Starting to make sense:
Further developing a nonsense sign repetition task

Ulrika Klomp (University of Amsterdam)

U.Klomp@uva.nl

Abstract

The nonsense sign repetition task that was developed for Sign Language of the Netherlands in 2015 (the NGT-NSRT) is investigated further in this paper. Specifically, I look into differences in performances on the NGT-NSRT between deaf signers and CODAs, into the effect of movement complexity on the scores of the participants, and into the relationship between phoneme-based scores and binary correct/incorrect scores. It turns out that the deaf signers score significantly better than the CODAs, and that the participants score significantly worse on signs with a combined movement compared to signs with a single movement. Furthermore, phoneme-based scores and correct/incorrect scores are significantly correlated. No evidence was found for a difference in complexity between signs with a hand-internal movement and signs with a path movement. Suggestions for further research and an alternative analysis of phonological complexity, as adopted by Vink (2018), are discussed.

Keywords: nonsense sign repetition, Sign Language of the Netherlands, assessment, phonology, complexity

1 Introduction

1.1 Sign language assessment tools

In comparison to spoken languages, there are very few assessment tools available for sign languages. Moreover, a quick look at www.signlang-assessment.info (Haug 2018) shows that so far, tests have only been developed for a handful of languages: American Sign Language (ASL), British Sign Language (BSL), German Sign Language, Sign Language of the Netherlands (NGT), Australian Sign Language, French Sign Language and Italian Sign Language. According to this website and Klomp (2015), there are two main tests available for NGT, namely the NGT-OP (Baker and Jansma 2002) and the T-NGT (Hermans, Knoors, and Verhoeven 2007), and a few additional instruments that are used locally in schools for the deaf, in higher education programmes such as the interpreters programme or for research purposes only. For valid assessment of language fluency, however, this is far from sufficient. Ideally, assessment tools should be available for participants of all ages, to test both language production and comprehension at all linguistic levels (e.g. phonology, syntax, etc.), with sensitivity to language delays or impairments. At the same time, practical issues such as ease and duration of testing and scoring play an important role in the development of such tools. It is clear that the situation for sign languages is far from perfect at the moment. However, recently, this subject has received increased attention, and efforts have been made

to expand the available tools (e.g. by the SIGN-HUB project[1]). In this paper, I look into the Nonsense Sign Repetition Task (NSRT) that I developed[2] for NGT in 2015. At that time, my interests lay with the relationship between this new NSRT and a general non-word repetition task for spoken Dutch, for which I tested ten hearing (adult) children of deaf adults (CODAs). This was the first step to see how the NGT-NSRT worked out, and a summary of the results will be discussed below. The goal of the current research is to investigate the applicability of the test further. Firstly, it aims at validating the NGT-NSRT by providing a small set of norm scores for deaf signers; secondly, it looks further into phonological complexity, specifically, the complexity of movement. The outline of this article is as follows: Section 1.2 discusses background information on nonsense sign repetition tasks and 1.3 describes the research questions and hypotheses. Section 2 addresses the methodology of the current study. In Section 3 the results are presented, and in Section 4 they are discussed. Section 5 concludes this paper.

1.2 Nonsense Sign Repetition Tasks

The NGT-NSRT is based on the NSRT of BSL, developed by Marshall, Denmark, and Morgan (2006), which in turn is based on nonsense word repetition tasks for spoken languages (e.g. Children's Test of Nonword Repetition, Gathercole et al. 1994). This type of tests is generally seen as a measurement of phonological abilities (Chiat 2015). The general idea is as follows: participants hear/see nonsense words/signs and repeat them, and the better they do, the better their phonological skills. Crucial is that the items do not exist as lexical items in the language at stake, but do follow its phonotactic rules. The reasoning is that phonological skills are measured without using real lexicon, so that individual vocabulary sizes should not affect the outcome (Chiat 2015).

Marshall and colleagues were specifically interested in Specific Language Impairment (SLI) in sign language, and for this purpose, tests with nonsense words or nonsense signs are generally suited (e.g. Graf Estes, Evans, and Else-Quest 2007). They therefore decided to develop a BSL-NSRT and conducted several studies with it (Mason et al. 2010; Mann et al. 2010). After pilot-tests and adaptations, their test consisted of 40 items, which could be categorized into four levels of complexity, depending on the combination of an unmarked or marked handshape and a single or combined movement (see Table 1). A single movement consists of a path movement or hand-internal movement, whereas a combined movement is a combination of these two.

Movement	*Handshape*	
	Unmarked	*Marked*
Single	Level 0	Level 1b
Combined	Level 1a	Level 2

Table 1: Levels of phonological complexity of the items
in the BSL-NSRT (Marshall, Denmark, and Morgan 2006).

One of the results of Mann et al. (2010) was that signs with a hand-internal movement might be more complex than signs with a path movement. In 2015, the NGT-NSRT was developed, and one of the choices was therefore to add an extra distinction between these two movement types, resulting in six complexity levels (Table 2):

---

[1] http://www.unive.it/pag/33761/#c306292
[2] With stimuli of dr. Ellen Ormel and dr. Marcel Giezen.

Movement	Handshape	
	Unmarked	*Marked*
Path	Level 1a	Level 2a
Internal	Level 1b	Level 2b
Combined	Level 1c	Level 2c

Table 2: Levels of phonological complexity of the items
in the NGT-NSRT (Klomp 2015).

For the 2015 study, ten (adult) children of deaf adults (CODAs) were tested to compare the NGT-NSRT to a non-word repetition task for spoken Dutch. It turned out that the scores on both tests were highly correlated, which suggests that the tests indeed assess the same skills (phonological abilities). Furthermore, the effects of handshape and movement on the scores were analyzed. There was an effect of handshape on the scores, but only after two deviant items of the NSRT were left out of the analysis.[3] An effect of movement was only found for signs containing a combined movement, and there was no evidence for a difference in complexity for signs with only a path or only an internal movement.[4] This is different from the above-mentioned finding reported by Mann et al. (2010), and there are several possible explanations. Firstly, it could be that my study in 2015 was too small to find an effect between path and hand-internal movement. Secondly, it could be that such effects are only found with children, while the NGT-NSRT was conducted with adults. Thirdly, it might be the case that we are dealing with a language-specific difference, meaning that the complexity of specific movement types differs between BSL and NGT. To investigate this further, the effect of movement type is one of the main research questions. The effect of handshape is not part of the current study, partially to narrow the scope, and partially because the deviant items that influenced the analyses in 2015 are still part of the NGT-NSRT in its current form – simply because it makes more sense to first investigate what needs to be changed, instead of changing a few items per time.

As for the practical use of the NGT-NSRT, it is valuable to look into several ways of rating the participants' repetitions. In 2015, the scoring was done in a phoneme-based way, in which the rater decides whether the handshape, location, movement and orientation were repeated correctly. This method is time-consuming and perhaps not necessary if one is not interested in specific detailed analyses per parameter, but merely in a general idea of the participants' phonological abilities. Another way of scoring is therefore to look at the sign as a whole, and score whether it is repeated completely correctly or not; this is perhaps a more economical way. This study therefore also looks into these two methods of scoring (see Section 2). The next paragraph addresses the research questions and hypotheses.

---

[3]In the 2015 study, effects of scores were analyzed using percentages per phonemes correct (PPC), which were calculated per item for all signers together. The PPC for signs with an unmarked handshape (n = 16, Mdn = 93.75) was significantly higher than the PPC for signs with a marked handshape (n = 17, Mdn = 80.00; U = 71.5, z = −2.32, p = .019). It concerns a medium to large effect (r = −.40) (Klomp 2015, 65). Note that the current study uses different statistical analyses (see Section 3), and will also not further address the complexity of handshape.

[4]The PPC for signs with a combined movement (n = 12, Mdn = 74) was significantly lower compared to signs with one movement (H(2) = 12.00, p = .002). No evidence was found for a difference between signs with only a path movement (n = 12, Mdn = 85.00) or only an internal movement (n = 12, Mdn = 88.75; p = .977) (Klomp 2015, 66).

1.3   Research questions and hypotheses

The current study aims (i) at validating the NGT-NSRT by providing a small set of norm scores, and (ii) at looking further into phonological complexity. Specifically, the following research questions were investigated (each followed by a hypothesis):

1. Do deaf signers and CODAs perform differently on the NGT-NSRT?

For the goal of the study in 2015, it made sense to test NGT-Dutch bilinguals (i.e. CODAs). However, it turned out that not all participants were highly proficient in NGT. This was not ideal, but also not crucial for the purpose in 2015. For the potential use of the NGT-NSRT with deaf participants in educational or clinical settings, however, it is necessary to now test it with deaf signers, and, to be able to interpret these results, it is interesting to compare the performances of the deaf signers to the performances of the previously tested CODAs. My hypothesis is that the deaf signers will outperform the CODAs, as I expect the deaf signers to have more sign language experience.

2. Does movement type (path, internal or combined) have an effect on the scores?

Based on the previous study in 2015, I expect participants to make more mistakes on signs with a combined movement compared to signs with a single movement. In addition, I expect no evidence for a difference between scores on signs with a hand-internal movement and signs with a path movement.

3. Do phoneme-based scores yield the same results as binary correct/incorrect scores?

Phoneme-based scores and correct/incorrect scores are based on the same data. They are not independent from each other, in the sense that the correct/incorrect scores are directly derivable from the phoneme-based scores (correct when the maximum score is achieved; incorrect when the maximum score is not achieved). I therefore expect a (positive) correlation between these two methods.

4. What is the inter-rater reliability for two raters for phoneme-based scores and for correct/incorrect scores?

To gain more insight into the reliability of the scoring methods, another rater was asked to score 10% of the data (see Section 3). This knowledge is important for the applicability of the test, for instance, when it is used in educational or clinical settings. There was no hypothesis formulated prior to the research.

2   Methodology

2.1   The NGT-NSRT

The NGT-NSRT is loosely based on the BSL-NSRT and consists of thirty-six test items and four practice items. The stimuli were not developed for this specific task, but come from a larger dataset of about two hundred nonsense signs. More information about the dataset can be found here: http://signbank.science.ru.nl/datasets/detail/11 (Ormel and Giezen 2018). Importantly, these nonsense signs were based on signs from ASL. They were designed to be

meaningless in NGT, but not necessarily to fit the phonotactic rules of NGT. For the development of the NGT-NSRT, a phonological description of about ninety signs of this large dataset was made, to ensure that the stimuli were balanced in terms of phonological complexity. The final selection of forty signs (including four practice signs) was checked by a native signer to confirm that these signs were meaningless, yet possible signs in NGT.

The items were put together in a VLC[5] playlist, and every participant saw the signs in the same order. The NGT-NSRT is developed for adults, and since repetition of only one nonsense sign would be too easy for this target group, the stimuli are shown in pairs (which are always from the same complexity level). Every pair is followed by a blue screen, which lasts six seconds, during which the participant should repeat the two non-signs. Figure 1 shows a sequence of two practice items of the NGT-NSRT.
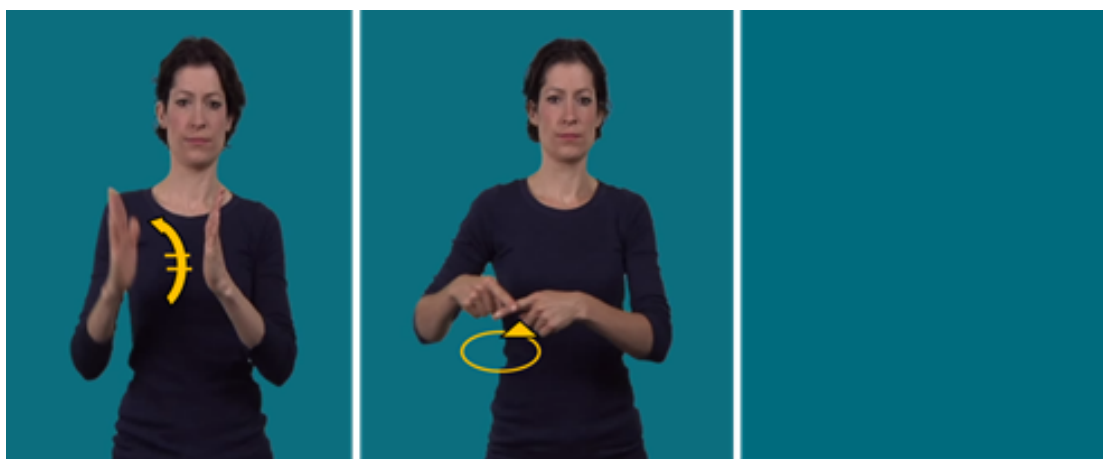

Figure 1: A sequence in the NGT-NSRT: non-sign 1, non-sign 2, blue screen.

2.2 Participants

The analyses in this study (see Section 3) were done with the scores of three groups of participants: a group of ten (adult) CODAs tested by Klomp (2015); a group of eleven (adult) CODAs tested by Zijlstra (2017); and a group of ten deaf (near-)native signers tested specifically for the current study. A table with an overview of the signers' characteristics can be found in Appendix A. Important to explain here is the motivation for including CODAs together with deaf signers in the participant pool. This has to do with the fact that the CODA data were already available from previous studies, and the wish to include as much data as possible for the current analyses. Given the potential use of the NGT-NSRT in practice, it is also needed to pilot-test it with deaf signers. Consequently, it makes sense to on the one hand compare the CODAs to the deaf signers to see whether they perform differently, and on the other hand to take both groups together to explore the general relationship between the complexity of signs and the scores of signers.

2.3 Procedure

Participants were instructed that they were about to do a computerized task in which they would see two non-existing signs, which they should watch carefully and repeat when a blue screen appeared. For the CODAs, this instruction was given in spoken Dutch; for the deaf participants, in NGT (by myself). In most cases, participants did other tasks before or after

---

[5]VideoLAN, www.videolan.org/vlc/.

doing the NGT-NSRT, but it differed per group which tasks they performed and in which order. They were given breaks in between the tasks and their understanding of the procedure was checked before the NGT-NSRT started. In total, the instruction and test took about five minutes. Most participants indicated spontaneously that they enjoyed performing the task.

2.4 Scoring

Two ways of scoring were adopted. The first one was phoneme-based: for each sign, it was noted whether the handshape, location, orientation and movement were repeated as in the target sign. For each correctly produced nonsense sign of level 1a, 1b, 2a and 2b (see Table 2 above), participants could receive 4 points. At level 1c and 2c, however, two movements were involved, so participants could receive 5 points. In total, there were 156 points to obtain.

The second scoring method was simply correct/incorrect scoring, i.e. whether the sign as a whole was repeated correctly or incorrectly. This could be done rather easily after the first way of scoring, and was done specifically in light of the third research question. In Table 3, an example of scoring is shown.[6] The columns represent the signs that were repeated; the rows show the scored parameters and the total scores for the two methods.

	118	108	88	175
Handshape	1	1	1	1
Hand-internal movement	1	1	1	1
Path movement			1	0
Location	0	1	1	0
Orientation	1	1	1	1
Phoneme-based total	3	4	5	3
Correct/incorrect total	0	1	1	0

Table 3: Example of scoring form displaying phoneme-based scores and correct/incorrect scores.

3 Results

The final scores per participant can be found in Appendix A, and the scores per item per participant can be found in Appendix B. To answer the research questions, I used R (R Core Team 2013) and applied the *glmer* function (*lme4* package, Bates et al. 2015) to the data to fit a logistic mixed effects linear model of scores on the NGT-NSRT as a function of group, handshape level and movement level. The full model[7] was: score ~ group * handshape * movement + (group|item) + (movement|participant).

The first result is that deaf signers score significantly better (mean score = 139) on the NGT-NSRT than the CODAs (mean score = 129) (odds ratio of scoring a point = 1.58, $p < 0.01$, z = 4.04, 95% CI from 1.27 to 1.98). See Figure 2 for a visual representation of the scores of the two groups.

---

[6]Note that signs 118 and 108 only have a hand-internal movement and no path movement, so this latter movement type is not scored for these signs.

[7]Note that scores were treated as binary, since the model calculates the chance that a specific parameter is repeated correctly (yes or no), depending on the given predictors.
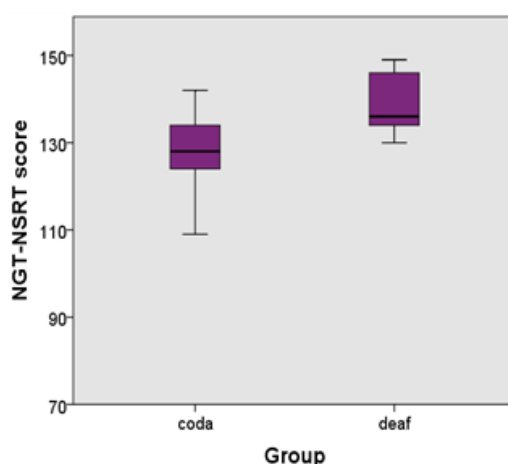
Figure 2: Boxplots of scores on the NGT-NSRT by CODAs and deaf signers.

On the one hand, this result might be surprising since all signers were considered to be early learners of NGT. On the other hand, as can clearly be seen in Figure 2, the variance in the CODA group is larger than in the group of deaf signers. This will be discussed further in the next section.

Secondly, regarding movement type, the scores on signs with one movement (either path or hand-internal, both groups together) were significantly better than the scores on signs with a combined movement (odds ratio = 1.94, $p < 0.01$, z = 2.95, 95% CI from 1.25 to 3.00). This indicates that the latter are more difficult to reproduce correctly. No evidence was found for a difference between scores on signs with a path movement and signs with a hand-internal movement (odds ratio = 1.25, $p = 0.42$). This result is not completely in line with the results of Mann et al. (2010), and will be discussed again in Section 4.

To assess the practical use of the NGT-NSRT, I was interested in the relationship between the phoneme-based scores and the binary correct/incorrect scores. It turns out that these two ways of scoring are significantly correlated (r = 0.92, $p < 0.01$, 95% CI from 0.84 to 0.97, effect size $R^2 = 0.84$). This suggests that, although phoneme-based scores are more detailed, binary scores could be sufficient to get an idea about the participants' phonological skills.

In addition, I looked into interrater reliability for two raters for the two methods. A second rater scored 10% of the data[8] and Cohen's kappa coefficient was used to determine interrater reliability. This was first done for the phoneme-based scores. There was moderate agreement between the two raters: $\kappa = .544$, $p < 0.01$, 95% CI from 0.43 to 0.66. The agreement was also moderate for the correct/incorrect scores: $\kappa = .579$, $p < 0.01$, 95% CI from 0.41 to 0.75. This will also be addressed in the next section.

4   Discussion

The first result to discuss is the better performance of the deaf signers compared to the CO-DAs. On the one hand, this instantly raises the question whether the CODAs were proficient enough to participate in the first place. As explained earlier, their data were already available from previous studies, and it is indeed important to keep in mind that the sign language experience differed per CODA. This might also be the reason why this group had a larger variance in their scores than the deaf signers (see Figure 2). On the other hand, with the

---

[8]This data was selected randomly out of Klomp's raw data, as Zijlstra's raw data was not available.

variation in language experience in mind, the difference in performance strongly suggests that differences in phonological skills are reflected in the scores. In other words, more fluent signers perform better on the NGT-NSRT than less fluent signers, and this is exactly what is needed for an assessment tool.

Another important aspect of assessment is defining the level of language proficiency (in this case phonological abilities) with the help of items of varying complexities. To that end, the effect of movement was tested. In line with the previous study in 2015, a difference in difficulty was found between signs with a combined movement and signs with a single movement. This seems to be the case simply because of the addition of movements, not because of the involvement of a specific movement type within these combined signs, since there is no evidence for a difference between signs with a path movement and signs with an internal movement. Crucially, this lack of evidence might render the distinction between these two movement types in the NGT-NSRT items irrelevant. As a practical consequence, the complexity levels in the NGT-NSRT should be reconsidered.

Around the same time that this study was conducted, Vink (2018) looked into alternative classifications of phonological complexity for the NGT-NSRT. She adopted the Dependency Model of Kooij (2002), which is a phonological model developed for NGT, to re-categorize the items of the NGT-NSRT. She shows that the items can be classified into five levels of complexity, which are related to the number of the item's feature specifications within the Dependency Model, and that these levels can predict the number of errors in signers' repetitions. This could indeed be a better way of analyzing the phonological complexity of the nonsense signs, although Vink also points out that this is only one of many ways, and that, in fact, her categories of signs are not strikingly different from those originally proposed by myself (Vink 2018, 43). Another finding of Vink is that, as a result of another way of scoring, in her analysis, three participants reached a score of 92.5% or higher correct (one participant even reached 100%). This result implies a ceiling effect, whereas this effect has not been found in the current scoring methods (see Section 3). Since a ceiling effect should be avoided, as a potential effect of independent variables might not show anymore, this might mean that the NGT-NSRT should be extended with extra complex items. Now, the most important question of her thesis is: how to define phonological complexity? In the current study, I can only speculate further on the complexity of movement.

Although the lack of evidence of a difference in complexity of signs with a path movement and signs with a hand-internal movement does not mean there is no difference, and although one should not interpret null results, let us hypothetically consider that these movement types are equally complex in NGT, at least for fluent signers. If we want to explore why this contradicts the results of Mann et al. (2010), a next step would be to let signing children perform the NGT-NSRT, so that the target groups of the two NSRTs are better comparable.[9] If children perform significantly better on signs with a certain movement type this would suggest that complexity is experienced differently by children than by adults. To rule out the possibility that this difference merely reflects language experience, fine motor skills or perhaps other factors, it would be interesting to also test adult second language learners of NGT. When children do not perform better on signs with a certain movement type, it could mean that NGT and BSL are structurally different in the complexity of their movement types.

As for the practical use of the NGT-NSRT, the results have shown that phoneme-based scoring is highly correlated to binary correct/incorrect scoring. The first way of scoring is

---

[9]Note, however, that the NGT-NSRT is developed for adults and should be adapted to use with children. The sequence of two nonsense items should be reduced to only one item per time, for example, and one could consider adding a game-like introduction. Other recommendations are mentioned in Klomp (2015).

definitely more precise, but also more time-consuming and asks for more linguistic insight. It is therefore valuable to know that correct/incorrect scoring yields similar results. Depending on the purpose of the use of the NGT-NSRT, one could decide to apply the one or the other method.

In addition, I was curious whether the instructions on scoring were sufficient and whether there would be good agreement between two raters. It turned out that the agreement between the ratings of myself and a colleague was moderate, for both methods. This shows that rating the data is not clear-cut and not completely reliable. This result should be taken into account when the NGT-NSRT is used in practice, and the instructions should be improved. At a first glance, it is clear that my ratings were generally higher than my colleague's ratings, and that ratings seemed to differ mostly in the movement parameter. Although it would be valuable to look into these differences in more detail, to discuss them and solve them, this was not done for the current study.

5   Conclusion

In this study, I collected data from deaf signers on the NGT-NSRT and compared their results to previously collected data from CODAs. In addition, I looked into the reliability of two scoring methods. It turns out that the deaf signers outperformed the CODAs; that non-signs with a single movement yield higher scores than signs with a combined movement; that there is a strong correlation between the two scoring methods and that there was moderate agreement between two raters.

The NGT-NSRT still has room for improvement. Klomp (2015) showed that some deviant items should be removed from the test; the current study shows that there might not be as much variety in the complexity of the items as expected; Vink (2018) even showed that, depending on one's analysis, there is the risk of a ceiling effect in the scores. Still, both Vink and the current study show that complexity does predict scores on the NGT-NSRT and that the NGT-NSRT could thus in principal function as an assessment tool for phonological skills in signing adults. In addition, the test only takes five minutes to conduct, and scoring can be done quickly and effectively by rating the participants' repetitions simply as correct or incorrect. These are positive results for the quest for more assessment tools for sign languages. The current study furthermore adds new insights into the phonological complexity of movement types in NGT.

To follow the suggestions of my previous study and of Vink (2018), it could be useful to revise the NGT-NSRT at least slightly; specifically, to remove deviant items and add more complex items. Note, however, that this means that new norm scores should be collected. Furthermore, for the questions of complexity it could be valuable to do additional research with the current version of the test, for example with other target groups, such as children.

Acknowledgments

Olthof and Marloes Oomen for general feedback on this paper. Last but not least, I am thankful for the participation of all the signers that were involved.

References

Baker, Anne E., and Sonja Jansma. 2002. "Handleiding NGT-Observatielijst voor peuters (NGT-OP)". University of Amsterdam.

Bates, Douglas, Martin Mächler, Ben Bolker, and Steve Walker. 2015. "Fitting linear mixed-effects models using lme4". *Journal of Statistical Software* 67 (1): 1–48. doi:10.18637/jss.v067.i01.

Chiat, Shula. 2015. "Non-word repetition". In *Assessing multilingual children: Disentangling bilingualism from language impairment*, 125–150. Edited by Sharon Armon-Lotem, Jan de Jong and Natalia Meir. Bristol: Multilingual Matters.

Gathercole, Susan E., Catherine S. Willis, Alan D. Baddeley, and Hazel Emslie. 1994. "The children's test of nonword repetition: A test of phonological working memory". *Memory* 2 (2): 103–127. doi:10.1080/09658219408258940.

Graf Estes, Katharine, Julia L. Evans, and Nicole M. Else-Quest. 2007. "Differences in the nonword repetition performance of children with and without specific language impairment: A meta-analysis". *Journal of Speech Language and Hearing Research* 50 (1): 177–195. doi:10.1044/1092-4388(2007/015).

Haug, Tobias. 2018. "Sign language assessment services". `www.signlang-assessment.info`.

Hermans, Daan, Harry Knoors, and Ludo Verhoeven. 2007. "An assessment instrument for Sign Language of the Netherlands". Sint-Michielsgestel: Via-taal.

Klomp, Ulrika. 2015. "Making sense of a nonsense sign repetition task". MA thesis, Vrije Universiteit.

Kooij, Els van der. 2002. *Phonological categories in Sign Language of the Netherlands: The role of phonetic implementation and iconicity*. Utrecht: LOT.

Mann, Wolfgang, Chloe R. Marshall, Kathryn Mason, and Gary Morgan. 2010. "The acquisition of sign language: The impact of phonetic complexity on phonology". *Language Learning and Development* 6 (1): 60–86. doi:10.1080/15475440903245951.

Marshall, Chloe R., Tanya Denmark, and Gary Morgan. 2006. "Investigating the underlying causes of SLI: A non-sign repetition test in British Sign Language". *Advances in Speech Language Pathology* 8 (4): 347–355. doi:10.1080/14417040600970630.

Mason, Kathryn, Katherine Rowley, Chloe R. Marshall, Joanna R. Atkinson, Rosalind Herman, Bencie Woll, and Gary Morgan. 2010. "Identifying specific language impairment in deaf children acquiring British Sign Language: Implications for theory and practice". *British Journal of Developmental Psychology* 28 (1): 33–49. doi:10.1348/026151009X484190.

Ormel, Ellen, and Marcel Giezen. 2018. "NGT non-signs". `http://signbank.science.ru.nl/datasets/detail/11`.

R Core Team. 2013. *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing. `http://www.R-project.org/`.

Vink, Lianne. 2018. "Looking for complexity: An analysis of NGT nonsense signs in the Dependency Model". MA thesis, University of Amsterdam.

Zijlstra, Denise. 2017. "Iconicity and phonological priming in production of Sign Language of the Netherlands (NGT) by hearing signers". MA thesis, University of Groningen.

Appendix A: An overview of the signers' characteristics

Participant	Coda/Deaf	Age	Phoneme-based total	Correct/ incorrect total	Study
1	coda	61	118	12	Klomp (2015)
2	coda	32	139	26	Klomp (2015)
3	coda	41	109	13	Klomp (2015)
4	coda	26	130	21	Klomp (2015)
5	coda	21	120	14	Klomp (2015)
6	coda	24	126	16	Klomp (2015)
7	coda	53	124	16	Klomp (2015)
8	coda	27	128	19	Klomp (2015)
9	coda	29	134	21	Klomp (2015)
10	coda	25	128	16	Klomp (2015)
11	coda	28	142	23	Zijlstra (2017)
12	coda	31	127	16	Zijlstra (2017)
13	coda	65	125	12	Zijlstra (2017)
14	coda	44	141	26	Zijlstra (2017)
15	coda	33	137	20	Zijlstra (2017)
16	coda	42	121	13	Zijlstra (2017)
17	coda	27	134	21	Zijlstra (2017)
18	coda	42	124	15	Zijlstra (2017)
19	coda	54	130	18	Zijlstra (2017)
20	coda	55	142	24	Zijlstra (2017)
21	coda	20	127	16	Zijlstra (2017)
22	deaf	70	133	16	current
23	deaf	73	130	20	current
24	deaf	25	145	25	current
25	deaf	58	136	20	current
26	deaf	28	147	27	current
27	deaf	37	149	29	current
28	deaf	44	135	22	current
29	deaf	53	142	24	current
30	deaf	39	134	18	current
31	deaf	52	141	23	current

Appendix B: Scores per participant per item

item	max	p1	p2	p3	p4	p5	p6	p7	p8	p9	p10	p11	p12	p13	p14	p15	p16
118	4	4	4	4	4	4	4	4	4	4	4	3	4	3	4	4	2
108	4	2	2	2	3	0	4	1	2	3	2	3	4	3	2	3	2
88	5	5	5	4	3	3	4	3	5	5	4	5	3	5	4	5	3
175	5	3	5	2	5	4	3	3	3	5	4	4	4	4	5	4	4
129	4	3	4	1	4	4	4	3	4	4	3	4	3	3	4	4	2
182	4	3	4	0	4	4	3	4	4	4	3	3	4	2	4	3	3
116	5	3	4	3	5	5	3	4	2	4	3	5	4	5	5	5	5
4	5	3	5	4	5	2	2	4	4	5	3	4	5	4	5	4	4
202	4	3	4	4	3	4	4	4	4	4	4	4	4	4	4	4	4
24	4	3	3	3	4	3	4	4	3	3	4	4	4	3	4	4	3
2	4	3	2	3	2	3	3	3	3	4	3	4	2	3	4	4	4
132	4	4	4	0	4	3	4	4	4	4	4	3	3	3	2	4	4
20	4	4	4	4	4	4	4	4	4	4	4	4	4	3	3	4	3
48	4	4	4	4	3	4	4	4	3	4	3	4	4	4	4	4	4
79	4	3	4	4	4	3	3	4	4	3	3	4	4	4	3	3	3
144	4	3	4	2	3	1	3	2	2	4	3	4	2	3	4	2	3
162	4	3	4	3	4	3	2	4	4	3	2	3	2	3	4	3	3
204	4	3	3	3	3	2	3	3	3	3	3	3	1	4	4	3	4
180	5	5	3	4	5	4	5	4	5	2	5	5	5	5	5	5	5
136	5	3	5	2	3	4	4	4	4	4	4	5	4	3	4	5	4
171	5	3	5	5	5	4	2	4	5	4	5	5	4	4	5	5	4
93	5	3	3	3	4	3	4	3	3	3	3	5	3	3	3	4	3
36	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	3
50	4	4	4	4	4	4	4	3	4	4	4	2	2	3	4	4	3
83	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4
35	4	4	4	3	4	3	4	4	4	4	4	4	4	4	4	4	4
166	4	2	4	3	3	4	4	4	4	3	3	3	3	3	4	4	4
131	4	4	4	3	4	4	4	4	4	4	4	4	4	4	4	4	4
34	4	3	4	4	4	3	4	3	3	4	4	4	4	4	4	3	2
21	4	2	4	4	0	4	3	3	3	4	3	3	3	3	3	3	3
102	5	4	5	4	5	4	4	3	4	2	5	5	5	3	5	4	3
111	5	3	4	2	4	3	3	2	2	3	3	5	4	4	5	3	3
195	5	2	2	0	1	2	4	2	5	4	3	5	4	3	3	3	3
172	5	2	3	2	2	3	2	3	2	3	3	4	3	2	5	4	1
87	4	4	4	4	4	4	3	4	4	4	4	3	3	3	2	3	4
189	4	3	4	4	3	3	3	4	2	4	4	4	3	3	4	4	4
total		118	139	109	130	120	126	124	128	134	128	142	127	125	141	137	121

The *max* column represents maximal possible score per item.
See the next page for the continuation of the table.

item	max	p17	p18	p19	p20	p21	p22	p23	p24	p25	p26	p27	p28	p29	p30	p31
118	4	4	1	4	4	4	4	2	4	4	4	4	4	4	4	4
108	4	2	3	3	4	2	4	1	3	3	3	3	4	3	3	3
88	5	5	1	4	4	5	5	5	4	4	4	5	3	5	5	5
175	5	5	5	5	5	4	4	4	5	5	5	5	5	4	5	5
129	4	3	3	4	3	4	3	4	4	3	4	4	3	3	4	4
182	4	4	2	4	3	3	2	4	4	4	4	4	3	4	3	3
116	5	3	5	3	5	3	5	5	5	4	5	5	5	5	4	4
4	5	4	4	4	5	4	4	4	5	5	5	4	5	4	5	3
202	4	4	4	4	4	3	3	4	4	4	3	4	4	4	4	4
24	4	4	3	3	4	4	3	3	4	4	3	3	4	3	3	4
2	4	4	4	4	3	4	4	4	4	3	4	4	4	4	4	4
132	4	3	2	4	2	4	4	1	4	4	4	4	4	3	3	4
20	4	4	4	4	4	4	3	4	3	3	3	4	4	4	4	4
48	4	4	4	4	4	4	4	4	4	3	4	4	4	4	3	4
79	4	4	4	3	4	4	3	4	4	4	4	4	4	4	4	4
144	4	4	3	2	4	2	3	3	3	4	4	4	3	4	3	3
162	4	2	3	3	4	2	2	4	4	2	4	4	3	4	2	4
204	4	4	3	2	3	3	3	4	3	3	4	4	3	4	3	4
180	5	4	4	4	5	5	5	5	5	5	5	4	5	5	4	4
136	5	3	5	5	5	4	5	4	4	5	4	5	5	4	3	4
171	5	5	5	5	4	4	3	3	4	3	5	5	5	5	4	4
93	5	5	4	4	4	5	4	2	5	4	5	5	2	5	5	4
36	4	4	3	1	4	4	3	4	4	4	4	4	4	4	3	4
50	4	2	3	2	3	3	4	4	4	4	4	4	4	4	3	4
83	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4
35	4	4	4	4	3	4	4	4	4	4	4	4	4	4	3	4
166	4	3	4	4	4	3	3	4	4	4	4	3	2	4	4	3
131	4	4	4	4	4	3	3	4	4	3	4	4	3	4	4	4
34	4	4	3	3	4	3	4	4	4	3	4	4	4	3	3	4
21	4	3	2	3	3	3	3	3	4	4	4	4	3	4	4	4
102	5	4	4	5	5	5	4	4	5	5	5	4	5	5	5	4
111	5	5	4	4	5	3	4	3	4	2	4	5	4	2	2	4
195	5	3	2	4	5	3	5	3	5	5	5	5	3	4	5	5
172	5	3	3	2	3	2	4	4	4	4	5	5	2	4	4	3
87	4	3	4	4	4	4	4	3	3	3	3	3	4	3	4	4
189	4	4	4	4	4	2	4	4	3	4	3	4	3	4	4	4
total		134	124	130	142	127	133	130	145	136	147	149	135	142	134	141