The more you move, the more action you construct – A motion capture study on head and upper-torso movements in constructed action in Finnish Sign Language narratives

Tommi Jantunen, Danny De Weerdt, Birgitta Burger, Anna Puupponen

University of Jyväskylä, Finland

Abstract

This paper investigates, with the help of motion capture data processed on corpus principles, the characteristics of head and upper-torso movements in constructed action and regular narration (i.e. signing without constructed action) in FinSL. Specifically, the paper evaluates the validity of two arguments concerning constructed action: that constructed action forms a continuum with regular narration, and that constructed action divides into three subtypes (i.e. overt, reduced and subtle). The results presented in the paper support the first argument but not directly the second one. Because of the ambiguous position of reduced constructed action in between subtle and overt constructed action, we argue in the paper that the present three-part typology of constructed action may need revising. As an alternative way of subcategorizing the phenomenon we propose a division between strong and weak constructed action.
Keywords: constructed action, narration, motion capture, kinematics, sign language, Finnish Sign Language

1. Introduction

In this paper we combine a corpus approach and motion capture technology to investigate the characteristics of head and upper-torso movements in constructed action and regular narration in Finnish Sign Language (FinSL). Constructed action is depictive gestural enactment in which signers (and speakers alike) use their hands and other parts of the body to show (as opposed to telling about) the actions, feelings, thoughts and speech of characters they are referring to in the discourse (e.g. Cormier et al. 2015; Ferrara & Hodge 2018). Reglar narration, on the other hand, refers to signing (or speech) without constructed action (i.e. pure telling) in a narrative context. Narration is here understood in a broad sense of a discourse type in which a signer (or a speaker) is active in reporting a story of some kind.

In works using or relying on sign language corpus data, constructed action has been argued (i) to form a continuum with regular narration, so that the articulatory border between the two is fuzzy and never clear cut (e.g. Cormier et al. 2015; Jantunen 2017). Additionally, on the basis of the number of enacting articulators and the prominence of character perspective, it has also been argued (ii), that constructed action – forming a continuum with regular narration – can be further divided into three subtypes, which have been labeled overt, reduced and subtle constructed action (e.g. Cormier et al. 2015; Jantunen et al. 2020). These two arguments are not competing ones. Instead, they complete each other from two different perspectives, the first of which focuses on the relation between constructed action and regular narration as a whole and the second on the internal variation of the constructed action. Moreover, the argument (ii) is particularly important from the methodological perspective as, for example, empirical corpus studies often need to refer to the subtypes explicitly even though their nature is elusive or fuzzy compositionally (i.e. a token may not fall neatly into any of the three types, see Cormier et al. 2015: 199–200).

In this paper we focus on these two arguments and evaluate their validity from a kinematic perspective. The motivation for this evaluation is that, although the two arguments have been well established from a semiotic corpus perspective, empirical work on sign languages has to regularly make reference to what is seen, that is, to the actual movements of the signers. The alignment of visually observed physical activities with semiotically established theoretical claims is thus highly important. In the kinematic evaluation of the two arguments we use motion capture data which is synchronized with video data processed by the application of corpus principles. There has been very little previous work exploiting motion capture technology in research on constructed action (see Jantunen et al. 2020) although there has been a gradual increase in the number of motion capture studies on sign languages in general (e.g. Tyrone & Mauk 2010; Jantunen 2013; Puupponen et al. 2015).

Theoretically, our work follows the guiding assumptions of the cognitive–functional framework and builds on the presumptions that language (signed or spoken) is a physical, mental and social entity all at the same time, and that it emerges gradiently from our bodily interactions with the environment. This premise has both a methodological consequence and a conceptual consequence, both of which we want to underline. The methodological consequence is that in order to fully understand any linguistic phenomenon – such as, in our case, constructed action in FinSL – it has to be studied with multidimensional data and a variety of methods. This is the motivation for us to integrate the corpus approach with motion capture. The conceptual consequence, on the other hand, is that we cannot assume any categorical distinction between concrete everyday activity and linguistic activity. By extension this means, for example, that actions traditionally referred to as with terms such as gesture or gesturality have to be considered part of language, not something outside of it. In the present paper we set out to test the validity of this claim explicitly (see the first argument above), as constructed action is gestural enactment by definition.

In their recent paper, Ferrara & Hodge (2018) approach linguistic utterances as multimodal composites (see Enfield 2009) and point out that they not only signal description but also indicate and depict in different degrees, depending on the context. We fully concur with this. In addition, we would like to add that while describing, indicating and depicting, the utterances also vary in their degree of conventionality and discreteness (see Jantunen 2017, 2018; Puupponen 2019). In other words, we presume throughout the paper that with linguistic utterances we not only tell about meanings but we also indicate and show meanings, and that ultimately each utterance may position differently on a continuum between a conventional, discrete end and an unconventional, gradient end (for more, see Jantunen 2017; Puupponen 2019).

It should be noted that utterances which are relatively unconventional and gradient and which primarily depict and indicate have been traditionally referred to as gestures of some kind. We want to reiterate (see above) that in this paper we indeed treat these types of utterances – manual gestures as well as constructed action – as part of language, not as systems existing outside language (for a similar view, see Kendon 2004; Enfield 2009) – although at the same time we set out to investigate the validity of this approach with kinematic data. Moreover, we also want to emphasize that this view of language is applicable not only to sign languages but also to spoken languages – as long as spoken language is not defined with reference to speech only. Consequently, although our focus in this paper is on FinSL, we assume that the findings we present in this paper are relevant to the investigation of enactment not only in the domain of sign language in general but also in the domain of spoken language (defined to include gestures and gesturality).

2. Constructed action, its degrees and articulation

As stated above in Section 1, we approach constructed action as a depictive gestural activity in which signers enact their individual mental constructions of someone else's actions, feelings, thoughts or speech for discourse purposes. This is a fairly broad definition and not all researchers have approached constructed action in this way. Most notably, the showing of the speech of the enacted referent (cf. quotation) has often been excluded from the definition of constructed action and approached as a phenomenon in its own right, namely, constructed dialogue (e.g. Hodge & Ferrara 2013; Ferrara & Johnston 2014; Jantunen 2017). However, as argued, for example, by Hodge & Cormier (2019), no clear distinction can be made between the enactments of actions (i.e. traditional or even prototypical constructed action) and the enactments of speech (i.e. constructed dialogue), and it is for this reason that we here include constructed dialogue in our definition of constructed action.

In some studies the physical phenomenon constituting constructed action has been analyzed under the notion of role shift. This is particularly true in some older research as well as in work carried out in the formal tradition, in which role shift is defined as a grammatical mechanism by which signers can shift into the role of a character and thus convey information from that character’s perspective (Lillo-Martin 2012). In our work we will occasionally speak about role shift too. However, unlike the definition of role shift in the formal tradition, we will use the notion only generally to describe the punctual process of actually 'shifting roles' for the purpose of starting constructed action, changing character roles within constructed action or returning from constructed action to regular narration (Cormier et al. 2015).

The early work on constructed action also approached it as an articulatorily holistic phenomenon (e.g. Hodge & Ferrara 2013; Ferrara & Johnston 2014; Jantunen 2017). However, recent work by Cormier et al. (2015) has shown that constructed action has degrees, on the basis of which it is possible to classify constructed action on a continuum into three prototypes, which Cormier et al. refer to as overt, reduced and subtle constructed action (id., 188–192). These three types together with their compositional characteristics are demonstrated with data from FinSL in Figure 1.

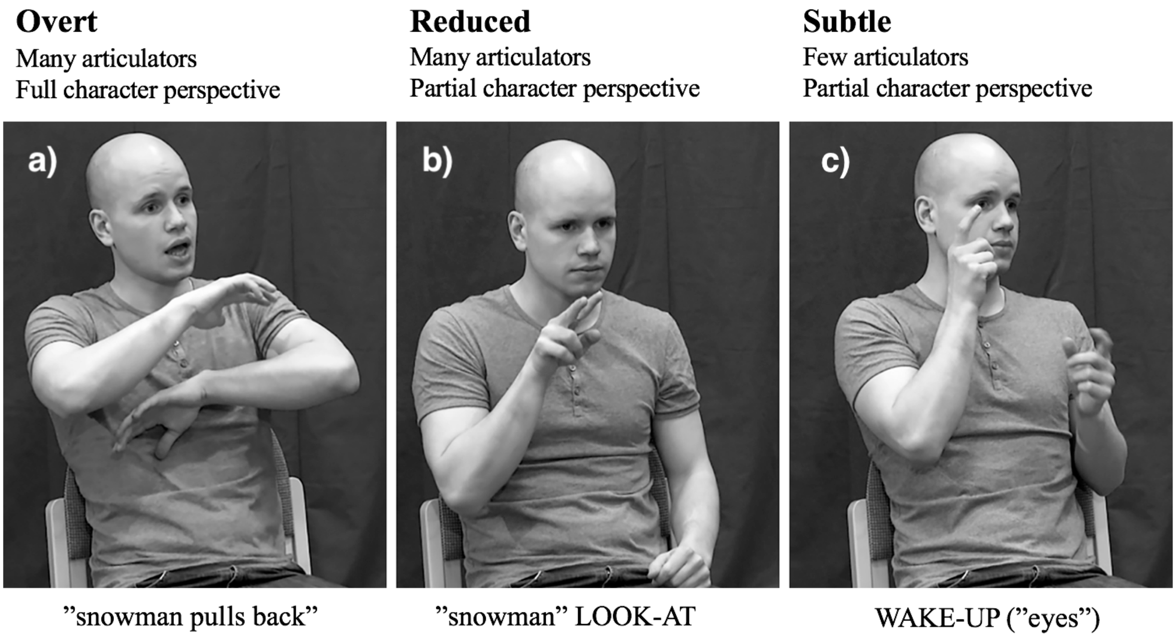

Figure 1. Examples of the three constructed action types and their compositional characteristics in FinSL. The video underlying the still images is part of the open access ProGram data and can be accessed online with annotations at http://hdl.handle.net/11113/00-0000-0000-0000-32DB-8@view.

The defining feature of overt constructed action is that a relatively high number of articulators are involved in the enacting and that the signer is fully in the role of a character. In Figure 1a this means that the movements and posture of the signer represent the movements and posture of the enacted referent, a snowman pulling back from a hot stove, not those of the narrator. In reduced constructed action, the number of enacting articulators is still relatively high but the character perspective is only partial, that is, the signer takes the roles of both the character and the narrator. In Figure 1b, this can be seen in the fact that while most of the activity of the signer is still portraying the snowman looking forwards, the dominant hand is producing the partly lexical sign LOOK-AT, which is a contribution from the narrator's perspective. In subtle constructed action, as the name suggests, there are only a few enacting articulators, and only a partial character perspective. In Figure 1c, this means that most of the visible activity is that of the signer-narrator, including the lexical sign WAKE-UP, and it is only the depictive activity of the upper face (i.e. the eyes and the eye brows) that portray some features of the character, the boy waking up and opening his eyes.

Definitionally, utterances involving overt constructed action are considered to be more unconventional, gradient and depictive than utterances involving subtle constructed action, which are in their nature relatively conventional, discrete and descriptive – also the prototypical characteristics of regular narration (see Ferrara & Hodge 2018). In studies where constructed action has been approached as an articulatorily holistic phenomenon (e.g. Ferrara & Johnston 2014; Jantunen 2017), the most subtle instances of constructed action have typically been excluded from the analysis.

Constructed action is an optional discourse strategy. Moreover, when constructed action is used, there is plenty of individual variation in its production (Ferrara & Johnston 2014; Jantunen 2017). This variation is captured in Figure 2, which shows the share of constructed action types and regular narration in Corpus FinSL data from ten FinSL signers performing narration from picture books that have no text.

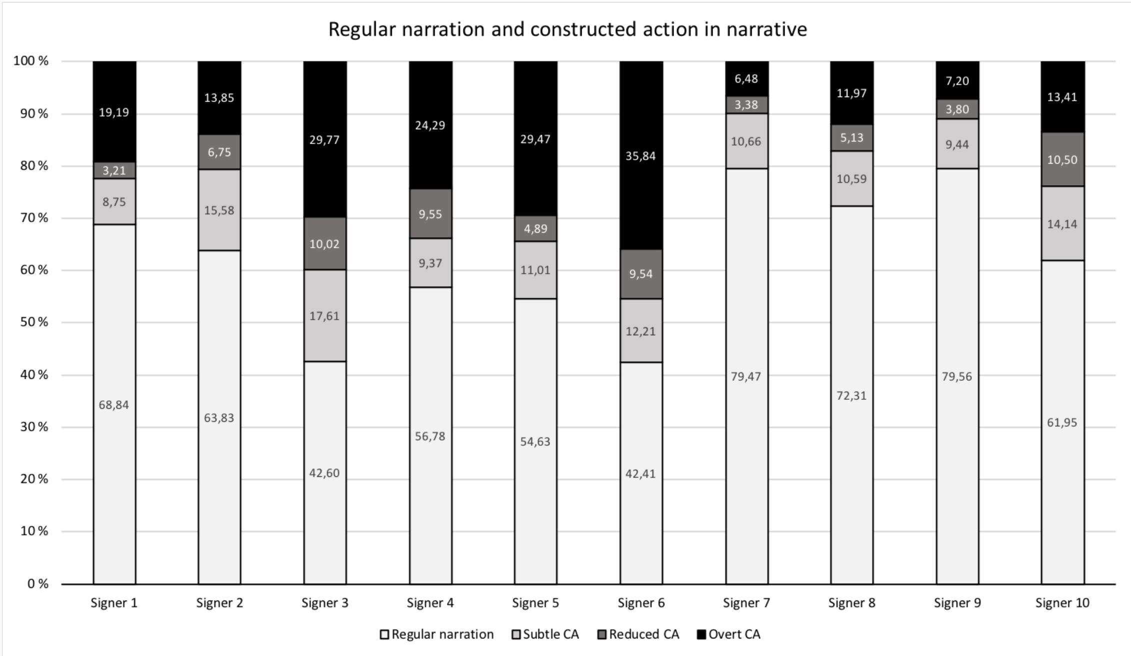

Figure 2. The individual variation in constructed action (CA) types and regular narration in FinSL stories signed on the basis of two textless picture books, the *Snowman* and *Frog, where are you?*. Data from ten signers included in Corpus FinSL, accessible at http://urn.fi/urn:nbn:fi:lb-2019012321.

Despite the individual variation, constructed action has been found to be used in fairly similar proportions across sign languages. For example, using data from Auslan and FinSL, the relative amount of (overt and reduced) constructed action in narratives has been found to be nearly 40 percent (Ferrara & Johnston 2014; Jantunen 2017). On the basis of this finding we assume that the results we will present in this paper are not FinSL-specific but apply equally well to other sign languages.

Concerning the articulation of constructed action explicitly, Ferrara & Johnston (2014: 200) state that "constructed action often involves shifts and changes in a person’s head and body posture, stance and facial expression, including the mouth, forehead, eyes and eyebrows." A bit later they continue that "there is not a formal list of features that identify constructed action, because their use depends on what is being enacted in any given instance" (id., 201). However, despite this, many studies on constructed action have given methodological priority to eye gaze: it is assumed that constructed action typically begins with a shift in eye gaze away from an addressee, and when constructed action ends, the gaze often returns toward an addressee (e.g. De Beuzeville et al. 2009; Herrmann & Steinbach 2012; Ferrara & Johnston 2014; Cormier et al. 2015; Engberg-Pedersen 2015; Jantunen 2017).

In our recent study (Jantunen et al. 2018), we used corpus annotations and eye-tracking data to investigate eye behavior at the beginning of stretches (total n=274) of different types of constructed action in FinSL. We found that although eye gaze shift occurred (with 81% frequency at the beginning of overt constructed action (n=105), the frequency of its occurrence at the beginning of reduced constructed action (n=109) decreased to 72% and at the beginning of subtle constructed action (n=60) it was only 58%. Our conclusion was that eye gaze shift away from the addressee cannot be considered a reliable indicator of constructed action, that is, we interpreted the results as supporting Ferrara & Johnston's (2014) claim about the lack of formal features identifying constructed action. In general, we interpreted the results of the eyetracking study to provide support for the three-way typology of constructed action.

In another study (Jantunen et al. 2020), we investigated the temporal order in which the dominant hand, the head, the upper torso and the eyes begin articulating overt constructed action from regular narration, that is, the role-shift from regular narration to overt constructed action. The study was based on a small sample (n=10) of computationally synchronized motion capture and eye-tracking data (Burger et al. 2018) and showed that the order in which the articulators began their overt constructed action-initial movement was not random but tended to follow the pattern, first the head and eyes and then the hand. In general, the result indicates that the change from regular narration to overt constructed action – the role shift – is indeed fairly systematic and, consequently, that there is perhaps more control in the articulation of (overt) constructed action than has often been acknowledged in previous studies. The issue awaits further study.

The characteristics of the movements of the head and the upper torso in constructed action and in regular narration, the focus of the present paper, have not been properly studied before. Our previous corpus-based analyses on FinSL (Jantunen 2017; Puupponen 2018) have shown that narration with constructed action tends to involve relatively more activity of the whole body (because the characters enacted often perform large whole body movements), while narration without constructed action tends to be associated with increased activity of the head (because the head has an important role in signalling various discourse-related functions such as affirmation, negation, utterance borders and contours). However, apart from these general findings, we still know very little about what the activity of the body and head is like in regular narration and constructed action. This lack of basic knowledge is also one of the reasons for carrying out the present study.

3. Data and methodology

Our work in this paper is based on synchronized motion capture and video data that have been annotated in ELAN for the three constructed action types and regular narration, according to the guidelines presented in Cormier et al. (2015). In the recording sessions, signers wore a set of 25 reflective markers whose locations were tracked with an eight-camera optical motion capture system. The task of the signers was to re-tell the content of textless cartoon strips to an addressee standing in front of them. For the present work we used data from five signers, who each participated in three tasks. On the basis of ELAN annotations, we extracted a total sample of 137 tokens belonging to the types of overt, reduced and subtle constructed action as well as regular narration. We then processed the motion capture data of these tokens in MATLAB so that we ended up with type-specific means across six parameters that measured the horizontal movement area as well as the speed and acceleration of the head and upper torso. In order to find out the relation between each type within each parameter, we imported all the data into SPSS for statistical analysis and also investigated the data qualitatively on the basis of graphic descriptors.

In the following sections we explain in more detail the procedures we followed for collecting and processing the motion capture data, annotating the video material synchronized with the motion capture data, as well as analyzing the data both quantitatively and qualitatively.

3.1 motion capture data

The data for the present study were recorded in the Motion Capture Laboratory of the Department of Music, Arts and Culture Studies at the University of Jyväskylä in the winter of 2017. The recording was done using one Full HD video camera synchronized with eight infrared motion capture cameras, physically attached to the laboratory's ceiling in a symmetrical arrangement. The cameras were controlled using the Qualisys Motion Tracking system. The operating speed of the video camera was 30 fps and of the motion capture cameras 120 Hz (120fps). Altogether, the data in the present study comprise 10 minutes and 45 seconds of short, continuous stories (n=15) produced by five FinSL signers aged between 30 and 60, 3 of them male and 2 of them female. The stories were based on textless, three to four-box Ferd'nand comic strips in which the main character is involved in various funny situations either alone or with some other character. The signers were instructed to retell the contents of the comic strips in FinSL "as vividly as they can". The stories were monologues, and they were told to another native FinSL signer positioned in front of the performer. Both the active signer and the addressee were standing.

25 small ball-shaped reflective motion capture markers were attached to each (active) signer in carefully determined body locations: Head front left, Head front right, Head back left, Head back right, Chin, Shoulder left, Shoulder right, Chest, Back, Hip front left, Hip front right, Hip back left, Hip back right, Elbow left, Elbow right, Radius left, Ulnar left, Radius right, Ulnar right, (index) Finger (knuckle) left, Finger right, Knee left, Knee right, Foot left, and Foot right. A visualization of the physical marker locations is presented in Figure 3a.

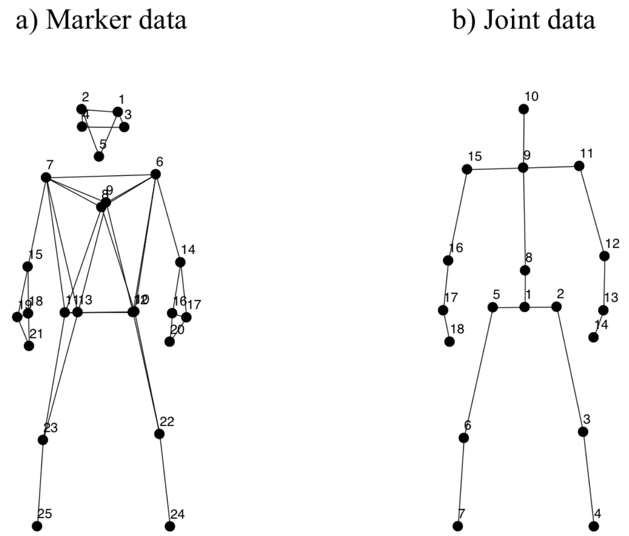

Figure 3. Visualization of physical marker locations (a) and derived joints (b).

The motion capture cameras recorded the three-dimensional (x, y, z) locations of the markers when the signers produced the stories. After the recordings, the data were first preprocessed in the Qualisys system by a member of the research team, who attached a body part identity to each marker. After this the data were transferred to MATLAB (Macintosh Intel version R2016b) software, in which any gaps that had occurred during the recording were filled. Gap fills were done by using a specific algorithm included in the MoCap [motion capture] Toolbox (ver. 1.5), developed by researchers in the Department of Music, Arts and Culture at the University of Jyväskylä for the purpose of kinematic analysis of motion capture data (Burger & Toiviainen 2013). In MATLAB, the data were also transformed into a joint representation, which in practice meant getting rid of redundant marker information by deriving computational joints on the basis of the marker locations. For example, a head joint is derived by calculating the centroid of the four upper-head marker locations. A visualization of joint locations is presented in Figure 3b.

Finally, all the numeric (joint) data were added to time codes and imported into ELAN (Max Planck Institute of Psycholinguistics, Nijmegen; see Crasborn & Sloetjes 2008) annotation software as track data, together with the video data. In ELAN, the numerical data and video were synchronized, annotated and further processed for the actual analysis.

3.2 Annotation

In ELAN, the data were annotated for signs, translations, constructed action and regular narration by a native FinSL signer, also fluent, for example, in Finnish. The annotation for signs and translations was done in Finnish and followed the conventions developed for the annotation of Corpus FinSL (see Salonen et al. 2018). In practice, signs were first tokenized – that is, the actual instances of signs on the video were identified and given a semantic gloss representing its contextual meaning – and then grouped into types identified by ID glosses (Johnston 2010) – that is, sign tokens were reanalyzed, according to their formal characteristics, into "lexical entries" similar to what are found in many dictionaries. The translation of the signing into Finnish was done on the sentence level meaning that the FinSL signing was first segmented into sentence-like utterances and then translated into Finnish by structurally corresponding Finnish sentences.

In the annotation of constructed action we followed the guidelines presented in Cormier et al. (2015). In these guidelines the point of departure is the identification of enacting articulators, which are annotated on independent tiers. We used altogether seven tiers, six of which were originally proposed by Cormier et al. (tiers for eye gaze, head, face, torso, and dominant and non-dominant hand). To these we added a tier for leg activity; the tier for legs was not included in Cormier et al. because their data only included signers who were sitting down. On the tiers we annotated when an articulator was enacting (symbol e).

In Cormier et al.'s (2015) constructed action annotation guidelines, the second step after the annotation of enacting articulators is determining the role of the signer (i.e. narrator or character) and its prominence (i.e. full, partial). As suggested by the guidelines, we did this by using two tiers: one for the primary role and the other for the secondary role. As in Cormier et al.'s data, our data too included a maximum of only two simultaneous roles per signer (these being always the narrator and a single charcter). It is possible for a signer to enact two or more character roles simultaneously but such instances did not occur in our data.

As explained in Cormier et al. (2015), the actual annotations for constructed action types (overt, reduced and subtle) emerged from the articulatory and role annotations. In overt constructed action there was always a relatively high number of enacting articulators and the primary role was always that of a character; there was no secondary role. In reduced constructed action the number of enacting articulators was slightly lower than in overt constructed action but not as low as in subtle constructed action; the primary role was that of a character and the secondary role that of the narrator. Finally, in subtle constructed action the number of enacting articulators was relatively low; the primary role was always that of the narrator and the secondary role that of a character.

If the signing did not include any constructed action – that is, the signing was regular narration – there were no annotations on articulatory tiers and the primary role was always that of the narrator; there was no secondary role assigned in these instances. In practice, we annotated regular narration automatically with ELAN's Create annotations from gaps function on the basis of the constructed action summary-tier, which identified the continuous stretches of discourse representing the same character.

A summary of the (non-numerical, see 3.3) tiers in our data is given in Table 1. Table 2 summarizes the criteria for determining the constructed action types and the lack of constructed action, according to Cormier et al. (2015). Finally, Figure 4 shows what the annotated material with visualized motion capture data looks like in ELAN.

Tier name	**Description (and annotation cell values)**
Translation	Sentence level translation in Finnish.
Gloss	A gloss identifying the sign in Finnish.
RN	Regular narration, i.e. no constructed action.
CA-type	The type of constructed action based on the annotations on the constructed action summary and role tiers (overt, reduced, subtle).
Role1	The primary role the signer is taking when using constructed action (narrator, "character").
Role2	The secondary role the signer is taking when using constructed action ("none", narrator, "character").
CA-summary	A stretch of discourse where constructed action is continuously used with one or more articulator to represent the same referent (i.e. within the same character role) (enacting).
CA-eyegaze	Break of eye gaze with addressee for purpose of enacting referent (enacting).
CA-head	Signer's use of his/her head to represent head movement/posture of referent (enacting).
CA-face	Signer's use of his/her facial expression to represent face of referent (enacting).
CA-torso	Signer's use of his/her torso to represent torso movement/posture of referent (enacting).
CA-dom-arm/hand	Signer's use of his/her dominant arm/hand to represent arm/hand of referent (enacting, instrument).
CA-ndom-arm/hand	Signer's use of his/her non-dominant arm/hand to represent arm/hand of referent (enacting, instrument).

CA-legs	Signer’s use of his/her legs to represent legs of referent (enacting).
Story	The duration of the story (story).

Table 1. Summary of ELAN tiers used in the annotation of signs, translations, constructed action (CA) and regular narration (RN). The annotation of constructed action is based on the guidelines presented in Cormier et al. (2015).

CA type	**Primary role**	**Secondary role**	**Description**
none	narrator	none	Narration with no elements of constructed action (i.e. regular narration). Native signer intuition: out of character.
overt	character	none	Clear use of constructed action (strong/many articulators), possibly simultaneous quotation of an utterance of the character. Native signer intuition: fully in character.
reduced	character	narrator	Some use of constructed action (use of articulators for constructed action between overt and subtle), possibly simultaneous quotation of an utterance of the character. Native signer intuition: mostly in character.
subtle	narrator	character	Some elements of constructed action (weak/few articulators), possibly simultaneous quotation of an utterance of the character. Native signer intuition: mostly out of character but also a bit in character.

Table 2. Constructed action (CA) types and lack of constructed action based on roles. Adapted from Cormier et al. (2015).

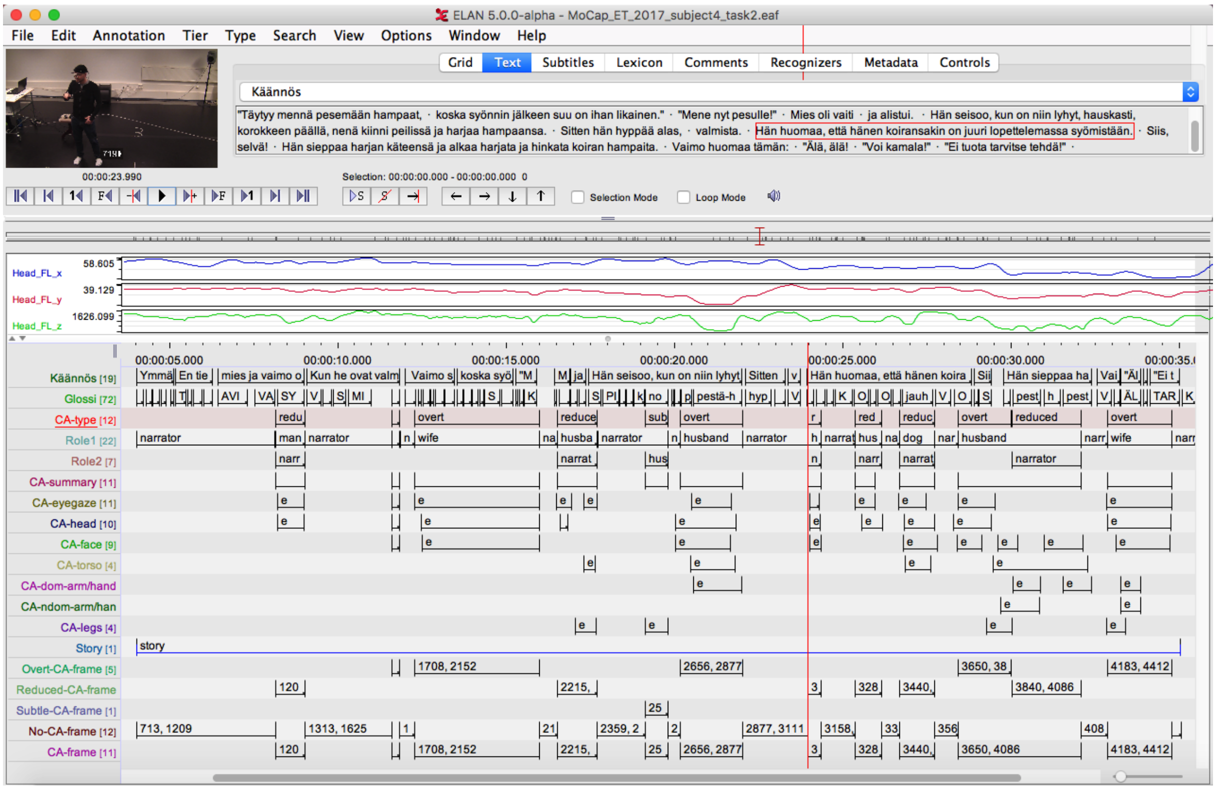

Figure 4. ELAN screenshot showing the visualized motion capture data and annotations.

3.3 Sampling, processing and analysis

After the annotation of constructed action and regular narration, we used ELAN's Extract track data function to associate each constructed action type and regular narration annotation cell with a beginning and ending frame number from the motion capture data. For this purpose we created four additional tiers labeled Overt-CA-frame, Reduced-CA-frame, Subtle-CA-frame and No-CA-frame (CA referring to constructed action). The extraction of this frame number information in ELAN was crucial because the use of frame numbers is the only way the annotated constructed action and regular narration sequences can be later referred to in MATLAB.

For the purpose of the actual analysis, we exported the frame number information of all of the constructed action and regular narration tokens into a tab-delimited text file. In doing this, we asked ELAN to add information about the duration of the annotation cell to each token. All this data was then imported into Excel and, on the basis of the annotation cell durations, divided into quartiles across the types of constructed action and regular narration. In order to exclude the durationally very short and very long tokens in each type, we deleted the information in the first and the fourth quartiles, which left us with a final sample of altogether 137 constructed action and regular narration tokens. The number of tokens in each type was as follows: overt constructed action n=28 (duration range 0,689–1,806 ms), reduced constructed action n=34 (duration range 0,698–1,591 ms), subtle constructed action n=19 (duration range 0,550–1,239 ms) and regular narration n=56 (duration range 0,796–3,175 ms). The sampling was done this way because we presumed that the information about the very short and very long annotation cells would skew the results.

The final sample was processed in MATLAB. First, by using the MoCap Toolbox, we reduced the amount of numerical motion capture data by reading in only the information concerning the head and the upper-torso joint (numbers 10 and 9 in Figure 3b). After this, we calculated six types of information for each constructed action and regular narration token: the horizontal movement area of the head, the horizontal movement area of the torso, the speed of the head movement, the speed of the torso movement, the acceleration of the head movement and the acceleration of the torso movement. All this information was exported into an Excel table, where it was also collapsed for type-specific means for the purpose of presenting the results as simply and clearly as possible.

The calculated kinematic information was decided on the basis of a previous study by one of the authors (Burger et al. 2013). In practice, the calculation exploited two existing functions available in the MoCap Toolbox: *mcboundrect* and *mctimeder*. The first function calculates the bounding rectangle that is the smallest rectangular area that contains the projection of the trajectory of a selected marker on the horizontal plane, that is, the floor (Burger & Toiviainen 2013: 76). The second function estimates time derivatives (i.e. velocity, acceleration) of motion capture data by using differences between two successive frames and a Butterworth smoothing filter (id., 139). In the present study, the second function (i.e. mctimeder) was followed by the use of yet another MoCap Toolbox function, *mcnorm*, which converts the three-dimensional motion capture data structure into a unidimensional norm structure by calculating the Euclidean norms of the kinematic vectors (Burger & Toiviainen 2013: 110). For example, with velocity this means the magnitude of the velocity vector, that is, speed.

The data were analyzed both quantitatively and qualitatively. Statistical analysis was conducted in SPSS with the Independent-Samples Kruskal-Wallis test with the significance level of .05. The dependent variables were the types of information calculated in MATLAB and the independent variables were the constructed action and regular narration types. Qualitative analysis relied on the visual observation of the Excel-based visual descriptors.

4. Results

The results of the MATLAB calculations are presented in Table 3 as type-specific means across the investigated kinematic variables. The Kruskal-Wallis test indicated significant differences (p=<.05) between constructed action types and regular narration with respect to all of the variables except those measuring the speed and acceleration of the upper torso. The four populations with significant differences are shown as graphical descriptors in Figure 5. The figure also shows the results of pairwise comparisons targeted at these populations, that is, between what types the significant differences actually existed.

	Horizontal movement area of the head (m^2)	**Horizontal movement area of the upper torso (m^2)**	**Speed of the head movement (mms)**	**Speed of the upper torso (mms)**	**Acceleration of the head movement (mms^2)**	**Acceleration of the upper-torso mov. (mms^2)**
Regular narration	0.0014	0.0007	127.0	93.2	1085.8	986.3
Subtle constructed action	0.0011	0.0006	136.2	88.1	1174.0	886.6
Reduced constructed action	0.0048	0.0023	213.8	133.6	1782.1	1160.1
Overt constructed action	0.0040	0.0027	237.7	165.3	1877.5	1561.8

Table 3. The results of the MATLAB calculations as type-specific means across the six investigated kinematic variables.

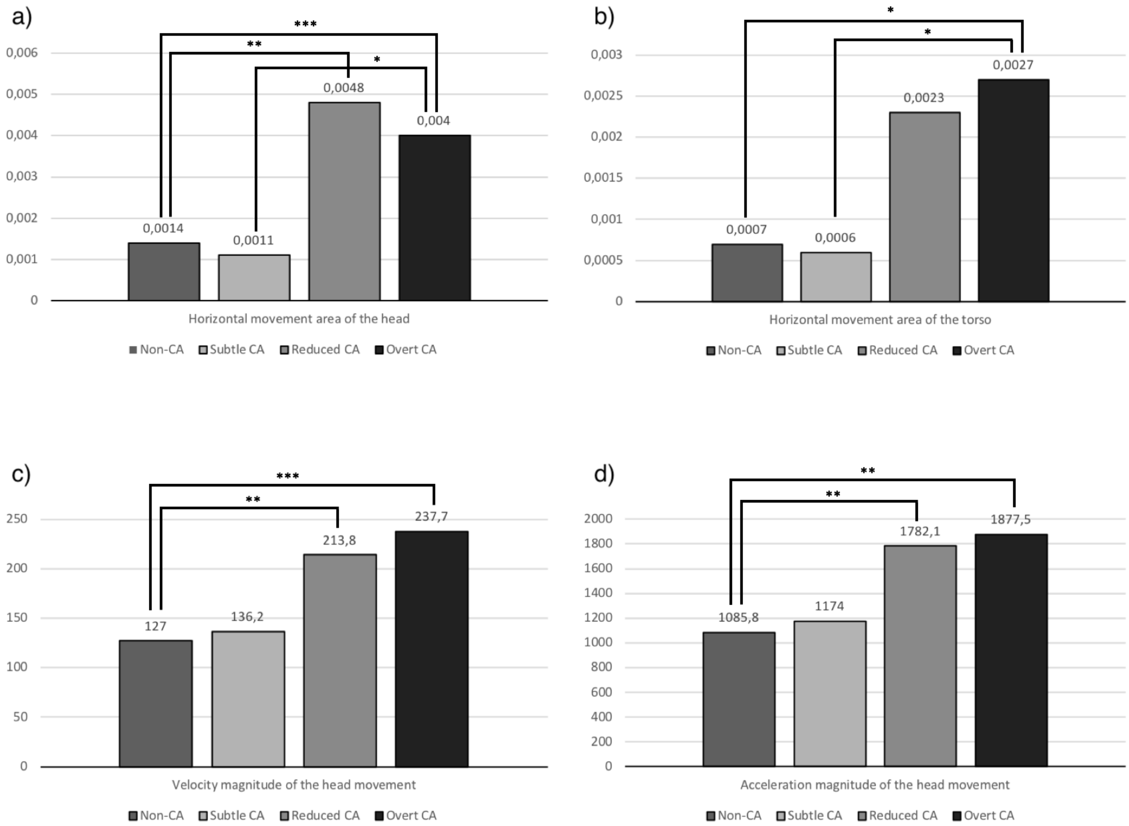

Figure 5. Summary of the significant ($p<.05$) results of the Independent-Samples Kruskal-Wallis test. In the legends, constructed action is labelled CA and regular narration is labeled non-CA.

The statistical results indicate two things. First, looking at a) and b) of Figure 5, we can say that the head and the upper torso move on a larger area with respect to the floor in reduced and overt constructed action than in regular narration and subtle constructed action. Second, looking at c) and d) of Figure 5, we can generalize that the movements of the head are faster and more rapid in reduced and overt constructed action than in regular narration and subtle constructed action.

When considering also the qualitative dimension, we can make two further generalizations. First, looking at the results as a whole, we can say that significant differences are found only between the extremities, that is, between regular narration and reduced/overt constructed action, or between subtle constructed action and overt constructed action. In other words, in terms of the investigated parameters, subtle constructed action is not significantly different from regular narration, nor is reduced constructed action significantly different from overt and subtle constructed action. Second, the results also indicate that the size of the movement area of the head and the upper torso, as well as the speed and acceleration of the head, tend to correlate positively with the amount of constructed action: the more constructed action there is, the larger tends to be the horizontal area on which the head and the torso move, and the faster and more rapid are the head movements.

5. Discussion

As stated in Section 2, the characteristics of the movements of the head and the upper torso in constructed action and regular narration have not been properly studied before. Our previous work with corpus data (Jantunen 2017; Puupponen 2018) showed that narration with constructed action tends to involve relatively more activity of the whole body, while narration without constructed action tends to be associated with increased activity of the head. The positive correlation between constructed action and the whole body movement is related to the fact that the body of the character being enacted is moving too and the signer wants to show this activity. The fact that regular narration is associated with increased activity of the head, in turn, tells that the head has more functions to fulfill in this type of signing compared to the body as a whole: these functions range from affirmative and negative ones to boundary and countour signalling ones (Puupponen 2019).

In terms of the torso movements, the results of the present study both corroborate and further explicate these previous results from the kinematic perspective: the fact that more constructed action associates with more activity of the whole body means, in the light of the current work, that the upper torso moves on a larger horizontal area. The fact that also the horizontal movement area of the head is larger and the speed and acceleration of head movements is faster and more rapid in overt constructed action than in regular narration may be linked to this previous corpus-based finding, too: in constructed action the head and the upper torso tend to move as a single unit, enabled by the anatomical connection between the two (Puupponen 2018).

The previous claim that narration without constructed action is associated with increased activity of the head is neither supported nor refuted by the present study. In order to address the claim from the kinematic perspective, different measures focusing, for example, on head rotation with respect to x, y and z axes would need to be developed.

The main objective of the present work was to evaluate two arguments concerning the nature of and relationship between constructed action and regular narration. These were (i), that constructed action forms a continuum with regular narration, so that the articulatory border between the two is fuzzy and never clear cut (e.g. Cormier et al. 2015; Jantunen 2017), and (ii), that constructed action is divisible internally into three subtypes (i.e. overt, reduced and subtle constructed action) (e.g. Cormier et al. 2015).

Concerning (i), we interpret the findings as supporting the proposed continuum-like, relatively fuzzy relationship between constructed action and regular narration. On the one hand, this interpretation is based on the fact that significant differences between the constructed action types and regular narration were found only between the extremities (i.e. between regular narration and reduced/overt constructed action, and subtle constructed action and overt constructed action; see Figure 5). On the other hand, the interpretation relies on the fact that the change from regular narration to overt constructed action in the speed and acceleration parameters is relatively gradual (i.e. the speed and acceleration of the head and the upper torso correlate positively with the amount of constructed action; see Table 3 and Figure 5c-d). However, this is not the case with the horizontal movement area of the head and the upper torso, although with the horizontal movement are of the upper torso (Figure 5b) the correlation is close to perfect. The fact that regular narration tends to be associated with increased activity of the head (see the discussion above) may be an explaining factor here. Further research on the topic is needed.

By extension, we also interpret the results as providing support for the ontological view (see Section 1) that gestural features such as gradience and unconventionality as well as, for example, depiction are an inseparable part of language. If constructed action forms a continuum with regular narration (as argued for from the kinematic perspective above), then – according to the definition of constructed action and regular narration (see Section 2) – the unconventional, gradient and depictive characteristics of utterances must also form a continuum with (i.e. cannot be separated from) their conventional, discrete and descriptive characteristics – characteristics which traditionally have been considered to form language. This is in no way an exceptional view in linguistics (see Kendon 2004; Enfield 2009; Jantunen 2017; Ferrara & Hodge 2018), and it was chosen as the ontological starting point also in the present study. However, as far as we are aware of, the fuzzy border between conventional, discrete and descriptive activity ("linguistic" activity in traditional terms) and unconventional, gradient and depictive activity ("gestural" and other everyday activity) has not been documented up to this precission before – if at all. In the future, investigating this continuum further with technology developed for spoken languages would be fruitful.

Concerning (ii), the interpretation of the results is this time not so straightforward. Obviously constructed action *can* be divided into three prototypes, and in the literature this view is relatively well established – the types of overt, reduced and subtle constructed action are existing categories especially from a methodological point of view (see Cormier et al. 2015), and the results of, for example, our eye-tracking study (Jantunen et al. 2018; see Section 2) seem to support the typology. However, the present study leads us to conclude that the validity of the three-part typology is nevertheless open to doubt. The reason for this is the ambiguous role of reduced constructed action in between subtle and overt constructed action.

As regards the role of reduced constructed action, our results show two things. First, as indicated by the statistical analysis, reduced constructed action is not distinguished significantly from its neighbors, subtle and overt constructed action: the kinematic characteristics of the head and the upper torso investigated here spread over all three constructed action types as if on a continuum. Second, as shown by the means in Figure 5 (as well as those in Table 3), reduced constructed action actually resembles overt constructed action in many respects: the two types may be distinct semiotically (see Cormier et al. 2015 who define the two types by referring to the presence/non-presence of lexical units), but with regard to the kinematic characteristics of head and upper-torso movements investigated here, drawing a clear line between the two may be impossible.

Our interpretation of the data is that they do not fully support the current three-part typology of constructed action, but instead one with only two subtypes. We propose that for methodological purposes these two types could be called strong constructed action and weak constructed action (for an overview of other two-part typologies of constructed action, see Cormier et al. 2015). Definitionally, the two types would be based on distributing the characteristics of reduced constructed action to both overt and subtle constructed action. Following this logic, strong constructed action would comprise the features of overt constructed action and some of those associated with reduced constructed action, while weak constructed action would comprise the remaining features of reduced constructed action as well as those associated with subtle constructed action. In practice, the border between the two types would be gradient, too.

We want to emphasize that the proposed two-part typology of constructed action does not need to be seen as replacing the three-part typology. Instead, it is suggested that it is an alternative way of subcategorizing the phenomenon of constructed action from a kinematic perspective. Such a possibility would give the researcher more flexibility in terms of choosing either a semiotic or a visual (kinematic) approach to the data. Adjusting and playing with the types is made possible even by Cormier et al. (2015) in their original paper: for example, they occasionally classify subtle and reduced constructed action as non-overt constructed action.

Figure 6 summarizes our proposal for the two-part typology of constructed action and contrasts it with the three-part typology proposed by Cormier et al. (2015). The figure places the types and their compositional semiotic and kinematic features on a continuum that ranges all the way from strong/overt constructed action to regular narration.

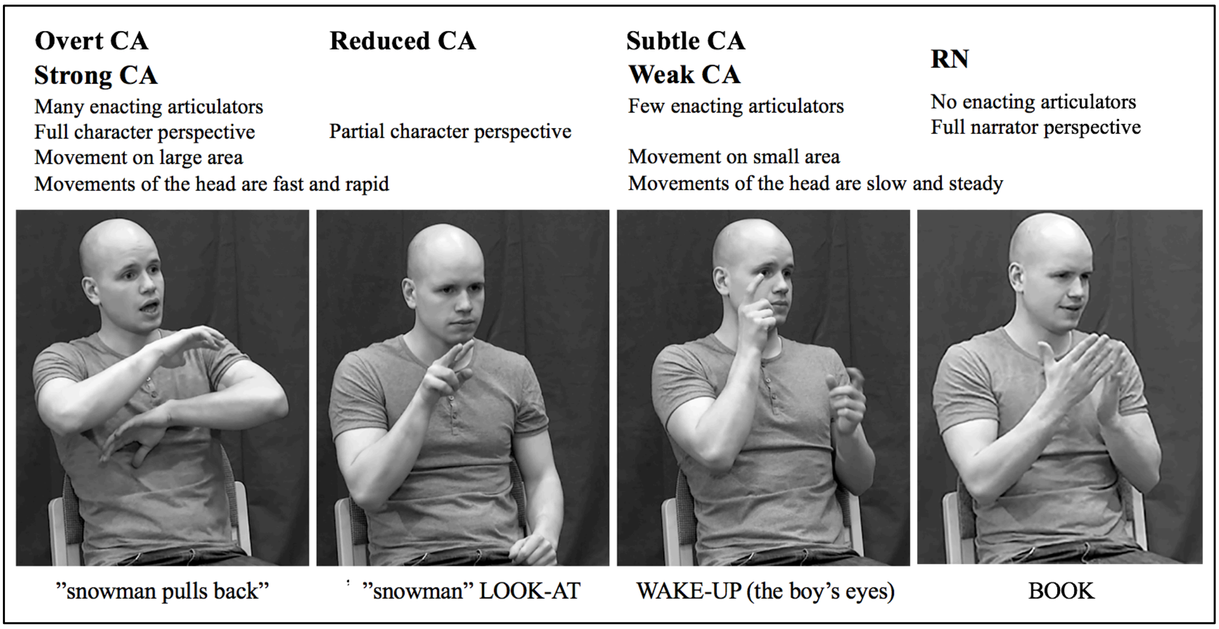

Figure 6. The continuum of discourse strategies with semiotic and kinematic features. Constructed action is labelled CA, regular narration RN. The video underlying the still images is part of the open access ProGram data and can be accessed online with annotations at http://hdl.handle.net/11113/00-0000-0000-0000-32DB-8@view.

In terms of video duration and the number of signers (i.e. less than 11 minutes of signing by only five signers), the data set is very small and is well below the norm of video duration and number of signers in any modern studies exploiting sign language corpora (i.e. several hours of signing by more than ten signers). However, when considering the size of the data and, consequently, the generalizability of the results, two things need to be remembered. First, the present work exploits motion capture data and the size of the numerical material is approximately half a billion characters. This makes the present data extremely deep, that is, it includes an enormous amount of information. Second, the motion capture data does not comprise isolated utterances but continuous signing. The use of continuous data is still very rare in motion capture studies on sign languages, the obvious reason for this being the challenges in the management of the data. All in all, the size of the data and the generalizability of the results are a compromise necessitated by the chosen approach.

The data for the present study came from only one sign language, FinSL. However, as discussed in Section 2, the fact that the relative share of constructed action has been found to be very similar in narratives of different sign languages encourages the belief that the findings are generalizable across sign languages, and are likely to shed light also to the enactment phenomena in spoken languages. Obviously, in order to address the generalizability issue specifically, a comparative study would need to be organized.

6. Conclusion

This paper has combined the corpus approach with motion capture data in order to investigate the kinematic characteristics of head and upper-torso movements in constructed action and regular narration in FinSL. Specifically, we have evaluated the validity of two arguments concerning constructed action, namely (i), that it forms a continuum with regular narration and (ii), that it divides into three subtypes. The results presented in the paper support the first argument and, by extension, also the ontological view according to which features such as gradience and unconventionality – that is, gestural features – are an inseparable part of language, together with features such as discreteness and conventionality. However, the second argument is not fully supported. Because of the ambiguous position of reduced constructed action between subtle and overt constructed action – for example, that it is not significantly distinguished from the two in terms of the horizontal movement area of the head and torso or the speed and acceleration of head movements – we have argued in the paper that the present three-part typology of constructed action may need some revision. As an alternative way of subcategorizing constructed action, we have proposed a division into strong constructed action and weak constructed action.

This study has been based on a conceptualization of language that encourages researchers to bring together data from different sources and to analyze data with a variety of methods. In the future we would like to see more studies on sign languages build on a similar setup. We also believe there is a need to apply modern technological resources to a more detailed exploration of the conventionality–discreteness/unconventionality–gradience continuum in spoken languages as well.

Acknowledgements

The authors wish to thank collectively all the wonderful people who have participated in this research. The financial support of the Academy of Finland under grants 269089 & 304034 (Jantunen) and 299067 (Burger) is gratefully acknowledged.

References

Burger, Birgitta, Saarikallio, Suvi, Luck, Geoff, Thompson, Marc & Toiviainen, Petri (2013). Relationships between perceived emotions in music and music-induced movement. *Music Perception* 30:5, 519–535.

Burger, Birgitta & Toiviainen, Petri (2013). MoCap Toolbox: A Matlab toolbox for computational analysis of movement data. In Roberto Bresin (ed.), *Proceedings of the 10th Sound and Music Computing Conference*, 172–178. Stockholm, Sweden.

Burger, Birgitta, Puupponen, Anna & Jantunen, Tommi (2018). Synchronizing eye tracking and optical motion capture: How to bring them together. *Journal of Eye Movement Research* 11:2, 16.

Cormier, Kearsy, Smith, Sandra & Sevcikova Sehyr, Zed (2015). Rethinking constructed action. *Sign Language & Linguistics* 18:2, 167–204.

Crasborn, Onno & Sloetjes, Han (2008). Enhanced ELAN functionality for sign language corpora. In Onno Crasborn, Elene Efthimiou, Thomas Hanke, Ernst D. Thoutenhoofd & Inge Zwitserlood (eds.), *Proceedings of the 3rd Workshop on the Representation and Processing of Sign Languages: Construction and Exploitation of Sign Language Corpora*, 39–43. Paris: ELRA.

De Beuzeville, Louise, Johnston, Trevor & Schembri, Adam (2009). The use of space with indicating verbs in Australian Sign Language: A corpus-based investigation. *Sign Language & Linguistics* 12:1, 52–83

Enfield, Nick (2009). *The anatomy of meaning: Speech, gesture, and composite utterances*. Cambridge: CUP.

Engberg-Pedersen, Elisabeth (2015). Perspective in signed discourse: The privileged status of the signer's locus and gaze. *Open Linguistics* 1, 411–431.

Ferrara, Lindsay & Johnston, Trevor (2014). Elaborating who's what: A study of constructed action and clause structure in Auslan (Australian Sign Language). *Australian Journal of Linguistics* 34:2, 193–215.

Ferrara, Lindsay & Hodge, Gabrielle (2018). Language as Description, Indication, and Depiction. *Front. Psychol*. 9:716.

Herrmann, Annika & Steinbach, Markus (2012). Quotation in sign languages: A visible context shift. In Ingrid van Alphen & Isabelle Buchstaller (eds.), *Quotatives: Cross-linguistic and cross-disciplinary perspectives*, 203–230. Amsterdam: John Benjamins.

Hodge, Gabrielle & Cormier, Kearsy (2019). Reported speech as enactment. *Linguistic Typology* 23:1, 185–196.

Hodge, Gabrielle & Ferrara, Lindsay (2013). Showing the story: Enactment as performance in Auslan narratives. In Lauren Gawne & Jill Vaughan (eds.), *Selected Papers from the 44th Conference of the Australian Linguistic Society*, 372-397.

Jantunen, Tommi (2013). Signs and transitions: Do they differ phonetically and does it matter? *Sign Language Studies* 13:2, 211–237.

Jantunen, Tommi (2017). Constructed action, the clause and the nature of syntax in Finnish Sign Language. *Open Linguistics* 3, 65–85.

Jantunen, Tommi (2018). Viittomakielet hybridisysteemeinä: hämärärajaisuus ja epäkonventionaalisuus osana viittomakielten rakennetta [Sign languages as hybrid systems: Gradience and unconventionality as a part of the structure of sign languages]. *Puhe ja kieli* 38:3, 109–126.

Jantunen, Tommi, Burger, Birgitta & Puupponen, Anna (2018). Constructed action types and eye behavior in Finnish Sign Language. Paper presented at the *8th Conference of the International Society for Gesture Studies* (ISGS 8). Cape Town, South Africa, July 4–8, 2018.

Jantunen, Tommi, Puupponen, Anna & Burger, Birgitta (2020). What comes first: Combining motion capture and eye tracking data to study the order of articulators in constructed action in sign language narratives. In *Proceedings of the 12th International Conference on Language Resources and Evaluation (LREC 2020)*, 6003–6007. Paris: ELRA.

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

Kendon, Adam (2004). *Gesture: Visible action as utterance*. Cambridge: CUP.

Lillo-Martin, Dianne (2012). Utterance reports and constructed action. In Roland Pfau, Markus Steinbach & Bencie Woll (eds.), *Sign language: An international handbook*, 365–387. Berlin: De Gruyter Mouton.

Puupponen, Anna (2018). The relationship between the movements and positions of the head and the torso in Finnish Sign Language. *Sign Language Studies* 18:2, 175–214.

Puupponen, Anna (2019). Towards understanding nonmanuality: A semiotic treatment of signers' head movements. Glossa: A Journal of General Linguistics 4:1, 39.

Puupponen, Anna, Wainio, Tuija, Burger, Birgitta & Jantunen, Tommi (2015). Head movements in Finnish Sign Language on the basis of Motion Capture data: A study of the form and function of nods, nodding, head thrusts, and head pulls. *Sign Language & Linguistics* 18:1, 41–89.

Salonen, Juhana, Wainio, Tuija, Kronqvist, Antti & Keränen, Jarkko (2018). *Suomen viittomakielten korpusprojektin (CFINSL) annotointiohjeet.* 1. versio. Kieli- ja viestintätieteiden laitos, Jyväskylän yliopisto.

Tyrone, Martha & Mauk, Claude (2010). Sign lowering and phonetic reduction in American Sign Language. *Journal of Phonetics* 38, 317–328.

Author information

Tommi Jantunen: University of Jyväskylä, Department of Language and Communication Studies / Sign Language Centre, P.O. Box 35, 40014 University of Jyväskylä, Finland, E-mail: tommi.j.jantunen@jyu.fi

Tommi Jantunen is Professor of Finnish Sign Language at the Sign Language Centre in University of Jyväskylä, Department of Language and Communication Studies, and in his research he investigates all aspects of FinSL grammar and phonetics as well as sign language technology.

Danny De Weerdt: University of Jyväskylä, Department of Language and Communication Studies / Sign Language Centre, P.O. Box 35, 40014 University of Jyväskylä, Finland, E-mail: danny.deweerdt@jyu.fi

Danny De Weerdt is University Teacher and PhD student at the Sign Language Centre in University of Jyväskylä, Department of Language and Communication Studies, and in his doctoral research he investigates ways of expressing spatial relationships in Finnish Sign Language and Flemish Sign Language.

Birgitta Burger: University of Jyväskylä, Department of Music, Art and Culture Studies, P.O. Box 35, 40014 University of Jyväskylä, Finland, E-mail: birgitta.burger@jyu.fi

Birgitta Burger is Adjunct Professor at the University of Jyväskylä, Department of Music, Art and Culture Studies, and a postdoctoral researcher at the Institute for Systematic Musicology at the University of Hamburg.

Anna Puupponen: University of Jyväskylä, Department of Language and Communication Studies / Sign Language Centre, P.O. Box 35, 40014 University of Jyväskylä, Finland, E-mail: anna.m.puupponen@jyu.fi

Anna Puupponen is Postdoctoral Researcher at the Sign Language Centre in University of Jyväskylä, Department of Language and Communication Studies, and in her research she studies sign language depiction, multimodality of signers' articulation and semiotic strategies in signed communication.