Proceedings of the 10th Workshop on the Representation and Processing of Sign Languages (sign-lang@LREC 2022), pages 110–117
Language Resources and Evaluation Conference (LREC 2022), Marseille, 20-25 June 2022

Making Sign Language Corpora Comparable: A Study of Palm-Up and Throw-Away in Polish Sign Language, German Sign Language, and Russian Sign Language

Anna Kuder
Sign Language Interpreting Section, Department of Special Education and Rehabilitation,
Faculty of Human Sciences, University of Cologne
Section for Sign Linguistics, Faculty of Polish Studies, University of Warsaw
Krakowskie Przedmieście 26/28, 00-927 Warsaw, Poland
anna.kuder@uw.edu.pl

Abstract

This paper is primarily devoted to describing the preparation phase of a large-scale comparative study based on naturalistic linguistic data drawn from multiple sign language corpora. To provide an example, I am using my current project on manual gestural elements in Polish Sign Language, German Sign Language, and Russian Sign Language. The paper starts with a description of the reasons behind undertaking this project. Then, I describe the scope of my study, which is focused on two manual elements present in all three mentioned sign languages: palm-up and throw-away; and the three corpora which are my data sources. This is followed by a presentation of the steps taken in the initial stages of the project in order to make the data comparable. Those steps are: choosing the adequate data samples from all three corpora, gathering all data within the chosen software, and creating an annotation schema that builds on the annotations already present in all three corpora.

Even though the project is still underway, and the annotation process is ongoing, preliminary discussions about the nature of the analysed manual activities are presented based on the initial annotations for the sake of evaluating the created annotation schema. I conclude the paper with some remarks about the performance of the employed methodology.

Keywords: gesture, sign, sign language corpus, corpus linguistics, annotation, Polish Sign Language (PJM), German Sign Language (DGS), Russian Sign Language (RSL), comparative studies

1. Introduction

For many years the standard of sign language (SL) research was based only on small samples of language material and/or the researcher's (and/or his/her informant's) own linguistic intuitions. This approach based on elicited data and linguistic judgements was used both in research regarding single SLs (e.g., Zeshan, 2006), and in comparative studies of multiple SLs (e.g., Pfau and Quer, 2004). In more recent years, since the creation of the Australian Sign Language (Auslan) Corpus (Johnston, 2009) and similar projects that have followed, studies based on corpus material are becoming more common for the analysis of individual SLs. For comparative studies of multiple SLs however, the approach utilizing elicited data and linguistic judgements is still more common.

But with the growing number of available resources, more and more cross-linguistic studies are being performed with the use of data coming from two of more separate corpora. Some examples include: the comparison of negation markers in Polish Sign Language (PJM) and Auslan (Kuder et al., 2018), the comparison of information structure in Russian Sign Language (RSL) and Sign Language of the Netherlands (NGT) (Kimmelman, 2019); the comparison of body-anchored verbs and argument omission in DGS and RSL (Oomen and Kimmelman, 2019); the comparison of discourse markers in French Belgian Sign Language (LSFB) and Catalan Sign Language (LSC) (Gabarró-López, 2020).

In line with this more recent trend, I set out to perform a cross-linguistic corpus-based study of two manual elements present in three European SLs: Polish Sign Language, German Sign Language, and Russian Sign Language. As the corpora from which I draw my data were primarily created using different standards for annotation, and in different software (PJM and DGS corpora – in iLex, and RSL corpus – in ELAN), this paper details the choices faced and decisions made in the preparation phase of the large-scale cross-linguistic corpus-based study.

2. Theoretical Background & Motivation

The topic of gesture and gesticulation has been tackled more often by spoken language (SpL) than SL linguists. Gestures, defined in SpLs as "visible actions of the hand, body, and face that are intentionally used to communicate" (Kendon, 1986, 2004, following: Özyürek, 2012, p. 627), are usually seen as integrated into the communication system, being another part of language, alongside speech (Özyürek, 2012).

This view of gestures is supported by the fact that, in SpLs, gestures are most often produced in a different modality than speech (e.g., Goldin-Meadow, 2003; Kendon, 2004; McNeill, 2005). They are easily distinguishable from fully syntactic elements just by being "shown" and not "said" (note the common notion of gestures as being "non-verbal"). Elements that are being "shown" while a spoken word/clause is being uttered are called co-speech gestures. However, this is not the case for gestures accompanying SLs, in which there is no modality difference between lexical and gestural elements. The fact that both signs and gestures in SLs are "shown" has led researchers to trying to establish a more prominent relationship between them than has ever been argued for SpLs. Namely, it has been claimed that some of the elements that in SpL linguistics are referred to as gestures, when present in SLs take on a grammatical function in a process known as grammaticalization, and instead are referred to as grammatical markers. This has been stated with respect to both non-manual elements, e.g., headshaking, and manual elements, e.g., palm-up (van Loon et al., 2014).

This approach to dealing with gestural elements in SLs stems from fact that SL researchers "naturally adopted the

theoretical and analytic tools that were established in spoken language linguistics" (Lepic, 2019, p. 3). Using these tools on SL data has led them to establish strict claims about lexicalization and grammaticalization of certain elements in some SLs (i.e., multiword expressions and morphologically complex signs (Lepic, 2019)).

However, some recent large-scale corpus-based studies provide evidence that contradicts these previous claims. It has been shown that elements serving as co-speech gestures in SpLs, when studied on the basis of SL corpus data, turn out to function in SLs in a similar way they do in SpLs (e.g., Johnston, 2018; Kuder, 2021 for headshaking), suggesting that they should not have been described as grammaticalized as previously stated. If claims must be made about the nature of these elements in SLs, then adopting a usage-based framework "alleviates the burden for sign language linguists to determine whether or not linguistic constructions have become <<lexicalized>>" (Lepic, 2019, p. 1) or, in this case, grammaticalized. Instead, by focusing only on the degree of analysability (Lepic, 2019) of an element, we can compare to what extent each element has been conventionalized (e.g., Schmid, 2020).

My current project follows the corpus-based approach and applies it to manual gestural elements present in SLs, to help gain a new perspective on the analysability of gestural elements in SLs and add to the discussion about the nature and role of gestural elements in SL discourse. The project is motivated by the need to conduct comparative studies of gestures across different sign languages which has been directly expressed by other authors (here with respect to palm-up): "there have already been several insightful corpus-based treatments of the palm-up in sign, but especially valuable would be further studies that compare use of the form in different sign languages using the same analytic criteria and theoretical framework. Such an approach would be critical in distinguishing cross-linguistic patterns from language-specific particulars" (Cooperrider et al., 2018, p. 12).

3. Scope of the Study and Data Sources

My current study focuses on two manual activities present both in SLs and SpLs:

- the **palm-up** – a multifunctional manual activity taking the form of rotating one's forearms so that the palms of the hand face upward (e.g., Cooperrider et al., 2018 among others; see fig. 1);
- the action of an open hand going downward having a common meaning of "never mind" or "not important" (Bressem and Müller, 2014; see fig. 2), which I will call the **throw-away**.

Throw-away has so far only been studied for co-speech gesture (Bressem and Müller, 2014, 2017, Francis et al., 2022).

Palm-up, on the other hand, is a manual form that has received a lot of scientific attention. It has been thoroughly studied in a number of sign languages: New Zealand Sign Language (McKee and Wallingford, 2011), Sign Language of the Netherlands (van Loon et al., 2014), Danish Sign Language (Engberg-Pedersen, 2002) and American Sign Language (Conlin et al., 2003). Small-scale studies of palm-up are also present for German Sign Language (Volk, 2016) and Russian Sign Language (Bauer, 2019). Preliminary comparative corpus-based studies of palm-up were also undertaken for French Belgian and Catalan Sign Languages (Gabarró-López, 2020). There are also analyses exploring the origin and relations of the element in signed and spoken communication (Cooperrider et al., 2018).

Figure 1: A palm-up (photo from the PJM corpus).

Figure 2: A throw-away (photo from the PJM corpus).

No large-scale and entirely corpus-based study has been conducted across multiple sign languages to compare the use of these two elements, which my study will provide. My current project is based on naturalistic corpus data extracted from the PJM, DGS and RSL corpora, all of which have open-access repositories. A substantial part of the PJM corpus is made publicly accessible as the "Open Repository of the Polish Sign Language Corpus" (Wójcicka et al., 2020; Kuder et al., this volume; https://www.korpuspjm.uw.edu.pl/en). The DGS corpus project is accessible as the "Public DGS Corpus" (with three different levels of access, Konrad et al., 2020; https://www.sign-lang.uni-hamburg.de/meinedgs/ling/start_en.html) and the RSL corpus as the "Online Russian Sign Language Corpus" (Burkova, 2015; http://rsl.nstu.ru).

4. Making Datasets Comparable

As all three corpora were created separately and published in different ways, the process of making my language material comparable involved 3 main questions:

(I) Which software(s) should be used for annotation?

(II) How to choose comparable data samples?

(III) How best to create an annotation schema that builds on the annotations already present in all three corpora?

4.1 Software

All three corpus projects were created and are published in different ways. Both PJM and DGS corpora were primarily created with the use of iLex (Hanke and Storz, 2008), while RSL corpus was made using ELAN (Crasborn and Sloetjes, 2008). Using two different tools throughout the project would make comparison difficult, if not impossible. However, all files in the repositories of PJM and DGS corpora are available to download both in iLex and ELAN formats. Therefore, I decided to work with only the ELAN files throughout my whole project. Importing the RSL annotation files into iLex would have been possible but was deemed unnecessary for a project conducted by an individual. If the study was conducted by a project team that needed to work on the annotation files at the same time, then using iLex would have been recommended instead.

4.2 Data Samples

To obtain comparable results, the data samples had to be chosen carefully, as each of the corpora features a different number of recorded informants and different lengths of recorded texts. A sample of 16 informants from each corpus was picked to be annotated. Each sample is balanced out with respect to gender (8 males & 8 females), and age (4 informants – 2 males & 2 females – from each of the age groups: 18-30; 31-45; 46-60; 60+).

As the geographical division of the data in the Polish Sign Language Corpus mirrors the distribution of Poland into 16 voivodeships[1], my sample includes one informant from each part of the country. The DGS corpus is also balanced geographically, following the division of the country into 13 regions[2] which correspond to the location of current and former Deaf schools. I thus decided to include one informant from each of the regular regions and two from the three biggest ones: Berlin, Leipzig and Nürnberg. The data from the Russian Sign Language corpus was collected in two places: Moscow and Novosibirsk[3]. Therefore, I decided to include 8 informants from each of the regions in the RSL sample.

The corpora differ also when it comes to the publication format of the publicly available files: approx. half of the files from the PJM Open Repository present signers talking in pairs and half of them present single signers. In the DGS files signers are almost always presented in pairs. Most of the RSL files only show one informant at a time. Due to the different formats of the three corpora, only the material coming from a single signer will be used in the study. For the dialogical tasks which show people signing in pairs, only data coming from one informant will be annotated per task.

The next decision was to choose suitable text produced by the informants so that the final samples would be as similar as possible with respect to text types and length. This was the most challenging part of the preparation phase, as here more than elsewhere I was limited to the material present in the open access corpora repositories. My final choices are presented in table 1 below.

	dialogue	narrative/ monologue	retelling
PJM	14 texts	24 texts	37 texts
DGS	5 texts	38 texts	3 texts
RSL	1 text	42 texts	27 texts

Table 1: The distribution of text types in the data samples from three corpora.

4.3 Existing Annotations

The biggest obstacle faced in the data preparation is the fact that the annotation schemas used in the original files from all three repositories are not identical, albeit similar.

As none of the present schemas was detailed enough to provide a good template to the study of gestural elements, a new schema had to be created. It had to be developed in such a way that would make use of the existing annotations and at the same time grasp all features of the articulatory elements important from the point of view of my study. This new schema needed to be developed in such a way that it could be applied in the files coming from all three corpora.

Only the tiers appearing consistently in all three datasets could have been consistently used in the study. These were limited to: tiers for glosses for dominant and non-dominant hand, and free translation. A comparison of all tiers existing in the files prior to starting the study is presented in table 2 below.

	PJM	DGS	RSL
Glosses for dominant and non-dominant hand (signer A[4]; written in the native language)	✓	✓	✓
Glosses for dominant and non-dominant hand (signer B; written in the native language)		✓	
Glosses for dominant and non-dominant hand (signer A; written in English)		✓	
Glosses for dominant and non-dominant hand (signer B; written in English)		✓	
Free translation into the native language	✓	✓	✓
Free translation into the English		✓	
HamNoSys notation	✓		
Mouthing/mouth gesture (signer A)		✓	✓
Mouthing/mouth gesture (signer B)		✓	
Non-manual features present on head, body & face			✓

Table 2: Overview of the annotation schemas used in the open repositories of all the three corpora prior to starting the current study.

[1] https://www.korpuspjm.uw.edu.pl/en

[2] https://www.sign-lang.uni-hamburg.de/meinedgs/ling/start_en.html

[3] http://rsl.nstu.ru/site/data

[4] As the files from the DGS corpus show two informants at the same time the relevant tiers are doubled to present the annotations for both signers separately.

The tiers for glosses and translations were used in the study in their present form. No alterations were made to the glossing and translating conventions. Even though they were not identical in all three datasets, they are similar enough from the point of view of the study which is not targeted to research purely lexical elements. Some tiers present in single datasets were important from the point of view of the current study (e.g., tiers for coding mouthing and non-manual elements). In such cases, existing data could already be used as is, but needed to be annotated from scratch for the remaining datasets.

4.4 The New Annotation Schema

The new schema was build based on the reports present in the literature concerning the elements important during studying the manual gestural elements in SLs (e.g., Cooperrider et al., 2018) and my own experience in building and using SL corpora (e.g., Kuder et al., 2018). The annotation process consists of four steps: (1) identifying all occurrences of palm-up and throw-away and defining their manual form; (2) defining the non-manual features associated with a given occurrence; (3) defining the function of the occurrence; (4) delineating the clauses that the occurrences are contained in.

As all three corpora feature pre-existing glosses for the two targeted manual elements (even though they are glossed differently in each of the corpora), the base of step (1) was already pre-prepared in all three datasets. After identifying each occurrence, I coded for[5]:

a) manual type (is it a palm-up or throw-away and is it one- or two-handed),
b) manual subtype (following Kendon, 2004 and Cooperrider et al., 2018 four subtypes of palm-up were distinguished: lateral, presentational, addressed and pointing).

In the (2) step I marked:

c) placing in the signing space,
d) handshape assimilation (if present),
e) nonmanual elements on the body,
f) nonmanual elements on the head,
g) nonmanual elements on the face,
h) gaze of the signer (if distinguishable by bare eye),
i) mouthing/mouth gesture.

If it was needed any additional information was added on the tier called:

j) "comment".

Step (3) consisted of tagging for:

k) function of the palm-up,
l) function of the throw-away,
m) lexical meaning of palm-up (if present),
n) lexical meaning of throw-away (if present).

Even though the files are equipped in pre-existing glosses for both palm-up and throw-away and in written translations, during the annotation process the whole video files are inspected sign by sign. This is needed to properly grasp the context of signing, which is crucial for establishing the function of the given manual element. Ambiguous cases are discussed with signers of each language.

The functional analysis was conducted based on pre-existing corpus annotations, my knowledge of the languages, observed context of signing, and consultations with users of the three target languages. The initial set of function tags was based on literature and then later augmented while studying the data, as not all of the functions I observed were previously reported on in the literature. I ended up with approx. 50 detailed function tags, which were later grouped into four broader categories (see section 5.3 for details).

Coding each occurrence with respect to the 14 listed tiers constitutes the first round of annotation for any given file. Annotations from these tiers are being used for cross-linguistic frequency counts and analyses of correlation of form and functions of the manual elements in question (see sections 5.2 & 5.3 for preliminary results).

Step (4) of annotation (the sentential annotation) serves the purpose of distinguishing "basic articulatory chunks of propositional meaning" (Johnston, 2019). It follows the protocol for clause like units (CLUs) tagging proposed by Johnston (2019) and adapted during the creation of the Polish Sign Language Corpus. This part of annotation consists of defining the boundaries of CLUs and then distinguishing their predicates, main arguments, and peripheral elements. The predicates and arguments are tagged for the macro roles and semantic roles they exhibit in the clauses. They are also marked with tags for parts of speech and in this process, I take into consideration all issues connected with distinguishing parts of speech (PoS) in sign languages (Schwager and Zeshan, 2008) and employ a usage-based notion of PoS (Linde-Usiekniewicz and Rutkowski, 2016) which focuses on the usage of a given sign in a given context. The types of the CLUs and dependencies between the clauses are then marked for, before adding the English translation.

Therefore, this subsection of my annotation schema contains eight tiers (see also fig. 3):

o) CLU (used for marking the scope of the clause),
p) arguments in the CLU (used for marking the predicate, its arguments, and peripherals),
q) macro roles in the CLU,
r) semantic roles in the CLU,
s) part of speech,
t) sentence type
u) type of CLU,
v) CLU within CLU (used for marking dependencies between the clauses),
w) English translation (on the basis of the written translations already present in the corpora).

Data collected in this round of annotation will be used in the future stages of the project for establishing what is the position of the manual elements in question within the sign languages clauses and whether there is a correlation between the position in the clause and a specific function or meaning of palm-up and throw-away.

5. Current State of the Project

5.1 Annotated Data Sample

As the project is still ongoing, so far the material coming from 9 informants from each of the corpora was annotated with the first round of annotation. The overview of the annotated sample is presented in the tables below.

[5] Each letter corresponds to a single tier in the annotation schema – see fig. 3.

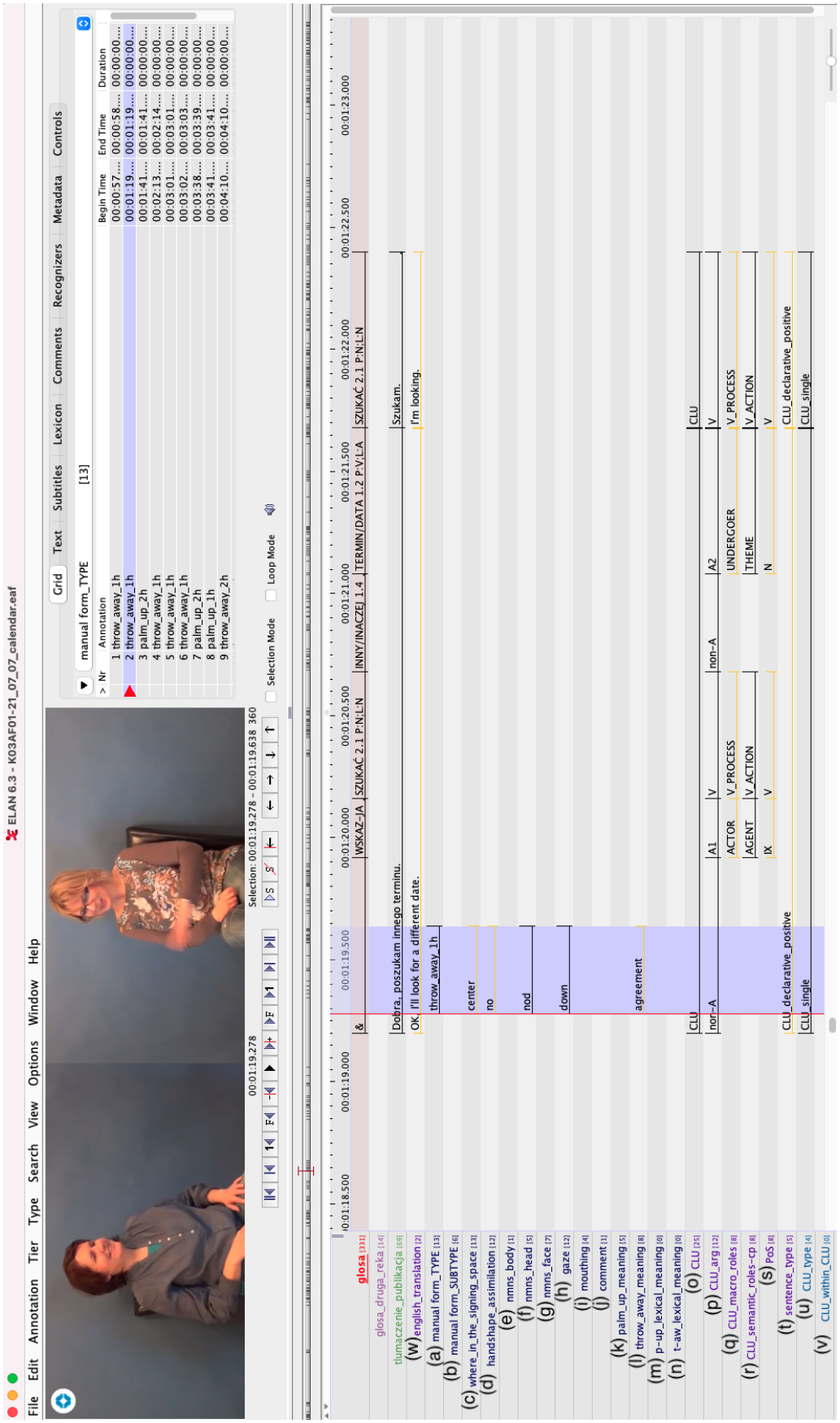

Figure 3: Annotation schema (photo from the PJM corpus).

	no. of texts	text types	length	no. of produced signs
PJM	43	retelling: 26 narratives: 9 dialogues: 8	05:37:32	20,851
DGS	27	narrative: 16 dialogues: 11	02:47:03	11,048
RSL	35	retelling: 16 narratives: 18 dialogue: 1	01:00:59	6,313

Table 3: The overview of the annotated dataset.

	gender/ age	18-30	31-45	46-60	60+
PJM	F	1	1	-	2
	M	2	1	1	1
DGS	F	1	1	-	2
	M	1	2	1	1
RSL	F	1	1	2	-
	M	2	1	1	1

Table 4: Age and gender of informants.

5.2 Preliminary Findings: Quantitative Analysis

As previously mentioned, the frequency analysis was based mostly on the pre-existing glosses present in all three corpora. However, aside from just targeting the existing glosses, I also examined the videos sign by sign, so as not to miss any instances of the manual forms (which may have been tagged with different labels than the anticipated ones). This also was needed for the functional analyses I will describe below. Fully understanding what is being signed was crucial for properly determining the functions of the manual elements, as they are heavily context based.

The frequency of the occurrence of palm-up and throw away in all three data samples is summarized in the table 5 below.

	no. of palm-ups (and as a % of all manual signs annotated)	no. of throw-aways (and as a % of all manual signs annotated)
PJM	729 (3.49%)	310 (1.49%)
DGS	734 (6.64%)	133 (1.20%)
RSL	269 (4.26%)	86 (1.36%)

Table 5: The frequency of both manual elements in the datasets.

The findings are consistent with the literature reports about the frequency of palm-up in other SLs of the world. For example, in the study of lexical frequency in British Sign Language (BSL), Fenlon and colleagues (2014) found that the percentage of palm-up occurrence stays at 5.5% making palm-up the second most frequent type of manual activity in the BSL data. They compared it to the Australian Sign Language (Auslan) data, in which the occurrence rate stays at 3.6% (Fenlon et al., 2014). In New Zealand Sign Language (NZSL), palm-up comprises 5% of all manual signs in the corpus and is the second most frequent sign type in the studied sample (McKee and Wallingford, 2011). In the next phases of the project, I will investigate the slightly higher occurrence rate of palm-up in DGS than in the other two languages.

When it comes to throw-away I have less possibilities for cross-linguistic comparison, but the percentages seem to be similar across studied languages.

What is more, these figures are consistent when checked against the whole of the PJM corpus, which currently comprises of approx. 706,233 glosses, of which palm-up is the second most common manual activity with approx. 30,558 occurrences (4.33%). Following this is throw-away with 7,134 occurrences, which put its frequency percentage at 1.01%.

The fact that the used method yields results comparable with the literature report about similar elements in other SLs shows that the chosen apparatus is working as planned.

5.3 Preliminary Findings: Qualitative Analysis

If the data prepared with the use of the newly formed annotation schema is adequate, then it will allow for a cross-modal comparison with what has been reported about palm-up and throw-away in co-speech gesture.

This can be done on the basis of the step (3) in the annotation process – the analysis of the elements' functions. As mentioned previously, all the detailed functions of the studied manual elements were grouped into four categories based on the type of function. The first three categories (van Loon et al., 2014; Bauer, 2019), which are also used to describe the functions of palm-up in co-speech gesture (cf. Ferré, 2012) are:

- Expressing modal meanings:
  - positive (e.g., agreement; revelation; surprise);
  - negative (e.g., lack of knowledge, lack of understanding, lack of interest, lack of ability; negation, surprise; annoyance; disappointment);
  - neutral (e.g., hesitation; hypotheticality; reinforcement of the stance);
- Discourse regulation: e.g., turn/topic opening or ending; response to the interlocutor's question/stance; connecting sentences;
- Conveying coherence: e.g., meta-comment; rhetorical question; self-correction.

My data suggests that all the functions performed by throw-away in all three SLs also fit into this categorisation.

The last category, labelled as "conveying lexical meaning", features all occurrences of both manual activities that were coded with lexical glosses by the original annotators. This tag was inserted in the "function" tier and the lexical meaning was specified on another annotation level (see the tiers labelled "lexical meaning of palm-up" and "lexical meaning of throw-away" in the fig. 3). The consistency of co-occurrence of palm-up and throw-away with particular lexical glosses raises an important question about the conventionalization level of the elements in question and the reports of palm-up functioning as a grammatical marker (van Loon et al., 2014). Some of the meanings consistently co-occurring with palm-up and throw-aways in the three SLs also possess different, fully lexicalized, manual forms in their lexicons (e.g., NOT-HAVE; NOT-BE; NOT-KNOW in PJM which I found to be associated with palm-up or BAD; TO-LET; DROP in DGS which I found to be

associated with throw-away). But the signers occasionally chose to substitute them with palm-up or throw-away and were understood by both the interlocutor and later by the annotators who chose to gloss the occurrence with a lexical gloss rather than a gestural marker. Future efforts within the study will be targeted towards explaining this issue within the usage-based framework (Lepic, 2019) and towards explaining the similarity of the functions of palm-up and throw-away observed in both signed and spoken modality.

6. Conclusion

The aim of this paper was primarily to show the preparation phase of a comparative corpus-based project when dealing with multiple SL corpora. The chosen methodology and annotation schema appear to be working well enough to provide adequate data to already allow preliminary conclusions about the nature of the analysed manual activities to be drawn.

The three issues connected to the topic of data comparability raised in the section 4 can be assessed as follows.

(I) Performing the annotations in ELAN was a good decision due to the very powerful search engine that is built into the software. Searching throughout annotated files is a key element of calculating the results. Searching in ELAN is more straightforward for a researcher without a programming background than searching within iLex, which requires the knowledge of SQL queries. The central database functionality of iLex was not needed for this project but would make iLex the preferred tool in any multi-annotator setting.

(II) The chosen data sample seems to be representative of the language usage as the obtained quantitative results are consistent with existing literature reports about palm-up in other SLs.

(III) The developed annotation schema, when applied to the chosen data sample, is providing adequate information about the frequency, form, and function of the two studied manual elements in all three SLs and allows for both cross-language and cross-modal comparison with the previous literature reports about the same topic in both signed and spoken languages. If anything, the schema might be too detailed. When it comes to coding for eye-gaze for example, it is unclear at this point if the corpus material is providing adequate data. It is hard to delineate the features that affect the signer's eye-gaze in the conversational data. Probably eye-gaze studies should be mainly based on the data obtained with the use of an eye tracker.

As mentioned previously, the current project is still ongoing. In order to gain a better understanding of the actual usage of the manual elements in question and to better understand the level of their conventionalisation, the next stages of my project will be devoted to conducting:

- analysis of co-occurrence of both gestures' types and subtypes with specific nonmanual markers;
- analysis of the correlation between the gestures' types and subtypes and their function;
- sociolinguistic analyses of the usage of the gestures across genders and age groups;
- CLU (sentential) coding and analysis;
- more detailed comparison of the gestures' usage between SLs and co-speech gesture.

The annotation schema has been prepared in a way that should make it possible to tackle all of these topics. However, assessing the choices and decisions made along the way will have to be done again, upon completion of the project (in the next 12 months). With the results of this further analysis, I hope to be able to add more direct claims to the discussion about the conventionalisation of palm-up and throw-away in the three studied SLs, as previously discussed in the theoretical background. If the assessment will yield positive results, in the future this project might serve as a basis for creating a blueprint for other comparative corpus-based studies.

7. Acknowledgements

My current research is financially supported by the Polish National Agency for Academic Exchange within the Bekker Programme (edition 2020): PPN/BEK/2020/1/00254.

Proofread and edited by Oliver B. Woof.

8. Bibliographical References

Bauer, A. (2019). Nonmanual components with palm-up in Russian SL. Paper presented at the *93rd LSA Annual Meeting*, New York, NY, USA.

Bressem, J. and Müller, C. (2014). The family of Away gestures: Negation, refusal, and negative assessment. In C. Müller et al. (Eds.), *Body – Language – Communication*, Berlin: De Gruyter Mouton, pp. 1592–1604.

Bressem, J. and Müller, C. (2017). The "Negative-Assessment-Construction" – A multimodal pattern based on a recurrent gesture? *Linguistic Vanguard*, 3:1–9.

Conlin, F., Hagstrom, P. and Neidle, C. (2003). A particle of indefiniteness in American Sign Language. *Linguistic Discovery*, 2(1):1–21.

Cooperrider, K., Abner, N. and Goldin-Meadow, S. (2018). The Palm-Up Puzzle: Meanings and Origins of a Widespread Form in Gesture and Sign. Frontiers in Communication, 3(23), https://doi.org/10.3389/fcomm.2018.00023.

Crasborn, O. and Sloetjes, H. (2008). Enhanced ELAN functionality for sign language corpora. In Proceedings of the 3rd Workshop on the Representation and Processing of Sign Languages: Construction and Exploitation of Sign Language Corpora. 6th International Conference on Language Resources and Evaluation (LREC 2008), pages 39–43, France, Paris, May. European Language Resource Association (ELRA).

Engberg-Pedersen, E. (2002). Gestures in signing: The presentation gesture in Danish Sign Language. In R. Schulmeister & H. Reinitzer (Eds.), *Progress in Sign Language Research: In Honor of Siegmund Prillwitz*. Hamburg: Signum, 143–162.

Fenlon, J., Schembri, A., Rentelis, R., Vinson, D. and Cormier, K. (2014). Using conversational data to determine lexical frequency in British Sign Language: The influence of text type. *Lingua*, 143:187–202.

Ferré, G. (2012). Functions of three open-palm hand gestures. *Multimodal Communication*, 1(1):5–20.

Francis, N., Grosz, P.G., Patel-Grosz, P. (2022). Analyzing the throwing away gesture as a discourse management device. Paper presented at the *DGfS-Workshop: Visual Communication. New Theoretical and Empirical Developments*, Tübingen, Germany.

Gabarró-López, S. (2020). Are discourse markers related to age and educational background? A comparative account between two sign languages. *Journal of Pragmatics*, 156:68–82.

Goldin-Meadow, S. (2003). Hearing Gesture: How Our Hands Help Us Think. Cambridge, MA: Harvard University Press.

Hanke, T and Storz, J. (2008). iLex – A database tool for integrating sign language corpus linguistics and sign language lexicography. In *Proceedings of the 3rd Workshop on the Representation and Processing of Signed Languages: Construction and exploitation of sign language corpora. International Conference on Language Resources and Evaluation* (*LREC 2008*), pages 64–67, France, Paris, May. European Language Resource Association (ELRA).

Johnston, T. (2009). Creating a corpus of Auslan within an Australian national corpus. In M. Haugh, K. Burridge, J. Mulder and P. Peters (Eds.), *Selected Proceedings of the 2008 HCSNet Workshop on Designing the Australian National Corpus*: *Mustering Languages*. Somerville: Cascadilla Proceedings Project, pp. 87–95.

Johnston, T. (2018). A corpus-based study of the role of headshaking in negation in Auslan (Australian Sign Language): Implications for signed language typology. *Linguistic Typology*, 22(2): 185–231.

Johnston, T. (2019). Auslan Corpus Annotation Guidelines. August 2019 Version, manuscript, Macquarie University, Sydney, Australia.

Kendon, A. (1986). Some reasons to studying gesture. *Semiotica*, 62(1-2):3–28.

Kendon, A. (2004). Gesture. Visible action as utterance. Cambridge: Cambridge University Press.

Kimmelman, V. (2019). Information structure in Sign Languages. Evidence from Russian Sign Language and Sign Language of the Netherlands. Berlin, Boston: De Gruyter Mouton.

Kuder, A., Filipczak, J., Mostowski, P., Rutkowski, P. and Johnston, T. (2018). What Corpus-based Research on Negation in Auslan and PJM Tells Us about Building and Using Sign Language Corpora. In *Proceedings of the 8th Workshop on the Representation and Processing of Sign Languages: Involving the Language Community. Language Resources and Evaluation Conference* (*LREC 2018*), pages 101–106, Miyazaki, Japan, May. European Language Resource Association (ELRA).

Kuder, A. (2021). Negation markers in Polish Sign Language (PJM), Sign Language and Linguistics, 24(1):118–131.

Kuder, A., Wójcicka, J., Mostowski, P., Rutkowski, P. (this volume). Open Repository of the Polish Sign Language Corpus: Publication Project of the Polish Sign Language Corpus. *In Proceedings of the 10th Workshop on the Representation and Processing of Sign Languages: Multilingual Sign Language Resources. Language Resources and Evaluation Conference* (*LREC 2022*), Marseille, France, June. European Language Resource Association (ELRA).

Lepic, R. (2019). A usage-based alternative to "lexicalization" in sign language linguistics. *Glossa: a journal of general linguistics*, 4(1):23:1–30. https://doi.org/10.5334/gjgl.840.

Linde-Usiekniewicz, J. and Rutkowski, P. (2016). The division into parts of speech in the Corpus-based Dictionary of Polish Sign Language. In T. Margalitadze & G. Meladze (Eds.), *Proceedings of the XVII EURALEX International Congress: Lexicography and Linguistic Diversity*. Tbilisi: Ivane Javakhishvili Tbilisi University Press, pp. 375–388.

McKee, R. and Wallingford, S. (2011). "So, well, whatever": Discourse functions of palm-up in New Zealand SL. *Sign Language & Linguistics*, 14:213–247.

McNeill, D. (2005). Gesture and Thought. Chicago: University of Chicago Press.

Müller, C. (2004). Forms and uses of the Palm Up Open Hand: A case of a gesture family? In C. Müller, & R. Posner (Eds.), *The Semantics and Pragmatics of Everyday Gestures*. Berlin: Weidler, pp. 233–256.

Oomen, M. and Kimmelman, V. (2019). Body-anchored verbs and argument omission in two sign languages. *Glossa: a journal of general linguistics*, 4(1): 42.

Özyürek, A. (2012). Gesture. In R. Pfau, M. Steinbach, & B. Woll (Eds.), *Sign Language. An International Handbook*. Berlin/Boston: De Gruyter Mouton, pp. 626–646.

Pfau, R. and Quer, J. (2004). On the syntax of negation and modals in LSC and DGS, paper presented at the *26. Jahrestagung der Deutschen Gesellschaft für Sprachwissenschaft*, Mainz, Germany.

Schmid, H.-J. (2020). The Dynamics of the Linguistic System: Usage, Conventionalization, and Entrenchment. Oxford Scholarship Online.

Schwager, W. and Zeshan, U. (2008). Word classes in sign languages. Criteria and classifications. *Studies in Language* 32(3): 509–545.

Van Loon, E., Pfau, R. and Steinbach, M. (2014). The grammaticalization of gestures in sign languages. In C. Müller et al. (Eds.), *Body – Language – Communication*. Berlin: De Gruyter Mouton, pp. 720–730.

Volk, E. (2016). Discourse functions of palm-up in German SL (DGS). Paper presented at the *FEAST conference*, Venice, Italy.

Zeshan, U. (Ed.) (2006). Interrogative and Negative Constructions in Sign Languages. Nijmegen: Ishara Press.

9. Language Resource References

Burkova, S. (2015). Russian Sign Language Corpus. Novosibirsk. http://rsl.nstu.ru/site/signlang.

Konrad, R., Hanke, T., Langer, G., Blanck, D., Bleicken, J., Hofmann, I., Jeziorski, O., König, L., König, S., Nishio, R., Regen, A., Salden, U., Wagner, S., Worseck, S., Böse, O., Jahn, E. and Schulder, M. (2020). MY DGS – annotated. Public Corpus of German Sign Language, 3rd release [Dataset]. Universität Hamburg. https://doi.org/10.25592/dgs.corpus-3.0.

Wójcicka, J., Kuder, A., Mostowski, P. and Rutkowski, P. (2020). Open Repository of the Polish Sign Language Corpus. Warsaw: Faculty of Polish Studies, University of Warsaw. https://www.korpuspjm.uw.edu.pl/en