

Research Article

Carl Börstell*, Tommi Jantunen, Vadim Kimmelman, Vanja de Lint, Johanna Mesch, and Marloes Oomen

Transitivity prominence within and across modalities

<https://doi.org/10.1515/opli-2019-0037>

Received May 23, 2018; accepted Dec 23, 2019

Abstract: We investigate transitivity prominence of verbs across signed and spoken languages, based on data from both valency dictionaries and corpora. Our methodology relies on the assumption that dictionary data and corpus-based measures of transitivity are comparable, and we find evidence in support of this through the direct comparison of these two types of data across several spoken languages. For the signed modality, we measure the transitivity prominence of verbs in five sign languages based on corpus data and compare the results to the transitivity prominence hierarchy for spoken languages reported in Haspelmath (2015). For each sign language, we create a hierarchy for 12 verb meanings based on the proportion of overt direct objects per verb meaning. We use these hierarchies to calculate correlations between languages – both signed and spoken – and find positive correlations between transitivity hierarchies. Additional findings of this study include the observation that locative arguments seem to behave differently than direct objects judging by our measures of transitivity, and that relatedness among sign languages does not straightforwardly imply similarity in transitivity hierarchies. We conclude that our findings provide support for a modality-independent, semantic basis of transitivity.

Keywords: transitivity, corpus linguistics, sign languages, valency, typology

1 Introduction

Transitivity has traditionally been recognized as a central notion in the organization of sentences in natural languages (see e.g., Dixon, 1979; Hopper and Thompson, 1980; Dowty, 1982; Næss, 2007; Kittilä, 2010). While a wealth of research into the topic has been published over the years, the work has primarily focused on languages in the spoken modality (see Section 1.1). An exception is Kimmelman (2016), who investigated transitivity in Russian Sign Language (RSL; Section 1.2). This study builds on Kimmelman's work but extends the scope to five sign languages, thus enabling comparisons both within and across modalities. It is also an extension of previous work by including a comparison of transitivity prominence between dictionary and corpus data, within the spoken modality. Section 1.3 outlines the objectives of this study. Details about the data and annotation procedure can be found in Section 2. Results are reported in Section 3 and discussed in Section 4. Finally, Section 5 concludes the paper.

Article note: The order of authors for this paper is alphabetical.

***Corresponding Author: Carl Börstell:** Radboud University, Erasmusplein 1, 6525HT Nijmegen, Netherlands; Email: c.borstell@let.ru.nl

Tommi Jantunen: University of Jyväskylä, Seminaarinkatu 15, 40014 Jyväskylä, Finland; Email: tommi.j.jantunen@jyu.fi

Vadim Kimmelman: University of Bergen, Sydnesplassen 7, 5010 Bergen, Norway; Email: vadim.kimmelman@uib.no

Vanja de Lint: University of Amsterdam, Spuistraat 134, 1012VB Amsterdam, Netherlands; Email: v.delint@uva.nl

Johanna Mesch: Stockholm University, 10691 Stockholm, Sweden; Email: johanna.mesch@ling.su.se

Marloes Oomen: University of Amsterdam, Spuistraat 134, 1012VB Amsterdam, Netherlands; Email: m.oomen2@uva.nl

1.1 Transitivity prominence

In their seminal paper, Hopper and Thompson (1980) construe transitivity as a multifactorial and gradable notion: co-varying properties, each representing a different facet of an action that is “transferred from one participant to another” (Hopper and Thompson, 1980, p. 253), determine the degree of transitivity of a clause. Examples of such properties mentioned by the authors are telicity, volitionality, and agency of the Agent argument. For any language that has a way of marking more than one of the properties they identify, Hopper and Thompson (1980) hypothesize that in any clause A with a feature that scores higher in transitivity than the corresponding feature in clause B, all concomitant features should also be associated with a higher degree of transitivity.

Tsunoda (1981, 1985) and Malchukov (2005) claim that the same sort of properties discussed for clauses by Hopper and Thompson (1980) also govern the transitivity-encoding of lexical verbs. Verbs of perception, for instance, are cross-linguistically more likely to select a transitive case-frame than verbs of emotion. That is, some verbs are claimed to simply be more “transitivity prominent” than others by virtue of their semantics.

Haspelmath (2015) takes a quantitative approach to the investigation of the transitivity prominence of verbs across a large number of typologically diverse languages.¹ Haspelmath’s (2015) study is based on data from 36 languages of 23 language families around the world. The data, accessible online at <http://valpal.info>, consist of valency information on 80 basic “verb meanings” and have been gathered by a team of language experts in the context of the *Valency Patterns Leipzig* (ValPaL) project (Hartmann et al., 2013). Haspelmath (2015) includes 70 of the 80 verb meanings in his study. Since the present study also makes use of ValPaL, more details about this database can be found in Section 2.

In contrast to Hopper and Thompson (1980), Tsunoda (1981), and Malchukov (2005), the definition by Haspelmath (2015) brings transitivity back to its core and classifies lexical verbs as transitive whenever they contain an A and a P argument. A and P are defined as “argument types”, which means that they must be overtly marked as such (e.g., by means of case). Haspelmath (2015) measures the transitivity prominence of each of the verb meanings in the ValPaL list by calculating the proportion of languages that realize each verb transitively and ranking the verbs according to these proportions.² The results show that verb meanings such as BREAK (1.00), KILL (1.00), and PUT (.98) are – almost without exception – realized transitively across languages, while DIE (.00) and BURN (.00) are at the opposite end of the scale. In between, verb meanings such as HELP (.78), LIKE (.78), THINK (.52), and SING (.38) reflect more varying degrees of transitivity prominence. These findings broadly corroborate Tsunoda’s (1981) and Malchukov’s (2005) qualitative analyses. In a similar fashion, Aldai and Wichmann (2018) and Wichmann (2015, 2016) look at valency and transitivity by using various quantitative statistics to identify cross-linguistic patterns of transitivity hierarchies and semantics. While these studies acknowledge some of the general patterns of the implicational hierarchies of Tsunoda (1985) and Malchukov (2005), their statistical analyses do not arrive at any clearly distinguishable semantic categories of verbs corresponding to transitivity rankings. As such, Aldai and Wichmann (2018) conclude that while cross-linguistic similarities in the interaction between semantics and transitivity are visible, there is also a fair amount of diversity and variation between languages and groups of verbs.

All of the works cited above have in common that they assume that real-world associations between events and their participants are often expressed with the encoding of transitivity across languages.

1.2 Transitivity prominence in Russian Sign Language

Given the apparent firm grounding of transitivity in real-world semantics, transitivity can be expected to be modality-independent. Inspired by Haspelmath’s (2015) study, Kimmelman (2016) tests this hypothesis on

¹ The implicit assumption is that languages differ with respect to how often they use transitive encodings. The idea that some languages are more transitivity prominent than others is not new and dates back to at least Hawkins (1986), but, to the best of our knowledge, Haspelmath (2015) is the largest quantitative study in this domain.

² For any individual language, there may be more than one verb *lemma* associated with a verb *meaning*.

the basis of corpus data from RSL. However, Haspelmath's (2015) methodology depends on dictionary data from a large number of languages. Kimmelman (2016) proposes a different measure of transitivity: instead of calculating, for each verb meaning in the ValPaL list, the proportion of languages that use it transitively, he calculates the proportion of overt direct objects that occur with each corresponding verb lemma in the corpus data of a single language. Thus, the measure rests on the assumption that verbs that are highly transitivity prominent occur with an overt object more frequently than verbs that are less transitivity prominent.

After excluding verb meanings from ValPaL with fewer than 25 tokens in the corpus, Kimmelman (2016) ends up with a transitivity ranking of 25 verb meanings in RSL. As expected, the ranking correlates with Haspelmath's (2015) cross-linguistic ranking of verbs: a Spearman's rank correlation test reveals a strong positive correlation ($\rho = .849$; $p = 8.081 \times 10^{-8}$) between the RSL data and the cross-linguistic transitivity prominence ranking from Haspelmath (2015). Kimmelman (2016) thus concludes that transitivity prominence is a modality-independent phenomenon.

1.3 Aims of the study

In this paper, we first test the validity of the practice of comparing valency dictionaries – which reflect binary values (i.e., *transitive* or *intransitive*) – and corpus data for estimating transitivity prominence. We investigate the matter by taking a sample of six spoken languages available in the ValPaL database (Hartmann et al., 2013) that are also included in the Universal Dependencies database (Nivre et al., 2017). We analyze whether the ValPaL valency categorization (transitive/intransitive) of verbs corresponding to 12 ValPaL verb meanings correlates with the observed transitivity scores of the same verbs in Universal Dependencies corpus data – see Section 2.1.

After confirming the validity of the comparison of data from the two types of sources – i.e., valency dictionary classification vs. corpus data – we address the primary aim of the study: to investigate whether the transitivity prominence of verbs across sign languages is correlated, and whether this can be extended across modalities to correlations with the transitivity prominence of spoken languages. If so, the results would once again provide support for the modality-independent nature of transitivity prominence. To this end, we add data from a set of four new sign languages to the sample, in addition to RSL: Sign Language of the Netherlands (NGT), German Sign Language (DGS), Finnish Sign Language (FinSL), and Swedish Sign Language (SSL). The last two are known to be historically related (Bergman and Engberg-Pedersen, 2010).³ NGT and RSL have both been claimed to be related to Old French Sign Language, but at least for RSL these claims have been debated (Bickford, 2005). No claims of relatedness that we know of have been reported for the other sign languages. The composition of the sample thus allows us to investigate possible relatedness effects – that is, whether FinSL and SSL pattern more similarly to each other than the other (unrelated) languages.

Thus, our research questions can be summarized as follows:

1. Are binary valency dictionary categorizations of verbs correlated with the observed transitivity of the same verbs in a corpus?
2. Is a transitivity prominence pattern visible *within* (comparing our sampled sign languages) and *across* modalities (comparing our sampled signed vs. spoken languages)?

³ The issue of relatedness between sign languages is complicated. Nonetheless, it has been established that there was early influence from SSL on FinSL and there is still considerable lexical overlap today (Mesch, 2006).

2 Data and methodology

2.1 Spoken language corpora

In comparing corpus data from sign languages with binary valency dictionary data from spoken languages, Kimmelman (2016) relies on the assumption that data from these two types of sources are, indeed, comparable. Our first objective is to evaluate whether the binary values of verbs in the valency database ValPaL align with overt object proportions with these same verbs in actual use. To investigate this, we select corpus data from a subsample of the ValPaL languages and compare the results to those reported in the ValPaL database. Six languages from ValPaL are included in the *Universal Dependencies* (UD) dataset of annotated language corpora, version 2.1 (Nivre et al., 2017): Chinese (Mandarin), English, German, Italian, Japanese (standard), and Russian (Malchukov and Jahraus, 2013; Zhang, 2013; Kishimoto and Kageyama, 2013; Cennamo and Fabrizio, 2013; Haspelmath and Baumann, 2013; Goddard, 2013).⁴ These languages make up our sample. Some of these languages are represented by more than one annotated corpus. For instance, for our 12 target verb meanings, we have verb tokens for English from four different corpora (out of five available), but from only one for German (out of two available).

For each language, we take the 12 verb meanings from ValPaL that could be collected and compared across our five sign languages based on our sampling procedure (see Section 2.3), and apply Haspelmath's definition of transitivity-encoding to classify a verb as either transitive or intransitive: "A verb is considered transitive if it contains an A and a P argument. A and P are defined as the arguments of a verb with at least two arguments that are coded like the 'breaker' and the 'broken thing' micro-roles of the 'break' verb." (Haspelmath, 2015, p. 136). All other verbs are coded as intransitive, even if they may take other arguments. Thus, the classification of transitive vs. intransitive is based directly on the micro-role coding per verb lemma in the downloadable data from the ValPaL online database (Hartmann et al., 2013). With the verbs from the ValPaL database, we calculate how many of their occurrences are linked to a dependent *obj* argument in the UD corpus data, i.e., how often they take an overt direct object.⁵ For each verb meaning in ValPaL, there are one or more verb lemmas associated with that meaning. For the most part, our six spoken languages have one verb lemma per verb meaning, but in some cases there are several lemmas. The verb meaning LEAVE, for instance, is represented by two verb lemmas in Italian: *lasciare* (transitive) and *partire* (intransitive). For such cases, each individual verb lemma is recorded as an instance of the ValPaL verb meaning for that language, and we conflate them – unless stated otherwise – by calculating the average for each verb meaning for each language (as we do with the sign language data, see Section 2.3). We then compute each verb's transitivity prominence in each language based on the total number of *obj* objects of the verb divided by the total number of tokens for that verb.⁶

2.2 The sign language corpora

Our sign language data come from preexisting corpora of RSL, NGT, DGS, FinSL, and SSL. All five corpora aim at representing the language of deaf signers of the respective communities and reflect variation in, for instance, regional area, gender, and age of signers. However, as work on sign language corpora is still in its early stages, not all corpora are equally comprehensive (e.g., in terms of the amount of collected or annotated

⁴ The motivation for choosing to use UD is three-fold: first, it is a dataset using a standardized annotation framework across languages, thus allowing for cross-linguistic comparisons; second, it features dependency relation annotations, enabling us to look specifically at the dependency role *obj* associated with specific verb lemmas; and third, it contains several of the languages featured in ValPaL.

⁵ Note that we rely on the UD label *obj* to identify overt direct objects in the UD data, thus referring to their definition of this function (see <https://universaldependencies.org/u/dep/obj.html>).

⁶ We thank Robert Östling for providing us with data extracted from UD v2.1.

Table 1: The corpus datasets for each of the five sampled sign languages in this study by number of tokens.

| Language | Acronym | Tokens |
|----------------------------------|---------|--------|
| Sign Language of the Netherlands | NGT | 80,000 |
| German Sign Language | DGS | 58,000 |
| Russian Sign Language | RSL | 25,000 |
| Finnish Sign Language | FinSL | 18,500 |
| Swedish Sign Language | SSL | 15,200 |

data), nor do they contain exactly the same type of data. The content of the sampled corpus data from each of the five sign languages is described below; sample sizes are summarized in Table 1.

The corpus of RSL (Burkova, 2015), accessible online at <http://rsl.nstu.ru>, was constructed between 2012 and 2015 at the Novosibirsk State Technical University. The corpus contains annotated narratives (e.g., retellings of cartoons, personal stories on different subjects) and a small number of dialogues (e.g., on topics related to deaf culture and sign language) from 43 signers of RSL, mainly from the areas of Moscow and Novosibirsk. The total length of the video material in the RSL corpus is 5 hours and 30 minutes and the corpus includes approximately 25,000 sign tokens, which make up the sample of RSL data in the present study.

The Corpus NGT (Crasborn et al., 2008; Crasborn and Zwitserlood, 2008) was first published in 2008 as a result of a long-term project carried out at Radboud University Nijmegen. The corpus, accessible online via The Language Archive of the Max Planck Institute of Psycholinguistics at <http://archive.mpi.nl>, contains over 15 hours of elicited as well as (semi-)spontaneous narrative and dialogue data produced by 92 signers from different areas of the Netherlands. The total number of annotated sign tokens in the most recent version of Corpus NGT amounts to approximately 97,000. In the present study, the basic sample of NGT data comprises approximately 80,000 sign tokens.

The DGS Corpus project commenced in 2009 at the Institute for German Sign Language and Communication of the Deaf at Hamburg University. Once completed in 2023, the corpus will include data from 330 deaf signers from thirteen different regions in Germany (Langer, 2012) but at the time at which the present study was carried out, material was available from 104 signers from eleven regions. This material, accessible online at <http://ling.meine-dgs.de> and constituting the sample of DGS data, comprises 58 spontaneous dialogues covering topics related to significant international events, deaf culture, and personal experiences; elicited material and narratives are not part of the dataset. The total length of the video recordings in this material is 8 hours and 30 minutes, and the material includes approximately 58,000 sign tokens.

The work on the FinSL corpus (Salonen et al., 2016) began at the Sign Language Centre of the University of Jyväskylä in 2013 and is still in progress. The collection of the recordings, representing 104 signers from all over Finland, was completed in 2017. Thus far, data are available for 22 signers. From this material, we selected a subset with descriptive narratives (retellings of the stories *Snowman* and *Frog, where are you?*) and semi-spontaneous conversations, involving 20 signers. The duration of video material in the sample, accessible online at <http://lat.csc.fi> (Jantunen et al., 2016; Salonen et al., 2019), is 2 hours and 30 minutes. The sample includes approximately 18,500 sign tokens.

The corpus of SSL (Mesch et al., 2012; Mesch, 2018), accessible online via <http://www.ling.su.se/teckensprakskorpus>, was published in 2011 at Stockholm University. The corpus includes 24 hours of recordings containing presentations, narratives (retellings of the stories *Snowman* and *Frog, where are you?*), and conversations by 42 signers from different areas of Sweden. Currently, 12 hours of video data have been annotated and from this material, 2 hours make up our sample of SSL data. Altogether, the sample comprises data from 28 signers and includes approximately 15,200 sign tokens.

2.3 Sampling, annotating, and processing sign language data

2.3.1 Sampling procedure and validity simulations

In his study on RSL, Kimmelman (2016) analyzed the tokens of 25 ValPaL verb meanings (see Section 1.2). In our additional data from NGT, DGS, FinSL, and SSL, we first identified verb lemmas corresponding to these same 25 verb meanings. Usually, the core verb meanings correspond to a single verb lemma in each language. However, some meanings (e.g., LEAVE) are associated with more than one sign in a language (e.g., the signs LÄHTEÄ ‘leave₁’ and MENNÄ ‘leave₂’ in FinSL). In RSL, DGS, and NGT, we selected the most frequent verb lemma in such cases. The procedure for FinSL and SSL was slightly different due to the smaller datasets from these languages. Tokens of multiple signs associated with the same meaning were therefore collapsed into a single group.⁷

Furthermore, we exclude so-called *classifier predicates* from our study to only focus on lexical signs. This choice is motivated by the variation found in the argument structure of these signs, although certain types of classifiers have been shown to be associated with particular argument structure alternations cross-linguistically (Benedicto and Brentari, 2004; De Lint, 2018; Kimmelman et al., 2019).⁸

Because of the different natures of the sign language corpora (Section 2.2), we do not find the same amount of tokens for all 25 ValPaL verb meanings in all languages. Kimmelman (2016) set the token threshold to 25 for RSL, but we find that following this same limit decreases the number of comparable verb meanings considerably across our sampled sign languages. In order to avoid this problem, we set the token threshold to 5, thus excluding all verb meanings with fewer than five tokens in any of the sign languages.

We had concerns that having such a small number of token per meaning, and also a relatively small number of meanings compared might affect the results by producing Type I errors: mistakenly finding significant correlations. In order to assess this risk, we ran a simulation in R (R Core Team, 2015) – see Appendix 1: Simulations for the full code and a sample result table. The simulation works as follows: we randomly generate data that is similar to our data and then calculate rank correlations between these randomly generated vectors and assess the significance of the correlation 10,000 times. Since the vectors are created randomly and independently, we expect the rank correlation to be significant (at $p < .05$) in 5% of the simulations (by definition of the identically set significance level).

More specifically, we generated the vectors using the `rbinom` functions, which produce a vector of random values from a sample for the binomial distribution. It has three parameters: the number of observations (which is parallel to the number of verbs in our case), the sample size (which is parallel to the number of tokens per verb in our case), and the probability (which is the probability of a verb occurring with an overt object in our case). We then divide each number in the vector by the sample size, which produces the vector of proportions (parallel to proportions of overt objects per verb).

We ran the simulations varying two parameters: the number of verbs (or observations in `rbinom`) and the sample size (the number of tokens), keeping the third parameter constant (see Appendix 1 for the details). For each of the parameters, we had three possible values (25, 10, and 5). It turned out that what crucially affects the results is the number of verbs, and not the number of tokens. Specifically, only with the number of verbs = 5 does the percentage of accidentally found significant correlations rise to around 6–8% (i.e., resulting in false positives). With 10 verbs, irrespective of the number of tokens, the percentage of significant correlations is around 5.5%. This is higher than the desired 5%, but only marginally. We therefore conclude that we can use

⁷ Using the same procedure for FinSL and SSL as for the other sign languages has no clear statistical impact. A Spearman’s rank correlation test shows that – within FinSL and SSL, respectively – there is a high correlation ($\rho > .8$) between the proportions collapsing verb lemmas and selecting the most high-frequent ones. Thus, we choose to collapse verb lemmas to have as much data as possible from the two sign languages with smaller datasets.

⁸ It should be noted that several studies, across different sign languages, have pointed to form–meaning mappings also in lexical verbs having an effect on the expression (form or omission) of arguments and argument marking (e.g., Meir et al., 2007; Oomen, 2017; Oomen and Kimmelman, 2019; Hou and Meier, 2018). This is, however, beyond the scope of the current paper.

our methods in assessing correlations between the languages with our sample sizes without a substantially increased risk of a Type I error.

2.3.2 Data annotation and analysis

For each language, we annotate in ELAN (Wittenburg et al., 2006) whether a token of a verb sign (V) associated with any of the 12 meanings occurs with an overt object. The type of object – direct, indirect, or clausal – is also annotated. *Direct object* (O) is defined as the Patient of a two-argument verb or the Theme/Patient of a three-argument verb. *Indirect object* (O2) is the Goal or Source argument of a three-argument verb, and *clausal object* (CO) is an argument manifested as a clause. Clause boundaries are identified on semantic grounds, i.e., on the basis of the verb and all its arguments, in combination with prosodic cues such as pauses or marked changes in non-manuals (cf. Crasborn, 2007; Fenlon et al., 2007; Börstell et al., 2014).

Examples of the different object types in SSL, NGT, DGS, and FinSL are given in (1–3).⁹ Examples (1a–b) represent clauses with a single (direct) object in SSL (1a) and NGT (1b).

- (1) (a) (Swedish Sign Language)
 #IN WANT [TAKE]_V [EGGS]_O
 ‘[I] want to take those eggs inside!’ (SSLC02_169 00:22.000)
- (b) (Sign Language of the Netherlands)
 INDEX₃ [SEE]_V [BIRD]_O
 ‘There [he] sees a bird.’ (CNGT0601 00:23.520)

Example (2) is an illustration of a clause with two overt objects from SSL.

- (2) (Swedish Sign Language)
 INDEX.PL₁ MEMBER INDEX.PL₃ [GIVE]_V [OBJPRO.PL₁]_{O2} [DISCOUNT]_O
 ‘[If] we are members, they give us a discount.’ (SSLC01_302 03:35.583)

Finally, examples (3a–b) illustrate clausal objects in DGS (3a) and FinSL (3b).

- (3) (a) (German Sign Language)
 POSS₁ HEAVY MOTHER INDEX₃ [SAY]_V [KENNEDY ALREADY DEATH SHOOT PU]_{CO}
 ‘I learned through my mother-in-law that Kennedy was shot dead.’ (dgskorpus_fra_12 01:27.365)
- (b) (Finnish Sign Language)
 INDEX₁ [THINK]_V [BRAVE GRANDFATHER INDEX₃]_{CO}
 ‘I think that the [snow]man was brave!’ (CFINSL2013_006_05-2 06:55.840)

In our study, we only focus on overtly expressed *direct objects* (i.e., those coded as “O”), as Kimmelman (2016) argues these are likely to best reflect transitivity prominence. Direct objects are also relatively easy to identify. Yet, with verbs expressing the meanings LEAVE, LIVE, and RUN in particular, the categorization of objects is ambiguous: these verbs may accept semantically *locative arguments* (e.g., INDEX₁ LIVE HOUSE ‘I live in a house’), which can be treated as *obliques*. Obliques constitute a grey area between proper (direct) objects and adjuncts and they are difficult to analyze in syntactic terms from a cross-linguistic perspective (see, e.g., Arka, 2014; Wichmann, 2014; Haspelmath, 2014). In our study, we categorize the locative arguments in two ways for alternative analyses: first, we treat them as direct objects (since there is no difference in the

⁹ Following common practice in sign language research, we gloss signs in SMALL CAPS (see also the Glossing section).

encoding¹⁰ compared to direct objects, following Haspelmath's definition stated in Section 3.1); in a second step, we treat these as a separate locative category (based on semantics). After this two-step analysis, our results suggest that locative argument-taking verbs behave differently in the syntax of the five sign languages – see Section 4.3 for a discussion of this issue.

After annotating the data, we calculated the transitivity prominence of each verb meaning in every language by dividing the total number of direct object occurrences by the total number of verb tokens for each individual verb meaning. The languages can thus be compared to each other and to the grouped spoken language data from both ValPaL (Haspelmath, 2015) and UD (Nivre et al., 2017), in which all languages are collapsed and counted as a group for each verb meaning.

3 Results

3.1 Valency categorization and corpus data

As described in Section 2.1, we first compare the binary values (transitive vs. intransitive) from the ValPaL database with corpus data from UD for six spoken languages. The distribution of each of the 12 verb meanings we studied is plotted in Figure 1.¹¹ The graph shows the (expected) pattern that transitive verbs (green) take

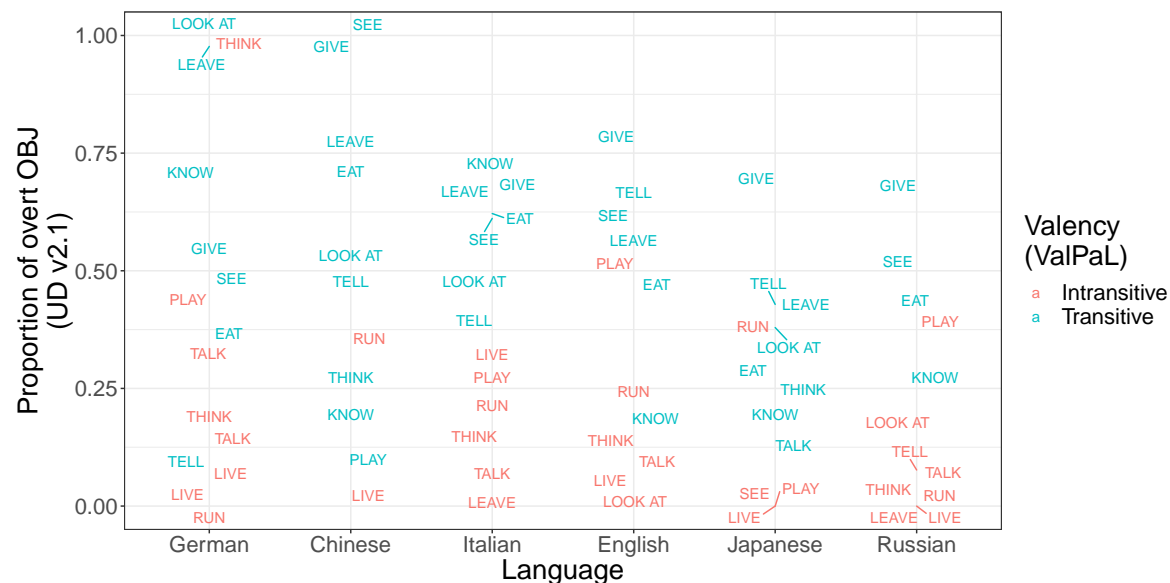


Figure 1: Distribution of verb meanings in six languages (x axis) based on overt object-taking occurrences in corpus data (UD) as a measure of transitivity prominence (y axis) set off against their classification in a valency database (color-coded). Languages are sorted from most to least transitive, averaged across verb meanings.

¹⁰ Argument-encoding definitions are complicated for sign languages as it has been observed that sign languages, as a group, tend to lack any overt marking (Gil, 2014). The few recorded cases of overt object marking to date appear to be instances of *differential object marking* (see, e.g., Bosson, 1985; Sinnemäki, 2014; Seržant and Witzlack-Makarevich, 2018), some of the cases involving languages in our sample (Meir, 2003; Pfau et al., 2018; Börstell, 2019).

¹¹ There were no occurrences of the verb meaning TALK in the Chinese data, hence there are only 11 verb meanings present for this language.

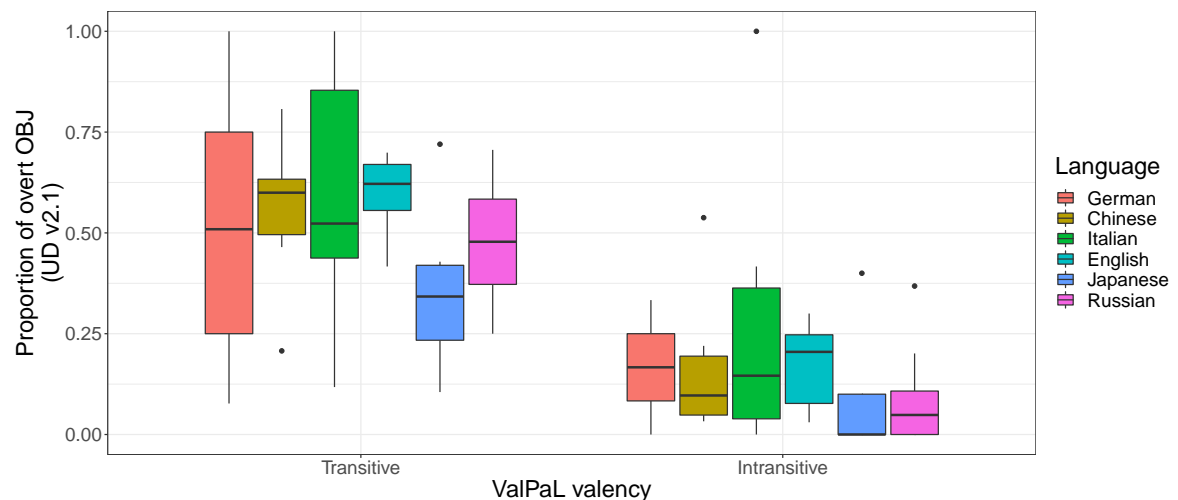


Figure 2: Distribution of verb meanings in six languages (color-coded) based on overt object-taking occurrences in corpus data (UD) as a measure of transitivity prominence (y axis) set off against their classification in a valency database (x axis). Languages are sorted from most to least transitive, averaged across verb meanings.

overt objects more often than intransitive verbs (red).¹² This pattern is observable across languages and verb meanings. Note that we plot each individual verb lemma, hence we have two instances of the verb meaning LEAVE for Italian: one for each lemma (which are also classified differently in ValPaL in terms of transitivity; see Section 2.1). The same pattern is visible in Figure 2, with all verbs being grouped per language, separated into the ValPaL categorization of transitive vs. intransitive. Here it is evident that across languages, verbs categorized as transitive are generally higher in UD transitivity prominence (occurring with overt objects) than verbs categorized as intransitive, even if a particular language can fall higher or lower on the overall transitivity prominence scale. For example, Japanese transitive verb lemmas (associated with the sampled verb meanings) occur with overt objects less than 50% of the time, on average, but the intransitive verbs are still even lower in transitivity prominence. This is a pattern visible for each transitive–intransitive comparison across the six languages, and a Wilcoxon signed rank test with continuity correction using the statistical language R (R Core Team, 2015) shows that the distribution of verb lemma transitivity prominence ranking (i.e., ranking each individual verb lemma based on its proportion of overt objects between 0 and 1), across languages, is significantly different between the transitive and intransitive groupings ($p = 5.332 \times 10^{-14}$).

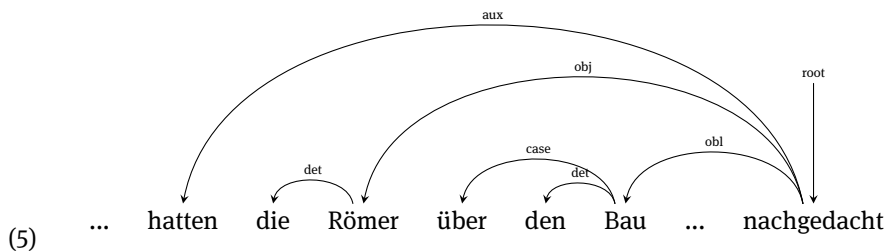
The distribution in Figure 2 is especially uniform for the verbs labeled as intransitive but less so for the verbs labeled as transitive – that is, the variation within intransitive verbs (within a specific language) is less extreme than the variation in some of the languages’ transitive verbs. This should come as no surprise, given that many transitive verbs are free to omit their P argument when it can be deduced from the context or when its specification is irrelevant, such as in *I eat (food)* in English, where the object *food* is optional. Such a construction is even acknowledged in the ValPaL database as an alternation called “Understood Omitted Object: A canonically transitive verb appears in this derived pattern without an overt object, usually implying a specific kind of expected object.” (Hartmann et al., 2013). This alternation can account for some of the notable exceptions to the general distributional tendency in Figure 1. For example, the verb *wán* in Chinese for the verb meaning PLAY has low transitivity prominence in the corpus data, even though it is classified as basically transitive in ValPaL. The optionality of the P argument is acknowledged in ValPaL, with the example sentence for this verb putting the object in brackets (see example 4). Thus, some differences between the

¹² Figures are created using the R packages *ggplot2* (Wickham, 2009) and *ggrepel* (Slowikowski, 2017).

valency database values and the corpus data can be attributed to the fact that the valency assignment in ValPaL may be based on whether the transitive encoding is *possible*, not *frequent*, for the basic verb.

- (4) *Didi zài cāochǎng shàng wán (pí qiú).*
 younger.brother PREP playground LOC play (ball)
 ‘My younger brother is playing (ball) on the playground.’ (Zhang, 2013)

Another case for which the ValPaL and corpus data differ is the German verb *nachdenken* ‘think, contemplate’, one of the two verbs listed for the ValPaL verb meaning THINK (the other being *denken* ‘think’). Here, it is a verb classified as intransitive that comes out as distinctly transitive in the corpus data. At closer inspection, it can be observed that there is only a single token for *nachdenken* in the UD dataset, and this token has an associated *obj* argument. However, the dependency analysis for the sentence in which this token appears is incorrect. As shown in example 5, the subject (*die*) *Römer* ‘(the) Romans’ has been labeled as an object, most likely due to incorrect automatic tagging or errors in a conversion between annotation frameworks.



Schon in der Antike, zur Zeit des Kaisers Augustus, hatten die Römer über den Bau eines Kanals in Südfrankreich zwischen Mittelmeer und Atlantik nachgedacht.

‘Already in the Classical era, in the time of Emperor Augustus, had the Romans thought about the building of a canal in Southern France between the Mediterranean and the Atlantic.’

(UD German-GSD; sentence ID: train-s14030)

Furthermore, we have no way of distinguishing different extended or polysemous meanings of verbs extracted from the UD datasets. For instance, a verb like *play* in English could be used in many different contexts (e.g., *playing*, *playing chess*, *playing the piano*, *playing Shakespeare*). In ValPaL, such contexts are listed as possible microroles of the verb meaning PLAY, but we do not know if the different contexts pattern similarly across languages, and which ones are more likely to occur with overt objects.

Overall, we still see that the expected pattern is clear and the difference between transitive and intransitive in their object-taking realization in use. That is, verbs classified as having a transitive encoding according to ValPaL are also more likely to take overt objects in corpus data, across languages. Though unsurprising, it provides a good indication that comparing valency dictionaries and corpus data is a feasible and valid method.

3.2 Transitivity proportions and correlations within and across modalities

Having justified a comparison between dictionary and corpus data, we move on to the main investigation of transitivity prominence within and across modalities. Following the methodology presented in Section 2.3, we obtain language-specific transitivity prominence values for each of the 12 ValPaL verb meanings from each of the five corpora of our sampled sign languages. These values, together with the combined spoken language values from Haspelmath (2015), are presented in Figure 3. The results show that certain verb meanings are very similar in terms of their transitivity prominence. For instance, GIVE is highly transitivity prominent (dark red color) in all languages and across modalities, whereas PLAY behaves more like an intransitive verb (light

Table 2: The statistical rank correlation across ValPaL verb meanings within each language dyad for our five sign languages and the combined spoken languages from Haspelmath (2015). Locative arguments are *included* for the sign languages.

| | FinSL | NGT | RSL | SSL | SpokenL |
|-------|---------------|---------------|--------------------|--------------------|-----------------|
| DGS | $\rho = .543$ | $\rho = .494$ | $\rho = .578^*$ | $\rho = .785^{**}$ | $\rho = .276$ |
| FinSL | | $\rho = .483$ | $\rho = .692^*$ | $\rho = .490$ | $\rho = .263$ |
| NGT | | | $\rho = .734^{**}$ | $\rho = .294$ | $\rho = .688^*$ |
| RSL | | | | $\rho = .357$ | $\rho = .302$ |
| SSL | | | | | $\rho = .057$ |

red color). However, the results also show that certain verb meanings (e.g., EAT, LIVE, TELL) may have very different transitivity prominence values across both languages and modalities.

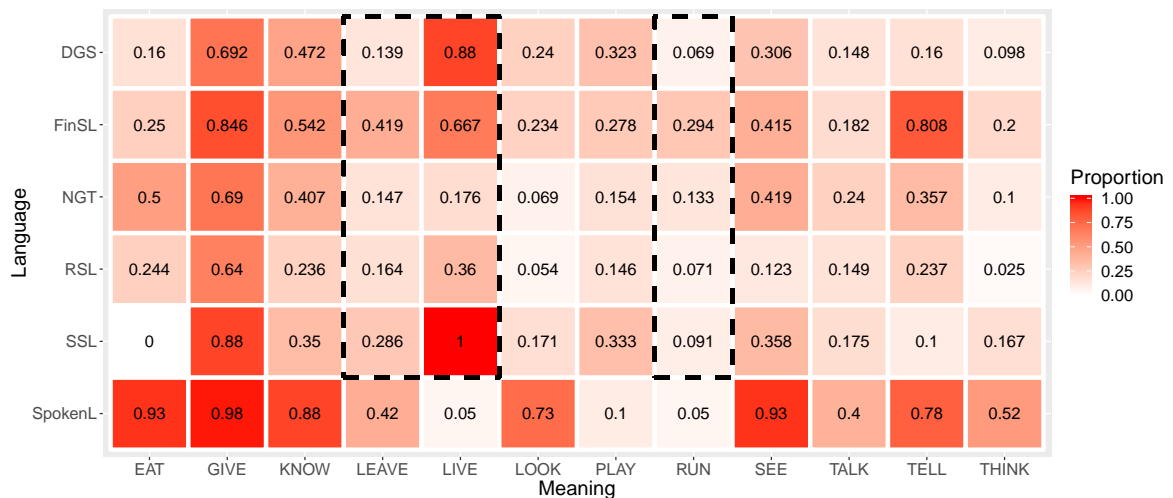


Figure 3: Transitivity prominence for 12 ValPaL verb meanings across five sign languages and the combined spoken languages. Degree of redness indicates the transitivity proportion (darker red corresponds to higher proportions). Values inside the dotted rectangles include locative arguments (verb meanings LEAVE, LIVE, and RUN).

In order to determine how the five sign languages correlate with each other as well as with spoken language data from Haspelmath (2015), we first calculate the statistical rank for the transitivity prominence values of each language – that is, the sequential order from least to most transitivity prominent (the proportions reported in Figure 3). The rank position of each verb meaning across each of the languages is visualized in Figure 4.¹³ We then estimate the relation between all languages by calculating the Spearman's rank correlation coefficient (ρ) for each language dyad using the statistical language R (R Core Team, 2015). A test of statistical significance for the rank correlations reveals significant positive correlations ($p < .05$) for the following language pairs: NGT–spoken languages ($\rho = .688$; $p = .013$), NGT–RSL ($\rho = .734$; $p = .007$), RSL–FinSL ($\rho = .692$; $p = .013$), DGS–SSL ($\rho = .785$; $p = .003$), and DGS–RSL ($\rho = .578$; $p = .049$) – see Table 2.¹⁴ All other correlations are statistically non-significant, but always positive.

¹³ In Figure 4, duplicate rankings are sorted alphabetically to avoid visual overlap.

¹⁴ Asterisks indicate significance level: ***: $p < .001$; **: $p < .01$; *: $p < .05$

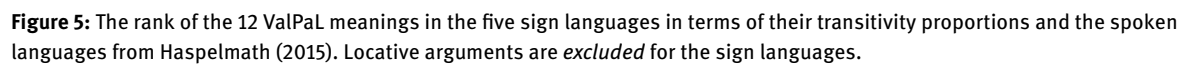
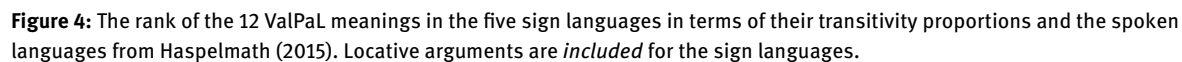


Table 3: The statistical rank correlation across ValPaL verb meanings within each language dyad for our five sign languages and the combined spoken languages from Haspelmath (2015). Locative arguments are *excluded* for the sign languages.

| | FinSL | NGT | RSL | SSL | SpokenL |
|-------|---------------------|---------------------|---------------------|---------------------|---------------------|
| DGS | $\rho = .888^{***}$ | $\rho = .771^{**}$ | $\rho = .742^{**}$ | $\rho = .860^{***}$ | $\rho = .767^{**}$ |
| FinSL | | $\rho = .844^{***}$ | $\rho = .860^{***}$ | $\rho = .713^{**}$ | $\rho = .845^{***}$ |
| NGT | | | $\rho = .915^{***}$ | $\rho = .627^{*}$ | $\rho = .934^{***}$ |
| RSL | | | | $\rho = .513$ | $\rho = .821^{**}$ |
| SSL | | | | | $\rho = .609^{*}$ |

In order to minimize the possible effect of verbs that take locative arguments on the results (see Section 2.3), we perform an additional correlation analysis in which all realizations of the secondary arguments of LEAVE, LIVE, and RUN in all sign languages are treated as obliques rather than direct objects (i.e., all transitivity prominence values in the dotted rectangle in Figure 3 are treated as zero). Recalculating the Spearman's rank correlations yields significant positive correlations in the transitivity ranking of the 12 verb meanings for all but one language pair: only the SSL–RSL pair has $p > .05$ ($\rho = .513$; $p = .089$) – see Table 3. The results of this additional analysis suggest that the classification of locative argument-taking verbs cause variation in the transitivity prominence ranking of verb meanings in the investigated sign languages (see Section 4.3 for further discussion).

In general, the results of the correlation study suggest that there are visible similarities in ranking of the 12 verb meanings across languages and modalities. However, it is equally evident from Figure 4 that certain verbs show significant variation in terms of the position they occupy in the ranking (but note that Figure 4 is based on the analysis in which we code locative arguments as direct objects). Some of these differences can be assumed to reflect true differences between languages, but we suspect that some can also be traced back to issues related to, for example, the type of data and object categorization. Furthermore, the exclusion of locative arguments but keeping the verbs in the analysis means that we are forcing these verb meanings to the bottom of the ranking across languages, thus resulting in a high correlation for these specific meanings by default across (sign) languages. As an alternative solution, taking these locative verbs out of the sample completely gives us a different picture with far fewer significant correlations across language pairs (see Table A2 in Appendix 2: Excluding locative verbs).¹⁵ We discuss these issues in more detail in Section 4.

Finally, calculating averaged transitivity prominence values across languages for which we have corpus data (i.e., five sign languages and six spoken languages) allows us to compare general transitivity prominence both across *modalities* (signed vs. spoken) and *data types* (corpus vs. dictionary). Figure 6 displays the ranking of the 12 verb meanings for each of our data types: sign languages (corpus data; with and without locatives); spoken languages (UD corpus data, from Nivre et al. 2017); and spoken languages (ValPaL, from Haspelmath 2015). The figure shows that one verb meaning in particular stands out by showing a lot of variation across modalities: LIVE. This is, of course, a consequence of the methodological issue previously discussed, namely that locative arguments have been included. Although adpositions are used with arguments in some sign languages – including SSL (see Börstell, 2017, 2019) – the sign language corpus data show that LIVE often takes an argument that is encoded no differently from those of other prototypical object-taking verbs. As a result, this verb meaning is the second highest ranked across our sign languages, while it is at the very bottom with both types of spoken language data. Still, Figure 6 illustrates the general pattern that the verb meanings that exhibit high transitivity prominence in each of the groups (i.e., those with a darker red

¹⁵ This alternative analysis should, however, also be viewed with some caution as it is reducing the number of verb meanings, which – as noted in the simulations (see Appendix 2: Excluding locative verbs) – could affect correlation tests.

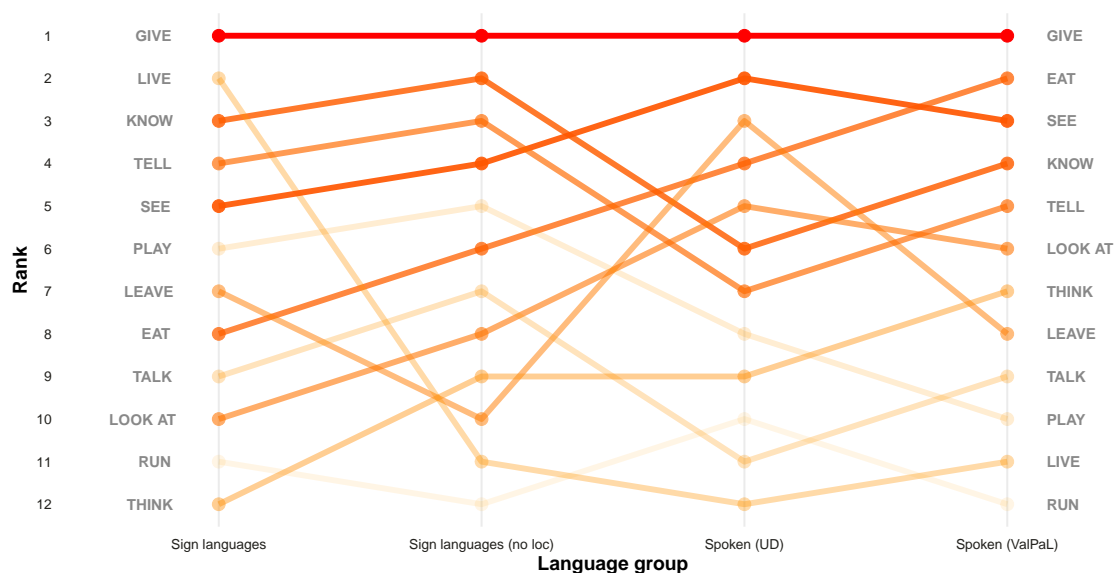


Figure 6: The rank of the 12 ValPaL meanings in terms of their averaged transitivity prominence across the five sign languages, the six spoken UD languages, and the ValPaL data from Haspelmath (2015). Locative arguments are both included and excluded for the sign languages (listed in separate columns). Color indicates the average rank across the three language groups (dark red means higher transitivity-prominence rank).

color) tend to end up in the upper half across the groups, and vice versa. Thus, although we observe some variation across language groups (especially between the modalities), a general pattern is still discernible.

Looking at the three possible pairings of languages/datasets – sign languages and ValPaL languages (SL–ValPaL); sign languages and UD languages (SL–UD); and ValPaL languages and UD languages (ValPaL–UD) – we can calculate correlations between transitivity rankings for each pair. Figure 7 shows the distribution of verb meanings for each of the three pairs, with added regression lines. From this, we do see a positive correlation across all three pairs, but the clearest pattern is seen for the ValPaL–UD pair (i.e., between the spoken language samples). We run a Spearman’s rank correlation test for each pair, which gives us the following: SL–ValPaL ($\rho = .364$; $p = .245$); SL–UD ($\rho = .291$; $p = .359$); and ValPaL–UD ($\rho = .821$; $p = .001$). Thus, although we do see a positive correlation for each pair, the correlation is only significant, and indeed very high, between the ValPaL data from Haspelmath (2015) and the UD data for six of the ValPaL languages.

If we again remove the locative arguments from the sign language data and re-analyze the language group pairings, we see a higher correlation for both pairs (Figure 8). The rank correlation is, however, statistically significant only in the ValPaL pairing (SL–ValPaL: $\rho = .756$; $p = .004$), but marginally significant ($p < .1$) in the UD pairing (SL–UD: $\rho = .507$; $p = .092$).

4 Discussion

4.1 Effects of data types

Our first research question concerned the correspondence between valency dictionary categories and observed corpus-based transitivity prominence. By comparing the transitive vs. intransitive classification of individual verbs from ValPaL with the corresponding languages and verbs in corpus data from UD, we could establish that these two types of data patterned similarly. We observed a statistically significant difference be-

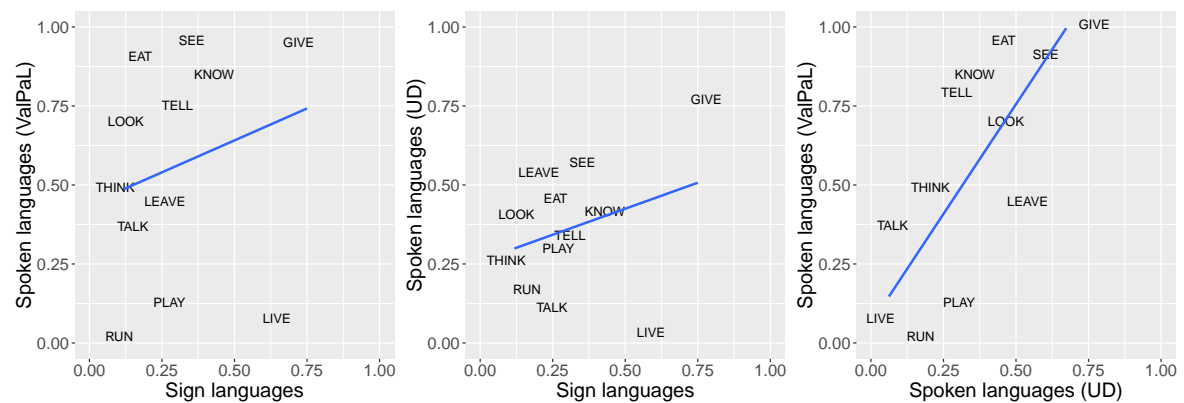


Figure 7: The transitivity proportion of the 12 ValPaL meanings compared for each of the three language group pairings: SL–ValPaL (left); SL–UD (middle); and ValPaL–UD (right). The blue lines show a fitted linear regression for each pairing.

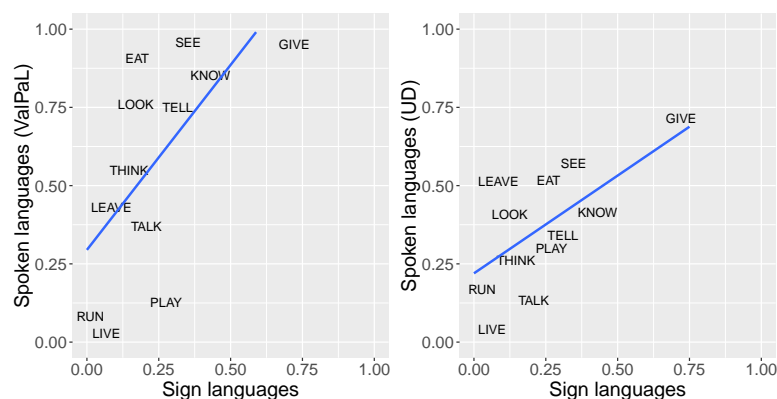


Figure 8: The transitivity proportion of the 12 ValPaL meanings compared for two language group pairings, SL–ValPaL (left); SL–UD (right), *without* locative arguments for the sign language group. The blue lines show a fitted linear regression for each pairing.

tween transitive and intransitive verbs in their respective transitivity prominence in corpus data, across languages, such that verbs labeled as transitive indeed exhibit higher transitivity prominence (i.e. co-occurrence with a direct object) in corpus data than intransitive verbs, which mostly do not take direct objects – as predicted by the binary classification. This is an important finding in its own right, since it points to the validity of *either* data type for the purpose of investigating transitivity patterns across languages. Unsurprisingly, transitive verbs show a higher degree of variation in transitivity prominence, since they may drop (object) arguments for semantic or pragmatic reasons – e.g., ambitransitive uses or context-dependent omissions. We observe this dictionary–corpus correlation as a general pattern across six different spoken languages, which leads us to believe that they are both valid measures of valency/transitivity. This, in turn, supports the findings of the previous study by Kimmelman (2016), comparing RSL corpus data to ValPaL dictionary data, and validates the methodology of this study, expanding the sample of sign languages.

While we can conclude that there is a general conforming pattern of transitivity prominence across sign languages and also cross-modally (as in, positive correlations), it is clear that the correlations are not perfect and often not statistically significant. We can identify several factors affecting these results. Firstly, although transitivity clearly has a semantic basis and this results in some general cross-linguistic patterns/preferences, it is trivially true that one cannot fully predict transitivity for individual languages and their verbs based on semantics alone, otherwise we would find little to no cross-linguistic variation – at least for verbs that

are close translational equivalents. That general patterns are visible but individual verbs/languages exhibit variation has also been argued by Aldai and Wichmann (2018) based on their statistical approach to cross-linguistic transitivity patterns.

Secondly, we suspect that some of the differences between individual sign languages analyzed are due to the method we used, namely the fact that we looked at corpus data and accepted a minimum of only five tokens per verb per language. Although our simulations support our five token minimum threshold, they also showed that the number of sampled verb meanings had a greater effect on the reliability. While we are still above the minimum number for verb meanings, too, in our simulation, zeroing out the locative verbs across sign languages also means we are “hard-coding” them to fall to the bottom of the ranking, with guaranteed cross-linguistic similarities by default, at least across sign language pairs. For this reason, an alternative solution is to exclude these verb meanings completely (see Appendix 2: Excluding locative verbs), which does result in fewer significant correlations, but also means other risks by reducing the number of verb meanings compared (cf. simulations in Appendix 1: Simulations). For the cross-modal comparisons, it appears that the location/movement verbs in our sample rarely pattern as transitives for the spoken languages, and the analysis of treating these as non-arguments for sign languages based on semantics consequently makes the data conform to the general spoken language pattern. The argument status of individual verbs in individual languages should be defined based on their behavior within those languages, but this requires further investigations and more data. We would also suggest that future work considers the total token sizes of sampled corpora, such that not only token thresholds, but also type (verb meaning) thresholds can be upheld with higher reliability.

Furthermore, we think that the content of the corpora might influence the occurrence of direct objects, and with small and unbalanced corpora we do not always get a good representation of a verb’s transitivity. As an example, consider the verb meaning EAT. In RSL, DGS, NGT, and FinSL it occurs – with varying frequencies – with overt objects, whereas it never occurs with a direct object in SSL. Looking at the content of the recordings in which EAT is used in SSL, it turns out that in each of those cases, the signer refers to the general activity of eating, not to the eating of, e.g., a specific food item. As such, the question is whether we would see a different pattern for SSL if we had, say, ten times the number of tokens, from a larger set of signers and conversational topics. Another example is the verb TELL: while it demonstrably can occur with a direct object in all of our sign languages, the verb is used with a direct object in an unusually high 80% of the cases in FinSL. Again, looking at the data, we find that TELL in the FinSL corpus is used mostly at the beginning of narratives as a means to start a story: INDEX₁ TELL STORY ‘Let me tell you a story’. In the other corpora, different (and more varied) contexts are used.

We conclude that larger and more balanced corpora are necessary to reliably establish transitivity based on the occurrence of direct objects in corpus data. Nonetheless, we also acknowledge the fact that we find cross-linguistic patterns and positive correlations between transitivity rankings across sign languages – despite limited datasets, and especially clear without locatives (see also Section 4.3 below) – and argue that this is an indication of a shared semantic basis for transitivity, as suggested by various proposed hierarchies in previous work.

4.2 Relatedness effects

One might wonder whether the similarities in transitivity patterns between the sign languages we analyzed could be explained in terms of historical relatedness. As discussed in Section 1.3, some of the sign languages in our sample have been claimed to be related (based on more or less reliable evidence). Specifically, SSL and FinSL are known to be related, whereas RSL and NGT both have a (debated) connection to Old French Sign Language. SSL and FinSL as a group are not related to DGS or to RSL and NGT as a group, and DGS is not related to RSL or NGT. If historical relatedness plays a role, we expect to see higher correlations between transitivity rankings for related pairs of languages than for unrelated pairs.

In Table 4, we summarize the correlations – first with and then without locatives as direct objects – for SSL and FinSL (related), RSL and NGT (possibly related), and the other pairs (unrelated). We see that the

Table 4: Rank correlations (ρ coefficient) in related and unrelated sign language pairs.

| | SSL–FinSL | RSL–NGT | unrelated pairs |
|-------|-----------|---------|------------------------------------|
| + loc | .490 | .734** | average: .521; range: [.251, .806] |
| – loc | .713** | .915*** | average: .757; range: [.491, .888] |

SSL–FinSL pair does not stand out with respect to the degree of correlation compared to the average across unrelated sign language pairs. On the contrary, the SSL–FinSL correlations are below the average correlation coefficients of unrelated pairs, regardless of whether or not locatives are included. In contrast, the RSL–NGT pair shows a higher correlation than the unrelated average and, if locatives are excluded, even the highest correlation across all possible language pairs.

Based on these results, we conclude that it is difficult to assess the effects of relatedness. Given that the correlations are not higher – but, in fact, lower – than average for the only pair of languages that are undoubtedly related, while RSL and NGT are not reliably established as being related but do show high correlations, we can tentatively say that relatedness does not seem to have a significant impact on the results. Further research is necessary to pinpoint the extent of the possible influence of relatedness on our methodology and measures, as it is possible that larger and cross-linguistically more comparable datasets would alter this picture.

4.3 Locative arguments

As discussed in Section 3, we looked at the transitivity rankings for each sign language in two steps, differing in the way locative arguments are treated. To recapitulate, the verb meanings *LEAVE*, *LIVE*, and *RUN* could be analyzed as having a direct object fulfilling the thematic role of Source, Location, and Goal, respectively. The crucial question is whether these potential direct objects should indeed be analyzed as such. The alternative is treating them as obliques (cf. Wichmann, 2014; Haspelmath, 2014).

From a semantic perspective, location descriptions are not normally considered arguments, except for when they belong to location/motion verbs. In such cases, it is reasonable to consider them semantic arguments since the meaning of the verbs is dependent on their presence and reference – that is, they are indeed part of the *core* meaning rather than being peripheral/secondary. In spoken languages, such arguments can be realized as direct objects: e.g., English *I left the room*.

At first glance, it appears that the three location/motion verbs indeed express locative arguments as direct objects in all five sign languages in our sample. For instance, in example (6) from RSL, the constituent INDEX CHIEF POLICE is a Goal and could be analyzed as a direct object given that there is no special marking (e.g., a preposition) to indicate that its status is rather that of an oblique/adjunct. This is based on the definition by Haspelmath, given in Section 2.1, repeated here: “A verb is considered transitive if it contains an A and a P argument. A and P are defined as the arguments of a verb with at least two arguments that are coded like the ‘breaker’ and the ‘broken thing’ micro-roles of the ‘break’ verb.” (Haspelmath, 2015, p. 136). That is, since we do not see any difference in the marking of the referent with the semantic role Goal from that of a prototypical Patient with a verb like ‘break’, we could argue for example (6) containing an A and a P argument, thus constituting a transitive construction.

- (6) (Russian Sign Language)
 RUN INDEX CHIEF POLICE
 ‘He ran to the chief policeman.’ (<http://rsl.nstu.ru/data/view/id/53/t/47520/d/50040>)

Thus, distinguishing between arguments and adjuncts by overt marking in sign languages is notoriously problematic. Sign languages generally lack case marking (cf. Gil, 2014), ruling out the use of accusative case as a criterion for direct object status (however, see Börstell, 2019; Meir, 2003). Directionality in verbs can potentially be used to distinguish arguments from adjuncts, but since only a subset of verbs are directional in sign languages, this is not a universally applicable – hence useful – criterion either. LEAVE, LIVE, and RUN in RSL are all non-directional verbs, for instance. Moreover, prepositions appear to be less frequent across sign languages (cf. Börstell et al., 2016), even when expressing spatial relations. Due to the visuo-spatial nature of the signed modality, spatial relations can be expressed using signing space. This is achieved by localizing/directing verbs, nominal signs, and pointing signs to reflect topographic relations between objects that are being described. Prepositions would be able to distinguish core locative objects from obliques, but since prepositions appear to be cross-linguistically infrequent for sign languages, this criterion is less useful for signed than for spoken languages.

In general, descriptions of place (and time) are realized identically with location verbs (with which they could be considered semantic arguments) as with all other verbs (with which one might analyze them as adjuncts). Thus, we have no good way of determining whether semantically locative arguments of verbs of location/motion are syntactic arguments or not.

Without making any *a priori* assumptions about the syntactic status of locatives, we conducted our analysis in two steps to assess whether the results would provide any support for either of the two possible definitions. As a first step, we treated locatives as direct objects, based on the fact that they indeed exhibit the same encoding as a prototypically transitive construction (as in, no overt marking). In a second step, we removed the locatives of these verbs on the basis of semantics, under the assumption that Goal arguments are syntactically different from Patient arguments, despite the identical encoding – that is, going against Haspelmath's definition. As described in Section 3, the former approach leads to a smaller number of significant (and overall lower) correlations between our sampled sign languages as well as between sign languages and spoken languages based on Haspelmath's (2015) ValPaL data.

The question thus arises as to how one should evaluate these two approaches. If one accepts the former approach as valid, it would indicate that locatives introduce more cross-linguistic variation by not patterning as uniformly across languages as semantically prototypical (Patient-like) direct objects do. This could in itself be a reason for treating them as obliques – i.e., reverting to the second approach – although such a definition would be based solely on semantics.

Conversely, if one assumes that these locatives are indeed direct objects, one consequence is then that one needs to account for the lower correlations across languages – within and across modalities. In this view, the variety could be attributed to some special status of such locative (direct) objects in sign languages, as a group. However, this interpretation would predict that the correlations between pairs of sign languages should remain high or even increase if it were a cross-linguistically stable property of the modality, while the correlations between sign languages and spoken languages (as represented by Haspelmath's ValPaL languages) should decrease. This is not what we find: if locatives are analyzed as direct objects, *all* correlations decrease, also within the group of sign languages. If one expects cross-linguistic and cross-modal comparisons to pattern similarly, the former approach – treating locatives as objects – tentatively seems more accurate, based on our data. For the purpose of our study, we adopted this definition as our starting point since it matched more closely to Haspelmath's definition of transitivity. We maintain that this is a reasonable assumption seeing as, e.g., Aldai and Wichmann (2018) only point to general patterns of semantics as a cross-linguistically identifiable correlate with transitivity (with variation within and between languages). That is, semantics alone cannot predict a universal cross-linguistically precise classification of transitivity, even if the general pattern appears to be pointing to many similar preferences on the basis of semantics.

5 Conclusions

In this paper, we had two overarching aims: first, we wanted to evaluate whether categorical valency dictionary classifications were correlated with corpus data, in order to validate previous cross-modal work (Kimmelman, 2016) and support the choice of replicating this methodology with our extended dataset; second, we wanted to look at possible correlations between languages, within and across modalities, to see to what extent transitivity prominence as a realization of semantic conceptualization exhibits a cross-linguistically visible pattern.

With regard to the first aim, we focused on spoken languages, for which we had both dictionary (Haspelmath, 2015) and corpus (Nivre et al., 2017) data. Here, we demonstrated that corpus-based transitivity correlates well with dictionary-based transitivity, validating both previous work and the choice of our current, similar approach. Despite limitations in terms of the size of datasets and language samples, the fact that two somewhat crude measures of transitivity are correlated is positive for further work in this domain. Whether one uses dictionary or corpus data, or both, the outcome should provide similar results. Also, both metrics pointed to the same general patterns across languages and we see this as another piece of support for a semantic motivation in the expression of transitivity.

As for the second aim, we compared transitivity prominence in five sign languages to each other (within the signed modality) and to transitivity prominence of corresponding verb meanings across spoken languages based on both categorical dictionary data and scalar (corpus) data from actual language use. We found positive correlations between our sampled sign languages as well as between the signed (corpus data) and spoken (dictionary and corpus data) languages treated as groups representing their respective modalities. We conclude that these results furthermore corroborate the idea of a cross-linguistically and cross-modally observable semantic basis of transitivity, albeit with considerable variation present in individual languages and verbs. Semantic patterns in transitivity hierarchies have been argued for previously on the basis of spoken language data (Hopper and Thompson, 1980; Tsunoda, 1981, 1985; Malchukov, 2005; Aldai and Wichmann, 2018), but now receives cross-modal support with this study.

Although we find some expected patterns across languages and modalities with regard to transitivity, the methods used here are based on limited data. There are, of course, limitations that come with relying on small-sized corpora, and we tentatively hypothesize that some of the cross-linguistic differences we observed are in fact caused by the sampled data, in terms of size or type of data. The most obvious example of this was the fact that semantically locative arguments (arguments of location/motion verbs) in sign languages should be defined as direct objects based on the definitions of transitivity used in previous studies (Haspelmath, 2015), but pattern differently across languages, both within and across modalities. To some extent, this could be a consequence of the visuo-spatial nature of the signed modality – e.g., by expressing location differently from spoken languages – but with the limited datasets we are using, this should not be seen as a conclusion of this study but rather an idea to be tested in future research.

Lastly, we found that relatedness between sign languages in our sample does not appear to have an effect on the degree of similarity in transitivity prominence ranking in any straightforward way. However, given the limited language sample and the small size of corpora here, this issue also warrants further research. Here, it would be useful if known cognate verb lemmas could be compared directly across related languages, rather than lumped together under overarching, semantically generalized verb meanings.

Acknowledgement: This study was carried out as part of the project *Argument structure in three sign languages: typological and theoretical aspects* funded by the Dutch Science Foundation (NWO), grant number 360-70-520, and has also partly been funded by the Academy of Finland, grant numbers 269089, 273408, and 304034.

We wish to thank Robert Östling for providing us with transitivity data from the Universal Dependencies corpora, and participants at the *Argument Structure Across Modalities (ASAM 2018)* workshop in Amsterdam, February 1–2, 2018, for comments on the earlier stages of the project.

We thank the editor and two anonymous reviewers for helpful comments on the paper.

Abbreviations

Acronyms

DGS German Sign Language (*Deutsche Gebärdensprache*)

FinSL Finnish Sign Language

NGT Sign Language of the Netherlands (*Nederlandse Gebarentaal*)

OBJ The label *obj* is used to refer to the UD label for the dependency of *direct object*, defined as “the noun phrase that denotes the entity acted upon or which undergoes a change of state or motion (the proto-patient)” (see <https://universaldependencies.org/u/dep/obj.html>)

RSL Russian Sign Language

SSL Swedish Sign Language¹⁶

UD Universal Dependencies (Nivre et al., 2017)

ValPaL The database *Valency Patterns Leipzig* (Hartmann et al., 2013)

Glossing

Signs are represented by approximate translation glosses in SMALL CAPS, as is tradition in sign language linguistics (see <https://benjamins.com/series/sll/guidelines.pdf>). The specific glosses used in this paper are listed below:

#SIGN: the prefix “#” marks a reduced fingerspelled sign

INDEX: represents a pointing sign

INDEX₃: the subscript “3” indicates 3rd person reference

OBJPRO: a dedicated object pronoun (see Börstell, 2017, 2019)

PL: plural form

POSS: a possessive marker

PU: palms up sign (a frequent sign/gesture with a variety of functions; see Cooperrider et al. 2018)

References

- Aldai, Gontzal, and Wichmann, Søren 2018. Statistical observations on hierarchies of transitivity. *Folia Linguistica*, 52 2 249–281.
- Arka, I Wayan 2014. Locative-related Roles and the Argument-Adjunct Distinction in Balinese. *Linguistic Discovery*, 12 2 56–84.
- Benedicto, Elena, and Brentari, Diane 2004. Where did all the arguments go? Argument-changing properties of classifiers in ASL. *Natural Language & Linguistic Theory*, 22 743–810.
- Bergman, Brita, and Engberg-Pedersen, Elisabeth 2010. Transmission of sign languages in the Nordic countries. In: Brentari, Diane (ed.), *Sign languages: A Cambridge language survey*, 74–94. New York, NY: Cambridge University Press.
- Bickford, J. Albert 2005. *The Signed Languages of Eastern Europe*. Tech. rep., SIL International.
- Börstell, Carl 2017. *Object marking in the signed modality: Verbal and nominal strategies in Swedish Sign Language and other sign languages*. Ph.D. thesis, Stockholm University.
- Börstell, Carl 2019. Differential object marking in sign languages. *Glossa: a journal of general linguistics*, 4 1 3.
- Börstell, Carl, Hörberg, Thomas, and Östling, Robert 2016. Distribution and duration of signs and parts of speech in Swedish Sign Language. *Sign Language & Linguistics*, 19 2 143–196.

¹⁶ As of April 2019, the Swedish Language Council (*Språkrådet*) and the Sign Language Linguistics group at Stockholm University have agreed to use the acronym “STS” for Swedish Sign Language, based on the Swedish name *svenskt teckenspråk*. We have decided to keep the English label here for continuity within our project, but acknowledge the change of preferred nomenclature for future papers.

- Börstell, Carl, Mesch, Johanna, and Wallin, Lars 2014. *Segmenting the Swedish Sign Language Corpus: On the possibilities of using visual cues as a basis for syntactic segmentation*. In: Crasborn, Onno, Efthimiou, Eleni, Fotinea, Evita, Hanke, Thomas, Kristoffersen, Jette, and Mesch, Johanna (eds.), *Proceedings of the 6th Workshop on the Representation and Processing of Sign Languages: Beyond the Manual Channel*, 7–10, Paris. European Language Resources Association (ELRA).
- Bossong, Georg 1985. *Differentielle Objektmarkierung in den Neuiranischen Sprache*. Tübingen: Gunter Narr Verlag.
- Burkova, Svetlana 2015. *Russian Sign Language Corpus 2012–2015*. Novosibirsk: Novosibirsk State Technical University.
- Cennamo, Michela, and Fabrizio, Claudia 2013. Italian Valency Patterns. In: Hartmann, Iren, Haspelmath, Martin, and Taylor, Bradley (eds.), *Valency Patterns Leipzig*. Leipzig: Max Planck Institute for Evolutionary Anthropology.
- Cooperrider, Kensy, Abner, Natasha, and Goldin-Meadow, Susan 2018. The Palm-Up Puzzle: Meanings and Origins of a Widespread Form in Gesture and Sign. *Frontiers in Communication*, 3.
- Crasborn, Onno, and Zwitserlood, Inge 2008. *The Corpus NGT: an online corpus for professionals and laymen*. In: Crasborn, Onno, Efthimiou, Eleni, Hanke, Thomas, Thoutenhoofd, Ernst D., and Zwitserlood, Inge (eds.), *Proceedings of the 3rd Workshop on the Representation and Processing of Sign Languages. [6th International Conference on Language Resources and Evaluation (LREC 2008)]*, 44–49, Paris. ELRA.
- Crasborn, Onno, Zwitserlood, Inge, and Ros, Johan 2008. *The Corpus NGT. An open access digital corpus of movies with annotations of Sign Language of the Netherlands*. Centre for Language Studies, Radboud University Nijmegen. ISLRN: 175-346-174-413-3.
- Crasborn, Onno A. 2007. How to recognise a sentence when you see one. *Sign Language & Linguistics*, 10 2 103–111.
- De Lint, Vanja 2018. NGT classifier constructions: an inventory of arguments. *Sign Language & Linguistics*, 21 1.
- Dixon, R. M. W. 1979. Ergativity. *Language*, 55 59–138.
- Dowty, David R. 1982. Grammatical relations and Montague grammar. In: Jacobson, Pauline, and Pullum, Geoffrey K. (eds.), *The Nature of Syntactic Representation*, 79–130. Dordrecht: Springer.
- Fenlon, Jordan, Denmark, Tanya, Campbell, Ruth, and Woll, Bencie 2007. Seeing sentence boundaries. *Sign Language & Linguistics*, 10 2 177–200.
- Gil, David 2014. Sign languages, creoles, and the development of predication. In: Newmeyer, Frederick J., and Preston, Laurel B. (eds.), *Measuring grammatical complexity*, 37–64. Oxford: Oxford University Press.
- Goddard, Cliff 2013. English Valency Patterns. In: Hartmann, Iren, Haspelmath, Martin, and Taylor, Bradley (eds.), *Valency Patterns Leipzig*. Leipzig: Max Planck Institute for Evolutionary Anthropology.
- Hartmann, Iren, Haspelmath, Martin, and Taylor, Bradley (eds.) 2013. *Valency Patterns Leipzig*. Leipzig: Max Planck Institute for Evolutionary Anthropology.
- Haspelmath, Martin 2014. Arguments and Adjuncts as Language-Particular Syntactic Categories and as Comparative Concepts. *Linguistic Discovery*, 12 2.
- Haspelmath, Martin 2015. Transitivity prominence. In: Malchukov, Andrej, and Comrie, Bernard (eds.), *Valency classes in the world's languages: A comparative handbook*, Vol. 1, 131–147. Berlin: De Gruyter.
- Haspelmath, Martin, and Baumann, Luisa 2013. German Valency Patterns. In: Hartmann, Iren, Haspelmath, Martin, and Taylor, Bradley (eds.), *Valency Patterns Leipzig*. Leipzig: Max Planck Institute for Evolutionary Anthropology.
- Hawkins, John A. 1986. *A comparative typology of English and German: Unifying the contrasts*. Austin, TX: University of Texas Press.
- Hopper, Paul, and Thompson, Sandra A. 1980. Transitivity in grammar and discourse. *Language*, 56 251–299.
- Hou, Lynn, and Meier, Richard P. 2018. The morphology of first-person object forms of directional verbs in ASL. *Glossa: a journal of general linguistics*, 3 1 114.
- Jantunen, Tommi, Pippuri, Outi, Wainio, Tuija, Puupponen, Anna, and Laaksonen, Jorma 2016. Annotated Video Corpus of FinSL with Kinect and Computer-Vision Data. In: Efthimiou, Eleni, Fotinea, Stavroula-Evita, Hanke, Thomas, Hochgesang, Julie, Kristoffersen, Jette, and Mesch, Johanna (eds.), *Proceedings of the 7th Workshop on the Representation and Processing of Sign Languages: Corpus mining [10th International Conference on Language Resources and Evaluation (LREC 2016)]*, 93–100. Paris: ELRA.
- Kimmelman, Vadim 2016. Transitivity in RSL: A corpus-based account. In: Efthimiou, Eleni, Fotinea, Stavroula-Evita, Hanke, Thomas, Hochgesang, Julie, Kristoffersen, Jette, and Mesch, Johanna (eds.), *Proceedings of the 7th Workshop on the Representation and Processing of Sign Languages: Corpus mining [10th International Conference on Language Resources and Evaluation (LREC 2016)]*, 117–120. Paris: ELRA.
- Kimmelman, Vadim, de Lint, Vanja, de Vos, Connie, Oomen, Marloes, Pfau, Roland, Vink, Lianne, and Aboh, Enoch O. 2019. Argument Structure of Classifier Predicates: Canonical and Non-canonical Mappings in Four Sign Languages. *Open Linguistics*, 5 1 332–353.
- Kishimoto, Hideki, and Kageyama, Taro 2013. Japanese (standard) Valency Patterns. In: Hartmann, Iren, Haspelmath, Martin, and Taylor, Bradley (eds.), *Valency Patterns Leipzig*. Leipzig: Max Planck Institute for Evolutionary Anthropology.
- Kittilä, Seppo 2010. Transitivity Typology. In: Song, Jae Jung (ed.), *The Oxford Handbook of Linguistic Typology*. Oxford: Oxford University Press.
- Langer, Gabriele 2012. *A Colorful First Glance at Data on Regional Variation Extracted from the DGS-Corpus: With a Focus on Procedures*. In: Crasborn, Onno, Efthimiou, Eleni, Fotinea, Evita, Hanke, Thomas, Kristoffersen, Jette, and Mesch, Johanna (eds.), *Proceedings of the 5th Workshop on the Representation and Processing of Sign Languages: Interactions between Corpus*

- and Lexicon. [8th International Conference on Language Resources and Evaluation (LREC 2008)], 101–108. Paris: ELRA.
- Malchukov, Andrej 2005. Case pattern splits, verb types and construction competition. In: Amberber, Mengistu, and de Hoop, Helen (eds.), *Competition and variation in natural languages. The case for case*, 73–117. Oxford: Elsevier.
- Malchukov, Andrej, and Jahraus, Alexander 2013. Russian Valency Patterns. In: Hartmann, Iren, Haspelmath, Martin, and Taylor, Bradley (eds.), *Valency Patterns Leipzig*. Leipzig: Max Planck Institute for Evolutionary Anthropology.
- Meir, Irit 2003. Grammaticalization and modality: The emergence of a case-marked pronoun in Israeli Sign Language. *Journal of Linguistics*, 39 1 109–140.
- Meir, Irit, Padden, Carol, Aronoff, Mark, and Sandler, Wendy 2007. Body as subject. *Journal of Linguistics*, 43 531–563.
- Mesch, Johanna 2006. Påminner nationella teckenspråk om varandra? In: Hoyer, Karin, Londen, Monica, and Östman, Jan-Ola (eds.), *Teckenspråk: Sociala och historiska perspektiv*, Nr. 6 in Nordica Helsingiensia – Teckenspråksstudier TS, 71–95. Nordica: Institutionen för nordiska språk och nordisk litteratur, Helsingfors universitet.
- Mesch, Johanna 2018. *Annotated files for the Swedish Sign Language Corpus. Version 6*. Sign Language Section, Department of Linguistics, Stockholm University.
- Mesch, Johanna, Wallin, Lars, Nilsson, Anna-Lena, and Bergman, Brita 2012. *Dataset: Swedish Sign Language Corpus Project (2009–2011). Version 1*. Sign Language Section, Department of Linguistics, Stockholm University.
- Nivre, Joakim, et al. 2017. *Universal Dependencies 2.1*. LINDAT/CLARIN digital library at the Institute of Formal and Applied Linguistics (ÚFAL), Faculty of Mathematics and Physics, Charles University.
- Næss, Åshild 2007. *Prototypical Transitivity*. Amsterdam/Philadelphia, PA: John Benjamins.
- Oomen, Marloes 2017. Iconicity in argument structure: Psych-verbs in Sign Language of the Netherlands. *Sign Language & Linguistics*, 20 1 55–108.
- Oomen, Marloes, and Kimmelman, Vadim 2019. Body-anchored verbs and argument omission in two sign languages. *Glossa: a journal of general linguistics*, 4 1 42.
- Pfau, Roland, Salzmann, Martin, and Steinbach, Markus 2018. The syntax of sign language agreement: Common ingredients, but unusual recipe. *Glossa: a journal of general linguistics*, 3 1 107.
- R Core Team 2015. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria.
- Salonen, Juhana, Puupponen, Anna, Takkinen, Ritva, and Jantunen, Tommi 2019. Suomen viittomakielten korpusta rakentamassa [Building Corpus FinSL]. In: Jantunen, Jarmo Harri, Brunni, Sisko, Kunnas, Niina, Palviainen, Santeri, and Västi, Katja (eds.), *Proceedings of the Research Data and Humanities (RDHum) Conference 2019: Data, Methods and Tools*, Vol. 17 of *Studia Humaniora Ouluensia*, 81–96. Oulu: University of Oulu.
- Salonen, Juhana, Takkinen, Ritva, Puupponen, Anna, Pippuri, Outi, and Nieminen, Henri 2016. Creating the corpus of Finland's sign languages. In: Efthimiou, Eleni, Fotinea, Stavroula-Evita, Hanke, Thomas, Hochgesang, Julie, Kristoffersen, Jette, and Mesch, Johanna (eds.), *Proceedings of the 7th Workshop on the Representation and Processing of Sign Languages: Corpus mining [10th International Conference on Language Resources and Evaluation (LREC 2016)]*, 179–184. Paris: ELRA.
- Seržant, Ilja A., and Witzlack-Makarevich, Alena (eds.) 2018. *Diachrony of differential argument marking*. Nr. 19 in *Studies in Diversity Linguistics*. Berlin: Language Science Press.
- Sinnemäki, Kaius 2014. A typological perspective on Differential Object Marking. *Linguistics*, 52 2 281–313.
- Slowikowski, Kamil 2017. *ggrepel: Repulsive Text and Label Geoms for 'ggplot2'*. R package version 0.7.0.
- Tsunoda, Tasaku 1981. Split case-marking in verb types and tense/aspect/mood. *Linguistics*, 19 389–438.
- Tsunoda, Tasaku 1985. Remarks on transitivity. *Journal of Linguistics*, 21 2 385–396.
- Wichmann, Søren 2014. Arguments and Adjuncts Cross-Linguistically: A Brief Introduction. *Linguistic Discovery*, 12 2.
- Wichmann, Søren 2015. Statistical observations on implicational (verb) hierarchies. In: Malchukov, Andrej, and Comrie, Bernard (eds.), *Valency Classes in the World's Languages: Introducing the Framework, and Case Studies from Africa and Eurasia*, Vol. 1, 155–181. Berlin: De Gruyter Mouton.
- Wichmann, Søren 2016. Quantitative tests of implicational verb hierarchies. In: Kageyama, Taro, and Jacobsen, Wesley M. (eds.), *Transitivity and Valency Alternations: Studies on Japanese and Beyond*, 423–444. Berlin: De Gruyter Mouton.
- Wickham, Hadley 2009. *ggplot2: Elegant Graphics for Data Analysis*. New York, NY: Springer-Verlag.
- Wittenburg, Peter, Brugman, Hennie, Russel, Albert, Klassmann, Alex, and Sloetjes, Han 2006. *ELAN: A professional framework for multimodality research*. In: *Proceedings of the 5th International Conference on Language Resources and Evaluation (LREC 2006)*, 1556–1559.
- Zhang, Guohua 2013. Mandarin Chinese Valency Patterns. In: Hartmann, Iren, Haspelmath, Martin, and Taylor, Bradley (eds.), *Valency Patterns Leipzig*. Leipzig: Max Planck Institute for Evolutionary Anthropology.

Appendix 1: Simulations

We used the code below to run simulations as described in Section 2.3. It produces a table with the results of the simulation (average correlations in each case, which we expect to be 0, and average proportion of significant correlations, which we expect to be 0.05). Note that since random numbers are generated, each time the simulation is ran, the results are different. The results of one such run are reported in Table A1.

Table A1: Results of a simulation.

| No of verbs | No of items | Average correlation | p<0.05 |
|-------------|-------------|---------------------|--------|
| 25 | 25 | −0.0013122615 | 0.0506 |
| 25 | 10 | −0.0001732446 | 0.0503 |
| 25 | 5 | 0.0011945194 | 0.0497 |
| 10 | 25 | −0.0008762900 | 0.0537 |
| 10 | 10 | 0.0018579192 | 0.0563 |
| 10 | 5 | −0.0010861701 | 0.0540 |
| 5 | 25 | −0.0119686383 | 0.0587 |
| 5 | 10 | 0.0012462338 | 0.0686 |
| 5 | 5 | 0.0040195351 | 0.0819 |

```

1  iterations=10000
2  all<-0
3  sig<-0
4  results<-data.frame(cbind(rep(0, each=9),rep(0, each=9),rep(0, each=9),rep(0, each=9)))
5  colnames(results)<-c("N of verbs", "N of items", "Corr", "P<0.05")
6  j=1
7  for (verbs in c(25,10,5)) {
8    for (size in c(25,10,5)) {
9      for (i in 1:iterations){
10         a<-(rbinom(verbs, size, 0.2))/size
11         b<-(rbinom(verbs, size, 0.5))/size
12         all[i]<-cor(rank(a), rank(b))
13         sig[i]<-cor.test(rank(a), rank(b))$p.value<0.05
14       }
15       results[j,1]=verbs
16       results[j,2]=size
17       results[j,3]=mean(na.omit(all))
18       results[j,4]=sum(na.omit(sig)+is.na(sig))/iterations
19       j=j+1
20     }
21   }
22  results

```

A couple of clarifications of the code are in order. First, the third value in the `rbinom` function (lines 10, 11) is the probability of success in each trial (that is, modelling whether the verb occurs with an overt object or not). These probabilities should not play a role in the simulation, but just in case we used 0.2 to imitate the comparatively low rates of overt objects in corpus data and 0.5 to imitate the comparatively higher numbers in the ValPaL data.

Second, the correlation test does not work if all the values in the two vectors are identical (this produces NAs as the result). Therefore, the code includes the `na.omit` command when calculating the average correlations (line 17) and the proportion of significant correlations (line 18). However, if we simply omit the cases where the vectors are accidentally identical, we are clearly underestimating the accidental correlations. So the proportion of significant correlations is calculated based on the sum of such correlations and the cases that produce NAs (identical vectors).

Appendix 2: Excluding locative verbs

Table A2 shows the pairwise correlations between languages with locative verb meanings completely removed.

Table A2: The statistical rank correlation across ValPaL verb meanings within each language dyad for our five sign languages and the combined spoken languages from Haspelmath (2015). Locative verb meanings are removed for all languages.

| | FinSL | NGT | RSL | SSL | SpokenL |
|-------|-----------------|---------------|--------------------|-----------------|--------------------|
| DGS | $\rho = .736^*$ | $\rho = .460$ | $\rho = .393$ | $\rho = .762^*$ | $\rho = .450$ |
| FinSL | | $\rho = .633$ | $\rho = .617$ | $\rho = .433$ | $\rho = .636$ |
| NGT | | | $\rho = .800^{**}$ | $\rho = .333$ | $\rho = .845^{**}$ |
| RSL | | | | $\rho = .067$ | $\rho = .577$ |
| SSL | | | | | $\rho = .276$ |