	NEX	enerzijds-anderzijds, beiden, of dit of dat, richting, ene kant andere kant	0.69
GELIJK-NIVEAU.mov	NEX	gelijkwaardig, bij elkaar aansluiten, op de zelfde hoogte, neutraal, gelijk niveau	0.81
ONTMOETEN.mov	NEX	ontmoeten, elkaar tegenkomen, treffen	0.69
OVEREENKOMST.mov	NEX	overeenkomst, akkoord, compromis	0.88
RUZIE-A.mov	NEX	ruzie, oorlog, vijandig, conflicten	0.81

GLOSS	FMU	translations	Accuracy
SAMEN-B.mov	NEX	samen, als partners samen, een stel zijn, twee personen zijn samen, twee partijen zijn samen	0.85
SYNONIEM.mov	NEX	synoniem	0.85
ZELFDE-A.mov	NEX	lijken op, alsof, ook, dezelfde, zelfde, zelfde, hetzelfde, hetzelfde als, vergelijkbaar, gelijk, consequent, gelijkwaardig	0.85
ZUS-C.mov	NEX	zus, broer	0.85
AFRONDEN.mov	ROT	afronden, afmaken, afwerken, afkrijgen, voltooiën, afgerond, beëindigen, afkappen, nu stoppen	0.58
BEGINNEN-A.mov	ROT	begin, beginnen, hup, laten we starten, bij aanvang, in het begin, toen het begon, begonnen	0.58
BESLUITEN-B.mov	ROT	vaststellen, beslissen, bepalen, besluit, besloten, bepaald, knoop doorhakken, besluiten, beslissing, zo doen	1.00
BREKEN.mov	ROT	breken, knakken, kapot, iets kapot maken, kapot breken	0.81
DRAAIEN.mov	ROT	draaien, draaiknop, kraan, omdraaien, aandraaien, opendraaien, dichtdraaien	0.58
OVERLIJDEN.mov	ROT	overlijden, sterven, dood, overleden, doodgaan, omkeren, bekeren, dopen, omdraaien, omkeren	0.92
SCHAKELEN-B.mov	ROT	afwisselen, om en om, schakelen, switchen, omschakelen, transfer, iets omzetten, overgang, overgaan	0.77
SLUITEN.mov	ROT	uitsluiten, afsluiten, sluiten, opsluiten, dicht, gesloten, beëindigen, op slot, op slot doen, uit doen	0.73
UIT-ELKAAR-VALLEN.mov	ROT	uit elkaar vallen	0.69
VERANDEREN-A.mov	ROT	andersom, veranderen, switchen, omzetten, verandering	0.77

GLOSS	FMU	translations	Accuracy
ERGENS-A.mov	WID	ergens, daar ergens	0.65
OMGEVING-A.mov	WID	omgeving, zij, hen, de mensen, milieu, natuur, jullie, allemaal, klimaat	0.81
PARASOL.mov	WID	parasol, koepel	0.58
REGIO-A.mov	WID	gebied, omgeving, regio, landelijk, overal	0.85
RUIMTE-A.mov	WID	gebied, familie, omgeving, algemeen, maatschappij, groep mensen, regio, situatie, milieu wijk, kamer, plaats, ruimte, ergens, buurt	0.81
SOMS-A.mov	WID	soms, af en toe	0.35
VERSPREIDEN-A.mov	WID	verspreiden, netwerk	0.65
WAAR-A.mov	WID	waarom, waar	0.50

C Linguistic Background Questionnaire

C.1 Long version

Radboud Universiteit Nijmegen



Vragenlijst – Achtergrond

pp. Nr.: ____

De volgende vragen gaan over uw algemene achtergrond en taalachtergrond. Probeer deze zo compleet mogelijk te beantwoorden.

Achtergrond (*Omcirkel wat van toepassing is)

Naam:

Geboortedatum: - - 19.....

Geslacht: m / v*

Dominante hand: rechtshandig / linkshandig*

Wat is uw opleidingsniveau en opleiding (bijv. B Commerciële Economie):

.....

Wat is uw beroep (baan/bijbaan; bijv. student/ondernemer):

Bent u geboren in Nederland? Ja / Nee*

Zo nee, waar bent u geboren?

Hoe oud was u toen u naar Nederland kwam? jaar

Hoe lang woont u al in Nederland? jaar

Zo ja, heeft u sinds uw geboorte in Nederland gewoond? Ja / Nee*

In welke stad (of dorp) heeft u gewoond?

In welke stad (of dorp) woont u momenteel?

Hoe vaak speelt u videogames: Zeer weinig / Weinig / Gemiddeld / Veel / Zeer veel*

Zo ja, wat voor soort games (bijv. schietspellen, racegames, enz.)?

.....

Speelt u een muziekinstrument of heeft u in het verleden een instrument gespeeld? Ja /Nee*

Zo ja, hoe lang en welk(e) instrument(en)?

.....

Heeft u dove ouders, dove grootouders, dove broertjes of zusjes? Zo ja, geef aan welke directe familieleden doof zijn:

Taalachtergrond

Wat is/zijn uw eerste taal/talen?

Geef alstublieft aan welke andere talen u nog meer kent. Noteer de talen die u nog veel in het dagelijks leven gebruikt of voor een langere periode in het verleden hebt gebruikt. Probeer een schatting te maken van de beheersing die u heeft over elke taal. Gebruik hierbij de volgende schaal:

|


 Radboud Universiteit Nijmegen

Niet goed 1 2 3 4 5 6 7 8 9 10 Heel goed

Taal	Spreken/ Gebaren	Luisteren/ Afzien	Schrijven (gesproken talen)	Lezen (gesproken talen)
1				
2				
3				
4				
5				

Geef alsjeblieft een indicatie van de leeftijd waarop u deze talen hebt geleerd, hoe vaak u de taal gebruikt (kies uit *dagelijks*, *wekelijks*, *maandelijks of aantal keer per jaar*) en het land waarin u deze talen hebt geleerd. Schrijf ook eventuele opmerkingen op waarvoor u de taal hebt geleerd indien u enige tijd in een ander land hebt gewoond (bijv. stage of studeren in het buitenland).

Taal	Leeftijd	Hoe vaak	Land	Opmerkingen
1				
2				
3				
4				
5				

Geef alsjeblieft aan welke taal/talen u gebruikt met de volgende personen:

Taal	Taal/Talen
Moeder	
Vader	
Broer(tje)/Zus(je)	
Andere familieleden	

Radboud Universiteit Nijmegen



Partner/ Vrienden	
Leraren/ Klasgenoten/ Collega's	

Geef alstublieft aan welke taal/talen u voor deze activiteiten gebruikt:

Activiteit	Taal/Talen
Lezen	
TV kijken	
Luisteren naar de radio/muziek	
Gamen	
E-mail/Internet	

Gebruik bij de volgende stellingen een schaal van 1 tot 10. Omcirkel uw antwoord.

Hoeveel houdt u ervan om nieuwe talen te leren?

Ik hou er totaal niet van

1 2 3 4 5 6 7 8 9 10

Ik hou er heel erg van

Hoe makkelijk vindt u het om nieuwe talen te leren?

Moeilijk

1 2 3 4 5 6 7 8 9 10

Makkelijk

Hoe vaak maakt u gebruik van meerdere talen tegelijkertijd?

Bijna niet

1 2 3 4 5 6 7 8 9 10

Heel vaak

Radboud Universiteit Nijmegen



De volgende vragen zijn van toepassing wanneer u de Tolken/Lerarenopleiding NGT volgt of NGT kent:

Hoeveel gebruikt u deze taal naar schatting gemiddeld per dag (sinds u in aanraking bent gekomen met NGT)?

..... uren per dag

Hoeveel blootstelling heeft u gemiddeld per dag aan deze taal (sinds u in aanraking bent gekomen met NGT)?

..... uren per dag

Welke vakken/modules NGT heeft u al afgerond?

.....

Geef van onderstaande stellingen aan op een schaal van 1-10 (1: bijna niet - 10: heel erg) in hoeverre de volgende stellingen van toepassing zijn uw beheersing van het NGT met betrekking tot uw **productie**.

	Beoordeling 1-10
Ik kan bekende dagelijkse uitdrukkingen en eenvoudige woorden of zinnen produceren	
Ik ben bekend met veelgebruikte uitdrukkingen en kan gesprekken voeren over alledaagse zaken	
Ik kan mijn eigen mening geven en ik kan ervaringen, gebeurtenissen, dromen en verwachtingen beschrijven	
Ik kan de hoofdlijnen van complexe teksten produceren; ik kan duidelijke, gedetailleerde tekst produceren en ik kan spontaan aan een gesprek deelnemen	
Ik kan mezelf vloeiend uitdrukken en ik kan de taal flexibel en efficiënt gebruiken voor sociale, academische en professionele doeleinden	
Ik kan mezelf spontaan, zeer vloeiend, precies en genuanceerd uitdrukken, ook in meer complexe situaties.	

Radboud Universiteit Nijmegen



Geef van onderstaande stellingen aan op een schaal van 1-10 (1: bijna niet - 10: heel erg) in hoeverre de volgende stellingen van toepassing zijn op uw beheersing van het NGT met betrekking tot uw **begrip**.

	Beoordeling 1-10
Ik kan bekende dagelijkse uitdrukkingen en eenvoudige woorden of zinnen begrijpen	
Ik ben bekend met veelgebruikte uitdrukkingen en ik kan met behulp van veelal hoogfrequente woorden gesprekken voeren over alledaagse zaken	
Ik kan de hoofdpunten begrijpen van onderwerpen over het dagelijks leven of onderwerpen die redelijk vertrouwd zijn wanneer er langzaam en duidelijk wordt gebaar.	
Ik kan de hoofdlijnen van complexe en langere teksten begrijpen wanneer het onderwerp redelijk vertrouwd is.	
Ik kan complexe en langere teksten begrijpen, zelfs wanneer deze niet duidelijk gestructureerd zijn en wanneer relaties impliciet zijn.	
Ik kan zonder moeite alles volgen wat ik waarneem, zelfs wanneer in een erg hoog tempo wordt gebaar, als ik tenminste ook enige tijd heb om vertrouwd te raken aan de gesprekspartner.	

Heeft u taalproblemen. Bijvoorbeeld stotteren of dyslexie?

.....

Heeft u een gehoorverlies?

Radboud Universiteit Nijmegen



.....
Beschouwt u zichzelf als lid van de Dovengemeenschap?
.....

Gaat u wel eens naar Dovenactiviteiten?
.....

Heb u andere opmerkingen over uw taalachtergrond die u denkt dat belangrijk zijn voor uw vaardigheden in de talen? Voel u vrij om deze hier op te schrijven:

.....
.....
.....
.....
.....

Hartelijk bedankt voor het invullen van deze vragenlijst!

Modificatie van de Language History Questionnaire (Source: Marianne Gullberg and Peter Indefrey (2003), Language Background Questionnaire. Developed in The Dynamics of Multilingual Processing. Nijmegen, Max Planck Institute for Psycholinguistics)

C.2 Short version (WDD, Leeuwaarden; Deaf club, Groningen)



Vragenlijst – Achtergrond

pp. Nr.: ____

De volgende vragen gaan over uw algemene achtergrond en taalachtergrond. Probeer deze zo compleet mogelijk te beantwoorden.

Achtergrond (*Omcirkel wat van toepassing is)

Leeftijd:

Geslacht: m / v*

Dominante hand: rechtshandig / linkshandig*

Wat is uw opleidingsniveau en opleiding (bijv. B Commerciële Economie):

.....

Wat is uw beroep (baan/bijbaan; bijv. student/ondernemer):

Heeft u dove ouders, dove grootouders, dove broertjes of zusjes? Zo ja, geef aan welke directe familieleden doof zijn:

Heeft u taalproblemen. Bijvoorbeeld stotteren of dyslexie?

.....

Heeft u een gehoorverlies?

.....

Taalachtergrond

Wat is/zijn uw eerste taal/talen?

Geef alstublieft aan welke andere talen u nog meer kent. Noteer de talen die u nog veel in het dagelijks leven gebruikt of voor een langere periode in het verleden hebt gebruikt. Probeer een schatting te maken van de beheersing die u heeft over elke taal. Gebruik hierbij de volgende schaal:

Niet goed 1 2 3 4 5 6 7 8 9 10 Heel goed

Taal	Spreken/ Gebaren	Luisteren/ Afzien	Schrijven (gesproken talen)	Lezen (gesproken talen)
1				
2				
3				
4				
5				



Heb u andere opmerkingen over uw taalachtergrond die u denkt dat belangrijk zijn voor uw vaardigheden in de talen? Voel u vrij om deze hier op te schrijven:

.....

.....

.....

.....

.....

Hartelijk bedankt voor het invullen van deze vragenlijst!

Modificatie van de Language History Questionnaire (Source: Marianne Gullberg and Peter Indefrey (2003), Language Background Questionnaire. Developed in The Dynamics of Multilingual Processing, Nijmegen, Max Planck Institute for Psycholinguistics)