

Multidimensionality in Sign Language Synthesis



Layout: typeset by the author using L^AT_EX.
Cover illustration: CC0 Public Domain

Multidimensionality in Sign Language Synthesis

Translation of Dutch into Sign Language of the Netherlands

Adriana Judina Corsel
11891173

Bachelor thesis
Credits: 18 EC

Bachelor *Kunstmatige Intelligentie*



University of Amsterdam
Faculty of Science
Science Park 904
1098 XH Amsterdam

Supervisor
Dr. F. Roelofsen

Institute for Logic, Language and Computation
Faculty of Humanities
University of Amsterdam
Science Park 904
1098 XH Amsterdam

Jun 26th, 2020

Abstract

This paper explores the implementation of a translator to sign language, using the JASigning avatar software. Specifically, translation of Dutch into Sign Language of the Netherlands (NGT). As Dutch has been translated to a textual version of NGT in previous research, this research translated from textual NGT to synthesised NGT, the output signed by an avatar. This is one of three papers, each focusing on a different aspect of the translator: Lexical Resources, the Signing Space, and Multidimensionality. This paper pertains to the latter, and thus attempts to implement multiple dimensions into the avatar: the manual and non-manual component, which are both necessary to properly articulate a sign. The non-manual component of a sign being, for example, a facial expression, a head shake, the posture of the signer, etc. Specifically, it attempts to implement the non-manual markers paired with interrogative and negative constructs in NGT. Although the scope of this project is limited, it could provide a decent foundation for further research.

Acknowledgements

I would like to give recognition to the developments of the projects VisiCast, eSign, Gebarennet, Dicta-Sign, and a project from TNO and NPO, that provided the material for the database, and without whom there would not have been enough data required to make one. Creating a database with thousands of entries is very laborious, so I would like to express gratitude to all organisations and project contributors involved.

I would also like to express my thanks to Dr. Inge Zwitserlood, Prof. dr. Onno A. Crasborn and Mr. Johan Ros, for their friendly cooperation and guidance towards the aforementioned projects.

Additionally, I would like to show gratitude to Dr. John Glauert and Dr. Richard Kennaway for further explaining how to handle the SiGML format and JASigning software.

Finally, I would like to thank Ms. Marijke C. Scheffener for providing the videos for the evaluation.

Preface

This paper describes an individual contribution to a larger project executed in close collaboration with Ms. Lyke D. Esselink and Ms. Shani E. Mende-Gillings (2020; 2020). The first parts of these three papers describe the overall project and were jointly written, making them largely identical. Section 1 describes the global research question, and Chapters 2 and 3 provide theoretical context and set up a hypothesis. Section 5.1 shows the overall program created, including the components of the other projects, and Section 6.1 evaluates the result. These previously mentioned Chapters and Sections and this preface have thus been written in joint collaboration with Esselink and Mende-Gillings. For more specific information about the other components of this project, the reader is advised to read the other two aforementioned papers. However, it is not necessary to have read them to understand this paper.

An introduction to and motivation for the overall project are provided at the start of Chapter 1. Chapter 4 discusses and addresses the goals specific to the research of this paper only, and Sections 5.2 and 6.2 pertain to the results and the evaluation of this research. Finally, in Chapter 7 a discussion and conclusion of this research is given. These Chapters and Sections regard individual efforts, and have therefore been written individually.

Contents

1	Introduction	6
2	Theoretical Foundation	7
2.1	History of Sign Language of the Netherlands	7
2.2	Grammar of Sign Language of the Netherlands	8
2.2.1	Phonology	8
2.2.2	Non-manuals	10
2.2.3	Syntax	12
2.2.4	Signing Space	13
2.3	Previous Research on Sign Language Translation and Synthesis	14
3	Hypothesis and Global Methodology	15
3.1	Avatar Software	15
3.1.1	The Hamburg Notation System	15
3.1.2	Signing Gesture Markup Language	17
3.1.3	JASigning	18
3.2	Lexical Resources	18
3.3	Grammatical Resources	19
3.4	Overview Components	19
4	Multidimensionality	20
4.1	Negation and Affirmation	20
4.1.1	with Non-manual Marker	20
4.1.2	with Manual Marker	20
4.2	Interrogation	21
4.2.1	Polar Questions	21
4.2.2	Content Questions	21
4.2.3	Alternative Questions	21
5	Results	23
5.1	Global Results	24
5.2	Results Multidimensionality	24
5.2.1	Negation and Affirmation	24
5.2.2	Interrogation	26
6	Evaluation	29
6.1	Global Evaluation	29
6.2	Evaluation Multidimensionality	30
6.2.1	Negation	30
6.2.2	Interrogation	30
7	Discussion	31
	Bibliography	33
	A Hamnosys Handshapes	35
	B HamNoSys Locations	36

C JASigning User Interface	37
D Evaluation Sentences	38
E Evaluation Form	39
F Global Results	41

Chapter 1

Introduction

Children have the ability to learn language from birth. They learn it subconsciously, when language is being used around them. This happens, for example, when their parents talk to them as a baby, when hearing their parents talk to friends, or through a television show in the background. Despite not being able to understand the words, they acquire language steadily regardless (Anderson, 2015).

When a child is born deaf or hearing impaired, however, this situation is interfered with. Not being able to hear sounds, Deaf children do not have access to spoken speech, and in turn are not able to benefit from oral input around them. They do, however, acquire language the same way as hearing children do (Meier, 1991). Consequently, when their parents communicate with them via sign language, this indirect language acquisition remains. The difficulty here is that most children are born into a hearing family (Peters et al., 2014). Only familiarised with spoken language, learning a new language so different from their native one poses quite the challenge for these parents. It can prove difficult for them to find sufficient resources to learn sign language. Allegedly, of all hearing parents trying their best to master *Sign Language of the Netherlands*, just five percent manage to become fluent signers.

Aim

Something to help parents to learn sign language, would be a translator. A translator in which a sentence of a spoken language is typed, and the respective sentence in sign language outputted. The goal of the project is to build the basis for a tool in which a user can enter a Dutch sentence and in return see a correct translation of this sentence in Sign Language of the Netherlands (NGT), by an animated avatar. This will prove a valuable resource for people attempting to learn a sign language and/or communicate with a Deaf person.¹ The project aims to answer the question “What are the necessary components for a system that uses an avatar to translate a Dutch sentence to Sign Language of the Netherlands?”

Part of the question, namely how to translate from Dutch to NGT, has been covered in previous research (Brinkhuijsen, 2019; Smit, 2019; Weille, 2019). Therefore, this research will focus on the production of NGT from glosses, which are textual representations of signs (e.g. HOUSE² is the gloss of the sign that means ‘house’). The goal of this project is to create a proof of concept with a limited vocabulary that allows a user to input a sentence in NGT glosses, whose translation will be signed by an avatar. This avatar would display all the attainable essential elements of the translation.

The remainder of this paper is structured as follows: Chapter 2 discusses the theoretical foundation of the workings of NGT and previous research done in the field of sign language translation and synthesis. In Chapter 3, a hypothesis for the research question is formed on the basis of Chapter 2, and a global methodology is described for the development of the necessary components. Chapter 4 focuses on the implementation of interrogation and negation non-manuals, by formulating specific goals for its development, and illustrates the implementation of these goals in detail. In Chapters 5 and 6, both global and non-manual results are presented, analysed, and evaluated. Finally, Chapter 7 delivers a summary and critical reflection on the approaches and results, and discusses open issues and future work.

¹It is common convention to use Deaf capitalised when referring to someone who is Deaf, whereas the lowercased version is used when referring to the medical condition.

²Glosses are indicated by capital letters to illustrate that these words are signed and not spoken.

Chapter 2

Theoretical Foundation

It is a common misconception that there is a universal sign language. Although similarities can be found in signs for words that are globally characterised in the same manner (e.g. the sign for ‘house’ is identical in multiple sign languages — see Figure 2.1), parallel to spoken languages, sign languages have evolved independently of each other and impose their own distinctive set of grammatical rules (Emmorey, 2001). This is clearly illustrated by American Sign Language and British Sign Language: “despite the fact that [they] are surrounded by the same spoken language, they are mutually unintelligible”, and therefore “most often named for the country or area in which they are used”, as they are so distinctively unique (Emmorey, 2001). Sign Language of the Netherlands (NGT) is no exception to this rule: it even has five official dialects (Schermer and Koolhoof, 1991).

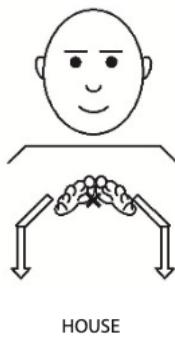


Figure 2.1: The sign for HOUSE (Baker et al., 2016)

2.1 History of Sign Language of the Netherlands

Five Deaf institutes have been established in the Netherlands between 1790 and 1911, in different regions of the country. However, the use of sign language in Dutch Deaf-education was banned between 1915 and 1980 (Schermer and Koolhoof, 1991). This ban was due to the concept of ‘oralism’, “the practice of teaching Deaf students through spoken language using amplification devices and lip-reading, to the exclusion of all sign language communication” (Morrissey and Way, 2005). It was believed that it was better for Deaf children to be educated in solely spoken language, as the usage of signs would inhibit the development of spoken language (Schermer and Koolhoof, 1991). Nevertheless, Deaf people still signed amongst each other. The combination of the previous and the fact that the five Deaf institutes in the Netherlands barely communicated with each other at the beginning of the 20th century caused the different dialects in NGT (Schermer and Koolhoof, 1991). Currently, NGT is used in Deaf-education and recognised as a language both politically and socially, but not legally (Schermer and Koolhoof, 1991). The Deaf community is taking action to legally recognise NGT as an official language in the Netherlands (de Groot, 2019).



Figure 2.2: The signing space (Baker et al., 2016)

2.2 Grammar of Sign Language of the Netherlands

NGT, like all other sign languages, is a visual-spatial language: it is “articulated by using the hands, face, and other parts of the body, and all these articulators are visible...signs are articulated on the body or in space close to the body” (Baker et al., 2016). Figure 2.2 shows the signing space used in NGT, which will be explained in Section 2.2.4.

2.2.1 Phonology

The phonology of NGT consists of four aspects: handshape, orientation, location, and movement. When discussing sign language, the term *phonology* is used “despite its sound-based etymology, in order to emphasise that the same level of structure exists in spoken language” (Morgan and Woll, 2002).

Handshape

For handshape, there is a distinction between which fingers are selected (active) and the position of these fingers (Baker et al., 2016). Examples are shown in Figure 2.3. In Figure 2.3a, all four fingers are selected; in Figure 2.3b, the index- and middle finger are selected; and in Figure 2.3c, the little finger, ring finger, and middle finger are not selected, as they are not the ‘active’ fingers in this case. Selected fingers “can make contact with the body, the head, or the other hand and arm; can adopt a special position (curved, bent, closed, spread); can move (open and close)” (Baker et al., 2016). The position of selected fingers describe: “curving of the fingers [Figure 2.3a]; spreading of the fingers [Figure 2.3b]; an aperture relation between the thumb and the selected fingers [Figure 2.3c]” (Baker et al., 2016).



(a) A ‘c’ handshape

(b) A ‘v’ handshape

(c) A ‘t’ handshape

Figure 2.3: Different kinds of handshapes (Baker et al., 2016)



EASY



SUPPOSE-THEAT

(a) The orientation for the sign EASY

(b) The orientation for the sign SUPPOSE-THEAT

Figure 2.4: Different kinds of orientations in NGT (Baker et al., 2016)

Orientation

The orientation of a sign can be described by “identifying the part of the hand that points towards the location of the sign”, the parts of the hand being the “palm, the back of the hand, the thumb side, the little finger side, the wrist side, and the tips of the fingers” (Baker et al., 2016). In Figure 2.4, the location of the two different signs is the same (the chin), but in Figure 2.4a, the palm points to the location, and in Figure 2.4b, the thumb side of the hand points to the location (Baker et al., 2016).

Location

The location of the sign is where in the signing space (Figure 2.2) the sign is articulated. There are four main locations: “the head, the upper body, the non-dominant (or weak) hand, and the neutral space”, the neutral space being the space in front of the body (Baker et al., 2016). The location of a sign is an important distinction for its meaning. An example can be seen in Figure 2.5, where Figure 2.5a displays the sign for CRUEL, and Figure 2.5b the sign for SWEET: two contrasting meanings, yet almost identical signs.



CRUEL

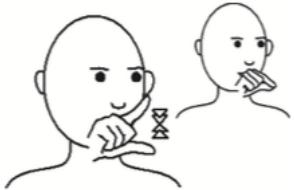


SWEET

(a) The location for the sign CRUEL

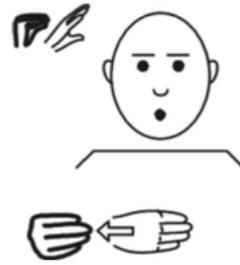
(b) The location for the sign SWEET

Figure 2.5: How location can affect the meaning of a sign — British Sign Language (Baker et al., 2016)



CHICKEN

(a) The movement for the sign CHICKEN



HUNDRED

(b) The movement for the sign HUNDRED

Figure 2.6: Different kinds of movements in NGT (Baker et al., 2016)

Movement

Movements can be divided into two types: “movements of the fingers and wrist (hand-internal movements and orientation changes) and movements of the entire hand (path movements)” (Baker et al., 2016). One example of a path movement is the sign SUPPOSE-THE in Figure 2.4b, where the whole hand moves towards the chin. Figure 2.6 shows two additional examples: Figure 2.6a illustrates hand-internal movements (only the fingers move), and Figure 2.6b displays a combination of a path- and a hand-internal movement: the fingers close while the entire hand moves (Baker et al., 2016).

2.2.2 Non-manuals

Non-manuals are composed of “form elements that relate to the posture of the body and the head, facial expressions, and certain movements or configurations of the mouth”, and are vital components of a sign (Baker et al., 2016). Figure 2.7 illustrates four different types of non-manuals: mouthings, mouth gestures, a head shake, and general facial expression. Figure 2.7a shows the identical manual component for the signs BROTHER and SISTER, the distinction lies in whether the signer uses the mouthing ‘broer’ (brother) or ‘zus’ (sister). Figure 2.7b displays the mouth gesture accompanying the sign IDIOT: “a lax



(a) The sign BROTHER and SISTER (Verlinden and Zwitserlood, 2002)



(b) Non-manuals for the sign IDIOT (Baker et al., 2016)



(c) Non-manuals for the sign WANT-NOT (Gebarencentrum, 2020)



(d) Non-manuals for the sign WHY (Verlinden and Zwitserlood, 2002)

Figure 2.7: Non-manuals in NGT

tongue hanging slightly out of the mouth while some air is being blown out” (Baker et al., 2016). Figure 2.7c accentuates the head shake for the sign WANT-NOT¹: without it the sign would not be valid. Lastly, Figure 2.7d illustrates the slightly raised chin and furrowed eyebrows that are tied to asking a content question (WH question) in NGT.

When asking a question in certain spoken languages, including Dutch and English, speakers use intonation to differentiate a question from a statement. A sentence like: ‘I do the dishes’, can be a statement when pitch is gradually lowered, but is a question if the pitch is gradually raised. NGT is similar, except instead of adjusting intonation, signers adjust non-manuals (Baker et al., 2016). To indicate a sentence is a polar question, eyebrows are slightly raised, and the head tilted slightly forwards (De Vos et al., 2009). Examples (1) and (2) show the difference between a statement and a polar question, to exemplify that solely the addition of non-manuals can completely change the meaning of a sentence. The non-manual marker changes Statement (1) into Question (2).

(1) ***Statement***

IK AFWASSEN
I do.the.dishes

‘I do the dishes’

(2) ***Polar Question***

IK AFWASSEN (accompanied by raised eyebrows and a head pushed slightly forward)
I do.the.dishes

‘I do the dishes?’ (accompanied by a raise in pitch)

The vital importance of non-manuals is also visible in negation. Regarding negative constructions, NGT is a non-manual dominant sign language. This means the non-manual marker is more important than the manual one, and the latter is usually omitted (Baker et al., 2016). When a manual marker is used, however, this is usually for emphasis, which can be seen in Examples (3) and (4). A signer would use Sentence (3) to clarify that they are not going to the zoo tomorrow. They would use Sentence (4), however, to clarify the same point more sternly, by the addition of the sign NOT (Koolhof and Schermer, 2009). As a final note, the non-manuals discussed in the previous paragraphs are not an exhaustive list. Other non-manuals include (but are not limited to): a change of posture to fit the situation, a head nod for affirmation, a puffing of the cheeks to indicate that the subject matter is large.

(3) ***Negative Construction***

MORGEN WIJ DIERENTUIN GAAN (with a head shake)
tomorrow we zoo go

‘We’re not going to the zoo tomorrow.’

(4) ***Negative Construction with emphasis***

MORGEN WIJ NIET DIERENTUIN GAAN (with a head shake)
tomorrow we not zoo go

‘We’re really not going to the zoo tomorrow.’

¹A gloss consisting of multiple words separated by a hyphen denotes a single sign.



(a) BOOK-FALLS



(b) PERSON-FALLS

Figure 2.8: Classifiers for the verb TO-FALL in NGT (Zwitserlood, 2003)

Classifiers

In NGT (and other sign languages), the handshape of a sign may vary with different subjects. Figure 2.8 illustrates this with the NGT sign TO-FALL. In Figure 2.8a, the handshape is flat, resembling the book that is falling. Figure 2.8b, on the other hand, displays a handshape resembling the legs of a person, to indicate that a human is falling. This process is called ‘classification’, and is restricted to localisation in the signing space and verbs of motion (Baker et al., 2016). Classifiers are an important aspect of sign languages as they simplify conversation. A signer may localize a character with the person-classifier, and later on refer to this character by simply indicating the same classifier in the previously defined location.

2.2.3 Syntax

The syntax of Dutch and NGT are inherently different. Example (5) below demonstrates the difference in constituent order, adjective order, and verb conjugation. Firstly, NGT has a basic sentence order of Subject-Object-Verb (SOV), which indicates that the subject, object, and verb usually appear in that order (Baker et al., 2016). Dutch, on the other hand, has a sentence order of Subject-Verb-Object (SVO), switching around the object and the verb. Likewise, modifiers are deployed in a different order, such as the adjective succeeding the noun in Dutch, whereas the inverse is common in NGT (Baker et al., 2016; Brunelli, 2011). Verb conjugation is not influenced by tenses, and only happens in the context of agreement (further explained in section 2.2.4). It only affects the manual component of the sign, the gloss and non-manuals do not change.

- (5) [MAN OUD]_S [HUIS]_O [LOOPEN]_V
 '[De oude man]_S [loopt naar]_V [huis]_O.'
 'The old man walks home.'

With negative and interrogative constructions, the location of the manual marker differs from the location of the respective Dutch word. A content question (shown in Example 6), has the WH-sign (signed counterpart of content question words) in the final sentence position. WH-doubling (the doubling of the WH-sign) is used often for emphasis, resulting in a WH-sign in initial and final sentence position (Baker et al., 2016). Furthermore, negative constructions (demonstrated in Example 7), have the manual marker positioned usually succeeding the subject, or at final sentence position (Oomen and Pfau, 2017).

- (6) *Content Question*

(WIE) BEURT _HEBBEN WIE
 (who) turn.has who

'Whose turn is it?'

- (7) *Negation*

IK NOOIT BROCCOLI ETEN / IK BROCCOLI ETEN NOOIT
 I never broccoli eat / I broccoli eat never
 'I never eat broccoli.'

2.2.4 Signing Space

In sign language the signing space (Figure 2.2) is used to keep track of the conversation and make communication smoother. When localising entities in the signing space, the signer will do this logically and most likely based on their own perspective. This means that if the signer sees a tree to the right of a house, the signer will localise the tree to the right of the house in the signing space. By localising entities in the signing space, many aspects of communication are simplified (Baker et al., 2016).

During conversations, people will often reference something or someone mentioned before, instead of repeating it. It is unlikely that someone would say the following: ‘Bella is a young woman and Bella is pretty. Bella had blue earrings, but Bella lost the blue earrings.’ Instead ‘Bella is a young woman and she is pretty. She had blue earrings, but she lost those.’ sounds much more natural. This is because of the use of personal and demonstrative pronouns, the former being words like ‘I’, ‘she’ and ‘it’, and the latter words such as ‘that’, ‘those’ and ‘these’. In sign languages a similar method is used, but those pronouns are replaced by a pointing gesture, called an INDEX (Baker et al., 2016). The pointing gesture is made towards the place (locus) in the signing space that the entity being referred to has been localised or where they are in the surrounding space (e.g. for ‘I’ the signer simply points towards themselves). Figure 2.9 shows the loci that are most commonly used in sign language. Loci 1 and 2 are always present (the signer and the interlocutor), while loci 3a and 3b are used for referents that do not have to be present. So the sentence ‘I like you’ would be translated in sign language to INDEX₁ INDEX₂ LIKE. An INDEX can also be used to distinguish between far away (*there*) and nearby (*here*) locations. The former is performed by making a pointing gesture short and down-ward and the latter by making the gesture arc-shaped, longer and forwards.

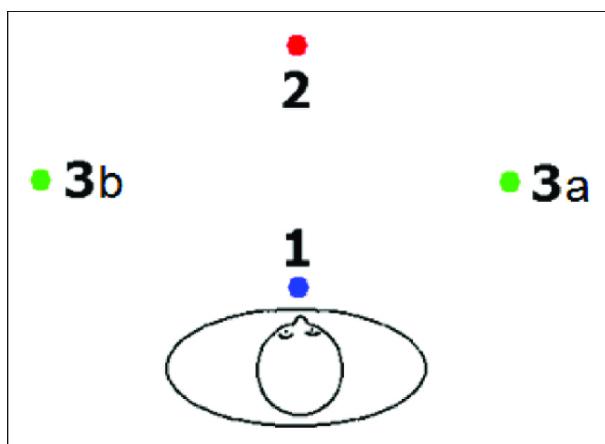


Figure 2.9: Indices in the signing space (Pfau et al., 2018)

Another important use of the signing space originates from the fact that in general, sign languages make little use of adpositions (prepositions and postpositions) to convey temporal (e.g. *before the meeting*), spatial (e.g. *on top of the house*) and abstract relations (e.g. *I did it for him*) between entities. Instead, these relations are mostly made clear using the signing space (Baker et al., 2016). For example, when signing ‘The boy walks to the cinema’, the sign for ‘walks’ is performed from the locus of BOY in the direction of the locus of CINEMA. So what is actually being signed is BOY INDEX_{3a} CINEMA INDEX_{3b} WALK_{3a,3b}, where the direction of movement and orientation of WALK makes the use of a sign for ‘to’ unnecessary. Whilst most sign languages have signs for expressing temporal relationships, the signing space can also be used instead. In NGT in order to sign ‘before the meeting’, a so-called ‘timeline’ is used. The two-handed sign for MEETING is made and then localised in the signing space by articulating INDEX with the non-dominant hand. The index finger of this hand remains at the locus of MEETING whilst the dominant hand moves from that location towards the signers body, indicating ‘before’. Had the movement been made in the other direction, so away from the signer’s body, then it would have meant ‘after the meeting’.

As shown earlier in the example of WALK, in sign language a verb might need to be modified to suit the sentence. Such verbs that are adapted to suit the subject and object, or rather their loci, are known as agreeing verbs (Baker et al., 2016). In the case of WALK its direction of movement and orientation are modified; in addition to the orientation of the sign, the start and end point of its movement indicate the subject and the object. Such verbs are known as directional verbs and allow the signer to identify the

subject, verb, and object of a sentence with one gesture. For example, if the signer performed the sign for ‘help’ moving away from themselves, this would mean ‘I help you’ (HELP_{1,2}). However, if they had performed the same sign in the opposite direction, towards the signer, it would be ‘Help me’ (HELP_{2,1}). Directional verbs are not limited to involving the signer and the interlocutor, though. If the signer placed BOB to their right and HOME to their left and subsequently signed WALK from right to left, they would be signing BOB INDEX_{3a} HOME INDEX_{3b} WALK_{3a,3b}. However, not all agreeing verbs are directional verbs, as the orientation alone of a verb can also show agreement. For instance, CALL (‘roepen’) has no path movement in NGT and only has a small repeated movement from the wrist. In this case the orientation of the sign decides the subject and object. The back of the hand indicates the caller and the fingertips point towards whom they are calling. In the examples mentioned, the signs move from the subject of the sentence to the object and are orientated towards the latter. There are however also verbs where this is the other way around, such as the sign for INVITE in NGT. These verbs, whose target is the subject, are known as ‘backwards verbs’.

2.3 Previous Research on Sign Language Translation and Synthesis

In the past two decades there has been increasing interest in the use of avatars for communicating in sign language as a result of a desire to make public spaces and services more accessible to the Deaf community (Wolfe et al., 2016). Avatars have the great advantage of allowing for more flexibility, as the signs being displayed can be easily changed. Furthermore, when combined with machine translation, the avatars can also be used for (automatic) translation between sign language and text or even speech, allowing for easier communication between hearing and Deaf people. For example, TESSA and her successor VANESSA were created in 2002 and 2004, to aid a Deaf person when they went to the post office, by translating the clerks speech into British Sign Language (Cox et al., 2002; Tryggvason, 2004). In another example, a system was developed that translated German train announcements into Swiss German Sign Language to be displayed on screens at train stations (Ebling and Glauert, 2013).

The use of sign language avatars is, however, not limited to the translation between spoken and sign language. PAULA is a computer-based sign language tutor that was created in 2006 to facilitate hearing adults in learning a limited vocabulary in American Sign Language (ASL) for use at the facility they all worked at (Davidson, 2006). The evaluation conducted with the staff on the use of PAULA showed that a sign language avatar can greatly improve the learning experience of hearing adults over other methods such as videos or even face-to-face lessons with a teacher. This is because the student can learn at their own pace and have the avatar repeat signs as little or as much as is necessary — and even adjust the speed at which the avatar signs. It is also much easier for the student return to a specific sign that they have difficulty with, and just repeat that one. More recent research in Geneva in 2017 lead to a similar conclusion (Rayner et al., 2017). Based on these conclusions, a sign language avatar would be an accessible and effective method for hearing adults to learn how to speak sign language. This would be particularly of use in situations where it is essential for a hearing person to learn to speak sign language.

Chapter 3

Hypothesis and Global Methodology

The background on NGT and the previous research into using an avatar to translate and synthesise sign language in the previous chapter led to the hypothesis that an effective system consists of three fundamental components:

1. software for sign language synthesis;
2. lexical resources for encoding signs;
3. grammatical resources for structuring sentences

The following Sections discuss each of these components in detail.

3.1 Avatar Software

When it comes to synthesising sign language, two research groups have produced the most promising results: PAULA and Java Avatar Signing (JASigning) (Davidson, 2006; Elliott et al., 2010). However, as the research for PAULA is very specific to ASL, and not open source, this project utilised the JASigning software.

In order to animate sign language using JASigning, the signs must be encoded in Signing Gesture Markup Language (SiGML), which is an XML application. SiGML is based on the Hamburg Notation System for Sign Languages (HamNoSys).

3.1.1 The Hamburg Notation System

HamNoSys is an alphabetic system for transcribing signs using the five components handshape, orientation, location, movement, and non-manuals, as described in Sections 2.2.1 and 2.2.2. An example of the HamNoSys notation for HOUSE can be seen in Figure 3.1.

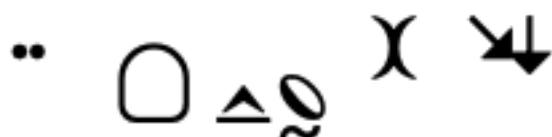


Figure 3.1: The HamNoSys Notation for HOUSE ((Hanke, 2004))

Handshape¹ is determined by the general shape of the hand (e.g. fist or open) and the position and optional bending of the thumb. Additionally, the position and optional bending of individual fingers can be specified (see Appendix A). Orientation describes the direction of the extended fingers² (or if they were to be extended) (see Figure 3.2a) and the direction of the palm³ relative to them (see Figure 3.2b). The location consists of two components, the first determines where in relation to the body⁴ and the

¹The second symbol in Figure 3.1.

²The third symbol in Figure 3.1.

³The fourth symbol in Figure 3.1.

⁴Omission signifies neutral space.

second⁵ determines at what distance from the body the sign is performed (see Appendix B). In the case of two-handed signs, the location can also describe the relation of the two hands to each other. Actions describe in-place and path movements of hands⁶, but can also describe the non-manual component of the sign (see Figure 3.2c). It can also be specified whether the actions are performed sequentially or simultaneously. Two-handed signs are indicated by the symmetry symbol⁷ at the start of the description and specifying exceptions if the hands do not copy each other exactly (Hanke, 2004).

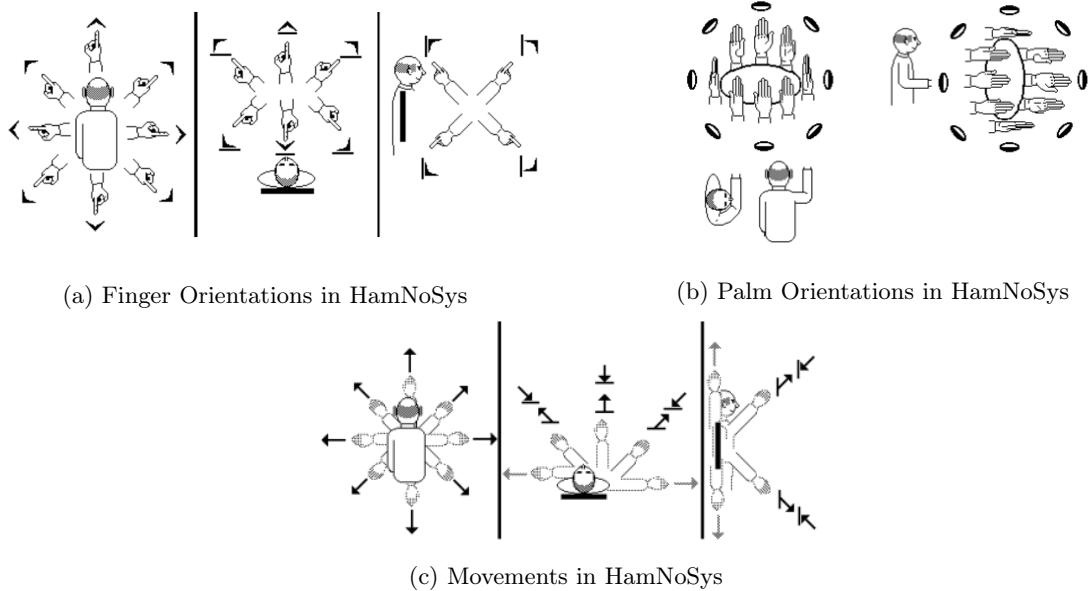


Figure 3.2: Orientations and movements in HamNoSys ((Hanke, 2004))

The HamNoSys notation of a single sign starts with a description of the initial posture followed by the possible actions that are performed sequentially or simultaneously in order to change that posture. A posture consists of a description of the aforementioned components in order, meaning that the initial handshape is first, followed by the orientation of the fingers and palm, followed by an optional body part and relative location to it, followed by an optional movement (Hanke, 2004).

The great advantage of using HamNoSys is that it does not rely on the conventions of a sign language, as these differ from country to country. This means that HamNoSys can be used to describe any sign language, which is why it is one of the most widely used transcription systems. A disadvantage of HamNoSys, however, is that it mainly focuses on the manual components of a sign. The non-manual aspect is underdeveloped, which means that the non-manual components of a sign cannot be controlled to the same extent as the manual components (Hanke, 2004).

⁵The fifth symbol in Figure 3.1.

⁶The last two symbols in Figure 3.1.

⁷The first symbol in Figure 3.1.

3.1.2 Signing Gesture Markup Language

Signing Gesture Markup Language (SiGML) is an XML based language used for “generation of sign language performances by a computer-generated virtual human, or avatar” (Elliott et al., 2004).

There are two different types of SiGML: HamNoSys SiGML (H-SiGML) and gestural SiGML (G-SiGML). H-SiGML is based directly on HamNoSys, while G-SiGML is an extension of H-SiGML with more precise controls over the signing features. This project used H-SiGML, as the database on which the corpus is based contained definitions of signs in this format. Listing 3.1 shows an example of SiGML code.

```
1 <?xml version="1.0" encoding="utf-8"?>
2 <sigml>
3
4 <hns_sign gloss="HUIS">
5   <hamnosys_nonmanual>
6     <hnm_mouthpicture picture="hYs"/>
7   </hamnosys_nonmanual>
8   <hamnosys_manual>
9     <hamsymmlr/>
10    <hamflathand/>
11    <hamextfingerul/>
12    <hampalmdr/>
13    <hamparbegin/>
14      <hamindexfinger/>
15      <hamfingertip/>
16      <hamplus/>
17      <hamindexfinger/>
18      <hamfingertip/>
19    <hamparend/>
20    <hamtouch/>
21    <hamshouldertop/>
22    <hamparbegin/>
23      <hammovedr/>
24      <hamsmallmod/>
25      <hamarcu/>
26      <hamreplace/>
27      <hamextfingero/>
28      <hampalml/>
29    <hamparend/>
30  </hamnosys_manual>
31 </hns_sign>
32
33 </sigml>
```

Listing 3.1: SiGML code for HOUSE

The start of a block of SiGML code is indicated by the notations in lines 1 and 2, and its end is indicated by the notation in line 33. Line 4 signals that the following lines encode the sign corresponding to the gloss. The non-manual component is given in lines 5 - 7, and in this case contains a mouth picture. The manual component is given in lines 7 - 30. As explained in the previous Section, line 9 denotes that the left hand must mirror the right and that this is therefore a two-handed sign. Lines 10, 11 and 12 define the handshape, finger orientation and palm orientation. Lines 14 - 18 and 23 - 28 describe the movements in the sign, as signalled by the notations on lines 13 and 22. In the second movement the hands first move down at an angle, shaping the roof of the house, and then the wrists rotate so that the fingers face forwards, shaping the walls. The first movement, however, does not actually describe a movement. It defines the touching in line 20 as only pertaining to the index fingers and not the whole hands. Line 21 simply states the location at which the sign should be made. Finally, the result of this code can be seen in Figure 3.3.



Figure 3.3: HOUSE signed by an avatar

3.1.3 JASigning

JASigning accepts either H- or G-SiGML as input, but internally converts the former into the latter (John Glauert, personal correspondence, May 28th 2020). Moreover, JASigning has additional functionality, including control of the duration of a sign, an option for a non-manual to over-arch onto multiple signs, or the addition of pauses between signs (Glauert and Elliott, 2011; Ebling and Glauert, 2013). JASigning is accessible via a website or an applet.⁸ The user interface (UI) provides several options, which include changing the signing speed of the avatar, displaying the sign frame by frame, and showing the gloss of a sign. An image of the JASigning UI is displayed in Appendix C.

3.2 Lexical Resources

In spoken languages, there is often not merely one single translation for a word from one language to another. The same issue arises when translating between a spoken language and a sign language. An extra layer of difficulty is added to a translation in sign languages due to the use of classifiers (2.2.2). As a result, it is important to account for the fact that a word or concept should be represented in various ways in a lexicon. A system should be able to use the context of a word while computing its correct representation when multiple options are available. If, on the other hand, there are no known representations of a word, the system should still have the ability to convey the message of the sentence. In real life, signers often either improvise or fingerspell the word. As fingerspelling is less ambiguous, it is desirable for the program to resort to this rather than improvisation — which would not be an efficient way to learn the language. Lastly, the program should be able to count correctly, as learning to do so is an important aspect of studying a new language. All these resources should be contained in a database with encoded machine-readable signs.

⁸<http://vhg.cmp.uea.ac.uk/tech/jas/vhg2020/index.html>.

3.3 Grammatical Resources

As mentioned in Section 2.1, the signing space and non-manuals are essential for creating grammatically correct NGT sentences. Therefore, implementations of these two components are indispensable in a translator from glossed NGT to NGT. Through use of the signing space, sentences in NGT become much more comprehensible (see section 2.2.4). The addition of non-manuals to gestures makes signs comprehensible,⁹ but also increases the naturalness of sign language (see Section 2.2.2).

3.4 Overview Components

Figure 3.4 gives an overview of the expected final product and its components. This paper in particular will cover the implementation of non-manuals. For an explanation of the implementation of lexical processing and the signing space see the paper of Esselink and Mende-Gillings (2020; 2020).

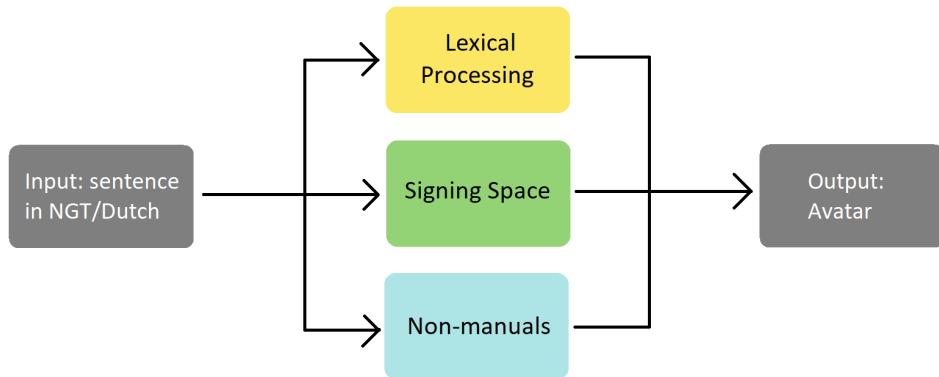


Figure 3.4: Outline Necessary Components

⁹Interesting to note is that previous research regarding sign language synthesis often neglects the implementation of non-manuals.

Chapter 4

Multidimensionality

This chapter focuses on the implementation of multidimensionality; the use of non-manual markers alongside manual markers. Specifically, the integration of non-manual markers and the JASigning avatar, which aimed to create interrogative, negative and affirmative constructions from user input. The use of non-manuals in NGT has been introduced in Section 2.2.2, and the following Sections specify manual and non-manual use in interrogative, negative and affirmative constructions, as well as explaining how this was realised in the avatar. The implementation utilised the JASigning avatar with H-SiGML input (see Section 3.1), since the database entries were based on HamNoSys, and consequently quickly and effortlessly converted to this SiGML type. Moreover, H-SiGML accepts a wide array of non-manuals, which allowed all attempts to integrate the necessary non-manuals.¹ Adding one code line containing a non-manual to the code once creates a slight movement, whereas adding it multiple times creates a larger movement.

4.1 Negation and Affirmation

Negation and affirmation go hand in hand, as they both mark the polarity of a sentence. In NGT they involve two cases: negation or affirmation solely using non-manual markers, and negation or affirmation with co-occurring manual and non-manual markers. The following Sections discuss negation and affirmation regarding both types, and how they were implemented.

4.1.1 with Non-manual Marker

As NGT is a non-manual dominant sign language (see Section 2.2.2), polarity of a sentence can be marked solely using non-manuals. Affirmation can be marked by repeated head nods, and negation is marked by a head shake (Oomen and Pfau, 2017). For the program to recognise these constructs, they needed to be able to be expressed in the user input, requiring negation and affirmation markers. When translating Dutch to NGT glosses, Brinkhuijsen (2019) indicated the scope of negation using a special bracket notation utilising the letter *n* for negation, demonstrated in Example (8). Continuing this idea, negation and affirmation have been represented identically, using an *n* for negation, and an *a* for affirmation. When encountering these brackets, everything enclosed within needed to be negated using a head shake, or affirmed using head nods.

(8) *negation indicators ‘n(’ and ‘)n’*

n(MORGEN WIJ DIERENTUIN GAAN)n
n(tomorrow we zoo go)n

‘We’re not going to the zoo tomorrow.’

4.1.2 with Manual Marker

Similar to the example WANT-NOT in Section 2.2.2 there are five other verbs with a mandatory head shake:

¹An overview of the inputtable non-manuals can be found in the file *sigmlnonmanual.dtd*.

- NOT-ABLE-TO ('kan-niet'),
- NOT-ALLOWED-TO ('mag-niet'),
- NOT-NECCESARY ('hoeft-niet'),
- NOT-SUCCEEDING ('lukt-niet'),
- DO-NOT ('doe-niet') (Koolhof and Schermer, 2009).

When using these verbs positively, two are paired with affirmative head nods: ABLE-TO, and SUCCEEDING (Koolhof and Schermer, 2009). When the program encountered these words they had to be immediately appropriately negated or affirmed.

Furthermore, the manual negative markers of NGT "include negative particles ('not'), negative adverbials ('never'), negative completives ('not yet')" as well as "n-words (e.g., 'nothing, nobody')". The avatar therefore needed to negate these words also, the scope of this negation extending to the end of the sentence (Oomen and Pfau, 2017).

If, however, a user entered the negation indicators manually, as discussed in the previous Section, the program forwent adding them a second time. This aimed to withhold interfering with the output the user expected.

4.2 Interrogation

There are three types of interrogative constructions: polar questions, content questions, and alternative questions. During the implementation of each of these, two components needed to be considered: the manual and non-manual component. The following Sections 4.2.1 - 4.2.3 discuss each question type, and the corresponding elements that needed avatar implementation.

4.2.1 Polar Questions

As described in Section 2.2.2, a polar question is paired with raised eyebrows, while pushing your head slightly forward. Moreover, there is no manual marker present (De Vos et al., 2009). The avatar therefore required an option to add the polar question non-manual markers (polar NMM) to the output. This option became a question mark ('?') at the end of the sentence, since it is the most intuitive.

4.2.2 Content Questions

As pointed out in Section 2.2.2, the facial expression during content questions is a slightly raised chin and furrowed eyebrows. Likewise, it stated that the WH-sign is in final sentence position, and possibly in initial sentence position as well (Baker et al., 2016). The avatar thus needed to be able to accept both of these options, while adding content question non-manual markers (WH NMM). The program had to integrate the WH NMMs when it encountered a question mark in conjunction with a WH-sign as the final sign, since the WH-sign is assuredly in the final sentence position.

Furthermore, a content question is often succeeded by a general question article: the PALM-UP movement (PU movement) (De Vos et al., 2009). The avatar therefore needed to be able to integrate this into the content question, by switching from the WH-sign to the PU movement in one smooth motion.² In the database, the PU movement is encoded with a set starting and end location, resulting in a convulsive movement when signed after a WH-sign. This problem was solved by fetching the final location, orientation and handshape of the WH-sign, and substituting it with the initial location, orientation and handshape of the PU movement, causing a smooth transitioning.

4.2.3 Alternative Questions

There is very little research into how alternative questions (as 'coffee or tea?') exert themselves in NGT, making substantiated decisions towards these difficult. To not disregard this type of question, some decisions were determined based on research, and some characteristics were assumed, to at least make a temporary implementation.

²Note, however, as PALM-UP is a general question article, it can also be used after a polar question. As the appendix examples of research of De Vos show, this is significantly less common, ergo disregarded during implementation in the interest of simplicity (De Vos et al., 2009).

When comparing Italian Sign Language and NGT, Brunelli (2011), using a limited database of alternative questions, states that NGT uses a polar facial expression during alternative questions. American Sign Language (ASL), on the other hand, uses both a content question expression when the alternative question is exclusive (asking someone if they want coffee, tea, or something else), and a polar question expression, which would turn the alternative question inclusive (if the question would be to choose either coffee or tea) (Zeshan, 2006). Brunelli argues NGT functions comparably, although her database is too limited to draw any conclusions. Since the avatar did not take into account context, it was especially difficult to make decisions taking this into account, hence for the purposes of the algorithm the avatar defaulted to polar question non-manuals.

Brunelli also states that if a signer would sign: COFFEE OR TEA?³, a signer would tilt their head left when signing coffee, and tilt their head right during the signing of tea, to make a clear distinction between the two. To increase this distinction, the avatar turned her body to locate the two options in the signing space (see Section 2.2.4), as this is more noticeable for a user than a head tilt.

Although nothing is established for NGT, certain Sign Languages make use of a manual marker when asking alternative questions. ASL has a WHICH sign, between the two options or in final sentence position, resembling ‘or’. Japanese Sign Language utilises PU movements, signed after an option at opposite sides of the signer (Zeshan, 2006). The avatar produced something similar to the latter, namely raising both hands simultaneously at both sides of the body, making an alternating up-and-down movement. This happened at the end of the sentence, comparable to the PU movement of content questions. Lastly, inadequate information is available about the presence of the sign OR in NGT. The avatar signed it nevertheless, aiming to clarify the alternative question.

³Here, a question mark is used to represent the non-manual markers to clarify it is a question, although NGT does not use a question mark.

Chapter 5

Results

The following Chapter discusses the resulting program. Section 5.1 gives a general overview of the complete pipeline, made conjointly with Esselink and Mende-Gillings, and subsequently discusses the specific results of the multidimensionality aspect.

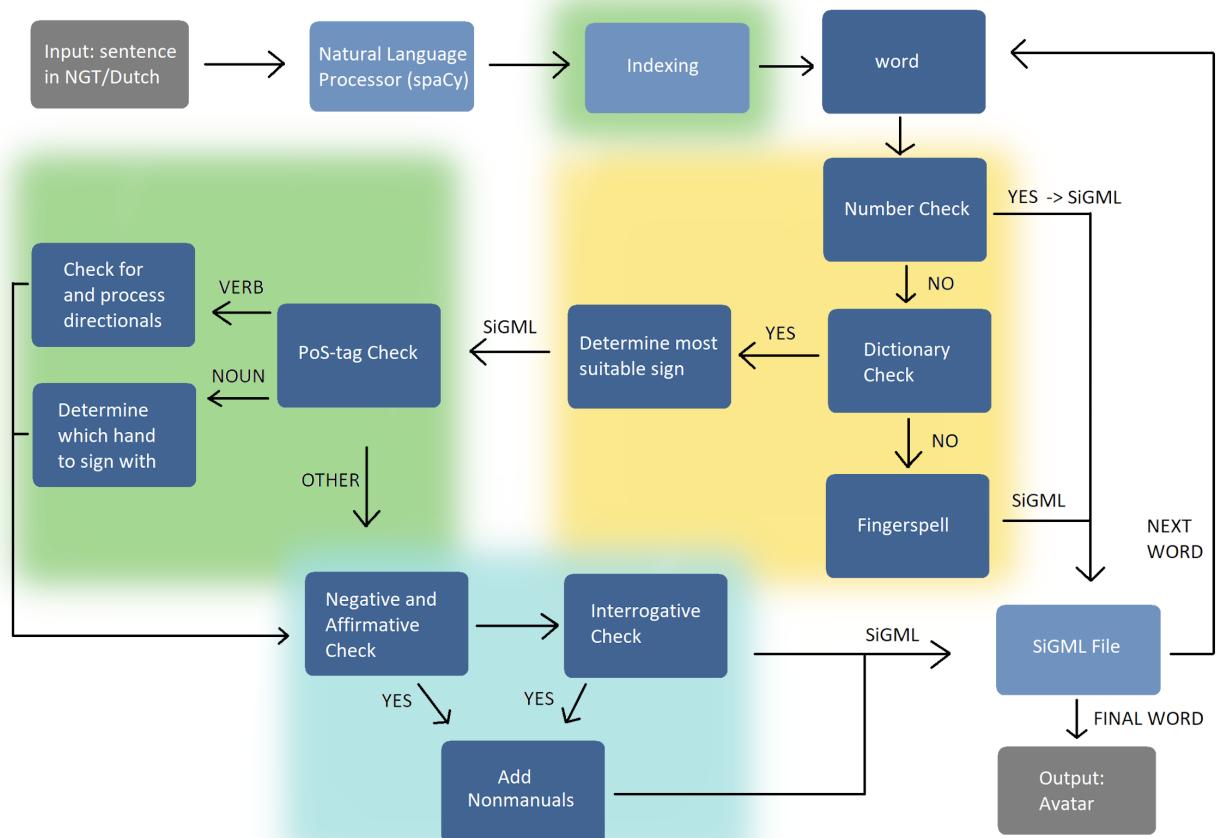


Figure 5.1: Pipeline of the entire program (the components part of lexical resources, the signing space, and non-manuals are indicated respectively by the yellow, green and light blue background squares)

5.1 Global Results

The pipeline in Figure 5.1 shows the overall mechanism of the final system.¹ The program starts with an input sentence that can either be in NGT or an intermediate form of Dutch and NGT. It also accepts sentences in Dutch — as it produces signs in the order of the sentence — however, the output might not be in grammatically correct NGT. In addition, the program is not able to process multiple sentences at a time. Table 5.1 explains which input will produce correct versus incorrect output sentences. The first incorrect input sentence, MAN HUIS LOOPT, is incorrect due to the conjugation of LOPEN to LOOPT. The second sentence MAN NAAR HUIS LOPEN is incorrect due to the word NAAR (to), since NGT makes little use of adpositions. Instead, indices may be used to indicate relations between entities, as can be seen in the second correct input sentence MAN INDEX_{3A} HUIS INDEX_{3B} LOPEN. Finally, the sentence DE MAN HUIS LOPEN contains the erroneous article DE, which should be omitted.

Correct Input	Incorrect Input
MAN HUIS LOPEN MAN INDEX _{3A} HUIS INDEX _{3B} LOPEN	MAN HUIS LOOPT MAN NAAR HUIS LOPEN DE MAN HUIS LOPEN

Table 5.1: Examples of correct and incorrect input of the Dutch sentence ‘De man loopt naar huis.’ (The man walks home.).

The program consists of three main steps: pre-processing, translation of words into SiGML, and communication with the avatar. The first step of the program is the pre-processing of the sentence. SpaCy² is a natural language processor used to acquire the Part-of-Speech (PoS) tags and dependencies of the words in the sentence. In the case that the sentence is not in NGT, indices are added where necessary. The next step is completed for each individual word in the sentence. In order to ensure that the avatar knows which signs to produce, words are assessed on type. The program first checks whether the current word is a number and, if so, adds its SiGML to the file. Next, in the case that the word is not in the dictionary, it will be fingerspelled by the avatar, and otherwise an algorithm is applied to determine which sign in the dictionary is the most suitable to use and retrieves SiGML. After the corresponding sign has been chosen, the PoS-tag of the word is evaluated. If the word is a noun, the program checks with which hand to sign the word, and if the word is a verb, the program checks whether it is a directional verb and thus needs to be adapted. The final check determines whether the sentence is interrogative, negative, or affirmative, so that the appropriate non-manuals can be added to the sign. Once the word has been processed, the sign is added to a file which collects the SiGML of all the words in the sentence. Finally, after every word in the sentence has been processed, the file is sent to the avatar which then produces the signs in order.

5.2 Results Multidimensionality

The following Section discusses the results of multidimensionality, namely when non-manuals are integrated and how this is achieved. As can be seen in Figure 5.1, they are added as the last steps in the pipeline. As a result, the sentence was processed normally, and non-manuals did not interfere with the manual component of the signs. On a different subject, spaCy has been used to retrieve PoS-tags when necessary, for example when distinguishing a verb from a noun.

5.2.1 Negation and Affirmation

Non-manual Marker

Examples of correct input sentences with non-manual negation are shown in Table 5.2. A user could put balanced negation brackets anywhere in the sentence, and the avatar negated the corresponding sentence parts by shaking her head, doing so with a tiny pause between each sign. This is because the signs are accompanied by the non-manuals, which are signed with a slight intermission before moving on to the next, thus causing a tiny pause between the non-manuals as well. Figure 5.2 demonstrates the avatar output during this negation: a head shake.

¹For instructions on how to install and use the program, see: <https://github.com/LykeEsselink/SignLanguageSynthesis>.

²<https://spacy.io>

Correct Input	Corresponding Avatar Output
n(MAN LOPEN)n	n(MAN LOPEN)n ('The man doesn't walk.')
MAN n(LOPEN)n	MAN n(LOPEN)n ('The man doesn't walk.')

Table 5.2: correct input sentences with negation



Figure 5.2: head shake during WALK

Non-manually affirming worked exactly the same as negating, only with affirmation brackets instead of negation brackets. Table 5.3 presents one correct input sentence, the parts to be affirmed enclosed in affirmation brackets. The resulting head nods can be seen in Figure 5.3. Likewise the head shake, the head nods were interrupted by tiny pauses between the signs.

Manual Marker

There were multiple options for user input in regards to negation with a manual marker, displayed in the first row of Table 5.4. Firstly, the user could input a sentence solely containing a manual marker like NEVER, after which the program proceeded to detect the manual marker and estimate the scope of

Correct Input	Corresponding Avatar Output
a(MAN LOPEN)a	a(MAN LOPEN)a ('The man does walk.')

Table 5.3: Example of a correct input sentence with affirmation



Figure 5.3: Headnod during WALK

negation. Alternatively, the user could define the scope manually using the negation brackets, since the program always considered manual brackets more important than automatically detected ones. Within the program there was a list containing manual negative markers, which the program used when detecting them in the input. Additionally, if user input contained both a verb that needs to be non-manually negated and NOT, the program automatically conjoined and negated them, as can be seen in the second row of Table 5.4. The six verbs with mandatory head shake, mentioned in Section 4.1.2, were processed like this, although only two were signed, since the database only contained NOT-NECESSARY and NOT-ALLOWED-TO.

Correct Input	Corresponding Avatar Output
MAN NOOIT LOPEN MAN n(NOOIT LOPEN)n	MAN n(NOOIT LOPEN)n (‘The man never walks.’)
MAN LOPEN MAG NIET MAN LOPEN NIET MAG MAN LOPEN MOGEN NIET MAN LOPEN NIET MOGEN MAN LOPEN MAG NIET MAN LOPEN n(MAG NIET)n MAN LOPEN MOGEN: _MAG NIET MAN LOPEN n(MOGEN: _MAG NIET)n	MAN LOPEN n(MOGEN: _MAG NIET)n (‘The man is not allowed to walk.’)

Table 5.4: Correct input sentences with negation

As for affirmation, the list of possible manual affirmation markers was significantly shorter. Furthermore, despite not conjoining them automatically, the program detected the two verbs with mandatory head shake, and affirmed them accordingly. An example of input containing one of these verbs can be seen in Table 5.5, showing the affirmation of ALLOWED-TO, and possible input.

Correct Input	Corresponding Avatar Output
MAN LOPEN MOGEN: _MAG_WEL MAN LOPEN a(MOGEN: _MAG_WEL)a MAN LOPEN MAG_WEL MAN LOPEN a(MAG_WEL)a	MAN LOPEN a(MOGEN: _MAG_WEL)a (‘The man is allowed to walk.’)

Table 5.5: Example of correct input sentences with affirmation

5.2.2 Interrogation

Polar Questions

The correct user input for a polar question, as shown in Table 5.6, is the query concluded with a ‘?’ . The ‘?’ served as a question indicator, to inform the program that it is dealing with a polar question. The corresponding output was the same query with partial polar NMM, as the avatar only produced a natural facial expression; she pushed her head in front of her body and put it back above her torso for every sign, resulting in a head bobbing effect. As this was a great distraction from the actual non-manuals, as well as the signs, this non-manual marker was discarded. Another problem was a glitch in the avatar program when using polar NMM, causing the glitched facial expression shown in Figure 5.4b. This glitch occurred whenever a query was not signed consecutive to neutral input, and caused the avatars eyes to disappear. In both pictures of Figure 5.4, the avatar is provided the exact same lines of code, the only difference being that one was signed right after a query without non-manuals, and one was not. When the first input was a query with a neutral facial expression, followed by a polar question query, the intended facial expression appeared on the avatar, as Figure 5.4a. However, if this was not the case, the glitched facial expression appeared, seen in Figure 5.4b.

Correct Input	Corresponding Avatar Output
MAN LOPEN?	MAN LOPEN (+NMM) (‘Does the man walk?.’)

Table 5.6: Correct input for a polar question



(a) Intended facial expression (b) Glitching facial expression

Figure 5.4: Different faces for polar question non-manuals

Content Questions

Table 5.7 shows the correct input for a content question: a WH-sign at final — and possibly initial — sentence position, closed with a ‘?’ . Similar to polar questions, the ‘?’ serves as a question indicator, and the WH-sign distinguished a content question from a polar question. The same Table shows the corresponding avatar output: the sentence signed with non-manual markers and a PU movement. The actualisation of the latter is presented in Figure 5.5; you can see a WH-sign in Figure 5.5a, and the PU gesture in Figure 5.5b. Modifying the PU sign as described in Section 4.2.2 worked for all WH-signs except WHO. The structure of the HamNoSys notation of WHO did not allow the modification to work, and it had to be treated separately from the other WH-signs. The correct PU movement for WHO was hard-coded and added to the database, causing the program to output the correct PU movement when encountering any WH-sign.

Correct Input	Corresponding Avatar Output
MAN LOPEN WAAROM?	MAN LOPEN WAAROM-PALM_UP (+NMM) (‘Why does the man walk?’)
WAAROM MAN LOPEN WAAROM?	WAAROM MAN LOPEN WAAROM-PALM_UP (+NMM) (‘Why does the man walk?’)

Table 5.7: Correct input sentences for a content question



(a) WHY (b) PALM- UP

Figure 5.5: Content Question

Alternative Questions

Table 5.8 shows the correct input for an alternative question: a query containing OR, closed by a ‘,’. Using a ‘,’ instead of a ‘?’ allowed the program to distinguish a polar question containing OR (e.g. ‘Do you have three or more children?’) from an alternative question. The corresponding output was the alternative question signed with polar NMM (the aforementioned polar NMM glitch occurring). Figure 5.6 illustrates the progression of the avatar when signing WALK OR RUN: in Figure 5.6a WALK is localised at the left side of the signing space, then RUN is localised at the right in Figure 5.6b, and lastly both options are presented alternatingly in Figures 5.6c and 5.6d. When the sentence contained a subject — possibly with an INDEX — in initial sentence position or an additional verb in final sentence position, these were signed in the middle.

Correct Input	Corresponding Avatar Output
MAN LOPEN OF RENNEN,	MAN LOPEN OF RENNEN (+NMM) (‘Does the man walk or run?’)

Table 5.8: Correct input for an alternative question



(a) WALK localised left (b) RUN localised right (c) Present option one (d) Present option two

Figure 5.6: Alternative Question WALK OR RUN

Chapter 6

Evaluation

The following Chapter is distributed the same way as Chapter 5. Firstly, a global evaluation is made. This is followed by an evaluation for multidimensionality.

6.1 Global Evaluation

In order to assess the overall performance of the program, eighteen test sentences (see appendix D) were constructed and given to the program to sign in NGT. The output of the program — the signing performed by the avatar — was recorded for each sentence and stored in a database along with its meaning. Two evaluators, one native speaker of NGT and one who learned it as an adult, were asked to watch the videos and fill out an evaluation form (Appendix E). The first step of the evaluation required them to interpret the meaning of the sentence based on the signing of the avatar. In the second step they were shown what the avatar was meant to sign, followed by various questions to assess the comprehensibility of the sentence and the naturalness of the signs.

The evaluators were asked to score the comprehensibility and naturalness of each sentence on a scale from 1 to 10. Figure 6.1 shows the average results of the scoring.¹ Overall, the naturalness scores higher than the comprehensibility, with average scores of 7.31 and 6.97 respectively. Furthermore, the average scores given by the native speaker, 7.33 and 7.56, are slightly higher than the average scores given by the non-native speaker, 7.28 and 6.38, for respectively naturalness and comprehensibility. It is important to note that the feedback revealed that not all signs in the database are correct, which influenced the scores considerably.

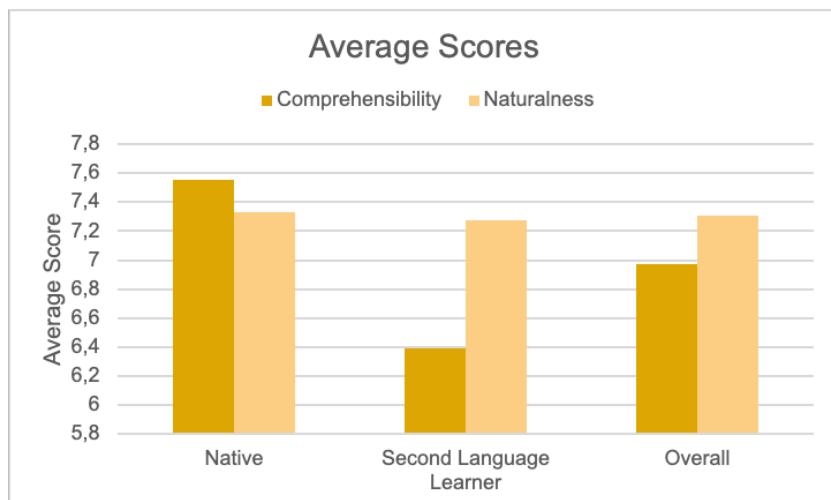


Figure 6.1: Average Scores of Comprehensibility and Naturalness

¹ Appendix F shows the overall results of the scoring.

6.2 Evaluation Multidimensionality

This next Section evaluates the integration of multidimensionality in the avatar, partially using the feedback form introduced in the previous Section (5.1), partially using a corpus signed in NGT. The corpus consists of a collection of 21 videos signed by a native signer, each video featuring a negative construct.² The corpus has thus been used to evaluate negative constructs. The feedback form contains six sentences regarding interrogative constructions, hence used to evaluate that type of construct. Although 21 videos and six sentences is not a great amount, it provides at least some basis for comparison and, in the long run, improvement.

Non-manuals appeared often in the feedback form, not exclusively regarding negation, affirmation or interrogation. Although outside the scope of this paper, the evaluators mentioned that in certain signs facial expression is crucial, and the current situation (neglecting them) makes the signs substantially less comprehensible. Furthermore, their remarks noted that various mouthings and gazes were slightly off, making the avatar appear slightly odd.

6.2.1 Negation

The following observations are made based on the corpus of videos, each video containing a manual marker, such as NEVER, NOTHING and NOBODY. In most, if not all, of the relevant videos, non-manual markers also include a ‘negative’ facial expression (e.g. pouted lips, frowned face), which is a familiar phenomenon in sign languages (Zeshan, 2006). Despite this fact, a negative facial expression was not present in the avatar. As for the non-manuals integrated in the avatar, the head shake appears in both the avatar and the videos. An obvious difference in the scope of the head shake, however, is that the avatar produced tiny pauses between signs, whereas the signer produced one continuous motion.

6.2.2 Interrogation

The Section below describes observations made based on the aforementioned feedback form. The polar NMM did not receive any critique from the evaluators, while the native speaker thought the WH NMM looked somewhat unnatural. Changing to alternative questions, neither of the evaluators pointed out the use of OR seemed out of place nor reduced the naturalness or comprehensibility of the sentence. As a final point, the sentence combining the alternative question, with adjectives (for more information see paper of Mende-Gillings (2020)) and negation, was scored with a very low comprehensibility. This was a result of the avatar signing the wrong sign, a wrong adjective placement confusing the evaluators, as well as superfluous negating, making it hard to pay attention to the actual sentence.

²An excerpt of the corpus used in the research of Brinkhuijsen (2019).

Chapter 7

Discussion

Summary

In summary, this paper explored the idea of creating a translator from Dutch to synthesised NGT. As there was previous research in translating Dutch to NGT glosses (Brinkhuijsen, 2019; Smit, 2019; Weille, 2019), this paper was mainly centered around the remaining step, namely translating from NGT glosses to synthesised NGT. Focusing on the use of dimensionality, it aimed to integrate manuals alongside non-manuals in interrogative, negative, and affirmative constructions, albeit on a basic level. For negation and affirmation this resulted in the addition of a head shake and head nods to the output. For interrogation this resulted in adding non-manuals, and keeping into account the possible manuals. These attempts were evaluated by two NGT speakers and a corpus of videos. The former remarked that the content question facial expression looked somewhat unnatural, while the latter revealed that a ‘negative’ facial expression was overlooked in negation implementation, as well as there being an undesirable pause between head shakes. Finally, during the implementation of polar question non-manuals, there were two complications. Firstly the occasional apparition of a glitch removing the eyes of the avatar, and secondly the inability to push the head of the avatar forward properly.

Open Issues

As this project was a basic exploration, there are a lot of things that could be improved upon in the future. For instance, as mentioned in the results (Section 6.2), the eyes of the avatar partially disappeared when she used polar question non-manuals. This is unwanted, considering this is a great distraction to the output, and it is not conveying the right non-manuals. Since this glitch did not seem to occur when signed right after a neutral facial expression, one solution would be to add a ‘resting sign’ at the start of the output: a sign during which the avatar would stand in a natural position, not conveying any non-manuals. This would then serve as a neutral sign in front of the output, and the avatar would create the right non-manuals.

Furthermore, negation requires some improvement, as mentioned in the evaluation. It is currently lacking a ‘negative’ facial expression, as this was overlooked during implementation. Making an ‘innocent’ expression while signing I NEVER LIE, or a ‘clueless’ expression during I NO-ONE SAW is something that happens very naturally when signing. When comparing the avatar output to the evaluation videos, it was prominently missing. In the 21 evaluation videos alone, however, there was a wide range of different facial expressions, so using the correct ‘negative’ facial expression for the correct situation might prove to be quite a challenge. Evidently, there needs to be looked into research as to when different ‘negative’ facial expressions are made. Nevertheless, even if the avatar would use one, general expression (for example frowning), it would assumably improve negation significantly, without much effort.

As became clear in the feedback form, content question non-manuals looked unnatural to the native speaker. Asking further questions why the native speaker thought this was the case, whereas the second language learner did not, will improve understanding of what caused the the unnaturalness, and how it can be resolved. Furthermore, the evaluators noticed the mouthings were not clear. If these two non-manual elements are refined, this will prove valuable progress for the naturalness of the avatar.

The current implementation utilises H-SiGML to instruct the avatar, instead of G-SiGML. The former was easy compatible with the database, whereas the latter has more functionality. One additional function, for example, is non-manuals being able to over-arch onto the multiple signs, which would get rid of the unwanted pauses between head shakes and nods, and the head bobbing effect. As previously

mentioned, JASigning works with G-SiGML internally already, hence it would not be very inconvenient for the program to convert it in advance, especially since it would bring about more functionality. Although for the remaining purposes of this implementation the extra functionality was not inadequate, it might prove useful to have access to whilst further developing the program.

Lastly, a limitation to the current taken approach is spaCy, the natural language processor used to acquire the PoS-tags and dependencies. Due to the fact that spaCy is created to process Dutch, it does not PoS-tag NGT glosses flawlessly. The current program is created assuming these tags are always correct, anticipating to switch to a better alternative in the future. This could be, for instance, a translator from Dutch to NGT glosses, which would most likely have access to the correct PoS-tags and dependencies after translation. When the program is connected to this, it would be able to receive more correct information, and, in turn, create output that is more correct.

Further Research

A natural progression is to attempt to implement more non-manuals, for example gazes and facial expressions. In the feedback form, the absence of these was mentioned often as unnatural. In the current situation the avatar looks blankly to the user at all times. This is uncommon in NGT (and sign languages in general), as they are such a visual language. Making the avatar look to what they are locating in the signing space, and creating a ‘sad’ facial expression when signing SAD would result in more natural looking output, and, in turn, an improved avatar.

As mentioned in Section 2.2.4, the signing space is used to keep track of and simplify conversation, doing so by placing entities in the signing space for future reference (Baker et al., 2016). As a result, context is very important to keep in mind, as the previously signed sentence influences the following. In the current situation, the avatar only allowing one input sentence, such use of the signing space is impossible. Therefore, increasing the amount of allowed input sentences, as well as keeping in mind the context, would move the program forward significantly, as it would mean to sign more effectively and naturally.

Three areas that have been attempted to implement in a translation system are multidimensionality, the signing space, and lexical resources. More specifically, interrogation, negation, directional verbs, classifiers, fingerspelling and counting (Esselink, 2020; Mende-Gillings, 2020). There are multiple areas these to not cover, for example role shifting, the majority of localisation, time lines, to name a few. It would be a great asset to the program if they would be integrated, as it would enhance the ability of the avatar to use grammatical constructs of NGT.

This paper mainly focused on translating NGT glosses to synthesised NGT, and it established a program that integrated basic grammatical constructs apparent in NGT. Previous research had established a program translating Dutch to NGT glosses, albeit with a smaller database (Brinkhuijsen, 2019; Smit, 2019; Weille, 2019). It would be interesting to connect the product of the researches, to create a translator going from Dutch to synthesised NGT, as per the research question of this paper.

Conclusion

In conclusion, this study has discussed the necessary components for a translator from Dutch to Sign Language of the Netherlands: an avatar, a database, and a grammar. A secondary aim of the study was to investigate implementation of the latter, by creating a basis for a translation program from NGT glosses to synthesised NGT. The part of the program this paper focused on was the implementation of interrogative, negative and affirmative constructions. Notwithstanding that there is much room for improvement, the current avatar output is comprehensible to NGT speakers, and provides a decent basis for further developments.

Bibliography

- John R. Anderson. *Cognitive Psychology and Its Implications*. Worth Publishers, 8 edition, 2015. ISBN 1464148910.
- Anne Baker, Beppie van den Bogaerde, Roland Pfau, and Trude Schermer. *The linguistics of sign languages: An introduction*. John Benjamins Publishing Company, 2016.
- Sander Brinkhuijsen. *Interlingual Machine Translation for Dutch to Sign Language of the Netherlands*. Bachelor's thesis, Universiteit van Amsterdam, 2019.
- Michele Brunelli. *Antisymmetry and Sign Languages: A comparison between NGT and LIS*. LOT, 2011. ISBN 9789460930676.
- Stephen Cox, Michael Lincoln, Judy Tryggvason, Melanie Nakisa, Mark Wells, Marcus Tutt, and Sanja Abbott. Tessa, a system to aid communication with deaf people. In *Proceedings of the fifth international ACM conference on Assistive technologies*, pages 205–212, 2002.
- Mary Jo Davidson. Paula: A computer-based sign language tutor for hearing adults. In *Intelligent Tutoring Systems 2006 Workshop on Teaching with Robots, Agents, and Natural Language Processing*, pages 66–72, 2006.
- Linda de Groot. Initiatiefwet: Nederlandse gebarentaal erkennen als officiële taal. *Nederlandse Omroep Stichting (NOS)*, 2019. URL <https://nos.nl/nieuwsuur/artikel/2300498-initiatiefwet-nederlandse-gebarentaal-erkennen-als-officiele-taal.html>.
- Connie De Vos, Els Van Der Kooij, and Onno Crasborn. Mixed Signals: Combining Linguistic and Affective Functions of Eyebrows in Questions in Sign Language of the Netherlands. *Language and Speech*, 52(2-3):315–339, 2009. ISSN 00238309.
- Sarah Ebling and John R. W. Glauert. Exploiting the full potential of jasigning to build an avatar signing train announcements. 2013.
- Ralph Elliott, J.R.W. Glauert, Vince Jennings, and J.R. Kennaway. An overview of the sigml notation and sigmlsigning software system. In *Sign Language Processing Satellite Workshop of the Fourth International Conference on Language Resources and Evaluation, LREC 2004*, pages 98–104, 2004.
- Ralph Elliott, Javier Bueno, Richard Kennaway, and John Glauert. Towards the integrationn of synthetic sl annimation with avatars into corpus annotation tools. In *4th Workshop on the Representation and Processing of Sign Languages: Corpora and Sign Language Technologies, Valletta, Malta*, page 29, 2010.
- Karen Emmorey. *Language, cognition, and the brain: Insights from sign language research*. Psychology Press, 2001.
- Lyke D. Esselink. *Lexical Resources for Sign Language Synthesis: The Translation of Dutch to Sign Language of the Netherlands*. Bachelor's thesis, Universiteit van Amsterdam, 2020.
- Nederlands Gebarencentrum. Mini gebarenwoordenboek, 2020.
- John Glauert and Ralph Elliott. Extending the SiGML Notation – a Progress Report. *Second International Workshop on Sign Language Translation and Avatar Technology (SLTAT)*, 2011. URL <http://vhg.cmp.uea.ac.uk/demo/SLTAT2011Dundee/>.

Thomas Hanke. Hamnosys-representing sign language data in language resources and language processing contexts. In *LREC*, volume 4, pages 1–6, 2004.

Corline Koolhof and Trude Schermer. *Van Dale Basiswoordenboek Nederlandse Gebarentaal*. Van Dale, 2009.

Richard P. Meier. Language Acquisition by Deaf Children. *American Scientist*, 79(1):60–70, 1991. ISSN 00030996.

Shani E. Mende-Gillings. *The signing space for the synthesis of directional verbs in NGT*. Bachelor’s thesis, Universiteit van Amsterdam, 2020.

Gary Morgan and Bencie Woll. *Directions in sign language acquisition*, volume 2. John Benjamins Publishing, 2002.

Sara Morrissey and Andy Way. An example-based approach to translating sign language. 2005.

Marloes Oomen and Roland Pfau. Signing not (or not): A typological perspective on standard negation in Sign Language of the Netherlands. *Linguistic Typology*, 21(1):1–51, 2017. ISSN 1613415X.

H.F.M. Peters, R Bastiaanse, J. Van Borsel, K. Jansonius-Schultheiss, P.H.O. Dejonckere, Sj. der Meulen, and B.J.E. Mondelaers. *Taalontwikkelingsstoornissen ten gevolge van doofheid*, pages 117–149. Bohn Stafleu van Loghum, Houten, 2014. ISBN 978-90-368-0796-8. URL https://doi.org/10.1007/978-90-368-0796-8{_}7.

Roland Pfau, Martin Salzmann, and Markus Steinbach. The syntax of sign language agreement: Common ingredients, but unusual recipe. *Glossa: a journal of general linguistics*, 3(1), 2018.

Manny Rayner, Irene Strasly, Nikolaos Tsourakis, Johanna Gerlach, and Pierrette Bouillon. Menusigne: A serious game for learning sign language grammar. pages 181–186, 08 2017. doi: 10.21437/SLATE.2017-32.

Trude Schermer and Corline Koolhoof. *De Nederlandse Gebarentaal*. Van Tricht Twello, The Netherlands, 1991.

B.W.H. Smit. *Rule-Based Machine Translation from Dutch to Dutch Sign Language using the Grammatical Framework*. Bachelor’s thesis, Universiteit van Amsterdam, 2019.

Judy Tryggvason. Vanessa: A system for council information centre assistants to communicate using sign language. *School of Computing Science, University of East Anglia*, 2004.

Margriet Verlinden and Inge Zwitserlood. Cki en gebarentaal. slideplayer.nl, 2002. URL <https://slideplayer.nl/slide/11272197/>.

Steven De Weille. *Automatic translation of Dutch into Dutch Sign Language with a focus on the use of tense*. Bachelor’s thesis, Universiteit van Amsterdam, 2019.

Rosalee Wolfe, Eleni Efthimiou, John Glauert, Thomas Hanke, John McDonald, and Jerry Schnepf. recent advances in sign language translation and avatar technology, 2016.

Ulrike Zeshan. *Sign Language Typology series: Interrogative and Negative Constructions in Sign Language*. Ishara Press, 2006. ISBN 9789086560011.

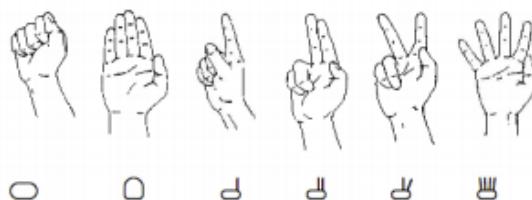
Ingeborg Elisabeth Petra Zwitserlood. *Classifying hand configurations in Nederlandse Gebarentaal (Sign Language of the Netherlands)*. PhD thesis, Universiteit Utrecht, 2003.

Appendix A

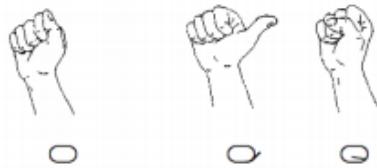
Hamnosys Handshapes

Handshapes

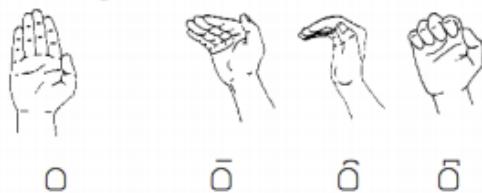
Base forms...



are combined with diacritical symbols for thumb position



and bending.



(Hanke, 2004)

Appendix B

HamNoSys Locations

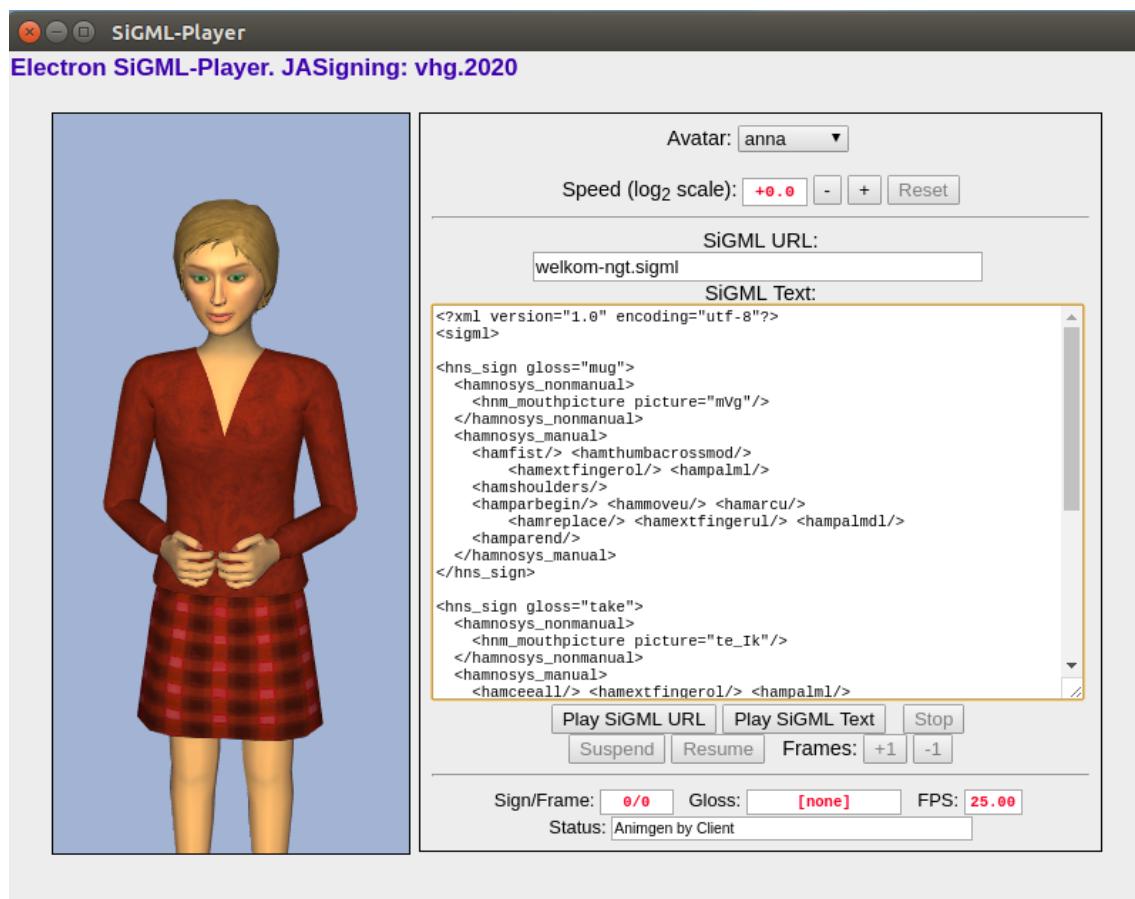
Locations - Head and Body

	left to	left side of	center of	right side of	right to
○ above head	□○	■○	○	○▪	○□
○ head	□○	■○	○	○▪	○□
□ forehead	□□	■□	□	□▪	□□
† nose	□†	■†	†	†▪	†□
○ mouth	□○	■○	○	○▪	○□
⊖ tongue	□⊖	■⊖	⊖	⊖▪	⊖□
⊖ teeth	□⊖	■⊖	⊖	⊖▪	⊖□
⌞ chin	□⌞	■⌞	⌞	⌞▪	⌞□
⌞ below chin	□⌞	■⌞	⌞	⌞▪	⌞□
⌞ neck	□⌞	■⌞	⌞	⌞▪	⌞□
⌞ shoulder line	□⌞	■⌞	⌞	⌞▪	⌞□
⌞ breast line	□⌞	■⌞	⌞	⌞▪	⌞□
⌞ belly line	□⌞	■⌞	⌞	⌞▪	⌞□
⌞ abdominal line	□⌞	■⌞	⌞	⌞▪	⌞□
⌞ eye brows	□⌞	■⌞	⌞	⌞▪	⌞□
∞ eyes	□∞	■∞	∞	∞▪	∞□
⌞ ears	□⌞	■⌞	⌞	⌞▪	⌞□
⌞ cheeks	□⌞	■⌞	⌞	⌞▪	⌞□

(Hanke, 2004)

Appendix C

JASigning User Interface



Appendix D

Evaluation Sentences

	NGT	Dutch
Lexical Resources	MAN 33 PLANT KOPEN VANAVOND KIND WIL SNOEPJES ETEN PAARD HOUDEN-VAN SPRINGEN MEISJE TENTAMEN WILLEN HALEN JONGEN BOOS BRUG OPEN KASTEEL 3571 JAAR OUD	De man koopt 33 planten Het kind wilt vanavond snoepjes eten Het paard houdt van springen Het meisje wilt het tentamen graag halen De jongen is boos want de brug staat open Het kasteel is 3571 jaar oud
Directional Verbs and Indexing	MAN LANG OUD HUIS MOOI GROOT KIJKEN IK OUD MAN LOPEN JIJ INDEX_3A GEVEN VROUW MOOI MANDARIJN HOUDEN VAN OP_AUX VRÖUW LANG OUD ZIJN VERDRIETIG MEISJE MOOI LANG VROUW MOOI BANG	De lange oude man kijkt naar het mooie grote huis Oude ik loopt naar de man Jij geeft aan haar/hem De mooie vrouw houdt van de mandarijn De lange oude vrouw is verdrietig Een mooi lang meisje en een mooie bange vrouw
Non-manuals	OUD-EN-NIEUW OLIEBOL ETEN WIE? VRACHTWAGEN BOTSEN WAAROM? NEEF NACHTMERRIE HEBBEN? JIJ 3 OF MEER DOCHTER HEBBEN? KEIZER LOPEN OF RENNEN? BOER NOOIT MELK VIES OF NOOIT EI VIES	Wie eet een oliebol op oud-en-nieuw? Waarom botste de vrachtwagen? Heeft de neef een nachtmerrie? Heb jij 3 of meer dochters? Loopt de keizer of rent de keizer? Heeft de boer nooit vieze melk, of nooit vieze eieren?

Table D.1: Test sentences provided for evaluation

Appendix E

Evaluation Form

Evaluatie Thesis

Graag invullen voor alle filmpjes

* Required

Email address *

Your email

1. Wat is de zin die U denkt dat de avatar uitbeeldde in NGT? *

Your answer

2. Wat is de zin die de avatar uitbeelde volgens het text bestand? (in NGT of Nederlands) *

Your answer

3. Vond u de zin die de avatar uitbeeldde duidelijk? *

1 2 3 4 5 6 7 8 9 10

Zeer onduidelijk Zeer duidelijk

4. Als het onduidelijk was, waar lag dat aan?

Het indexen was onduidelijk
 De vraagstelling was onduidelijk
 Het vingerspellen was onduidelijk
 Het getal was onduidelijk

Figure E.1: Evaluation Form

- Het klopte grammaticaal niet
 Er was een gebaar fout
 Het gebaar was niet goed te zien
 Er zat een gebaar in de zin die er niet in hoorde
 Other: _____

5. Vond u de zin die de avatar uitbeeldde natuurlijk? *

1 2 3 4 5 6 7 8 9 10

Zeer onnatuurlijk Zeer natuurlijk

6. Als het onnatuurlijk was, waar lag dat aan?

- De handvorm was onnatuurlijk
 De locatie was onnatuurlijk
 De nonmanuele component was onnatuurlijk
 Het klopte grammaticaal niet
 De gebaren liepen niet goed in elkaar over
 De pauzes tussen de gebaren waren te lang
 Other: _____

7. Heeft u nog andere feedback voor deze zin?

Your answer

Submit

 Page 1 of 1

Figure E.2: Evaluation Form

Appendix F

Global Results

Two figures showing the scores of comprehensibility and naturalness of the avatar.¹

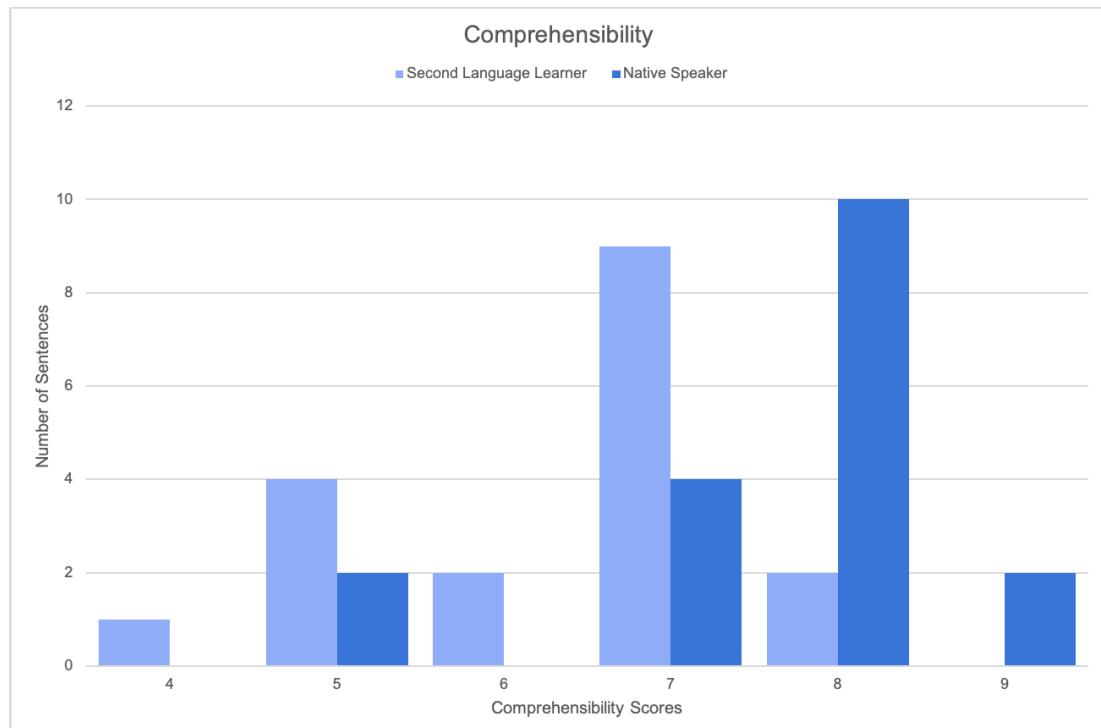


Figure F.1: Comprehensibility (Average Score of 6.97)

¹Scores that were never chosen by the evaluators are not shown.

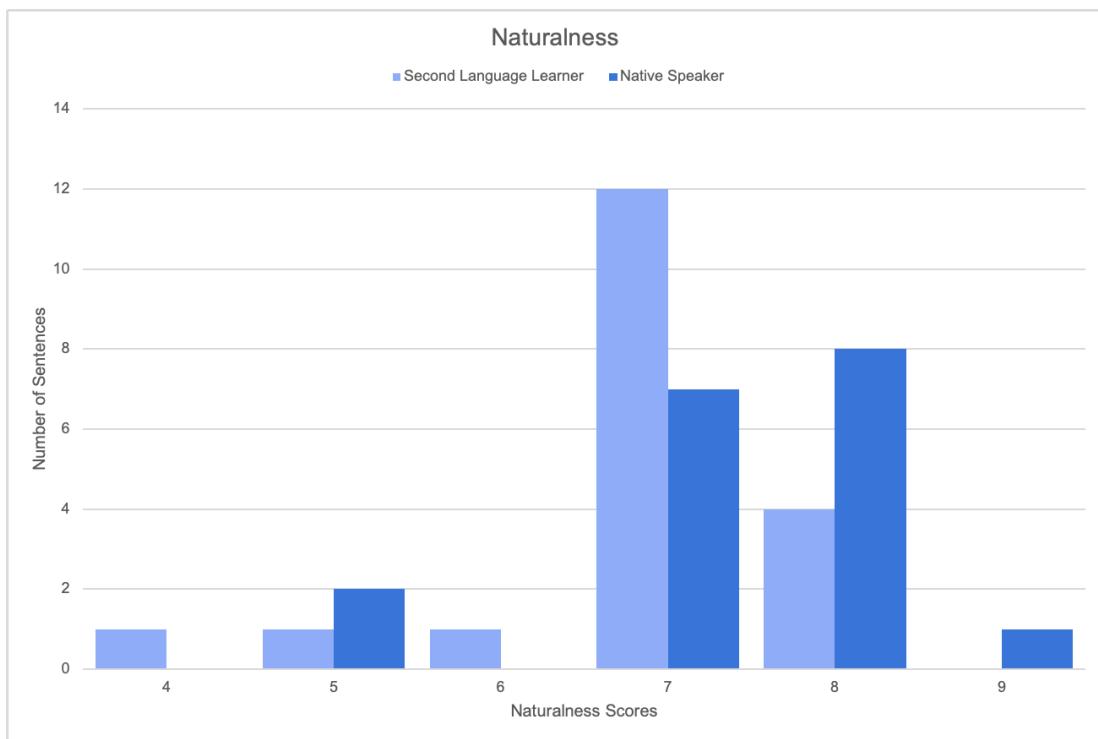


Figure F.2: Naturalness (Average Score of 7.31)