TOPICS IN COGNITIVE SCIENCE

Topics in Cognitive Science 7 (2015) 36–60

ISSN:1756-8757 print / 1756-8765 online
DOI: 10.1111/tops.12122

Visible Cohesion: A Comparison of Reference Tracking in Sign, Speech, and Co-Speech Gesture

Pamela Perniss,[a] Asli Özyürek[b]

[a]*School of Humanities, University of Brighton*
[b]*Radboud University Nijmegen & MPI for Psycholinguistics*

Received 28 February 2013; received in revised form 1 May 2014; accepted 20 July 2014

Abstract

Establishing and maintaining reference is a crucial part of discourse. In spoken languages, differential linguistic devices mark referents occurring in different referential contexts, that is, introduction, maintenance, and re-introduction contexts. Speakers using gestures as well as users of sign languages have also been shown to mark referents differentially depending on the referential context. This article investigates the modality-specific contribution of the visual modality in marking referential context by providing a direct comparison between sign language (German Sign Language; DGS) and co-speech gesture with speech (German) in elicited narratives. Across all forms of expression, we find that referents in subject position are referred to with more marking material in re-introduction contexts compared to maintenance contexts. Furthermore, we find that spatial modification is used as a modality-specific strategy in both DGS and German co-speech gesture, and that the configuration of referent locations in sign space and gesture space corresponds in an iconic and consistent way to the locations of referents in the narrated event. However, we find that spatial modification is used in different ways for marking re-introduction and maintenance contexts in DGS and German co-speech gesture. The findings are discussed in relation to the unique contribution of the visual modality to reference tracking in discourse when it is used in a unimodal system with full linguistic structure (i.e., as in sign) versus in a bimodal system that is a composite of speech and gesture.

Keywords: Reference tracking; Co-speech gesture; Sign language; Visual modality; Use of space; Pointing

1. Introduction

Establishing and maintaining reference to discourse entities is a crucial component of successful communication and necessary to achieving discourse cohesion. Studies across

Correspondence should be sent to Pamela Perniss, School of Humanities, University of Brighton, Brighton BN1 9PH, UK. E-mail: p.perniss@brighton.ac.uk

a range of spoken languages have shown that speakers consistently vary the quantity of marking material in referring expressions—choosing between noun phrases (e.g., *the man*), pronouns (e.g., *he*), and zero anaphora (*Ø*)—according to referent accessibility and the referential context of the clause (i.e., introduction, maintenance, or re-introduction) (Ariel, 1994; Chafe, 1994; Givón, 1984). However, our communicative ability to track referents in discourse is not specific to speech. Recent research has shown that discourse cohesion is also established in systematic ways in the visual modality, that is, in the use of co-speech gestures as well as in sign languages (Gullberg, 2006; McKee, Schembri, McKee, & Johnston, 2011). Yet compared to the vast amount of research on this topic in spoken languages, we know little about visual means of tracking referents and about how modality-specific features are exploited in maintaining referential cohesion. This paper investigates reference tracking in the visual modality, as it is used in sign language (in a unimodal system) and in gesture (integrated with speech in a bimodal system). In order to fully understand how gestures are recruited for discourse cohesion, we also include speech in our investigation.

Like referring expressions in speech, co-speech gestures, the movements of the hands that accompany speech, have been shown to be sensitive to referential context in terms of the quantity of marking material (Levy & Fowler, 2000; McNeill & Levy, 1993). Specifically, gestures are more likely to occur when referents are introduced or re-introduced into discourse, and are less likely to occur when a referent is maintained across consecutive clauses. For sign languages, the natural languages used by deaf communities around the world, studies have found that overt expression of referents in subject position is more likely to occur in contexts of switch reference, that is, re-introduction, compared to maintained reference, where null subjects are more likely (McKee et al., 2011; Wulf, Dudis, Bayley, & Lucas, 2002). Thus, the quantity of marking material varies as a function of referential context in both the vocal and visual modalities.

What, then, are the modality-specific features that sign and co-speech gesture contribute to reference tracking in discourse? One major feature is the availability of the space in front of and around the body as a visual-spatial medium for articulation. The affordance of spatial modification allows referents to be associated with specific locations in space, corresponding, for example, to referent locations in the speaker's conceptualization of a discourse event (Liddell, 2003; McNeill, 1992). For co-speech gesture, there is some evidence that the spatial affordances of the visual modality are exploited for marking referential context. In contexts of referent introduction and re-introduction, co-speech gestures accompanying referring expressions are more likely to be spatially modified, that is, produced at a particular location in space, than gestures occurring with referring expressions in contexts of referent maintenance (Gullberg, 2006).

In sign languages, spatial modification is an integral part of grammatical structure. In particular, pronominal reference and many kinds of predicates rely on spatial modification and allow the creation of referent-location associations that may be referred back to once established (Engberg-Pedersen, 1993; Liddell, 1996; Lillo-Martin & Klima, 1990; Senghas & Coppola, 2001; van Hoek, 1996). Pronominal reference and spatial modification of predicates in sign space are both highly relevant to reference tracking, as we will see.

The use of space on the level of discourse structure has been investigated in terms of achieving discourse cohesion through spatial mapping, where related discourse themes are mapped onto the same area of space (Mather & Winston, 1998; Winston, 1995). There has been little systematic investigation, however, of whether and how the use of spatial modification—as a primary affordance of the visual modality—is exploited on the level of discourse structure to mark referential context and referent accessibility.

Moreover, there has been no direct and systematic comparison of sign language and co-speech gesture in this domain.[1] Such a comparison is necessary because the two forms of expression share access to the affordances of the visual modality and have been assumed to use spatial modification in similar ways (Liddell, 1996, 2003; Schembri, Jones, & Burnham, 2005). This similarity may extend to the use of space for reference tracking. Furthermore, based on assumed similarities, both gesture and sign uses of space have been claimed to be generated directly from a common origin, namely signers'/speakers' mental imagery of event spaces. However, sign and gesture are notably different in communicative function—integrated with speech in the case of co-speech gesture, while exhibiting full linguistic structure in the case of sign languages—and thus may exhibit significant differences in the use of space for purposes of reference tracking.

The present study investigates how referential context is reflected in the use of space in the visual modality through a direct comparison of co-speech gesture and sign language, specifically comparing narratives produced in German and German Sign Language (*Deutsche Gebärdensprache*, DGS). We also analyse the speech produced in these contexts. Co-speech gesture occurs in tight semantic and temporal integration with speech and needs to be considered and interpreted as part of the speech signal—here, as part of the referential functions performed by speech. Before describing our study in detail, we first provide additional background, summarizing previous research on reference tracking in speech, co-speech gesture, and sign language.

2. Background

2.1. Reference tracking in speech

We know from research in spoken languages that, in discourse, speakers vary the linguistic means by which referents are referred to—choosing between full NPs, pronouns, and zero anaphora—according to the referential context and the accessibility of referents. Specifically, speakers choose a fuller referring expression when introducing (first mention) or re-introducing (subsequent mention) a referent into discourse, corresponding to lower accessibility of the referent in the mind of the addressee (Ariel, 1990), and choose less full referring expressions when reference to an entity is maintained (immediate subsequent mention), and the referent is highly accessible. Highly accessible referents require little linguistic marking in order to be correctly identified by the addressee (as exemplified by the pronoun subject in line b and the null subject in line d of example 1 below),

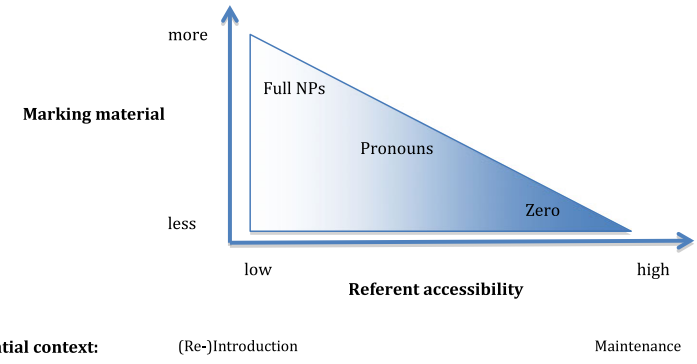

Fig. 1. Schematic representation of the *Principle of Quantity* for topic continuity.

while more linguistic marking is required when the referent is new or less accessible (as indicated by the full NPs in lines a, c, and d of example 1).

(1)
a. [*A man*$_1$] goes into a store. [Intro$_1$]
b. [*He*$_1$] wants milk and eggs. [Maint$_1$]
c. [*The store clerk*$_2$] points to aisle 3. [Intro$_2$]
d. [*The man*$_1$] smiles and [ϕ_1] heads there. [Re-intro$_1$][Maint$_1$]

These correspondences between linguistic device, referent accessibility, and referential context conform to the *Principle of Quantity* for topic continuity (Givón, 1984; see Fig. 1) and have been shown to hold across typologically different spoken languages, including pro-drop and non-pro-drop languages (Cameron, 1992, 1998; Hickmann & Hendriks, 1999; Yoshioka, 2008).

2.2. *Reference tracking in co-speech gesture*

Systematic correspondences between the amount of marking material and referent accessibility have been shown to exist not only in the choice of linguistic referring expression but also in the behavior of gestures that accompany these linguistic forms in speech. Gestures are more likely to co-occur with referring expressions used to introduce or re-introduce referents into discourse (low referent accessibility) and are less likely to occur with referring expressions used for maintenance of a referent across consecutive clauses (high referent accessibility) (Levy & Fowler, 2000; McNeill & Levy, 1993). In this way, co-speech gestures reflect the *Principle of Quantity* in parallel with speech (McNeill, 1992). More marking material in gesture (i.e., the presence of gesture) corresponds to more marking material (i.e., fuller referring expressions) in speech; less marking material in gesture (i.e., the absence of gesture) corresponds to less marking material (i.e., leaner referring expressions) in speech. This is consistent with other findings showing that gestures parallel patterns in speech in terms of referent specification (So, Kita, & Goldin-Meadow, 2009). In

the study by So et al. (2009), when referents were uniquely identified in speech, they were also identified in gesture. Conversely, when referents were underspecified in speech, with reference left ambiguous, gesture did not compensate for this underspecification by supplying disambiguating deictic or iconic information about the referent.

Co-speech gestures have also been found to exploit the spatial affordances of the visual modality to reflect referential context. Specifically, gestures accompanying reference to entities in contexts of referent introduction or re-introduction are more likely to be spatially modified, such that a referent gets associated with a particular location in space, than gestures accompanying referring expressions in contexts of referent maintenance (Gullberg, 2006). Furthermore, over a stretch of discourse, gesturers exhibit consistency in their use of spatial modification, that is, often using the same location when the associated referent is re-introduced into discourse (Gullberg, 2006; So et al., 2005, 2009). In this way, co-speech gestures create spatial anaphoric linkages (i.e., provide visible cohesion) that support the anaphoric linkages established in speech.

Previous research on the behavior of co-speech gestures with respect to reference tracking has focused only on gestures accompanying (or not) the referential form in speech identifying the subject of a clause. However, there has been almost no research investigating the influence of referential context on gestures accompanying predicates in speech—that is, the verbs or other forms (e.g., locatives or adjectives) used to predicate information about the referents of referring expressions. (An exception is Debrelioska, Özyürek, Gullberg, & Perniss, 2013, which looks at the influence of referential context on the use of viewpoint in action and motion gestures.) It is thus an open question, and one that we address in the current study, whether spatial modification occurs with gestures accompanying predicates, for example, gestures depicting referent actions like *stirring*, *twisting the lid off a jar*, etc., and whether such spatial modification contributes to discourse cohesion by marking referential context and referent accessibility.

2.3. Reference tracking in sign language

Compared to the wealth of research on the sensitivity of reference tracking devices to referential context in spoken languages and even gestures, sign languages have received little attention. Only few previous studies have directly addressed the presence versus absence of verb arguments in terms of reference tracking in discourse in sign languages (McKee et al., 2011 for New Zealand Sign Language and Australian Sign Language; Lucas, Bayley, & Valli, 2001; Wulf et al., 2002 for American Sign Language). More generally, the licensing of null arguments has been related to verb morphology in the sign language literature (de Quadros & Lillo-Martin, 2010; Lillo-Martin, 1986, 1991). In *directional* (or *agreement*) verbs (e.g., the sign for "give" in DGS), null arguments are licensed through spatial modification: the verb moves between locations in space previously associated with the subject and/or object arguments of the verb. Verbs that cannot be spatially modified in this way are called *plain* verbs (e.g., the sign for "like" in DGS), and have been said to license null arguments through topic-hood and discourse continuity (Lillo-Martin, 1986), an analysis that implicitly invokes the notion of referential context.[2]

Due to this propensity for null arguments especially in the class of *directional* verbs, American Sign Language (ASL) has been described as a pro-drop language (Lillo-Martin, 1991). This analysis may be extended typologically to sign languages in general, given that nearly all sign languages studied to date have been shown to exhibit similar verb class structure (Lillo-Martin & Meier, 2011; Padden, 1990). Little is known, however, about how the properties of different verb types are exploited for discourse cohesion.

From the studies that exist on reference tracking in discourse in sign languages (McKee et al., 2011; Wulf et al., 2002), we know that signers, like speakers, consistently vary the use of referring expression (i.e., choosing between full NPs, pronouns, and zero anaphora) according to the referential context of a clause. These studies have thus similarly provided evidence for the *Principle of Quantity*, finding that subject arguments are more likely to be overtly realized in contexts of switched reference, while null subjects are more likely to occur when reference is maintained.

As with co-speech gesture, the research on reference tracking in sign languages has focused on referring expressions, and on identifying whether the subject of a clause is overtly expressed or not. Moreover, even though the spatial modification of signs has been widely studied for morphosyntactic purposes, as mentioned above (Klima & Bellugi, 1979; Liddell, 2003; Sandler & Lillo-Martin, 2006), there has been no investigation of the spatial behavior of predicates (e.g., different verb types) in sign languages with respect to reference tracking. Such an investigation, however, is important for a full understanding of how modality-specific features are recruited in the service of referential cohesion on a discourse level.

3. The present study

The aim of the present study is to investigate reference tracking in the visual modality by directly comparing sign language (DGS) and co-speech gesture (with German), as well as the speech context in which gestures are embedded. We first compare the use of referring expressions across all three forms of expression—sign, speech, and co-speech gesture—seeking to replicate previous findings regarding the quantity of marking material. We then focus particularly on modality-specific features of reference tracking in the referring expressions and investigate similarities and differences between sign language and co-speech gesture in the use of spatial modification to mark referential context. As a novel contribution, we also look at the use of spatial modification in the predicates of DGS clauses and in the co-speech gestures accompanying predicates in German speech. This investigation beyond the referring expressions themselves is crucial both to achieving a more comprehensive comparison between sign and gesture and a full understanding of the contribution of the spatial affordances of the visual modality to reference tracking.

Direct comparisons between sign, co-speech gesture, and speech, are vital to understanding the influence of the visual modality in shaping communicative expression, and particularly, to understanding how this influence is modulated by use of the visual

modality within a bimodal system (as in co-speech gesture together with speech) versus a unimodal system (as in sign languages). Despite the differences in communicative function, the shared access of sign and gesture to the spatial affordances of the modality have led to assumptions of similarity between sign and gesture in the use of space to represent referents. Specifically, sign language structures in these domains have been argued to be "gestural" in nature, meaning that the spatial modifications observed in both signs and gestures are best described in terms of mental imagery and the conceptualization of an event space in the mind of the speaker/signer (Liddell, 1996, 2003; Quinto-Pozos, 2007; Schembri et al., 2005). However, few studies have tested these claims empirically, using actual sign and co-speech gesture data to investigate the influence of the visual modality in different domains of expression. While signs and gestures may indeed both base the creation of referent-location associations on an imagistic conceptualization of an event space, we have little understanding to date of how such associations are used and maintained beyond the single utterance level in sign (in a unimodal system) and gesture (in a bimodal system). Here, we provide an empirical investigation of how the spatial affordances of the visual modality contribute to marking referential context and referent accessibility through a direct comparison of sign, co-speech gesture, and speech.

We conduct this comparison on DGS (cf. Hansen & Heßmann, 2007; Herrmann, 2007; Hosemann, 2011; Perniss, 2007a,b; Pfau, 2008; Schwager & Zeshan, 2008) and German with co-speech gesture, two languages familiar to the authors.

4. Method

4.1. Participants

Participants in the study were eight deaf native signers[3] of German Sign Language (*Deutsche Gebärdensprache*, DGS) and eight hearing (non-signing) native speakers of German. Deaf participants were recruited from the German state of Nordrhein-Westfalen. DGS data collection sessions took place in Aachen or Essen, Germany, and were conducted by a deaf German research assistant. Hearing participants were university students at the Viadrina University in Frankfurt/Oder, Germany. Data collection sessions took place at the university and were conducted by the first author (a native speaker of German) and a German research assistant.

4.2. Materials and procedure

Participants were instructed to narrate a video vignette to an addressee seated opposite them. The vignette (68 sec long) was shown in two parts (30 sec; 38 sec) in order to minimize the cognitive load required to remember and narrate the whole video, and thus to insure that the narratives produced did not lack detail as a result of memory difficulty. The stimulus vignette featured three women engaged in a cooking activity in a kitchen.[4] The three women are at fixed locations within the scene and perform individual actions

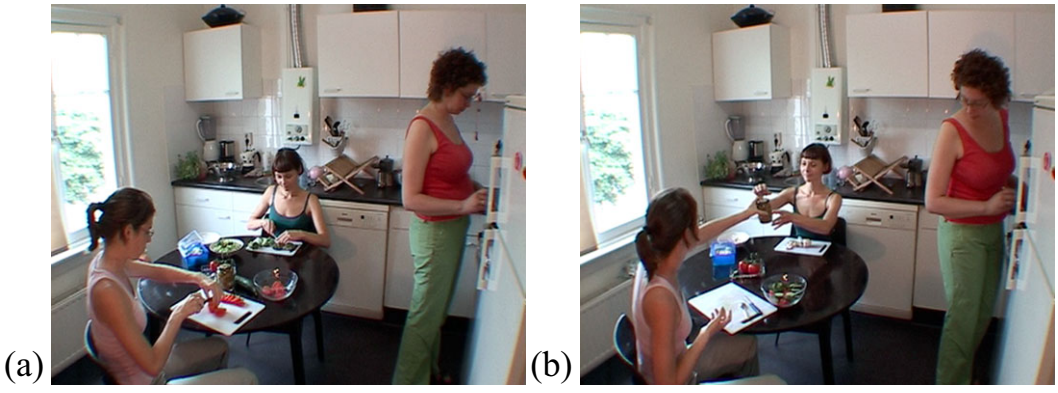

Fig. 2. Stills from the stimulus video vignette used for narrative elicitation.

(e.g., cutting, stirring; see Fig. 2a) and collaborative actions (e.g., giving, taking; see Fig. 2b). These aspects are important for coding of the signs and gestures in terms of reference tracking (detailed in the next section), as individual referents are associated with unique locations and actions that can be represented through the spatial modification of gestures and signs.

Participants watched the stimulus vignette on a laptop computer placed next to them. When finished, they turned forward to face their addressee and provided a narration of the vignette. In the DGS data collection, addressees were deaf DGS signers; in the German data collection, addressees were hearing, native speakers of German. In all cases, addressees were naive to the materials. Addressees were asked to renarrate the vignette back to the signer to demonstrate their comprehension and to enhance the communicative nature of the interaction.

4.3. Annotation and coding

4.3.1. Speech

German narratives were transcribed using German standard orthography and were divided into clauses. Following Berman and Slobin (1994), a clause was defined as any unit containing a predicate (e.g., a verb) that expresses a single activity, event, or state. We counted only those clauses whose subjects referred to the animate referents in the stimulus vignette.

For each relevant clause, we first coded the referring expression (RE) (i.e., the referential form identifying the subject) and the predicate (Pred) as separate constituents. For each constituent, we then coded the referential context as Introduction (I), Maintenance (M), or Re-Introduction (RI). Coding of referential context was based on the notion of local coreference (Hickmann & Hendriks, 1999). A Maintenance (M) context implies that the subject referent of the clause is the same as that of the immediately preceding clause.[5] A Re-Introduction (RI) context instead implies that the subject referent of the clause is different from that of the immediately preceding clause, but that the referent has

been previously mentioned in the discourse.[6] Introduction contexts, which correspond to first mentions of referents, were excluded from analysis.

For each referring expression (RE), we coded the type of expression as nominal (Nom), pronominal (Pron), or zero (Ø). Nominals included various types of nominal phrases: indefinite (e.g., *eine Frau* "a woman"), definite (e.g., *die Frau* "the woman"), and modified NPs (e.g., *die zwei Frauen* "the two women"). Pronominals included demonstrative, personal, relative, and indefinite pronouns. Example (2), which contains two clauses separated here by commas, illustrates these coding categories.

Ref. Context	**RI**	**RI**	**M**	**M**
Constituent type	**RE**	**Pred**	**RE**	**Pred**
Type RE	**Nom**		**Ø**	
	[*Die eine*]	[*ist dann fertig mit den Tomaten*] ,	[Ø]	[*nimmt sich noch 'ne Gurke*].
	the one	is then done with the tomatoes	Ø	takes self else a cucumber

"The one (woman) finishes the tomatoes and takes a cucumber"

4.3.2. Co-speech gesture

Gestures were coded on the basis of frame-by-frame analysis with the video annotation software ELAN.[7] We coded all gesture strokes that occurred in the relevant clauses for analysis. Gesture strokes are the expressive segments of the stream of manual production (Kita, van der Hulst, & van Gijn, 1998). For purposes of coding, the stroke is considered as the most effortful part, or peak, of the manual excursion and can be identified on the basis of changes in handshape, direction of movement, location, and tension of the hands (Kendon, 2004). For each gesture occurring with a referring expression (RE) or predicate (Pred), we identified the type of representation according to the following categories:

- *Enactment* (the hands, face, and/or body take on the role of an animate referent to enact that character's actions, e.g., manual manipulation of an object, or to display that character's affective state; e.g., stirring gesture, cutting gesture, showing surprise; occurred with activity and emotion predicates in speech)
- *Transfer* (the hand represents the motion trajectory of an object, possibly incorporating object shape; e.g., giving jar gesture, movement to table gesture; occurred with transfer and motion predicates in speech)
- *Point* (the location of an entity is represented by pointing to a location in space associated with that entity; the point may involve one or more fingers, or the thumb; occurred primarily with referring expressions in speech)
- *Beat* (the hand carries out a baton-like movement that is timed with rhythmic peaks of the concurrent speech to have an emphasis function; occurred with different types of referring expressions and predicates in speech)

In the relevant clauses, we considered only gestures that referred to the animate referents, as these were the referents of interest in terms of reference tracking. Gestures could

refer to the animate referents either by co-occurring with a referring expression denoting one or more of the women or by co-occurring with a predicate expressing an activity or state of one or more of the women. Gestures were counted as occurring with a referring expression or predicate, depending on their temporal synchrony with these constituents in the speech stream.[8] We excluded gestures that referred to other entities (e.g., the table at which the two women are seated), unless the gesture incorporated an entity being handled by one of the animate referents (e.g., the jar held by the women seated at the table). We also excluded gestures that used space for other purposes (e.g., a gesture occurring with the German word for "before" and functioning as a temporal deictic). We did include beat gestures, however, as these may play a role in giving prominence to a referent (e.g., to support re-activation of a referent in a re-introduction context). In addition, beat gestures may be produced at a location associated with a particular referent and may thus be informative about the way space is used to mark referential context.

For each relevant gesture, we coded whether it was spatially modified. Gestures were considered to be spatially modified if they were produced in a non-neutral location in space (i.e., not directly and centrally in front of the body), thereby associating a referent with a particular location in space (Senghas & Coppola, 2001; So et al., 2005). Spatial modifications could occur in different types of gestures, for example, points (indicating only location), enactments (indicating action at a location), or transfer (indicating motion toward or away from a location). Finally, we coded for consistency of use of referent-location associations. Consistency was determined based on whether gesturers used the same location for a particular referent throughout the course of a narrative. In addition, we determined whether the configuration of referent-location associations in the gesture space corresponded in an iconic way to the configuration of referents in the stimulus vignette as viewed by participants (e.g., an entity viewed on the left in the stimulus associated with a location on the participant's left).

4.3.3. Sign language

All signs were transcribed on a frame-by-frame basis in ELAN using glosses in both German and English. All signs that occurred in the relevant clauses of analysis (i.e., clauses whose subject referred to one or more of the animate referents in the stimulus vignette) were coded for further analysis. Clauses in DGS were determined in the same manner as for German. We identified units containing one predicate (e.g., a verb) and expressing a single activity, event, or state. We used semantic, syntactic, and prosodic cues to help determine the clause units (Hansen & Heßmann, 2007; Herrmann, 2010).

For comparability with speech, we likewise categorized clause constituents as referring expressions (RE) or predicates (Pred), and for each of these, we coded the referential context as Introduction (I), Maintenance (M), or Re-Introduction (RI). As for German, Introduction contexts were excluded from analysis. For each referring expression, we coded the type of expression as nominal (Nom), pronominal (Pron), or zero (Ø). Nominals included nominal phrases of various types: simple noun (e.g., WOMAN "a woman"), definite (e.g., WOMAN IX_{locL} "the/this/that woman"), and modified (e.g., TWO WOMAN "two women").[9] The functional equivalents of pronouns in sign languages are

pointing signs to locations associated with a referent (e.g., IX_{locL}). To the extent possible, categories of predicates in the DGS clauses were defined comparably to predicate categories for co-speech gesture. We coded for the following categories of predicates, based on their form and the type of information encoded:

- *Enactment* (the hands, face, and/or body take on the role of an animate referent to enact that character's actions, e.g., manual manipulation of an object, to display that character's affective state or to attribute dialogue to that character, as in role shift; e.g., stir, cut, show surprise, request vegetables; activity and emotion predicates)
- *Transfer* (the hand represents the motion trajectory of an object, possibly incorporating object shape; e.g., give jar, move to table; transfer and motion predicates)
- *Posture* (the hand represents the whole referent and depicts its body posture; e.g., sitting, standing; entity classifier predicates of location)[10]
- *Discourse/lexical* (signals a discourse comment or other general lexical predicate, e.g., want, get)

Example (3), containing three clauses that are separated by commas in the transcript, provides an example of these coding categories.

Ref. Context	**RI**	**RI**	**M**	**M**	**RI**	**RI**
Constituent type	**RE**	**Pred**	**RE**	**Pred**	**RE**	**Pred**
Ref. Expr. type	**Pron**		**Ø**		**Nom**	
Predicate type	**enactment**		**discourse**		**enactment**	

[IX_{locL}] [CUT-FINISH], [Ø] [CUCUMBER GET], [IX_{locR} WOMAN] [STIR].

she cut-PERF Ø cucumber gets this woman stir

"She (on the left) is finished cutting [a tomato] and takes a cucumber. This woman (on the right) is stirring."

Finally, as for the co-speech gestures, and according to the same criteria, we coded for spatial modification of signs and for consistency of use of referent-location associations over the course of a narrative. In addition, as for co-speech gesture, we determined whether the configuration of referent-location associations in the sign space corresponded in an iconic way to the configuration of referents as observed by participants in the stimulus vignette (e.g., an entity viewed on the left in the stimulus gets associated with a location on the participant's left).

4.3.4. Intercoder reliability

Thirty percent of the data were independently coded for reliability. For German speech and co-speech gesture coding, the proportion of intercoder agreement was 97% for referential context; 98% for type of referring expression in speech; and 93% for type of co-

speech gesture occurring with referring expressions or predicates in speech. For coding pertaining to the use of space, there was 96% agreement about the spatial modification of gestures and 85% agreement about the consistency of use of referent-location associations over the course of a narrative. For DGS coding, the proportion of agreement was 85% for referential context; 97% for type of referring expression; and 90% for type of predicate. Intercoder agreement for spatial modification in DGS was 98% for referring expressions and 97% for predicates, and coders agreed 100% of the time about the consistency of use of referent-location associations in DGS over time.

5. Analyses and results

5.1. Quantity of marking material: Referring expressions

In our analyses, we focus on reference tracking as it is managed after the first mention of a referent. We thus include only Maintenance and Re-Introduction contexts, and exclude Introduction contexts.[11] We start with referring expressions and look first at the amount of marking material used across referential contexts for sign (DGS), speech (German), and co-speech gesture. That is, we ask whether the *Principle of Quantity* holds across all forms of expression, independent of modality. If so, we should see a greater proportion of overt referring expressions in DGS and German in Re-Introduction contexts compared to Maintenance contexts, and we should see a greater proportion of gestures accompanying spoken referring expressions in Re-Introduction compared to Maintenance contexts.

For DGS and German, we calculated the proportion of overt marking material in referring expressions over all clauses (i.e., over all referring expression constituents, both overt and null). Overt marking material in DGS and German included all overt referring expression types (i.e., all nominal and pronominal forms). For German co-speech gesture, we calculated the proportion of overt speech referring expressions that occurred with a gesture (since gestures cannot occur without speech, but elements in speech can occur with or without an accompanying gesture). Overall, the DGS narratives contained 305 relevant clauses, of which 96 contained an overt referring expression. The spoken German narratives contained a total of 213 relevant clauses, of which 142 contained an overt referring expression. Of these 142 overt referring expressions in speech, a total of 44 were accompanied by a gesture.

We used mixed-effects linear model analyses to compare the quantity of overt marking material used across referential contexts between language modalities (DGS, German) and in German co-speech gesture. For DGS and German, proportions of referents that were marked overtly (NPs and pronominals) out of all referring expression (RE) constituents were calculated for each participant; for co-speech gesture, the proportion of NPs and pronominals in German that were accompanied by a gesture were calculated for each participant (see Fig. 3). We first compared DGS and German. We performed a mixed-effects linear model analysis (applying Bonferroni correction) treating Participants as a

random factor and treating Modality (sign, speech) and Referential Context (Maintenance, Re-Introduction) as fixed factors. We found main effects of context ($F(1, 28) = 99.266$, $p < .001$) and modality ($F(1, 28) = 36.637$, $p < .001$) and no interaction between the two ($F(1, 28) = .086$, $p = .772$). We performed a second mixed-effects linear model analysis to assess the use of co-speech gesture with referring expressions across referential contexts, treating Participants as a random factor and treating Referential Context (Maintenance, Re-Introduction) as a fixed factor. We found a main effect of context ($F(1, 9.238) = 10.643$, $p < .01$). Taken together, these analyses show that the quantity of marking material principle holds across all forms of expression—DGS, German, and German co-speech gesture.

Though less central to our initial hypotheses, we conducted pairwise comparisons on the proportion of overt marking material used in each of the referential contexts in DGS and German. We performed mixed-effects linear model analyses looking separately at Re-Introduction contexts and Maintenance contexts, with Participants as a random factor and Modality (sign, speech) as a fixed factor. These comparisons revealed that there were more overt expressions in both Re-Introduction and Maintenance contexts in German than in DGS (Re-Introduction: $p < .01$, $SE_{diff} = .08$, $df = 7.57$; Maintenance: $p < .001$, $SE_{diff} = .07$, $df = 8.79$). We discuss this finding with respect to the influences of language typology (specifically, pro-drop vs. non-pro-drop) and modality (visual vs. vocal) in Section 6.

When we look at the specific types of overt expression used, we see fuller forms in Re-introduction contexts compared to Maintenance contexts, again in keeping with expectations from the Principle of Quantity. The distribution of overt referring expression types used is shown in Table 1. In both DGS and German, we see that fuller forms (i.e., nominals) are used more in Re-Introduction contexts, while leaner forms (i.e., pronominals)

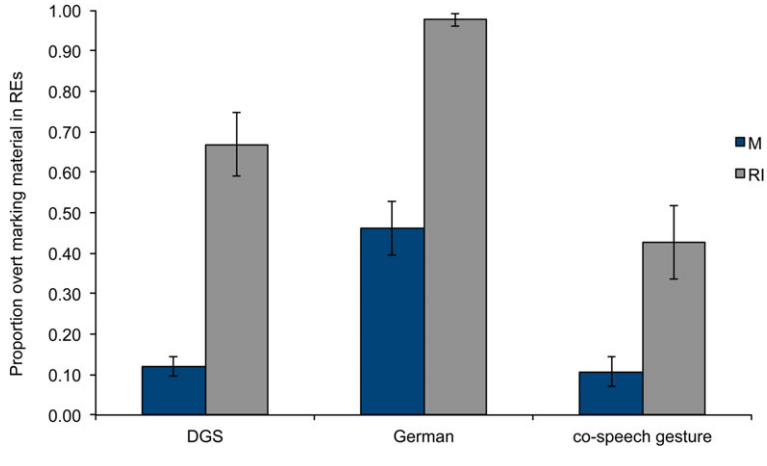

Fig. 3. Proportions of overt marking material in referring expressions (REs), in Maintenance (M), and Re-Introduction (RI) contexts for DGS, German, and German co-speech gesture. (Error bars represent SEs.)

Table 1
Distribution of overt referring expression types by referential context (Maintenance vs. Re-Introduction). For DGS and German, percentage of nominal versus pronominal referring expressions. For co-speech gesture, percentage of nominal versus pronominal speech forms accompanied by a gesture

Re-Introduction	DGS (N = 74)	German (N = 80)	Co-sp. Gesture (N = 30)
Nominal (%)	66	75	89
Pronominal (%)	34	25	11
Maintenance	**DGS (N = 22)**	**German (N = 62)**	**Co.-sp. Gesture (N = 14)**
Nominal (%)	40	7	8
Pronominal (%)	60	93	92

occur more in Maintenance contexts. In both contexts, however, the difference between the use of nominal and pronominal forms is more pronounced in German than in DGS. We see a similar pattern for co-speech gesture, such that gestures occurring with overt referring expressions in speech are more likely to accompany nominal forms in Re-Introduction contexts, and more likely to accompany pronominal forms in Maintenance contexts. Referential context did not influence the type of gestures used, however, with points and beats occurring equally often in Maintenance and Re-Introduction contexts.

We now turn to the modality-specific nature of reference tracking, namely the use of spatial modification in DGS and German co-speech gesture.

5.2. Spatial modification: Referring expressions and predicates

In the remaining analyses, we present comparisons between gestures and signs, zeroing in on the use of spatial modification as a visual–spatial device to mark referential context. We look at the use of this device in referring expressions as well as in predicates.

5.2.1. Referring expressions

Of all overt referring expressions, we calculated the proportion of signs and gestures that were spatially modified in Maintenance (M) and Re-Introduction (RI) contexts for each participant (see Fig. 4). We performed a two-factor mixed factorial ANOVA (2 × 2) with referential context as within-subjects (referential context: Maintenance, Re-Introduction) and modality as between-subjects (modality: gesture, sign) factor.[12] Results showed no main effect of context ($F(1, 14) = .44$, $p = .51$, partial $\eta^2 = .03$), a non-significant but marginal effect of modality ($F(1, 14) = 4.14$, $p = .06$, partial $\eta^2 = .22$), and an interaction between the two ($F(1, 14) = 9.65$, $p < .05$, partial $\eta^2 = .40$). Simple main effects tests showed that even though gestures occurring with referring expressions in speech were more likely to be spatially modified in Re-Introduction ($M = .49$, $SE = .14$) compared to Maintenance contexts ($M = .16$, $SE = .11$; $F(1, 7) = 7.27$, $p < .05$, partial $\eta^2 = .50$), there was no difference in spatial modification between Re-Introduction ($M = .46$, $SE = .07$)

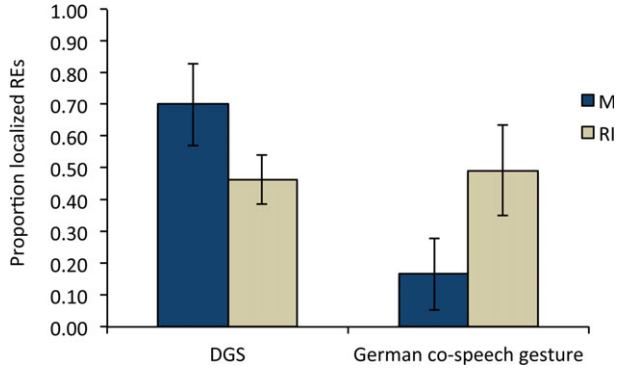

Fig. 4. Proportions of spatially modified overt referring expressions (REs) in DGS and German co-speech gesture in Maintenance (M) and Re-Introduction (RI) contexts. (Error bars represent SEs.)

and Maintenance contexts (M = .70, SE = .12) for signs ($F(1, 7) = 2.83$, p = .13, partial η^2 = .29).

In terms of the consistent use of referent-location associations over the course of narratives, we observed a high degree of consistency for referring expression constituents in both German co-speech gesture (87% consistent) and DGS (94% consistent) across referential contexts. Furthermore, in both sign and gesture, the locations used always corresponded in an iconic way to the fixed referent locations observed in the stimulus vignette.

5.2.2. *Predicates*

We now compare spatial modification in predicates between DGS and German co-speech gesture. For co-speech gesture, we counted all gestures that temporally overlapped with the predicate constituent in speech. Of a total of 213 predicates in speech, 126 were accompanied by a relevant gesture type and were thus included in the analysis. For DGS, 305 predicates were included in the analysis.

Out of the total (i.e., total number of predicates in sign; total number of predicates in speech accompanied by a gesture), we calculated the mean proportion that were spatially modified in Maintenance (M) and Re-Introduction (RI) contexts for each participant (see Fig. 5). We performed a two-factor mixed factorial ANOVA (2 × 2) with referential context as within-subjects (referential context: Maintenance, Re-Introduction) and modality as between-subjects (modality: gesture, sign) factor. Results showed main effects of context ($F(1, 14) = 7.84$, $p < .05$, partial η^2 = .36) and modality ($F(1, 14) = 13.19$, $p < .01$, partial η^2 = .48), but no significant interaction ($F(1, 14) = 3.23$, p = .09, partial η^2 = .18). Because the interaction may be said to be marginally significant, we performed further analyses to look more closely at differences between levels. Further simple main effects show that within DGS, the use of spatially modified predicates in Maintenance (M = .49, SE = .05) and Re-Introduction contexts (M = .70, SE = .08) revealed a signifi-

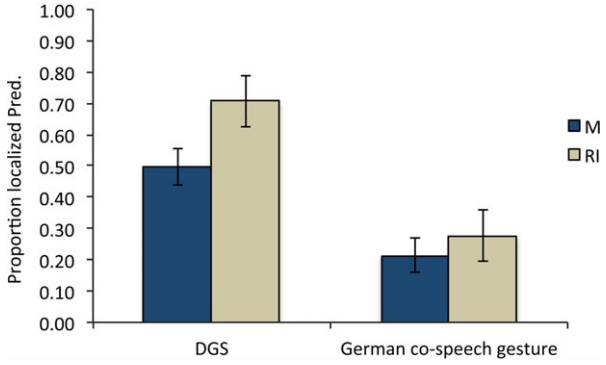

Fig. 5. Proportion of spatially modified predicates in DGS and co-speech gestures accompanying predicates in German speech in Maintenance (M) and Re-Introduction (RI) contexts. (Error bars represent SEs.)

cant difference ($F(1, 7) = 9.20$, $p < .05$, partial $\eta^2 = .56$), with predicates more likely to be spatially modified in Re-Introduction contexts compared to Maintenance contexts. However, this difference was not significant for co-speech gesture ($F(1, 7) = .56$, $p = .47$, partial $\eta^2 = .08$; Re-Introduction: $M = .27$, $SE = .08$; Maintenance: $M = .21$, $SE = .05$).

Again, in the use of referent-location associations, both DGS (96%) and German co-speech gesture (94%) displayed a very high degree of spatial consistency, and locations for referents in space always corresponded in an iconic way to the fixed locations of referents observed in the stimulus vignette.

Finally, we compare the distribution of types of predicates in DGS and types of co-speech gestures accompanying predicates in German speech by percentage of spatial modification in both referential contexts (see Table 2).

Looking at the likelihood of spatial modification for specific types, the difference between sign and gesture in the use of *Enactment* predicates is particularly striking. In both referential contexts, DGS signers perform *Enactment* predicates depicting manual manipulation (e.g., unscrewing the lid of a jar) at the location associated with the referent performing that action nearly half the time (see Fig. 6a). In contrast, German co-speech gesturers very rarely localize these types of *Enactment* predicates (see Fig. 6b).

Table 2
Distribution of types of predicates in DGS and types of co-speech gestures accompanying predicates in German speech by percentage of spatial modification, for Maintenance and Re-Introduction contexts

	Spatial Modification			Spatial Modification	
Sign Type	Maintenance	Re-Introduction	Gesture Type	Maintenance	Re-Introduction
Enactment	54/129 (42%)	44/81 (54%)	*Enactment*	1/26 (4%)	0/9 (0%)
Transfer	28/28 (100%)	19/19 (100%)	*Transfer*	16/16 (100%)	3/3 (100%)
Posture	4/5 (80%)	36/39 (92%)	*Point*	1/1 (100%)	3/3 (100%)
Discourse	0/10 (0%)	1/6 (17%)	*Beat*	4/38 (11%)	9/30 (30%)

Fig. 6. Examples of (a) spatially modified predicate (*unscrewing lid of jar*) in DGS; (b) non-spatially modified predicate (*unscrewing lid of jar*) in German co-speech gesture; and (c) still image from stimulus vignette of the scene being depicted in DGS and German co-speech gesture—the referent whose action is being depicted is displayed additionally in a close-up zoom.

6. Summary and discussion

In this study, we investigated how modality-specific features of co-speech gesture and sign contribute to reference tracking in discourse, presenting a systematic comparison between German co-speech gesture and German Sign Language (DGS) in this domain. The study aims to understand similarities and differences between sign language and co-speech gesture, and how the shared affordances of the visual modality are

differentially constrained or modulated through use within a unimodal (as in sign) versus a bimodal system comprising speech and gesture. Going beyond previous research, we looked at the use of spatial modification in the service of reference tracking not only with referring expressions but also with predicates to provide a comprehensive comparison between sign and gesture in terms of the use of space to mark referential context. We also analyzed speech to understand how gestures mark referential context and referent accessibility in the context of speech. In the following, we discuss our findings first with respect to the quantity of marking material and then with respect to spatial modification.

6.1. Quantity of marking material

We compared the use of overt referring expressions between DGS, German, and German co-speech gesture. Across all three forms of expression, we found that referents in subject position were more likely to be overtly marked in Re-Introduction contexts compared to Maintenance contexts. In further support of the quantity of marking principle, we found that fuller forms (i.e., nominals) were more likely to be used than leaner forms (i.e., pronominals) in Re-Introduction contexts compared to Maintenance contexts in both German and DGS. This confirms previous findings that the quantity of marking principle is a general, modality-independent principle of referent accessibility (e.g., Gullberg, 2006; McKee et al., 2011), and that gesture works in parallel here with concurrent speech (So et al., 2009).

A difference exhibited between German and DGS with respect to the quantity of marking material was in the overall use of overt forms. Compared to DGS, referring expressions in German were more likely to be overt (vs. zero) in both Maintenance and Re-Introduction contexts. The high occurrence of null subjects in DGS may not seem surprising given the morphosyntactic properties of space in sign languages (Meier, 2002; Sandler & Lillo-Martin, 2006). As we saw in our third analysis, and as we will discuss further below, the function of referent identification in DGS might be fulfilled to a substantial degree through the spatial modification of predicates, which may license null arguments in subject position. However, this claim needs further research to reach definitive conclusions.

Under the tentative assumption that DGS (like ASL; Lillo-Martin, 1991) is a pro-drop language, one may wonder whether we would have found the same difference in overall use of overt subject marking if we had compared DGS with a pro-drop spoken language, rather than with German. In a separate study that investigated reference tracking and discourse cohesion in Turkish narratives, using the same stimulus materials and coding procedure, it was found that overt referring expressions in speech were used 80% of the time in re-introduction contexts and 20% of the time in maintenance contexts (Azar, 2013). Comparing these numbers roughly to our own results (in Fig. 3), this suggests that overt subject encoding is less likely in a spoken pro-drop language (Turkish) than in a spoken non-pro-drop language (German), but that overt encoding is nevertheless more likely in a spoken pro-drop language than in a signed (pro-drop) language like DGS. Further research is needed in this direction.

6.2. Spatial modification

We compared the use of spatial modification to mark referential context in DGS and German co-speech gesture, for both referring expressions and predicates. First, focusing on similarities, both DGS and German co-speech gesture used spatial modification to mark referential context. Secondly, once locations for referents were established, these were referred back to in a highly consistent manner, using the same location for a particular referent throughout a narrative. The degree of consistency we observed in co-speech gesture use of referent-location associations is notably higher than found in a previous study on the consistent use of space by So et al. (2005). However, in that study, individual stimulus vignettes were narrated one at a time, and the vignettes included motion of referents in different settings, whereas our stimulus vignette featured referents at fixed locations in space and may thus have been more conducive to the consistent use of referent-location associations. In addition, the configuration of referents established in the gesture or sign space corresponded in an iconic way to the configuration of referents as they appeared in our stimulus vignette. This finding echoes a growing body of literature that recognizes the interface between linguistic and imagistic elements in signs that rely on spatial modification (e.g., Johnston, 2013; Liddell, 2003; Lillo-Martin & Meier, 2011; Schembri et al., 2005).

However, we also found crucial differences between DGS and German co-speech gesture in how the affordance of spatial modification functions on a discourse level. That is, gestures accompanying referring expressions in German were more likely to be localized when referents were re-introduced compared to when referents were maintained across consecutive clauses (corroborating previous research; Gullberg, 2006). However, in DGS, referring expressions were spatially modified equally in both contexts. This is due primarily to the use of pointing signs in DGS, which figures prominently into referring expressions in both maintenance contexts (as pronouns) and re-introduction contexts (as part of definite or demonstrative noun phrases), and reflects the fact that, in a unimodal system, the visual modality must carry the full load both of identifying referents and marking referential context. In contrast, when the visual modality is used in a bimodal system, it is speech that does the main work of referent identification, and the gestures accompanying speech contribute to discourse cohesion by marking referential context through spatial modification.

For spatial modification of predicates, the pattern was reversed. There was no difference between re-introduction and maintenance contexts with respect to localization of gestures accompanying predicates in German speech. In DGS, on the other hand, predicates were more likely to be localized when referents were re-introduced into discourse compared to when they were maintained. It is not news that the spatial modification of verbs in sign languages can serve to sufficiently and uniquely identify arguments (Padden, 1990). A striking aspect of our findings, however, is that spatial modification occurred to a high degree with enactment predicates in re-introduction contexts. Enactment predicates consist of representations that mimic, or enact, the (real-world) actions of a character (also called *constructed action*; Quinto-Pozos, 2007). When a sign that depicts an action like cutting vegetables is spatially displaced, it loses a central feature of enactment,

namely production of the action as it is performed by the character whose role is assumed. The fact that we do see such marking is testimony to a function of spatial modification at the discourse level. Though co-speech gesturers used enactment predicates of a very similar type, their use exhibited enactment in the strict sense of the word. Thus, here we see the systematic use of space in the service of reference tracking and discourse cohesion when the visual modality is used unimodally within a linguistic system. Future research may determine whether the spatial modification of such enactment predicates as a discourse-level reference-tracking device is a robust modality-specific feature of all sign languages, or whether it is specific to certain sign languages (e.g., see Özyürek and Perniss, 2011 for spatial modification of enactment-type predicates in Turkish Sign Language).

7. Conclusion

This study has shown that the use of the visual modality for communicative expression, as it occurs in co-speech gesture and sign, reflects referential context and referent accessibility. Co-speech gesture and sign exhibit broad similarities in the use of spatial modification to create spatial anaphoric linkages and in the use of the hands/body as articulators to represent referents. But the differences observed between the two forms of expression make evident the differential influence of using the visual modality within a unimodal (sign) versus a bimodal (co-speech gesture) system.

The study contributes to our understanding of similarities and differences between sign and gesture—important because of assumed similarities in the use of the spatial affordances of the modality—and has looked at a core domain in which the use of space is crucial. We have gone beyond previous research investigating properties of reference tracking in the visual modality by providing a direct comparison between sign, co-speech gesture, and speech, as well as by investigating both referring expressions and predicates, allowing a broader comparison between the two.

Our study also demonstrates that the ability to mark the referential status of referents in discourse extends beyond the principle of quantity of marking material and is reflected also in modality-specific affordances of communicative expression in the visual modality. Contributing to discourse cohesion by making referents more "visible" (i.e., making them stand out) through spatial modification is a modality-specific feature that can be exploited in both sign and gesture. Whether it is the referring expression or the predicate that is made more "visible" seems to be related to whether the visual modality is used within a unimodal or bimodal system. Of course, we are aware that our study has looked at only one sign language (DGS) and one spoken language (German) with co-speech gesture, such that we must be cautious about making generalizations. A larger cross-linguistic comparison between different sign languages and different spoken languages (pro-drop and non-pro-drop) with co-speech gestures, for example, would be informative to further investigate possible differential effects of language typology versus modality (specifically, the use of space) in reference tracking and achieving discourse cohesion. It would also be

interesting to investigate the development of discourse-level uses of space in emerging sign languages (cf. Meir, Padden, Aronoff, & Sandler, 2007; Senghas & Coppola, 2001), particularly with respect to the spatial modification of predicates as we have seen in DGS enactment predicates, but not in German co-speech gesture.

Acknowledgments

This research was supported by a Dutch Science Foundation (NWO)-VIDI grant, number 276-70-009, awarded to the second author and supporting the first author as a postdoc, and by the Max Planck Institute for Psycholinguistics. For collection and coding of DGS data, we thank Uwe Zelle; for collection of German data, we thank Dorothea Böhme. We are also grateful to Prof. Cornelia Müller for providing room facilities and access to German speakers at the Viadrina University in Frankfurt/Oder, Germany. We thank Lian van Hoof and Zeynep Azar for help with coding of the German co-speech gesture data. Finally, we thank four anonymous reviewers for their comments on an earlier version of this manuscript.

Notes

1. One study has compared the use of space in reference tracking in co-speech gesture and pantomime (i.e., speakers silently gesturing; So, Coppola, Liccidarello, & Goldin-Meadow, 2005).
2. In addition to manual subject/object marking, some analyses have proposed the use of eye gaze and head tilt as non-manual subject/object markers that would license null arguments with both directional/agreement verbs and plain verbs in ASL (Bahan, Kegl, Lee, MacLaughlin, & Neidle, 2000; Neidle, Kegl, MacLaughlin, Bahan, & Lee, 2000; but for a refutation of this claim, see Thompson, Emmorey, & Kluender, 2006 for ASL; Hosemann, 2011 for DGS).
3. Native signers are born to deaf parents and acquire a sign language (natively) from birth. Native signers constitute only a small portion of the signing population, as most deaf people are born to hearing parents who do not know a sign language.
4. In the first part (30 sec), two women, seated opposite each other at a table, are cutting vegetables and transferring the cut vegetables to bowls. The third woman is standing to the right of the table, cooking at a stove. Upon request, indicated by the standing woman's movement to the table, the seated woman at the back passes a bowl of cut vegetables to the standing woman. The standing woman empties the contents of the bowl into what she is cooking on the stove and passes the empty bowl back to the seated woman. In the second part (38 sec), the seated woman at the front tries unsuccessfully to open a jar of pickles. She passes the jar to the other seated woman, who also tries to open the jar, but who also fails. The two seated women pass the jar back and forth between them a few times, but nei-

ther of them can get the jar unscrewed. Finally, the standing woman, becoming aware of the commotion behind her, turns and reaches for the jar, which is passed to her by the seated woman at the back. The standing woman unscrews the lid easily and passes the jar back.

5. In some clauses, a different animate referent was mentioned as the indirect object of a predicate (i.e., with predicates of transfer, as in *She gave the jar to the other woman*). In these cases, we acknowledged the object referent as signaling the switch in reference and coded the following subject referent as occurring in a Maintenance context. We included only subject referents in our count of referring expressions.
6. A change from a singular to plural (e.g., *one woman* to *both women*), or from a plural to singular (e.g., *both women* to *one of the women*) referring expression signaled a Re-Introduction context (cf. Debrelioska et al., 2013). Thus, Maintenance contexts required full sameness of the subject referent in the preceding clause.
7. ELAN is the linguistic annotation tool developed at the MPI for Psycholinguistics in Nijmegen, Netherlands. It is available for free download at www.lat-mpi.eu/tools/elan/.
8. Note that this means it is impossible for a gesture to co-occur with a zero referring expression in speech. This is not an artefact of our coding decisions; no gestures related to animate referents in subject position were excluded due to coding criteria. Simply stated, (referential) gestures did not occur if there was no accompanying speech.
9. According to standard practice, sign glosses are provided in English with capital letters. Points are glossed as IX for "index," with a subscript indicating the location in sign space to which the point is directed (e.g. "IX_{locL}" for a point to a location on the left side of sign space).
10. In these predicates, the entity classifier refers to the handshape that depicts a seated or standing human. In the entity classifier used for a standing human, the index and middle fingers are extended and separated (forming a V-hand), and held upside down (inverted V-hand); for a seated human, the fingers are bent at the middle joint (bent V-hand) and the palm is oriented downward (see Emmorey, 2003, on the use of classifier predicate constructions across different sign languages).
11. We exclude Introduction contexts for narrations of both parts of the video vignette, thereby also accounting for an effect of the break in stimulus presentation.
12. The reported ANOVAS used arcsine-transformed proportions; figures present actual proportions and the significance level for Bonferroni corrections was set to $p = .025$ to control for multiple comparisons.

References

Ariel, M. (1990). *Accessing noun-phrase antecedents*. London: Routledge.

Ariel, M. (1994). Interpreting anaphoric expressions: A cognitive versus a pragmatic approach. *Journal of Linguistics*, *30*, 3–42.

Azar, Z. (2013). Discourse management: Reference tracking in speech and gesture in Turkish narratives. Unpublished Masters Thesis. Tilburg University.

Bahan, B., Kegl, J., Lee, R., MacLaughlin, D., & Neidle, C. (2000). The licensing of null arguments in American Sign Language. *Linguistic Inquiry*, *31*, 1–27.

Berman, R. A., & Slobin, D. I. (1994). *Relating events in narrative: A crosslinguistic developmental study*. Hillsdale, NJ: Lawrence Erlbaum Associates.

Cameron, R. (1992). Pronominal and null subject variation in Spanish: Constraints, dialects, and functional compensation. PhD thesis, University of Pennsylvania.

Cameron, R. (1998). A variable syntax of speech, gesture, and sound effect: Direct quotations in Spanish. *Language Variation and Change*, *10*, 43–83.

Chafe, W. (1994). *Discourse, consciousness, and time*. Chicago: Chicago University Press.

de Quadros, R. M., & Lillo-Martin, D. (2010). Clause structure. In D. Brentari (Ed.), *Sign languages* (pp. 225–251). Cambridge, UK: Cambridge University Press.

Debrelioska, S., Özyürek, A., Gullberg, M., & Perniss, P. (2013). Gestural viewpoint signals referent accessibility. *Discourse Processes*, *50*(7), 431–456.

Emmorey, K. (Ed.) (2003). *Perspectives on classifier constructions in sign languages*. Mahwah, NJ: Lawrence Erlbaum Associates.

Engberg-Pedersen, E. (1993). *Space in Danish Sign Language: The semantics and morphosyntax of the use of space in a visual language*. Hamburg, Germany: Signum Press.

Givón, T. (1984). Universals of discourse structure and second language acquisition. In W. E. Rutherford (Ed.), *Language universals and second language acquisition* (pp. 109–136). Amsterdam: John Benjamins.

Gullberg, M. (2006). Handling discourse: Gestures, reference tracking, and communication strategies in early L2. *Language Learning*, *56*, 155–196.

Hansen, M., & Heßmann, J. (2007). Matching propositional content and formal markers. Sentence boundaries in a DGS text. *Sign Language & Linguistics*, *10* (2), 145–175.

Herrmann, A. (2007). The expression of modal meaning in German Sign Language and Irish Sign Language. In P. Perniss, R. Pfau, & M. Steinbach (Eds.), *Visible variation. Comparative studies on sign language structure* (pp. 245–278). Berlin: Mouton de Gruyter.

Herrmann, A. (2010). The interaction of eye blinks and other prosodic cues in German Sign Language. *Sign Language & Linguistics*, *13*(1), 3–39.

Hickmann, M., & Hendriks, H. (1999). Cohesion and anaphora in children's narratives: A comparison of English, French, German, and Mandarin Chinese. *Journal of Child Language*, *26*, 419–452.

Hosemann, J. (2011). Eye gaze and verb agreement in German Sign Language. A first glance. *Sign Language & Linguistics*, *14* (1), 76–93.

Johnston, T. (2013). Towards a comparative semiotics of pointing actions in signed and spoken languages. *Gesture*, *3*(2), 109–142.

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

Kita, S., van der Hulst, H., & van Gijn, I. (1998). Movement phases in signs and co-speech gestures, and their transcription by human coders. In I. Wachsmuth & M. Fröhlich (Eds.), *Gesture and sign language in human-computer interaction* (pp. 23–35). Berlin: Springer.

Klima, E. S., & Bellugi, U. (1979). *The signs of language*. Cambridge, MA: Harvard University Press.

Levy, E., & Fowler, C. (2000). Grounding references in perception. In D. McNeill (Ed.), *Language and gesture* (pp. 215–234). New York: Cambridge University Press.

Liddell, S. K. (1996). Spatial representations in discourse: Comparing spoken and signed language. *Lingua*, *98*, 145–167.

Liddell, S. K. (2003). *Grammar, gesture and meaning in American Sign Language*. Cambridge, UK: Cambridge University Press.

Lillo-Martin, D. (1986). Two kinds of null arguments in American Sign Language. *Natural Language & Linguistic Theory*, *4*, 415–444.

Lillo-Martin, D. (1991). *Universal grammar and American Sign Language: Setting the null argument parameters*. Boston: Kluwer.

Lillo-Martin, D., & Klima, E. S. (1990). Pointing Out Differences: ASL Pronouns in Syntactic Theory. In S. D. Fischer, & P. Siple (Eds.), *Theoretical issues in sign language research*. Vol. *1* (pp. 191–210), *Linguistics*. Chicago: University of Chicago Press.

Lillo-Martin, D., & Meier, R. P. (2011). On the linguistic status of "agreement" in sign languages. *Theoretical Linguistics*, *37*(3/4), 95–141.

Lucas, C., Bayley, R., & Valli, C. (2001). *Sociolinguistic variation in American Sign Language*. Washington, DC: Gallaudet University Press.

Mather, S., & Winston, E. A. (1998). Spatial mapping and involvement in ASL storytelling. In C. Lucas (Ed.), *Pinky extension and eye gaze: Language use in deaf communities* (pp. 183–210). Washington, DC: Gallaudet University Press.

McKee, R., Schembri, A., McKee, D., & Johnston, T. (2011). Variable "subject" presence in Australian Sign Language and New Zealand Sign Language. *Language Variation and Change*, *23*, 375–398.

McNeill, D. (1992). *Hand and mind: What gestures reveal about thought*. Chicago: Chicago University Press.

McNeill, D., & Levy, E. (1993). Cohesion and gesture. *Discourse Processes*, *16*, 363–386.

Meier, R. P. (2002). Why different, why the same? Explaining effects and non-effects of modality upon linguistic structure in sign and speech. In R. P. Meier, K. Cormier, & D. Quinto-Pozos (Eds.), *Modality and structure in signed and spoken languages* (pp. 1–25). Cambridge, UK: Cambridge University Press.

Meir, I., Padden, C., Aronoff, M., & Sandler, W. (2007). Body as subject. *Journal of Linguistics*, *43*, 531–563.

Neidle, C., Kegl, J., MacLaughlin, D., Bahan, B., & Lee, R. (2000). *The syntax of American Sign Language: Functional categories and hierarchical structure*. Cambridge, MA: MIT Press.

Özyürek, A., & Perniss, P. (2011). Event representation in sign language: A crosslinguistic perspective. In J. Bohnemeyer & E. Pederson (Eds.), *Event representation in language: Encoding events at the language-cognition interface* (pp. 84–107). Cambridge, UK: Cambridge University Press.

Padden, C. (1990). The relation between space and grammar in ASL verb morphology. In C. Lucas (Ed.), *Sign language research: Theoretical issues* (pp. 118–132). Washington, DC: Gallaudet University Press.

Perniss, P. (2007a). Space and Iconicity in German Sign Language (DGS). Doctoral dissertation. MPI Series in Psycholinguistics, 45.

Perniss, P. (2007b). Achieving spatial coherence in German Sign Language narratives: The use of classifiers and perspective. *Lingua*, *117*(7), 1315–1338.

Pfau, R. (2008). The grammar of headshake: A typological perspective on German Sign Language negation. *Linguistics in Amsterdam*, *1*, 37–74.

Quinto-Pozos, D. (2007). Can constructed action be considered obligatory? *Lingua*, *117*(7), 1285–1314.

Sandler, W., & Lillo-Martin, D. (2006). *Sign language and linguistic universals*. Cambridge, UK: Cambridge University Press.

Schembri, A., Jones, C., & Burnham, D. (2005). Comparing action gestures and classifier verbs of motion: Evidence from Australian Sign Language, Taiwan Sign Language, and nonsigners' gestures without speech. *Journal of Deaf Studies and Deaf Education*, *10*(3), 272–290.

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

Senghas, A., & Coppola, M. (2001). Children creating language: How Nicaraguan Sign Language acquired a spatial grammar. *Psychological Science*, *12*, 323–328.

So, W. C., Coppola, M., Liccidarello, V., & Goldin-Meadow, S. (2005). The seeds of spatial grammar in the manual modality. *Cognitive Science*, *29*, 23–37.

So, W. C., Kita, S., & Goldin-Meadow, S. (2009). Using the hands to identify who does what to whom: Gesture and speech go hand-in-hand. *Cognitive Science*, *33*, 115–125.

Thompson, R., Emmorey, K., & Kluender, R. (2006). The relationship between eye gaze and verb agreement in American Sign Language: An eye-tracking study. *Natural Language & Linguistic Theory*, *24*, 571–604.

van Hoek, K. (1996). Conceptual locations for reference in American Sign Language. In G. Fauconnnier & E. Sweetser (Eds.), *Spaces, worlds, and grammar* (pp. 334–350). Chicago, IL: University of Chicago Press.

Winston, E. A. (1995). Spatial Mapping in Comparative Discourse Frames. In K. Emmorey & J. S. Reilly (Eds.), *Language, gesture, and space* (pp. 87–114). Hillsdale, NJ: Lawrence Erlbaum.

Wulf, A., Dudis, P., Bayley, R., & Lucas, C. (2002). Variable subject presence in ASL narratives. *Sign Language Studies*, *3*(1), 54–76.

Yoshioka, K. (2008). Gesture and information structure in first and second language. *Gesture*, *8*(2), 236–255.