

Using Open-Source Software for Sign Language Learning and Assessment

The Case of a Web-delivered Yes/No Vocabulary Test for Swiss German Sign Language

<https://doi.org/10.3991/ijet.v14i19.11123>

Tobias Haug (✉)

Interkantonale Hochschule für Heilpädagogik, Zurich, Switzerland
tobias.haug@hfh.ch

Sarah Ebling

University of Zurich, Zurich, Switzerland

Abstract—This study reports on the use of an open-source software for sign language learning and (self-) assessment. A Yes/No vocabulary size test for Swiss German Sign Language (*Deutschscheizerische Gebärdenprache*, DSGS) was developed, targeting beginning adult learners. The Web-based test, which can be used for self-assessment or placement purposes, was administered to 20 DSGS adult learners of ages 24 to 55 ($M = 39.3$). The learners filled out a background questionnaire, took the Yes/No test tests, and filled out a feedback questionnaire. The comments provided by the learners about the suitability of the Web-based DSGS vocabulary self-assessment instrument provided concrete feedback towards improvement of the system.

Keywords—Web-based sign language learning, self-assessment, open-source software, vocabulary learning, Swiss German Sign Language

1 Introduction

Web-based learning systems [1] have become more and more widespread in secondary and tertiary education worldwide. This is also the case for language learning in particular [2] [3]. In the last 20 years, the Common European Framework of Reference for Languages (CEFR) [4] has had a considerable impact on language education and assessment in Europe. More recently, the CEFR has also found its way into the field of learning sign languages as a second language in tertiary education [5]. As a result of this, sign language instructors are interested in providing new learning opportunities to their students. The concept of proficiency for any language, spoken or signed, is made up of many different components and sub-components as defined in theoretical models like the Communicative Language Ability [6]. When developing new learning opportunities for students in the classroom, it is common to focus on selected sub-components, e.g., vocabulary knowledge. In the second language acqui-

sition and assessment literature, vocabulary knowledge has been recognized as a crucial variable contributing to learners' overall language proficiency [7] [8].

The goal of the present paper is to report on the development of a Web-based vocabulary learning and assessment instrument for Swiss German Sign Language (*Deutschschweizerische Gebärdensprache*, DSGS) that relies on a self-report format (i.e., a Yes/No test) using the open-source software Lime Survey (<https://www.limesurvey.org>). The test can be used as both a self-assessment instrument and a placement or diagnostic test [8] in the context of sign language vocabulary learning. The paper at hand additionally presents feedback collected from learners following the application of the assessment instrument.

The remainder of this paper is organized as follows: Section 2 provides an overview of previous work, particularly Yes/No vocabulary size tests (Section 2.1) and Web-based instruments for sign language learning and assessment (Section 2.2). Section 3 introduces the DSGS YN test, reporting on its precise setup (Section 3.1), the participants chosen for an initial application of the test (Section 3.2), the testing procedure (Section 3.3), and the feedback questionnaire (Section 3.4). Section 4 reports on the results of the feedback questionnaire designed to gather insights, among other aspects, on the suitability of the Web-based assessment instrument. Section 5 finished with a discussion and conclusion.

2 Literature Review

2.1 Yes/No vocabulary size tests

Findings resulting from the application of Yes/No (YN) tests for spoken languages informed the development of the YN test for DSGS [9]. For example, the basic design of YN tests was adopted, consisting of a learner seeing a word and indicating whether he/she knows it [10]. In order to minimize the effect of guessing on the part of the learner, i.e., mitigating the risk of a test taker rating more words as familiar than he/she actually knows [8], [11] added pseudowords (phonological possible forms, e.g., for English *foggy* becomes *wuggy*; [12]) words to a YN test. The same approach was also used for the DSGS YN test, with nonsense signs, i.e., phonologically plausible forms of a sign language that bear no meaning in that language by manipulating the phonological parameters of that sign language [13]. These nonsense signs take the role of pseudowords. According to [10], there are no clear guidelines regarding the ratio of real words to pseudowords. For example, the ratios reported in different studies include 30:3 [14], 90:60 [15], 96:32 [16], 60:40 [17], and 40:20 [18]. For the current project, 25 nonsense signs were developed to supplement the existing 95 true test items.

2.2 Use of web-based instruments for sign language learning and assessment

In the past, Web-based sign language tests or learning tools were commonly purpose-built for a specific research project [19]. More recently, existing open-source

software solutions for sign language learning and assessment have allowed for managing large numbers of videos and streaming larger videos. As a result, learning platforms like Moodle have begun to be used for sign language learning scenarios.

3 Methodology

3.1 DSGS YN test

The format of a YN test lends itself to administration via a survey software tool, as the test is about investigating whether a learner knows a word/sign, which is equivalent to posing a yes/no question in a survey.

Lime Survey is an open-source software that has been developed for the purpose of conducting surveys, not specifically aiming at learning scenarios as in the case of this study. Since version 2.7, Lime Survey allows for more straightforward embedding of videos in different formats (e.g., .mp4 and webm format) directly through the graphical user interface. For the purpose of the current study, version 2.7. was used. Lime Survey was installed on the servers of the first author's university. The authors were the only persons who had access to the data.

The DSGS YN test implemented in Lime Survey consists of approximately 120 signs (including 25 nonsense signs, cf. Section 2.1) sampled from existing DSGS teaching materials [20] [21] [22] [23].

The test implemented in Lime Survey consisted of a welcome message and an introduction in written German and DSGS. Subsequently, each item was presented with the question “Do you know the form and the meaning of the sign?”, with “yes” and “no” as possible responses (omitting a response was not possible).

Since presentation of the items followed a uniform scheme, the items were inserted into Lime Survey by populating the underlying database (in this case, a PostgreSQL database), i.e., writing a Python script that automatically produced the corresponding insert statements for each item. The items were presented in a random order. A “continue” button allowed for proceeding to the next item (cf. Figure 1) and a “back” button for returning to the preceding item.

The response data was automatically sent to a secure server for storage. The test was self-administered, but the researcher was present during test taking in case of any technical issues. The results were later imported into SPSS for analysis.

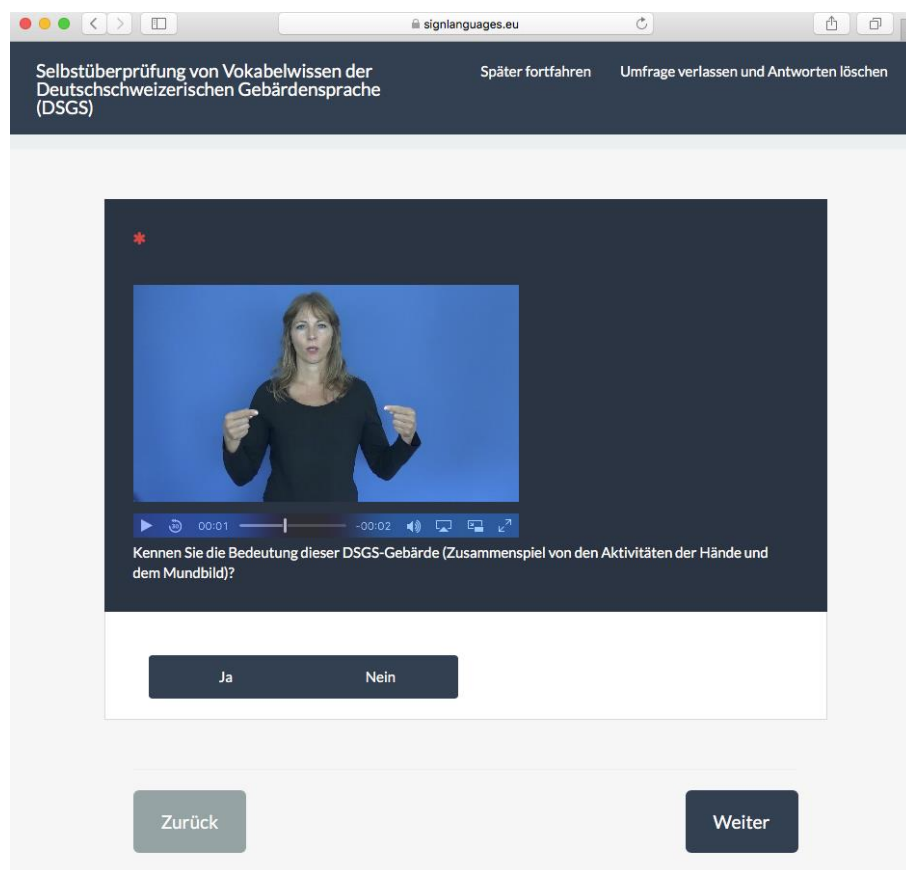


Fig. 1. Yes/No Test Practice Item 1: WIDERSPRUCH ('contradiction')

3.2 Participants

The DSGS YN test was initially carried out with 20 adult learners of DSGS in 2017. Of the 20 participants, 15 were female and 5 males. The learners were between 24 and 55 ($M = 39.3$) years old at the time of testing. 19 of the 20 learners were hearing; one had a Cochlear implant but had acquired German as a first language and was also learning DSGS as an adult. The majority of the learners had a spoken language (e.g., a Swiss German dialect or Standard German; $n = 18$) as their L1. Two participants reported having grown up with two spoken languages. All participants had learned DSGS as adults (range: 18-53 years, $M = 35.4$).

3.3 Procedure

The initial study took place in a quiet room at the first author's workplace. The researcher was present throughout the session. The session, which took approximately one hour, included the following steps for the learners:

- Receiving background information on the project, filling out the consent form, reading through the information sheets of the project
- Filling out the background questionnaire online on a laptop
- Taking the YN test online on a laptop
- Taking a second test (a translation test) using a separate assessment instrument, the results of which are reported in [9]
- Filling out the feedback questionnaire online on a laptop

3.4 Feedback questionnaire

The learners were asked to fill out a short online feedback questionnaire after using the YN test, also implemented in Lime Survey. The goal of this questionnaire was to gather general feedback on the YN test. Studies on test takers' feedback for spoken language tests informed the development of this questionnaire, for example, [24] and [25]. The learners were asked to indicate their agreement with the statements 1 to 4 below on a four-point Likert scale ranging from "strongly disagree" to "strongly agree", and to give one of the four possible responses with "negative", "rather negative", "rather positive", or "positive" to question 5:

- The wording of the test instruction is clear.
- The wording of what the test intends to measure is clear.
- The test matches my DSGS skills well.
- The length of the test is appropriate.
- How to you judge the potential of this test to evaluate your own DSGS vocabulary learning?

Test takers were given the possibility to leave a written comment. For the present paper, the optional free-text comment box is particularly relevant, as it provided an opportunity for learners to comment on the suitability of the Web-based assessment instrument.

4 Results of the Feedback Questionnaire

As part of the comment box, participants delivered a number of valuable comments on the format of the Web-based YN test that are capable of informing future Web-based assessment instrument development. The written German comments were translated into English by the first author.

For example, some of the participants left comments related to the suitability of the Web-based self-assessment system: "This self-assessment is an important instrument

to gain insights into signs that one does not remember anymore” (ID11). Another participant (ID17) stated: “I missed the immediate objective response as to whether I remembered the sign correctly or not. It would have been good to have a German translation to be able to validate my yes/no decision.” Still another participant’s (ID23) comment went into the same direction: “I would have preferred to be able to not only indicate whether I know or do not know the sign but also write out the German translation” [and get immediate feedback on it; in square brackets added by the authors]. Two participants did not perceive the Web-based self-assessment vocabulary instruments as very positive. One of these two participants (ID16) stated “I perceived the self-assessment as being more exhausting [than the second test, a translation test which was shorter]”, the other participant (ID21) emphasized the importance of the selected items to match the learners’ signing skills. He or she found it very frustrating that he/she had to hit the “No” button, indicating he did not know the sign, continuously. One participant (ID13) provided a useful comment related to the technical side of the self-assessment system. He or she stated that “it would be good for the survey to automatically advance once I have selected ‘yes’ or ‘no’, i.e., to not have to hit the ‘continue’ button myself”.

5 Discussion and Conclusion

Tests and assessment instruments can be evaluated from different perspectives, e.g., by collecting evidence of different validity arguments of test evaluation and use or collectively contributing to construct validity [26]. The focus of the present paper is different, it aimed to gather feedback on the Web-based self-assessment format.

Based on the learners’ comments, some changes will be implemented into the Lime Survey-based self-assessment system:

- Provide the option to see the German translation directly after indicating knowing or not knowing a sign.
- Provide the option of a report after completing the self-assessment entirely, aligned to the levels of the CEFR.
- Check if the number of items can be reduced while retaining sufficient information about the learners’ DSGS vocabulary size.
- Add information on the level of self-assessment (e.g., A1) in the beginning in order to avoid frustrating experience on the side of the learner.

The novelty of the work reported in this paper consists of using an existing open-source software as a basis for developing and evaluating a DSGS YN test for self-assessment purposes. A limitation of this study is that the sample is rather small and that feedback data on the Web-based format collected from the participants is rather scarce.

In the future, the self-assessment instrument will be used more regularly in DSGS learning contexts at the first author’s university with the goal of gathering more information about the usefulness of the YN test as well as gaining deeper insight into its current Web-based format using Lime Survey.

6 Acknowledgement

This study has been conducted as part of the Swiss National Science Foundation funded Sinergia project “[Scalable Multimodal Sign Language Technology for Sign Language Learning and Assessment](#)” (project number: 160811).

7 References

- [1] P. Chatwattana and P. Nilsook “A Web-based learning system using project-based learning and Imagineering,” *International Journal of Emerging Technologies in Learning*, vol. 12, no. 5, pp. 4-22, 2017. <https://doi.org/10.3991/ijet.v12i05.6344>
- [2] V. Persson and J. Nouri, “A systematic review of second language learning with mobile technologies,” *International Journal of Emerging Technologies in Learning*, vol. 13, no. 2, pp. 188-210, 2018. <https://doi.org/10.3991/ijet.v13i02.8094>
- [3] N. Yue, “Computer multimedia assisted English vocabulary teaching courseware,” *International Journal of Emerging Technologies in Learning*, vol. 12, no. 12, pp. 67-75, 2017. <https://doi.org/10.3991/ijet.v12i12.7955>
- [4] Council of Europe, *Common European Framework of Reference for Languages: Learning, teaching, assessment*. Cambridge: Cambridge University Press, 2009.
- [5] L. Leeson, B. Van den Bogaerde, C. Rathmann, and T. Haug, *Sign Languages and the Common European Framework of Reference for Languages Common Reference Level Descriptors*. Strasbourg: Council of Europe, 2016.
- [6] L. F. Bachman, *Fundamental Considerations in Language Testing*. Oxford: Oxford University Press, 1990.
- [7] J. A. S. Read and C. A. Chapelle, “A framework for second language vocabulary assessment,” *Language Testing*, vol. 18, no. 1, pp. 1-32, 2001. <https://doi.org/10.1177/026553220101800101>
- [8] J. A. S. Read, *Assessing Vocabulary*. Cambridge: Cambridge University Press, 2000. <https://doi.org/10.1017/CBO9780511732942>
- [9] T. Haug, S. Ebling, P. Boyes Braem, K. Tissi, and S. Sidler-Miserez, “Sign language learning and assessment in German Switzerland: Exploring the potential of vocabulary size tests for Swiss German Sign Language,” *Language Learning and Assessment*, vol. 2, no. 1, pp. 20-40, 2019 <https://doi.org/10.29140/lea.v2n1.85>
- [10] R. Beeckmans, J. Eyckmans, V. Janssens, M. Dufranne, and H. Van der Velde, “Examining the Yes/No vocabulary test: Some methodological issues in theory and practice,” *Language Testing*, vol. 18, no. 3, pp. 235-274, 2001. <https://doi.org/10.1177/026553220101800301>
- [11] R. C. Anderson and P. Freebody, “Reading comprehension and the assessment and acquisition of word knowledge,” in *Advances in reading/language research: A research annual*, B. Hudson, Ed. Greenwich, CT: JAI Press, 1983, pp. 231-256.
- [12] P. Sevigny and K. Ramonda, “Vocabulary: What should we test?,” in *JALT 2012 Conference Proceedings*, N. Sonda and A. Krau, ,vse, Eds. Tokyo: JALT, 2013, pp. 701-711.
- [13] W. Mann, C. R. Marshall, K. Mason, and G. Morgan, “The acquisition of sign language: The impact of phonetic complexity on phonology,” *Language Learning and Development*, vol. 6, no. 1, pp. 60-86, 2010. <https://doi.org/10.1080/15475440903245951>
- [14] N. Schmitt and C. B. Zimmermann, “Derivate word forms: What do learners know?,” *TESOL Quarterly*, vol. 36, no. 2, pp. 145-171, 2002. <https://doi.org/10.2307/3588328>

- [15] A. Mochida and M. Harrington, “The Yes/No test as a measure of receptive vocabulary knowledge,” *Language Testing*, vol. 23, no. 1, pp. 73-98, 2006. <https://doi.org/10.1191/0265532206lt321oa>
- [16] R. Stubbe, “Replacing translation tests with Yes/No tests,” *Vocabulary Learning and Instruction*, vol. 4, no. 2, pp. 38-48, 2015.
- [17] J. Eyckmans, *Measuring Receptive Vocabulary Size: Reliability and Validity of the Yes/No Vocabulary Test for French-Speaking Learners of Dutch*, doctoral dissertation, LOT, Utrecht, 2004.
- [18] P. Meara, *EFL Vocabulary Test*, Centre for Applied Language Studies, Swandea, UK, 1992.
- [19] T. Haug, “Use of information and communication technologies in sign language test development: Results of an international survey,” *Deafness & Education International*, vol. 17, no. 1, pp. 33-48, 2015. <https://doi.org/10.1179/1557069X14Y.0000000041>
- [20] P. Boyes Braem, *Gebärdensprachkurs Deutschschweiz, Stufe 1. Linguistischer Kommentar [Sign language course DSGS, level 1. Linguistic comments]*. Zurich: GS-Media/Schweizerischer Gehörlosenbund SGB, 2004a.
- [21] P. Boyes Braem, *Gebärdensprachkurs Deutschschweiz, Stufe 2. Linguistischer Kommentar [Sign language course DSGS, level 2. Linguistic comments]*. Zurich: GS-Media/Schweizerischer Gehörlosenbund SGB, 2004b.
- [22] P. Boyes Braem, *Gebärdensprachkurs Deutschschweiz, Stufe 3. Linguistischer Kommentar [Sign language course DSGS, level 3. Linguistic comments]*. Zurich: GS-Media/Schweizerischer Gehörlosenbund SGB, 2005a.
- [23] P. Boyes Braem, *Gebärdensprachkurs Deutschschweiz, Stufe 4. Linguistischer Kommentar [Sign language course DSGS, level 4. Linguistic comments]*. Zurich: GS-Media/Schweizerischer Gehörlosenbund SGB, 2005b.
- [24] A. Brown, “The role of test-taker feedback in the test development process: Test-takers’ reactions to a tape-mediated test of proficiency in spoken Japanese,” *Language Testing*, vol. 10, no. 3, pp. 277-301, 1993. <https://doi.org/10.1177/026553229301000305>
- [25] T. Sato and N. Ikeda, “Test-taker perception of what test items measure: a potential impact of face validity on student learning,” *Language Testing in Asia*, vol. 5, no. 1, pp. 1-16, 2015. <https://doi.org/10.1186/s40468-015-0019-z>
- [26] C. Weir, *Language Testing and Validation: An Evidenced-Based Approach*. New York, NY: Palgrave Macmillan, 2005.

8 Authors

Tobias Haug studied sign linguistics at Hamburg University and Deaf education at Boston University, where he received his Master in Education in 1998. In 2009 he earned his PhD at Hamburg University. From 1998 to 2004 he worked as a sign language interpreter and researcher. Since 2004 he has been the program director of the sign language interpreter program in Zurich, Switzerland. One of his research interests is sign language assessment in the context of new technologies. In 2017 he received his MA in Language Testing from Lancaster University. **Email id:** tobias.haug@hfh.ch

Sarah Ebling is a senior researcher at Institute of Computational Linguistics, the University of Zurich. With a background in computational linguistics, her focus in research and teaching is on the contribution of language technology to accessibility

for persons with disabilities. Her previous research includes sign language technology, specifically, machine translation into sign language and sign language synthesis, the generation of signing avatars. More recently, she has been involved in an interdisciplinary project centering around automatic sign language recognition.

Article submitted 2019-06-27. Resubmitted 2019-07-15. Final acceptance 2019-07-15. Final version published as submitted by the authors.