Nordic Journal of Linguistics (2026), **49**, 149–162
doi:10.1017/S0332586524000118

CAMBRIDGE UNIVERSITY PRESS

SHORT COMMUNICATION

Corpus research on signed languages in the Nordic countries

Tommi Jantunen[1], Johanna Mesch[2] and Lindsay Ferrara[3]

[1]Sign Language Centre, Department of Language and Communication Studies, PO Box 35, FI-40014 University of Jyväskylä, Finland, [2]Department of Linguistics, Stockholm University, SE-10691 Stockholm, Sweden and [3]c/o Department of Language and Literature, Norwegian University of Science and Technology, Mailbox 8900, NO-7491 Trondheim, Norway
Corresponding author: Tommi Jantunen; Email: tommi.j.jantunen@jyu.fi

(Received 18 December 2023; revised 9 August 2024; accepted 13 August 2024;
first published online 23 September 2024)

Abstract
This semi-systematic literature review examines signed language corpus research in the Nordic countries, with a quantitative and qualitative assessment of corpus research. The article first describes some critical components and functionalities of signed language corpora. It then outlines the evolution of Nordic corpus research, highlighting Sweden's pioneering role and subsequent developments in Finland and Norway. The findings suggest a progression from method-focused publications to those exploring linguistic phenomena within and across (signed) languages. Although the number of research publications is modest, there is a discernible shift towards comparative studies and applications in signed language teaching and learning.

Keywords: corpus research; Finnish Sign Language; literature review; Nordic signed languages; Norwegian Sign Language; signed language corpora; Swedish Sign Language

1. Introduction

Over the past 50 years, there has been a growing interest in how (deaf) signers use their languages in different interactional settings. To investigate this, researchers have already created signed language corpora for over ten (mostly Western) signed languages (see Börstell 2022, Fenlon & Hochgesang 2022) and have engaged corpus linguistic methodology to investigate a range of research questions from phonology to morphology, from syntax to interaction.[1] In this brief review article, we reflect on work that utilizes signed language corpora, specifically focusing on the Nordic context, in order to outline the historical development and contribution of this research and to provide insight into possible future research trajectories. By taking stock of where the field has been and where it might be going, we are better positioned to engage Nordic signed language corpora to improve our documentation and understanding of this group of minority languages.

© The Author(s), 2024. Published by Cambridge University Press on behalf of The Nordic Association of Linguists. This is an Open Access article, distributed under the terms of the Creative Commons Attribution licence (http://creativecommons.org/licenses/by/4.0/), which permits unrestricted re-use, distribution and reproduction, provided the original article is properly cited.

https://doi.org/10.1017/S0332586524000118 Published online by Cambridge University Press

In the following sections, we first introduce signed language corpora and their characteristics. Then we introduce the Nordic signed language corpora that are currently being created and used in the Nordic countries. Next, we detail how we carried out a semi-systematic literature review of Nordic signed language corpus research. Finally, we present our findings on the distribution of previous research across time and geography, along with the various topics of analysis. This will help us conclude with reflections on the historical development and hypotheses regarding future research trajectories in the field of signed language corpus linguistics in the Nordic countries.

2. Introducing (Nordic) signed language corpora

Before providing an overview of the major signed language corpora in the Nordic countries, we briefly introduce what constitutes a signed language corpus, as these characteristics were used to vet the data used in research publications about Nordic signed languages. Firstly, a linguistic corpus aims to be a large, representative sample of a language (McEnery & Wilson 2001, Johnston 2010, Fenlon & Hochgesang 2022). Such corpora work to document the language use of a relatively large group of signers/speakers, who exhibit different socio-linguistic characteristics (e.g. age, gender, level of education, etc.), as they interact in different contexts. Such representativeness is essential, because it allows research involving these language resources to make certain types of generalizations about the language, even generalizations about variation, a key feature of language (Biber, Conrad & Reppen 1998).

Second, a linguistic corpus must be machine-readable (McEnery & Wilson 2001, Johnston 2010, Fenlon & Hochgesang 2022). For signed language corpora, that means that video-recordings of the language must be accompanied by time-aligned, structured text annotations (Johnston 2010, Johnston & Schembri 2013). These text annotations are typically sign or translation level codes providing access to the continuous visual-gestural language produced in the videos. They are necessary to facilitate computer-assisted searches across the data. Signed language corpus annotation is mainly manual work, requiring time, and it is usually carried out in purpose-built software, such as ELAN (Max Planck Institute for Psycholinguistics; see Crasborn & Sloetjes 2008).

A third characteristic of (signed) language corpora is their availability to the language community and researchers (McEnery & Wilson 2001, Fenlon & Hochgesang 2022). In these modern times, a corpus should be accessible via the Internet. Different parts of a corpus may be regulated by different licenses, from Creative Commons licenses to more restricted academic licenses. Accessibility and availability of the corpus must also adhere to applicable privacy laws and ethical practices.

Over the years, many datasets of signed language materials have been collected across the Nordic countries. However, only four of these datasets – collected in Finland, Sweden, and Norway – meet the three criteria to be considered a fully fledged corpus.[2] These corpora are introduced in Table 1 along with details regarding their representativeness (number of signers), machine-readability (number of sign annotations), and accessibility (publication details). In Finland and Sweden, there exist also several smaller datasets that are referred to as a 'corpus'

https://doi.org/10.1017/S0332586524000118 Published online by Cambridge University Press

https://doi.org/10.1017/S0332586524000118 Published online by Cambridge University Press

Table 1. The four large signed language corpora in the Nordic countries

Country	Corpus & language name	Sign tokens annotated*	Total number of signers, region	Language settings	Access & availability
Finland	Corpus of Finnish Sign Language (SVK, *suomalainen viittomakieli*)	164,000; *in progress*	91 signers from all over Finland	Personal introductory presentations, dyad (elicited) narrative retellings, conversations about hobbies, work and events related to Deaf culture and community, free discussions	Material from 27 signers published in https://www.kielipankki.fi/aineistot under codes *snowfrog* and *cfinsl* (with access levels). Material from 12 signers to be published in 2024 in https://www.kielipankki.fi
Finland	Corpus of Finland-Swedish Sign Language (FSTS, *finlandssvenskt teckenspråk*)	20,600; *in progress*	12 signers from Swedish-speaking parts of Finland	The tasks of the FSTS corpus are the same as in the SVK corpus: see above	To be published in 2024 in https://www.kielipankki.fi
Sweden	Swedish Sign Language Corpus (STS, *svenskt teckenspråk*)	209,136; *finished*	50 signers from all over Sweden	Personal introductory presentations, dyad (elicited) narrative retellings, conversations about hobbies, work and events related to Deaf culture and community, free discussions. Parts of the ECHO corpus of fable narrations.	Published in https://teckenspraksk orpus.su.se (with access levels); also published in the Language Archive https://archive.mpi.nl/tla
Norway	Norwegian Sign Language Corpus (NTS, *norsk tegnspråk*)	65,700; *in progress*	100 signers (expected *c.*115)	Dyad, triad, and multiparty conversations, (elicited) narrative retellings, discussion of topics related to the Deaf community, public lectures, presentations	Published in https://repo.clarino.uib.no; *in progress*

*Individual signs, just like individual words in spoken language corpora, provide an indication of the size of a corpus. However, it should be kept in mind that this number does not comment on the total number of annotations across the corpora, as much annotation work focuses on various other (linguistic) features of the data.

but that either do not meet the above listed criteria for corpushood or include 'specialized' language data, e.g. from L2 users only. Our analysis here includes literature engaging the four corpora in Table 1.

3. Review questions and method

In order to reflect on the historical development of signed language corpus linguistics in the Nordic countries and to hypothesize about future research trajectories in the field, we conducted a semi-systematic literature review. This type of literature review is suitable for content analysis as well as for some types of quantitative investigations (e.g. Snyder 2019). In particular, semi-systematic literature reviews are used for topics that have been studied by different groups of researchers in different disciplines (Snyder 2019:335), such as linguistics and computer science in our case. Semi-systematic reviews also look at how research within a selected field has progressed over time and, for example, seek to identify and understand – with the help of meta-narratives rather than by measuring effect size – potentially relevant research traditions that have implications for the topic under study (Snyder 2019:336).

To carry out the review we collected all signed language corpus research publications from Finland, Sweden, and Norway that have been published or accepted for publication up until May 2024. We believed this a feasible goal, since the field is very small and the three authors have first-hand knowledge of most of the work. In fact, they are the ones who created and/or have been in charge of the corpora and their construction and have overseen (academically and/or as restricted access granters) all work that engages the corpora. Once a full list of publications was collated from the three countries they were vetted against three further criteria.

(i) The publication engaged primarily one or more of the Nordic signed language corpora presented in Table 1.
(ii) The publication is or will be associated with an ISSN or ISBN.
(iii) The publication reports original research in a broad sense, including original technology, methodology, and research infrastructure development reports.

Criterion (i) was the main determinant of the data. In practice, this criterion excluded all studies where a corpus listed in Table 1 was not used, or were so-called mixed-materials studies where the corpus was used only marginally to support other data and research conclusions. Criterion (ii) restricted the data to publications that had already been published or were accepted for publication. Research known to be submitted for publication or under review was excluded. The criterion also meant that unpublished theses (e.g. MA theses) were not considered. Criterion (iii) excluded reviews and, for example, position papers.

The resulting literature data (see the Appendix) were then scrutinized quantitatively for frequency and then also qualitatively for goals and topics. As part of the frequency analysis the data was initially observed to fall across three general functional categories. Some publications were found to describe the

https://doi.org/10.1017/S0332586524000118 Published online by Cambridge University Press

Table 2. Corpus publications and their main function per four Nordic signed languages

	SVK	FSTS	STS	NTS
'Methods'	7	1	19	0
'Single language'	7	0	6	5
'Comparative'	4	0	8	1
Total number of publications	18	1	33	6

construction of the corpus or the development of annotation methods (labeled as 'methods'). Others focused heavily on specific linguistic phenomena within an individual signed language, such as word order or turn-taking ('single language'). And some of the corpus publications emphasized the comparison of two or more signed languages ('comparative').

Next, the papers analyzed as either 'single language' or 'comparative' were then subjected to an iterated, detailed qualitative topic analysis. Upon a first parse, we assessed the publications against the stated goals of corpus linguistic research put forth by Biber, Reppen & Friginal (2012): describing linguistic features, describing linguistic varieties, and contributing to language learning and teaching. Then, in order to help us compare the work that has been done across the Nordics, we conducted another parse that classified the papers in terms of general topics of research. The topics were drawn from the publications' headings, abstracts, and keywords. Topics were summarized in single words or short phrases. The findings from this analysis work are reported below in Section 4.

4. Number and types of corpus research on signed languages in the Nordic countries

In total, there have been 53 publications using signed language corpora in the Nordic countries (the full list is reported in the Appendix). The distribution of the publications across the four signed languages (see Table 2) shows that STS (*svenskt teckenspråk*, Swedish Sign Language) has the most published research, followed by SVK (*suomalainen viittomakieli*, Finnish Sign Language) and then NTS (*norsk tegnspråk*, Norwegian Sign Language). For FSTS (*finlandssvenskt teckenspråk*, Finland-Swedish Sign Language), there is currently only one corpus publication.

Note that if the language-specific total numbers of Table 2 are added together, the total comes to 58 publications. This is because some publications are joint publications, dealing with multiple Nordic signed languages and thus falling under multiple language categories.

Table 2 also provides a breakdown of the main functional category of the publications for each signed language. For example, if we look at publications on SVK, we can see that seven fall into the 'methods' and seven into the 'single language' categories. The smallest category for SVK is 'comparative' with four publications.

In SVK and STS, there has been a relatively substantial number of publications describing the different phases of the corpus work and corpus construction (the

https://doi.org/10.1017/S0332586524000118 Published online by Cambridge University Press

Table 3. The primary goal of non-method papers

	SVK	FSTS	STS	NTS
Description of a single linguistic feature	7	0	11	6
Description of two or more varieties	3	0	0	0
Contributes to language learning and teaching	1	0	3	0
Total number of publications	11	0	14	6

'methods' category). Also, the only publication on FSTS is a 'methods' publication. In contrast, there are no such methods publications for NTS. Moving to the 'single language' publications, SVK, STS, and NTS are on an equal footing. Finally, comparative work (the third functional category) has been carried out the most for STS.

To remind the reader, the goals of corpus research, as presented by Biber, Reppen & Friginal (2012), involve describing linguistic features, describing linguistic varieties, and contributing to language learning and teaching. Each relevant (i.e. non-method) publication was next examined for these three goals and the results of this part of the analysis are summarized in Table 3. Overwhelmingly, corpus research publications have focused on describing specific linguistic features of SVK, STS, and NTS (remember that FSTS did not have any language-specific or comparative research publications). In Finland, there has also been an interest in language variation, such as differences in language use registers. Such a perspective has not been investigated in the other signed languages. Additionally, research on signed language learning and teaching has been carried out to some extent in Finland and Sweden, and Sweden in particular has been the forerunner in this field. Sweden's leading role in the field of language learning and teaching can also be seen in the fact that Sweden is the only Nordic country that has a dedicated L2 'corpus' (see Section 2), a dataset not included in this study.

Table 4 provides a closer qualitative look at the specific linguistic phenomena that have been examined in SVK, STS, and NTS using corpora. In SVK, research has investigated both the phenomena of traditional grammar and prosody as well as more non-conventional ways to make meaning. In STS, in addition to grammatical features, corpus research has been characterized by its interest in the use of the body in the context of discourse and its maintenance. NTS corpus research also has a strong discursive focus and is characterized by a particular interest in different semiotic resources, such as indexicality and depiction. So far, reference is the only common topic investigated across all three languages.

5. Development of corpus research on signed languages in the Nordic countries

Findings presented in the previous section showed that STS has the most corpus research publications ($n = 33$), followed by SVK ($n = 18$) and NTS ($n = 6$). So far, FSTS has only one publication. We can explain this difference in publication

https://doi.org/10.1017/S0332586524000118 Published online by Cambridge University Press

Table 4. The main topics of papers that describe single linguistic features or varieties. Bold text indicates common research themes

SVK	STS	NTS
reference	***reference***	***reference***
nods	**nods**	vantage point
transitivity	**transitivity**	finger pointing
rhythm	**rhythm**	turn-taking
word order	content **questions**	**questions**
depicting signs	parts of speech	fingerspelling
constructed action	iconicity	
head and torso movements	mouth actions	
	differential object marking	
	mouthings	
	backchannel signs	

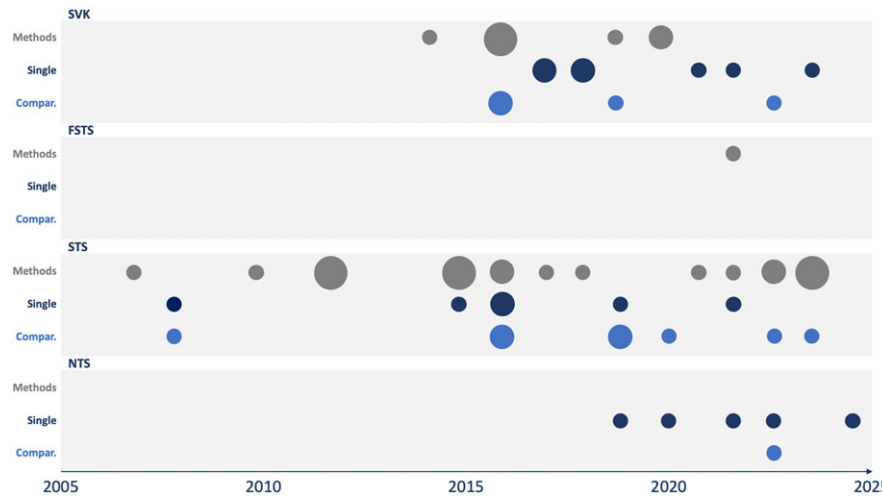


Figure 1. Timeline of corpus publications per four Nordic signed languages. The bigger the circle the more publications that year. The smallest circle corresponds to one publication, the largest to three.

numbers by looking at the timeline of corpus research for each language, presented in Figure 1.

The timeline shows the significant head start Sweden has had in the field of signed language corpus research. The work engaging signed language corpora in Sweden began soon after the turn of the millennium, resulting in a first publication in 2007. Finland began corpus work in 2013 and Norway in 2015. Initial publications came afterwards in 2014 and in 2019, respectively. The work on the

https://doi.org/10.1017/S0332586524000118 Published online by Cambridge University Press

FSTS corpus is as an extension of the SVK corpus work, as the videos were recorded as a part of the same recording process between 2014 and 2017 (see Salonen et al. 2022). However, the actual work of annotating the materials for the FSTS corpus began only in 2021, which explains the low number of dedicated publications to date.

The distribution of the three functional publication categories along the timeline reflects the early need to document corpus construction methods, technologies, and processes. The majority of STS publications before 2015 ($n = 8$) are categorized as 'methods' and in Sweden the reporting of methods has continued to the present date. Focus on methods papers in one form or another is also evident in SVK and FSTS. NTS is the exception with no publications focusing on methods. While this is likely to change in the future, the authors' first-hand knowledge (see Section 3) suggests that multiple 'methods' publications are not expected for NTS. This is mostly because to date (May 2024) there is already sufficient work published in the Nordics and internationally that addresses the various aspects of building and annotating signed language corpora. As a result, work using the NTS corpus is expected to continue on the path of doing empirical research on the language – an obvious direction of corpus work for other signed languages too (see Börstell 2022).

Overall, the timeline in Figure 1 shows that there has been a gradual shift in numbers from method-focused publications to language-specific and comparative research of Nordic signed languages. The timeline also shows that although the number of corpus research publications has increased over the past twenty years, the total number of publications ($n = 53$) is still relatively low. One factor explaining this relates to the capacity of each country and the researchers working with signed language corpora. For example, in Norway, there is a general lack of signed language linguists trained in corpus annotation, methods, and analysis. This limits how much research can be conducted on the language. Another factor relates to the slow speed of the corpus building process. Because of the amount of data and share of manual work, signed language corpus construction is a slow process. Evidence for this is found in Table 1, where it is reported that basic annotation work for all large corpora except the one for STS, after twelve years of annotation work, is still in progress.

The slow pace of annotation is clearly reflected in publication numbers and types. In moving forward, we wonder if we should be exploring calls for the development of assistive technology or even considering a change in research culture. The question of how much manual work needs and ought to be done in constructing a signed language corpus could be reconsidered. This applies to annotation work in particular: because annotation work is the slowest part of corpus building, it may be that the way annotation work is carried out needs revision. Obviously, a signed language corpus will still need to be machine-readable (see Section 2), but perhaps we can reconsider what basic annotation work (i.e. typically the identification and labeling of signs as well as the creation of sentence-level translations) must be done before other annotation work can be carried out – it is typically only the latter, research topic-focused round of annotation (i.e. the research question-guided annotation done on top of the basic annotation) that is needed when investigating a particular research question.

https://doi.org/10.1017/S0332586524000118 Published online by Cambridge University Press

Overall, we hypothesize that the future of signed language corpus research will likely emphasize the role of comparative research. This trend is already evident in the Nordic corpus signed language work (Figure 1). Now that signed language corpora are more established and have a certain level of basic annotation, comparative studies will be more feasible. Another prediction is that corpus research contributing to signed language learning and teaching will increase in the near future. So far, Sweden has led the way in this area, while in the other Nordic countries the use of corpora for teaching and learning is mainly in its infancy or still in planning stages. This hypothesis is further supported by the fact that the Nordic signed language corpora are all overseen by researchers whose institutions have signed language-dedicated educational programs.

6. Conclusion

This article has been a semi-systematic literature review of corpus research on signed languages in the Nordic countries. In general, the review has shown how research has progressed in Sweden, Finland, and Norway alongside corpus-building processes which partly have kept the total number of corpus research publications relatively modest. However, at the same time, corpora have been built successfully, and over the past twenty years, research exploiting corpora has transitioned from method-focused publications to topics focusing on various linguistic phenomena. Undeniably, the topics investigated have already added to our understanding of (signed) language structure and variation within the signing communities. In the future, the role of comparative research is expected to increase, as are corpus studies linked to signed language learning and teaching. In these fields, because of the small scale of the signed language corpus research field, Nordic research has potential to have a global impact as well.

Acknowledgements. This article is based on a talk presented for the first time at a symposium on FSTS held at the University of Helsinki in Finland on 3–4 November 2022. A second and updated version of the talk was presented at the 3rd Nordic Signed Language Corpus Network workshop, funded by the NOS-HS (The Joint Committee for Nordic research councils in the Humanities and Social Sciences), in Trondheim, Norway, on 12–13 October 2023. The authors thank the anonymous referees of the journal for their comments. Moreover, the authors gratefully acknowledge the funding from the NOS-HS Workshop call 2020 under the Research Council (Academy) of Finland grant 335095 and the NordForsk grant 126546. In addition, the work in Finland (TJ) has been financially supported by the Research Council of Finland under project 339268. In Sweden (JM), the work has been supported by the Swedish national research infrastructure Språkbanken and Swe-Clarin, funded jointly by the Swedish Research Council (2018–24, contract 2017-00626) and the ten participating partner institutions. The work in Norway (LF) has been financially supported by the Norwegian Research Council under project 287067.

Notes

1 The exact number of signed language corpora in the world is unknown, partly because the criteria for corpushood (see Section 2) are treated flexibly. The most well-known and referenced corpora exist in Europe and Australia (Western countries). However, types of corpora exist also, for example, in Brazil, South Africa, Hong Kong, and Japan (see Wehrmeyer 2023).

2 Currently, Denmark and Iceland do not have signed language corpora or related publications, although work for a signed language corpus was planned to begin in Denmark a few years ago (see Troelsgård &

https://doi.org/10.1017/S0332586524000118 Published online by Cambridge University Press

Kristoffersen 2018). There has also been a desire to build a corpus in Iceland. The absence of corpora in these countries is the reason why their signed languages, Danish Sign Language and Icelandic Sign Language (and Faroese and Greenlandic Signed Language; see Mesch 2022), are not covered in this review.

References

Biber, Douglas, Susan Conrad & Randi Reppen. 1998. *Corpus linguistics: Investigating language structure and use*. Cambridge: Cambridge University Press.

Biber, Douglas, Randi Reppen & Eric Friginal. 2012. Research in corpus linguistics. In Robert B. Kaplan (ed.), *The Oxford handbook of applied linguistics*, 2nd edn, 548–568. New York: Oxford University Press.

Börstell, Carl. 2022. Searching and utilizing corpora. In Jordan Fenlon & Julie A. Hochgesang (eds.), *Signed language corpora*, 90–127. Washington, DC: Gallaudet University Press.

Crasborn, Onno & Han Sloetjes. 2008. Enhanced ELAN functionality for signed language corpora. In Onno Crasborn et al. (eds.), *Proceedings of the LREC 2008 3rd Workshop on the Representation and Processing of Signed Languages: Construction and exploitation of signed language corpora*, 39–43. European Language Resources Association. https://www.sign-lang.uni-hamburg.de/lrec/pub/08022.html

Fenlon, Jordan & Julie A. Hochgesang. 2022. Introduction to signed language corpora. In Jordan Fenlon & Julie A. Hochgesang (eds.), *Signed language corpora* (Sociolinguistics in Deaf Communities 25), 1–17. Washington, DC: Gallaudet University Press.

Johnston, Trevor. 2010. From archive to corpus: Transcription and annotation in the creation of signed language corpora. *International Journal of Corpus Linguistics* 15(1). 106–131.

Johnston, Trevor & Adam Schembri. 2013. Corpus analysis of signed languages. In Carol A. Chapelle (ed.), *The encyclopedia of applied linguistics* [online publication, accessed 30 July 2024]. John Wiley. https://doi.org/10.1002/9781405198431.wbeal0252

Mesch, Johanna. 2022. Teckenspråken i Norden [Signed languages in the Nordic countries]. In *Framgång för små språk: En översikt om varför små språk i Norden behöver stärkas och vad som bidrar till ett lyckat språkstärkande arbete. Innehåller en checklista med framgångsfaktorer* [Success for small languages: An overview of why small languages in the Nordic region need to be strengthened and what contributes to successful language-strengthening work. Includes a checklist of success factors], 22–25. Uppsala: Institutet för språk och folkminnen [Institute for Language and Folklore].

McEnery, Tony & Andrew Wilson. 2001. *Corpus linguistics*. Edinburgh: Edinburgh University Press.

Salonen, Juhana, Maria Andersson-Koski, Karin Hoyer & Tommi Jantunen. (2022). Building the Corpus of Finland–Swedish Sign Language: Acknowledging the language history and future revitalization. In Jarmo Harri Jantunen et al. (eds.), *Diversity of methods and materials in digital human sciences: Proceedings of the Digital Research Data and Human Sciences Conference 2022*, 187–199. Jyväskylä: University of Jyväskylä. http://urn.fi/URN:ISBN:978-951-39-9450-1

Snyder, Hannah. 2019. Literature review as a research methodology: An overview and guidelines. *Journal of Business Research* 104. 333–339.

Troelsgård, Thomas & Jette Kristoffersen. 2018. Improving lemmatisation consistency without a phonological description: The Danish Sign Language corpus and dictionary project. In Mayumi Bono et al. (eds.), *Proceedings of the LREC 2018 8th Workshop on the Representation and Processing of Sign Languages: Involving the language community*, 195–198. European Language Resources Association. https://www.sign-lang.uni-hamburg.de/lrec/pub/18009.html

Wehrmeyer, Ella. 2023. (ed.), *Advances in sign language corpus linguistics*. Amsterdam/Philadelphia: John Benjamins.

Appendix: Full list of reviewed publications

1. Arnold, Brittany & Lindsay Ferrara. 2024. 'Your turn!' Using finger pointing and PALM-UP actions to ask questions in Norwegian Sign Language. *Sign Language Studies* 24(3). 621–651.
2. Börstell, Carl. 2019. Differential object marking in sign languages. *Glossa: A Journal of General Linguistics* 4(1). 3.

https://doi.org/10.1017/S0332586524000118 Published online by Cambridge University Press

3. Börstell, Carl. 2022. Searching and utilizing corpora. In Jordan Fenlon & Julie A. Hochgesang (eds.), *Signed language corpora*, 90–127. Washington, DC: Gallaudet University Press.
4. Börstell, Carl. 2024. Evaluating the alignment of utterances in the Swedish Sign Language Corpus. In Eleni Efthimiou et al. (eds.), *Proceedings of the LREC-COLING 2024 11th Workshop on the Representation and Processing of Sign Languages: Evaluation of sign language resources*, 22–31. European Language Resources Association & International Committee on Computational Linguistics. https://www.sign-lang.uni-hamburg.de/lrec/pub/24003.html
5. Börstell, Carl. 2024. How to approach lexical variation in sign language corpora. In Eleni Efthimiou et al. (eds.), *Proceedings of the LREC-COLING 2024 11th Workshop on the Representation and Processing of Sign Languages: Evaluation of sign language resources*, 222–229. European Language Resources Association & International Committee on Computational Linguistics. https://www.sign-lang.uni-hamburg.de/lrec/pub/24026.html
6. Börstell, Carl & Robert Östling. 2016. Visualizing lects in a sign language corpus: Mining lexical variation data in lects of Swedish Sign Language. In Eleni Efthimiou et al. (eds.), *Proceedings of the LREC 2016 7th Workshop on the Representation and Processing of Sign Languages: Corpus mining*, 13–18. European Language Resources Association. https://www.sign-lang.uni-hamburg.de/lrec/pub/16004.html
7. Börstell, Carl, Thomas Hörberg & Robert Östling. 2016. Distribution and duration of signs and parts of speech in Swedish Sign Language. *Sign Language & Linguistics* 19(2). 143–196.
8. Börstell, Carl, Tommi Jantunen, Vadim Kimmelman, Vanja de Lint, Johanna Mesch & Marloes Oomen. 2019. Transitivity prominence within and across modalities. *Open Linguistics* 5(1). 666–689.
9. Börstell, Carl, Johanna Mesch & Lars Wallin. 2014. Segmenting the Swedish Sign Language Corpus: On the possibilities of using visual cues as a basis for syntactic segmentation. In Onno Crasborn et al. (eds.), *Proceedings of the LREC 2014 6th Workshop on the Representation and Processing of Sign Languages: Beyond the manual channel*, 7–10. European Language Resources Association. https://www.sign-lang.uni-hamburg.de/lrec/pub/14023.html
10. Börstell, Carl, Mats Wirén, Johanna Mesch & Moa Gärdenfors. 2016. Towards an annotation of syntactic structure in the Swedish Sign Language Corpus. In Eleni Efthimiou et al. (eds.), *Proceedings of the LREC 2016 7th Workshop on the Representation and Processing of Sign Languages: Corpus mining*, 19–24. European Language Resources Association. https://www.sign-lang.uni-hamburg.de/lrec/pub/16025.html
11. Crasborn, Onno, Johanna Mesch, Dafydd Waters, Els van der Kooij, Bencie Woll & Brita Bergman. 2007. Sharing sign language data online: Experiences from the ECHO project. *International Journal of Corpus Linguistics* 12(4). 535–562.
12. Crasborn, Onno, Els van der Kooij, Dafydd Waters, Bencie Woll & Johanna Mesch. 2008. Frequency distribution and spreading behavior of different types of mouth actions in three sign languages. *Sign Language and Linguistics* 11(1). 45–67.
13. Ferrara, Lindsay. 2020. Some interactional functions of finger pointing in signed language conversations. *Glossa: A Journal of General Linguistics* 5(1). 1–26.
14. Ferrara, Lindsay. 2022. Indexing turn-beginnings in Norwegian Sign Language conversation. *Gesture* 21(1). 1–27.
15. Ferrara, Lindsay & Torill Ringsø. 2019. Spatial vantage points in Norwegian Sign Language. *Open Linguistics* 5. 583–600.
16. Ferrara, Lindsay, Benjamin Anible, & Lena Mei Kalvenes Anda. 2023. Exploring sign-writing contact and multilingualism in the Norwegian Deaf community. In Ella Wehrmeyer (ed.), *Advances in sign language corpus linguistics*, 66–89. Amsterdam/Philadelphia: John Benjamins.
17. Ferrara, Lindsay, Benjamin Anible, Gabrielle Hodge, Tommi Jantunen, Johanna Mesch, Lorraine Leeson & Anna-Lena Nilsson. 2023. A cross-linguistic comparison of reference across five signed languages. *Linguistic Typology* 27(3). 591–627.
18. Gavrilescu, Robert, Carlo Geraci & Johanna Mesch. 2024. Content questions in sign language: From theory to language description via corpus, experiments, and fieldwork. In Eleni Efthimiou et al. (eds.), *Proceedings of the LREC-COLING 2024 11th Workshop on the Representation and Processing of Sign Languages: Evaluation of sign language resources*,

https://doi.org/10.1017/S0332586524000118 Published online by Cambridge University Press

298–306. European Language Resources Association & International Committee on Computational Linguistics. https://www.sign-lang.uni-hamburg.de/lrec/pub/24037.html
19. Jantunen, Tommi. 2017. Constructed action, the clause and the nature of syntax in Finnish Sign Language. *Open Linguistics* 3(1). 65–85.
20. Jantunen, Tommi. 2017. Fixed and NOT free: Revisiting the order of the main clausal constituents in Finnish Sign Language from a corpus perspective. *SKY Journal of Linguistics* 30. 137–149.
21. Jantunen, Tommi, Johanna Mesch, Anna Puupponen & Jorma Laaksonen. 2016. On the rhythm of head movements in Finnish and Swedish Sign Language sentences. In Jon Barnes et al. (eds.), *Speech Prosody 2016: Proceedings of the 8th International Conference on Speech Prosody*, 850–853. International Speech Communication Association. https://doi.org/10.21437/SpeechProsody.2016-174
22. Jantunen, Tommi, Outi Pippuri, Tuija Wainio, Anna Puupponen & Jorma Laaksonen. 2016. Annotated video corpus of FinSL with Kinect and computer-vision data. In Eleni Efthimiou et al. (eds.), *Proceedings of the LREC 2016 7th Workshop on the Representation and Processing of Sign Languages: Corpus mining*, 93–100. European Language Resources Association. https://www.sign-lang.uni-hamburg.de/lrec/pub/16006.html
23. Kankkonen, Nikolaus Riemer, Thomas Björkstrand, Johanna Mesch & Carl Börstell. 2018. Crowdsourcing for the Swedish Sign Language Dictionary. In Mayumi Bono et al. (eds.), *Proceedings of the LREC 2018 8th Workshop on the Representation and Processing of Sign Languages: Involving the language community*, 171–174. European Language Resources Association. https://www.sign-lang.uni-hamburg.de/lrec/pub/18022.html
24. Keränen, Jarkko, Henna Syrjälä, Juhana Salonen & Ritva Takkinen. 2016. The usability of the annotation. In Eleni Efthimiou et al. (eds.), *Proceedings of the LREC 2016 7th Workshop on the Representation and Processing of Sign Languages: Corpus mining*, 111–116. European Language Resources Association. https://www.sign-lang.uni-hamburg.de/lrec/pub/16016.html
25. Leeson, Lorraine, Jordan Fenlon, Johanna Mesch, Carmel Grehan & Sarah Sheridan. 2019. The uses of corpora in L1 and L2/Ln sign language pedagogy. In Rosen S. Russell (ed.), *The Routledge handbook of sign language pedagogy*, 339–352. New York: Routledge.
26. Mesch, Johanna. 2010. Viittomien glossit ja ajalliset pituudet: Annotointityöskentelyyn liittyviä kysymyksiä [The glosses and temporal durations of signs: Questions relating to sign language annotation]. In Tommi Jantunen (ed.), *Näkökulmia viittomaan ja viittomistoon* [Theory and practice in applied linguistics], 43–55. Jyväskylä: University of Jyväskylä.
27. Mesch, Johanna. 2012. Swedish Sign Language Corpus. *Deaf Studies Digital Journal* 3. http://dsdj.gallaudet.edu/index.php?issue=4&section_id=2&entry_id=128
28. Mesch, Johanna. 2015. *Svensk teckenspråkskorpus: Dess tillkomst och uppbyggnad* (Forskning om teckenspråk XXIV) [Swedish Sign Language Corpus: Its creation and structure (Research on sign language XXIV)], 1–25. Stockholm: Section for Sign Language, Department of Linguistics, Stockholm University. https://urn.kb.se/resolve?urn=urn:nbn:se:su:diva-123713
29. Mesch, Johanna. 2016. Manual backchannel responses in signers' conversations in Swedish Sign Language. *Language & Communication* 50. 22–41.
30. Mesch, Johanna. 2023. Creating a multifaceted corpus of Swedish Sign Language: Visual, tactile, and L2 signing. In Ella Wehrmeyer (ed.), *Advances in sign language corpus linguistics*, 242–261. Amsterdam/Philadelphia: John Benjamins.
31. Mesch, Johanna & Lars Wallin. 2008. Use of sign language materials in teaching. In Onno Crasborn et al. (eds.), *Proceedings of the LREC 2008 3rd Workshop on the Representation and Processing of Sign Languages: Construction and exploitation of sign language corpora*, 134–137. European Language Resources Association. https://www.sign-lang.uni-hamburg.de/lrec/pub/08025.html
32. Mesch, Johanna & Lars Wallin. 2012. From meaning to signs and back: Lexicography and the Swedish Sign Language Corpus. In Onno Crasborn et al. (eds.), *Proceedings of the LREC 2012 5th Workshop on the Representation and Processing of Sign Languages: Interactions between corpus and lexicon*, 123–126. European Language Resources Association. https://www.sign-lang.uni-hamburg.de/lrec/pub/12002.html

https://doi.org/10.1017/S0332586524000118 Published online by Cambridge University Press

33. Mesch, Johanna & Lars Wallin. 2015. Gloss annotations in the Swedish Sign Language Corpus. *International Journal of Corpus Linguistics* 20(1). 102–120.
34. Mesch, Johanna, Thomas Björkstrand, Eira Balkstam, Patrick Hansson & Nikolaus Riemer Kankkonen. 2024. Swedish Sign Language resources from a user's perspective. In Eleni Efthimiou et al. (eds.), *Proceedings of the LREC-COLING 2024 11th Workshop on the Representation and Processing of Sign Languages: Evaluation of sign language resources*, 54–61. European Language Resources Association & International Committee on Computational Linguistics. https://www.sign-lang.uni-hamburg.de/lrec/pub/24007.html
35. Mesch, Johanna, Elisabet Cortes, Thomas Björkstrand, Nikolaus Riemer Kankkonen, Joel Bäckström & Patrick Hansson. 2023. Teckenspråkslexikografi: Utmaningar i en annan modalitet [Sign language lexicography: Challenges in a different modality]. In Louise Holmer et al. (eds.), *Nordiska studier i lexikografi 16: Rapport från 16:e konferensen om lexikografi i Norden*, 225–240. Lund/Göteborg: Nordiska föreningen för lexikografi i samarbete med Meijerbergs institut för svensk etymologisk forskning.
36. Mesch, Johanna, Krister Schönström & Sebastian Embacher. 2021. Mouthings in Swedish Sign Language: An exploratory study. *Grazer Linguistische Studien* 93. 107–135.
37. Mesch, Johanna, Lars Wallin & Thomas Björkstrand. 2012. Sign language resources in Sweden: Dictionary and Corpus. In Onno Crasborn et al. (eds.), *Proceedings of the LREC 2012 5th Workshop on the Representation and Processing of Sign Languages: Interactions between corpus and lexicon*, 127–130. European Language Resources Association. https://www.sign-lang.uni-hamburg.de/lrec/pub/12001.html
38. Öqvist, Zrajm, Nikolaus Riemer Kankkonen & Johanna Mesch. 2020. STS-korpus: A sign language web corpus tool for teaching and public use. In Eleni Efthimiou et al. (eds.), *Proceedings of the LREC 2020 9th Workshop on the Representation and Processing of Sign Languages: Sign language resources in the service of the language community, technological challenges and application perspectives*, 177–180. European Language Resources Association. https://www.sign-lang.uni-hamburg.de/lrec/pub/20014.html
39. Östling, Robert, Carl Börstell & Lars Wallin. 2015. Enriching the Swedish Sign Language Corpus with part of speech tags using joint Bayesian word alignment and annotation transfer. In Beáta Megyesi (ed.), *Proceedings of the 20th Nordic Conference on Computational Linguistics (NODALIDA 2015)* (NEALT Proceedings Series 23), 263–268. Linköping: ACL Anthology & Linköping University Electronic Press. http://www.ep.liu.se/ecp/109/ecp15109.pdf
40. Östling, Robert, Carl Börstell, Moa Gärdenfors & Mats Wirén. 2017. Universal dependencies for Swedish Sign Language. In Jörg Tiedemann (ed.), *Proceedings of the 21st Nordic Conference on Computational Linguistics (NODALIDA 2017)* (NEALT Proceedings Series 29), 303–308. Linköping: ACL Anthology & Linköping University Electronic Press. http://www.ep.liu.se/ecp/131/043/ecp17131043.pdf
41. Puupponen, Anna. 2018. The relationship between movements and positions of the head and the torso in Finnish Sign Language. *Sign Language Studies* 18(2). 175–214.
42. Puupponen, Anna, Gabrielle Hodge, Benjamin Anible, Juhana Salonen, Tuija Wainio, Jarkko Keränen, Doris Hernández & Tommi Jantunen. 2024. Opening up Corpus FinSL: Enriching corpus analysis with linguistic ethnography in a study of constructed action. *Linguistics*. https://doi.org/10.1515/ling-2023-0196
43. Puupponen, Anna, Tommi Jantunen & Johanna Mesch. 2016. The alignment of head nods with syntactic units in Finnish Sign Language and Swedish Sign Language. In Jon Barnes et al. (eds.), *Speech Prosody 2016: Proceedings of the 8th International Conference on Speech Prosody*, 168–172. International Speech Communication Association. https://doi.org/10.21437/SpeechProsody.2016-35
44. Puupponen, Anna, Tommi Jantunen, Ritva Takkinen, Tuija Wainio & Outi Pippuri. 2014. Taking non-manuality into account in collecting and analyzing Finnish Sign Language video data. In Onno Crasborn et al. (eds.), *Proceedings of the LREC 2014 6th Workshop on the Representation and Processing of Sign Languages: Beyond the manual channel*, 143–148. European Language Resources Association. https://www.sign-lang.uni-hamburg.de/lrec/pub/14009.html

https://doi.org/10.1017/S0332586524000118 Published online by Cambridge University Press

45. Puupponen, Anna, Laura Kanto, Tuija Wainio & Tommi Jantunen. 2022. Variation in the use of constructed action according to discourse type and age in Finnish Sign Language. *Language & Communication* 83. 16–35.
46. Salonen, Juhana, Maria Andersson-Koski, Karin Hoyer & Tommi Jantunen. 2022. Building the Corpus of Finland–Swedish Sign Language: Acknowledging the language history and future revitalization. In Jarmo Harri Jantunen et al. (eds.), *Diversity of methods and materials in digital human sciences: Proceedings of the Digital Research Data and Human Sciences Conference 2022*, 187–199. Jyväskylä: University of Jyväskylä. http://urn.fi/URN:ISBN:978-951-39-9450-1
47. Salonen, Juhana, Antti Kronqvist & Tommi Jantunen. 2020. The Corpus of Finnish Sign Language. In Eleni Efthimiou et al. (eds.), *Proceedings of the LREC 2020 9th Workshop on the Representation and Processing of Sign Languages: Sign language resources in the service of the language community, technological challenges and application perspectives*, 197–202. European Language Resources Association. https://www.sign-lang.uni-hamburg.de/lrec/pub/20004.html
48. Salonen, Juhana, Anna Puupponen, Ritva Takkinen & Tommi Jantunen. 2019. Suomen viittomakielten korpusta rakentamassa [Building Corpus FinSL]. In Jarmo Harri Jantunen et al. (eds.), *Proceedings of the Research Data and Humanities (RDHum) 2019 Conference: Data, methods and tools*, 83–98. Oulu, Finland: University of Oulu. http://urn.fi/urn:isbn:9789526223216
49. Salonen, Juhana, Ritva Takkinen, Anna Puupponen, Henri Nieminen & Outi Pippuri. 2016. Creating corpora of Finland's sign languages. In Eleni Efthimiou et al. (eds.), *Proceedings of the LREC 2016 7th Workshop on the Representation and Processing of Sign Languages: Corpus mining*, 179–184. European Language Resources Association. https://www.sign-lang.uni-hamburg.de/lrec/pub/16017.html
50. Schönström, Krister & Johanna Mesch. 2022. Second language acquisition of depicting signs: A corpus-based account. *Language, Interaction and Acquisition* 13(2). 199–230.
51. Sipronen, Suvi & Laura Kanto. 2021. Utterance fluency in Finnish Sign Language L1 and L2 signing. *Finnish Journal of Linguistics* 34. 149–177.
52. Takkinen, Ritva, Jarkko Keränen & Juhana Salonen. 2018. Depicting signs and different text genres: Preliminary observations in the corpus of Finnish Sign Language. In Mayumi Bono et al. (eds.), *Proceedings of the LREC 2018 8th Workshop on the Representation and Processing of Sign Languages: Involving the language community*, 189–194. European Language Resources Association. https://www.sign-lang.uni-hamburg.de/lrec/pub/18038.html
53. Takkinen, Ritva, Juhana Salonen, Anna Puupponen & Henri Nieminen. 2020. Miten viittomakielen korpusta luodaan ja mihin sitä tarvitaan? Viittomakielten korpukset ja niiden tehtävät [How is a sign language corpus created and for what?]. *Puhe ja kieli* 40(1). 61–82.

Cite this article: Jantunen T, Mesch J, and Ferrara L (2026). Corpus research on signed languages in the Nordic countries. *Nordic Journal of Linguistics* **49**, 149–162. https://doi.org/10.1017/S0332586524000118

https://doi.org/10.1017/S0332586524000118 Published online by Cambridge University Press