Response to commentaries: Gesture, language, and directionality

DIANE LILLO-MARTIN and RICHARD P. MEIER

We are grateful to the commentators for their thoughtful replies. We had two goals in writing the target article: (i) to bring the issue of sign language directionality to the attention of a wider audience of linguists; and (ii) to move forward the debate regarding the nature of this phenomenon by changing the focus somewhat, specifically by distinguishing between person-marking and agreement. We are pleased that the commentaries indicate that we have achieved these goals to some extent.

In our response to the commentaries, we organize our reactions according to four questions aimed towards the second goal listed above. The various commentaries address these four questions to different degrees, and we feel that organizing our response in this way will allow us to integrate the different approaches taken by the commentators and also to emphasize our take-home message. The four questions ask:

- What is the relationship between language and gesture?
- Is there a grammatically-constrained process deriving the directionality of verbs?
- If so, is this process properly considered person marking?
- Is *agreement* the optimal explanation for the mechanics of this process?

The first question allows us to discuss more explicitly what we assumed in the target article regarding the relationship between language and gesture. We hope that clarifying this point will help make our overall claims clearer. In the remaining sections, we highlight some of the points of the commentators and connect them to our own views.

I. What is the relationship between language and gesture?

We start by highlighting a fundamental distinction in the approaches to directionality that is made clear in Quer's commentary. He distinguished two conceptions of language – one which approaches language as a human faculty, (at least some of) whose operations can be profitably explored independently from other cognitive modules; and another conception which emphasizes the interactions between language and other cognitive systems. Along with Quer's own work, our work takes the first approach; however, we acknowledge that this approach has its limits, and we find that sign language research may help us to define these limits in ways somewhat differently from linguistic studies focusing exclusively on spoken languages. Thus, research on sign languages could lead to the conclusion that it is impossible to study language narrowly, without consideration of its interactions with other components; or this research could suggest that certain aspects of language can be profitably examined for their behavior independent of other components. If gesture is inextricably intertwined with sign language, it would be impossible to come to any solid generalizations about the nature of sign language without including the gestural component. However, if gesture interfaces with language, there may be aspects of language – for example, the narrow syntax – which can be explained independent of gesture. Quer clearly thinks that this is so and suggests that we have ascribed too large a role to gesture.

It is clear that speech and the gestures that accompany it are different in important ways. As summarized by Goldin-Meadow and McNeill (1999), speech is analytic, segmented and combinatorial, whereas gestures are analogue and mimetic. Yet because of the tight connections between speech and gesture, they adopt the view that speech and gesture form a unified system, following principles of gesture-speech integration proposed in McNeill (1992). Still, Goldin-Meadow and McNeill claim that sign languages are like speech in showing the crucial segmented/combinatorial characteristics. They then ask, "Can the manual modality at the same time also be used for holistic and mimetic expression? In other words, do signers gesture along with their signs . . . ?" (Goldin-Meadow and McNeill 1999: 166). Goldin-Meadow and McNeill suggest that certain mouth movements accompanying signs might constitute a gestural component (cf. Sandler 2009), but they do not elaborate this point.

It seems clear that signers do, in fact, gesture along with their signs. As Emmorey (1999) shows, sometimes the signer will stop signing and insert a

gestural element. However, on our view, sometimes the gesture combines with the sign, as in the examples of pronouns and directionality under discussion here. In this, we agree with Liddell that gesture must be invoked to explain the actual locations towards which directional signs point. We take Rathmann & Mathur and Aronoff & Padden to agree with us on this. Rathmann & Mathur spell out their view of how the relation between language and gesture permits a view of directionality that is very similar to ours. What is important to all of us, however, is whether the patterning of the non-gestural elements can be described independently of the gestures.

In the following paragraphs, we highlight the essential aspects of our view regarding the separability of certain grammatical components from gesture. For us, what is crucial to establish this separability is that some linguistic phenomena must refer to the fact that verbs have directionality, but not to the specific locations used for the various (non-first) forms. For example, the word order facts from Libras discussed in the target article indicate that syntactic structural differences are found for sentences containing verbs that are marked with directionality as compared to those not so marked. These word order differences are not dependent on specific differences among the non-first forms. This indicates the presence of a linguistic process, independent of the gestural realization of various forms.

When we claim that directionality is linguistic, we mean that the distribution of directional verbs is grammatically constrained in ways that are familiar from the analysis of agreement in spoken language. Further evidence for this claim comes from the observations that directional verbs license null arguments in ways that seem to mirror what we know about the syntax of Romance languages; and that directional verbs exhibit a subset of the grammatical categories that characterize person-marking systems in other languages. Thus, from the point of view of the morpho-syntax, the gestural elements are irrelevant.

Our view regarding the 'narrow' morpho-syntax is thus compatible with Quer's. We do not claim that the different loci used for different non-first referents are represented morpho-syntactically. In the morpho-syntax we see only the distinction between first and non-first. At the syntactic level, what is relevant are the featural distinctions (e.g., first vs. non-first), but not the actual physical locations. Accordingly, we would expect to see no grammatical rule that refers to the actual physical locations used by different non-first referents. No rule applies differently to referents on the right side of a signer compared

to those on the left side.[1] In this sense, it doesn't matter at this level what the precise form of the verb is. This is similar to Quer's point about the 'categorical' interpretation of loci as being more important than the actual physical points. In fact, we think this part of our view separates language from gesture even more than Quer's does because of our claim that 2nd and 3rd persons are not grammatically distinct. As far as we know, no grammatical process applies to 2nd person forms differently from 3rd person forms. Quer refers to the need for a 2nd person category to identify imperatives, but we do not know of a morpho-syntactic marker of imperatives in ASL. We would be very interested in knowing whether other sign languages (such as LSC) have such a process.

Of course, even at this more abstract level it is important to keep different non-first referents distinct, as they will be relevant for interpretation and for constraints such as binding. It is for such reasons that we maintain the notion of a referential index. We consider the index to be syntactically/semantically represented. Rathmann & Mathur worry that maintaining the notion of an R-index as relevant to the grammatical process under investigation forces the grammar to contain a non-finite number of R-loci. We prefer to separate the abstract R-indices from their physical realization in R-loci. On views of grammar which admit indices (for example, that adopted by Schlenker), there are an indeterminate number of such abstract entities available.

At another (phonetic) level, actual physical locations will need to be specified. Here is where we adopt part of Liddell's solution by invoking gesture. We are convinced by his descriptions of the actual physical form of indicating verbs. In order to describe the physical form, we must appeal to Real or Mental spaces, in Liddell's terms.

Thus, our view of the interface between language and gesture allows for a gesture-independent computational component, with an articulation between language and gesture permitting the realization of certain linguistic elements to be influenced by gesture. In speech, such gestures would include not only those iconic and metaphoric manual gestures that Goldin-Meadow and McNeill discuss, but also speech gestures such as the extended production of a word like

---

1   There are some restrictions that are clearly based in articulation. With some verbs (e.g. ASL ASK) it is difficult to articulate left-to-right movement using the right hand – or right-to-left movement using the left hand (Mathur 2000; Mathur & Rathmann 2001). Even such constraints are not based on the actual physical locations used (e.g., right side), but the particular arrangement of the hands in space.

'long' for effect (Okrent 2002). In sign languages, this interface permits verbs and pronouns to be co-articulated with ostensive indication (among other effects). Aronoff & Padden, in their commentary, speculate that the nature of the language/gesture interface in sign languages means that signed languages will not lexicalize phonologically-constant second-person forms, whether pronouns or verbal inflections. By dint of the indexicality of gestural reference to addressees and non-addressed participants, the forms used to refer to addressees are, on their view, too variable to allow the emergence of fixed linguistic forms.

Using Liddell's Figure 1, our abstract conceptualization of the R-locus used to refer to Billy's Mother is the location occupied by the shape representing Billy's Mother, which we will call R-locus$_{Mot}$. Neither R-locus$_1$ nor R-locus$_3$ is our intended locus, as claimed by Liddell. We could say that R-locus$_{Mot}$ is the target towards which the verbs TELL and ASK in his example are directed, even though their actual production only moves partway to this goal. Padden (1990) already made this point by discussing the movement of verbs as occurring along a vector formed from the R-locus of the subject referent and the R-locus of the object referent. Neither we nor Liddell claim that the actual physical form of the sign must reach the physical position of Billy's Mother, but this is the position toward which the signs are directed. This clarification should help with understanding why the four claims Liddell attributes to us are not contradictory.

Liddell's view of the relationship between language and gesture is different from ours. In this context we must clarify our intention in bringing up the differences between his views of pointing signs and the 'agreement' analysis. Certainly it was not our purpose to dismiss his proposals (and we are quite sorry if our comments suggested that we were dismissive of them); on the contrary, we felt that his views were significant enough to warrant considerable attention in the target article and in the subsequent discussion in these pages. We hoped that our debt to Liddell's arguments and proposals would be obvious – clearly we adopt many of his insights (though not all) in our own views on the phenomenon.

We took as our starting point Liddell's statements about verbs published in his 1995 paper, where he describes three types of mental spaces within which pointing signs and indicating verbs are directed: Real Space, Surrogate Space, and Token Space. In this paper, Liddell came to one conclusion about the

directionality of verbs in all three types of space – it is not determined by a linguistic process. He found "there are no linguistic features identifying the location the hands are directed toward. Instead, the hands are directed toward the specific part of the referent's body in Real Space by non-discrete gestural means" (Liddell 1995: 26). "I concluded that those aspects of the phonetic details of signs that are used to refer to Real Space are therefore nonlinguistic features. . . . the location toward which signs are directed would not be described through linguistic features. . . . the same conclusion holds in describing the underlying structures for Surrogate Space. Linguistic features are used to describe handshape, movement, and certain aspects of the orientation of the hand, but not the locations signs are directed toward" (Liddell 1995: 31). "There is still no basis for an agreement analysis because the way verbs are directed in space does not reflect any type of correspondence between linguistic elements" (Liddell 1995: 31). "Once again, there is no basis for an agreement analysis because, even with tokens, the locations that signs are directed toward is not a reflection of any type of correspondence between linguistic elements" (Liddell 1995: 35).

These strong claims led to a great deal of thought in the sign linguistics field; the issues that Liddell raised in 1995 still resonate loudly. Of course, the field has developed since Liddell's 1995 paper was published, and naturally enough Liddell's own views have evolved. As we noted in the target article, in his 2003 book Liddell's stance is substantially different from the earlier work. He says, "A few years ago the issue of whether something was 'linguistic' or not seemed like an important question to me. It no longer does" (Liddell 2003: 138). Importantly, it is not his assessment of the nature of directionality that has changed; rather, he incorporates gesture into the domain of language, and in this way his later work treats directionality as 'linguistic'.

II. Is there a grammatically-constrained process deriving the directionality of verbs?

Having in mind the different views held about the relationship between language and gesture, we turn to considering whether the directionality of verbs is due to a grammatically-constrained process.

We understand Liddell's position on this question to be 'no'. In his 2003 book, he discusses the different forms available for different verbs, and concludes as follows: "the idea that a single lexical unit underlies all the possible

directional uses associated with a particular verb (e.g. INFORM, SAY-NO-TO) is unable to account for the distribution of $^{x{\rightarrow}y}$, $^{{\rightarrow}y}$, and first person object forms that actually occur" (Liddell 2003: 114).[2] His discussion of these forms leads us to think that he considers the lexicon to include whichever variants a verb permits. He does discuss some generalizations across forms (for example, the existence of a $^{-1}$ form implies the existence of the $^{{\rightarrow}y}$ form). On the basis of such generalizations, he concludes, "the $^{{\rightarrow}y}$ form is probably the most basic form. From it the $^{x{\rightarrow}y}$ form or the $^{-1}$ form can develop" (Liddell 2003: 116). We take this to indicate that he believes a diachronic process may lead to the development of new forms, but he does not discuss this in terms of an active synchronic process.

One reason for the difference between Liddell's view and ours has to do with the level of detail thought to be required as the output for a rule. In his commentary, Liddell calls for a precise formulation of the spatial behavior of indicating verbs. We have been more focused on the distribution of signs – including indicating signs – in sentences.

We see two relevant levels of analysis. At one level, what we see as a grammatical process abstractly matches features between a verb and its arguments. At this abstract level, a precise formulation of the form of the verb is not necessary. The question is not 'where exactly does the verb move', but 'is there a process accounting for the verb's movement?' Our claim is that there is such a process. As we illustrated in the target article and discussed in section I of this response, because the process interacts with other morpho-syntactic processes, it must take place in the morpho-syntactic component.

At some point, of course, the details of the morphological output of this process do need to be spelled out. At the least, these details will include lexically-specified height marking for individual verbs (cf. Schlenker's commentary on this point), a default first-person location (midline of signer's chest), and a list of the idiosyncratic first-person forms that must have additional lexical specifications. Exactly how these are worked into the grammar will depend in part on details of the analysis we have left open (see section IV of this response).

---

2   On Liddell's notation, $^{x{\rightarrow}y}$ represents a verb that indicates both a trajector and a landmark – in other words, a subject and an object; $^{{\rightarrow}y}$ represents a verb that indicates only a landmark – in other words, only an object.

Although we take Liddell to disagree with our conclusion that a grammatical process is involved, it seems that all the other commentators agree with us, either implicitly or explicitly, that there is. Each makes a different contribution to our understanding of this process, and to its grammatical nature. Our next question addresses in broad terms the nature of this grammatical process.

III.   Is this process properly considered person marking?

The primary goal of our target article was to show why we conclude that the grammatical process under consideration is indeed person marking. Various commentators agreed with us and provided additional data relevant to this question. The commentaries by Quer and by Rathmann & Mathur incorporate the notion of person-marking into proposals that treat directionality as agreement. Since we discuss the agreement analysis in more detail in section IV, we will return to their commentaries there.

Aronoff & Padden support the person-marking analysis. They provide an interesting discussion about how arbitrariness provides a way to identify person categories in sign, specifically first person. As we've noted already in section I, they also discuss the reasons why signed languages may lack a second person (see also Rathmann & Mathur on this). In addition, they address the question of where directionality comes from. They propose that the reason sign languages use directionality to make a distinction between first and non-first person comes from the nature of pointing. Non-first forms are deictic and constantly shift; but the first-person form is fixed at the locus of the signer.

Schlenker also supports our conclusion, providing additional evidence from Weak Crossover that (in his words) "directional verbs have an agreement/ pronominal component." We consider his use of the hedge 'agreement/ pronominal' to indicate his endorsement of our claim that directionality marks 'person', while (like us) refraining from a specific morphological analysis.

We are grateful to Schlenker for providing a possible semantic analysis that would allow for the interpretation of person marking even with the iconic elements that the incorporation of gesture entails. Schlenker's comments address one of Liddell's frequent observations: for a referent present in the discourse situation, that referent's actual physical location (usually) determines the R-locus associated with him/her, both in terms of the location on the horizontal plane and, in some cases, with respect to the relative height of the refer-

ent (i.e., the placement of the referent in three dimensions). Schlenker captures these observations by use of *presuppositional* rules. Such rules should account for the unacceptability of using some locus other than the locus of the physically present referent even in certain contexts in which the referent might be imagined to be elsewhere. However, we wonder what the rationale is for using a presuppositional rule in particular – does the failure to use the appropriate locus have the status of a presuppositional failure? In this regard, we consider Quer's claim that use of the wrong (non-first) locus results not in ungrammaticality, but in infelicity. Is this a type of presuppositional failure? Or is it that such an utterance simply has a different meaning by virtue of picking out a different referent?

   In sum, the target article and the commentaries point rather forcefully toward an analysis in terms of *person marking*. We hope that establishing this characterization of directionality will enable the focus of debate to shift. Finally, we turn to the more precise question to which we did not offer a definitive answer: is this process in some technical sense *agreement*? The comments permit us to emphasize some final points for consideration.

IV.   Is *agreement* the optimal explanation for the mechanics of this process?

With our focus on characterizing (certain types of) directionality as person marking, we deliberately put to the side an extended discussion of location marking, and thus we avoided a direct answer to the question of whether one or two (or three) processes are involved. Rathmann & Mathur argue specifically for three types of agreement: person/number, locative, and class (gender). Leaving aside the latter completely, let us consider Rathmann & Mathur's argument that person marking and locative marking are distinct and compare their view to Quer's claim of the opposite. To some degree, the different claims of these authors are due to their differing focuses of attention: Rathmann & Mathur include morpho-phonological realization among the phenomena they hope to account for, while Quer focuses on morpho-syntactic properties. However, the differences between the proposals are deep and must be considered further.

   There are some empirical disagreements between Rathmann & Mathur and Quer, particularly concerning the possibility of combining person/number

marking and locative marking. Furthermore, Quer argues (following Quadros & Quer 2008, 2010; see also Steinbach's commentary) that verb classes cannot be straightforwardly predicted on existing thematic accounts – something that Rathmann & Mathur contest.

This is not the place to settle the dispute, but we would add some observations to the mix. We agree with Quer that verbs cannot be lexically classified simply as 'person' or 'location' agreeing verbs, because they appear with different types of markings in different syntactic environments. However, it is not clear that this leads to the conclusion that locative marking can be analyzed syntactically as 'agreement' equivalent to person marking, since person 'agreement' is generally controlled by verbal arguments, but locative marking is controlled by adjuncts. Furthermore, differences in the behavior of auxiliaries and backwards verbs raise questions for whether a unified account is even possible for person marking. Perhaps what Quer has in mind is that locative agreement is simply the case of the more general agreement process that appears when the feature agreed with is a locative feature rather than a person feature. It will be interesting to see how this approach can capture the similarities and differences between person and location marking that arise in differing levels of linguistic analysis, from syntactic to phonological.

The 'agreement' analysis includes number as an agreeing feature. Rathmann & Mathur point out important interactions between number and person marking, providing nice evidence for the relevance of the number feature. However, Aronoff & Padden find our arguments for number marking to be weaker than those for person marking. Apparently, they take the compositionality of the non-first plural forms as insufficient evidence for a morphosyntactic category of plural. In regard to first person plural, they bring up the Associative Plural Generalization (cf. Wechsler 2010), an observation based on examination of hundreds of spoken languages. Universally, the semantics of first- and second-person plural forms is special – 'associative'. On the assumption that sign language first person plural forms have a similar semantics (an assumption we do not contest), Aronoff & Padden find the existence of first-person 'plural' forms in sign languages a 'misnomer', and insufficient evidence for the morphosyntactic reality of a true plural. Even if this semantic concern raises doubts regarding the nature of the marking called plural, it clearly interacts with person marking in the way that Rathmann & Mathur describe.

This issue is beyond what we can deal with in the present discussion, though we refer readers to Bobaljik (2008) for an account of the APG which includes

sign language data. Wechsler's proposals make a clear distinction between second and third person forms, but may – as Aronoff & Padden suggest – be compatible with the sign language facts. Certainly this is another ripe area for additional research.

Steinbach further supports the agreement analysis by discussing in some detail the nature of the auxiliary signs found in some sign languages. As Steinbach argues, in some languages auxiliary signs can be analyzed as pure subject/object agreement markers. In this sense they are much less problematic for the agreement analysis than directional verbs are. For example, he notes (as does Quer) that auxiliary signs move from subject to object even when occurring with 'backwards' verbs (see section 7.3 of the target article, where we mention this observation as made by Quadros & Quer 2010). Steinbach cites other interesting characteristics of the DGS auxiliary PAM: specifically, its interpretation as emphatic when co-occurring with a verb marked with directionality, and its use to extend the argument structure of verbs. We look forward to the continued contributions that the study of sign language auxiliaries will make to our understanding of sign language morpho-syntax. The presence versus absence of auxiliaries is an important typological difference within signed languages.

Finally, we arrive at the crux of the 'agreement' analysis – what exactly counts as an analysis in terms of 'agreement'? Cysouw provides an engrossing historical tour of the use of the term 'agreement' and related terms, and concludes that directionality can be characterized as person marking, but not as agreement, at least when the term agreement is used in its historically appropriate sense. It would be very helpful to have a more extended discussion of Cysouw's proposed alternative: cross-reference. As Cysouw recognized, our use of the term 'agreement' was meant within the 'agreement/inflection' approach; on the other hand, we intended to be somewhat vague about implementation, focusing our argument on the evidence for person marking.

We cited Siewierska's (2004) use of the term 'person marking' on purpose, since she collapses verbal morphology with weak or clitic pronouns. We intended to anticipate a possible analysis of sign language person marking using a grammatical mechanism of cliticization. We are pleased to see Nevins' attempt at just such an approach. As Nevins argues, the clitic analysis gets around many of the problems for the agreement (inflection) approach, including the subject/object asymmetry and the existence of a highly similar (the same?) process for location marking. We hope that Nevins' proposal will be followed up

with attempts to account for the wide range of observations that have been made by the many researchers exploring the nature of 'directionality' in sign languages.

References

Bobaljik, Jonathan D. 2010. Missing persons: A case study in morphological universals. *The Linguistic Review* 25, 203–230.

Emmorey, Karen. 1999. Do signers gesture? In L. Messing & R. Campbell (eds.), *Gesture, speech, and sign*, 133–159. New York: Oxford University Press.

Goldin-Meadow, Susan & David McNeill. 1999. The role of gesture and mimetic representation in making language the province of speech. In Michael C. Corbalis & Stephen E. G. Lea (eds.), *The descent of mind: Psychological perspectives on hominid evolution*, 155–172. New York: Oxford University Press.

Liddell, Scott K. 1995. Real, surrogate, and token space: Grammatical consequences in ASL. In Karen Emmorey & Judy Reilly (eds.), *Language, gesture, and space*, 19–41. Hillsdale, NJ: Lawrence Erlbaum Associates.

Liddell, Scott K. 2003. *Grammar, gesture, and meaning in American Sign Language.* Cambridge: Cambridge University Press.

Mathur, Gaurav. 2000. *Verb agreement as alignment in signed languages*. Cambridge, MA: Massachusetts Institute of Technology dissertation.

Mathur, Gaurav & Christian Rathmann. 2001. Why not GIVE-US: articulatory constraints in signed languages. In V. Dively, M. Metzger, S. Taub, & A. Baer (eds.), *Signed languages: Discoveries from international research*, 1–26. Washington, DC: Gallaudet University Press.

McNeill, David. 1992. *Hand and mind: What gestures reveal about thought.* Chicago: University of Chicago Press.

Okrent, Arika. 2002. A modality-free notion of gesture and how it can help us with the morpheme vs. gesture question in sign language linguistics (Or at least give us some criteria to work with). In Richard P. Meier, Kearsy Cormier & David Quinto-Pozos (eds.), *Modality and structure in signed and spoken languages*, 175–198. Cambridge: Cambridge University Press.

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

Quadros, Ronice Müller de & Josep Quer. 2008. Back to back(wards) and moving on: On agreement, auxiliaries and verb classes in sign languages. In Ronice Müller de Quadros (ed.), *Sign languages: Spinning and unraveling the past, present and future*, 530–551. Petrópolis, Brazil: Editora Arara Azul.

Quadros, Ronice Müller de & Josep Quer. 2010. A caracterização da concordância nas línguas de sinais. In Heloisa Maria Moreira Lima-Salles & Rozana Reigota Naves (eds.), *Estudos gerativos de Língua de Sinais Brasileira e de aquisição de Português (L2) por Surdos*, 33–58. Goiânia, Brazil: Cânone Editoração.

Sandler, Wendy. 2009. Symbiotic symbolization by hand and mouth in sign language. *Semiotica* 174, 241–275.

Siewierska, Anna. 2004. *Person*. Cambridge: Cambridge University Press.

Wechsler, Stephen. 2010. What 'you' and 'I' mean to each other: Person indexicals, self-ascription, and theory of mind. *Language* 86, 332–365.