TOPICS IN COGNITIVE SCIENCE

Topics in Cognitive Science 7 (2015) 2–11

ISSN:1756-8757 print / 1756-8765 online
DOI: 10.1111/tops.12127

The Influence of the Visual Modality on Language Structure and Conventionalization: Insights From Sign Language and Gesture

Pamela Perniss,[a†] Asli Özyürek,[b†] Gary Morgan[c]

[a]*School of Humanities, University of Brighton*
[b]*MPI for Psycholinguistics & Radboud University Nijmegen*
[c]*Deafness, Cognition and Language Research Centre & City University London*

Received 31 October 2014; received in revised form 13 November 2014; accepted 13 November 2014

Abstract

For humans, the ability to communicate and use language is instantiated not only in the vocal modality but also in the visual modality. The main examples of this are sign languages and (co-speech) gestures. Sign languages, the natural languages of Deaf communities, use systematic and conventionalized movements of the hands, face, and body for linguistic expression. Co-speech gestures, though non-linguistic, are produced in tight semantic and temporal integration with speech and constitute an integral part of language together with speech. The articles in this issue explore and document how gestures and sign languages are similar or different and how communicative expression in the visual modality can change from being gestural to grammatical in nature through processes of conventionalization. As such, this issue contributes to our understanding of how the visual modality shapes language and the emergence of linguistic structure in newly developing systems. Studying the relationship between signs and gestures provides a new window onto the human ability to recruit multiple levels of representation (e.g., categorical, gradient, iconic, abstract) in the service of using or creating conventionalized communicative systems.

Keywords: Visual modality; Sign language; Gesture; Language emergence; Conventionalization; Language structure; Cross-linguistic comparison; Cross-modal comparison

1. Introduction

For humans, the ability to communicate and use language is instantiated not only in the vocal modality but also in the visual modality. The main examples of this are sign languages and (co-speech) gestures. Sign languages, the natural languages of Deaf

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

[†]These authors have contributed equally to the issue introduction.

communities, use systematic and conventionalized movements of the hands, face, and body for linguistic expression (Brentari, 2010; Emmorey, 2002; Klima & Bellugi, 1979; Stokoe, 1960). Co-speech gestures, though non-linguistic, are produced in tight semantic and temporal integration with speech and constitute an integral part of language (Kendon, 2004; McNeill, 1992, 2005). Thus, language—in its primary instantiation as a system of communication in contexts of face-to-face interaction—is a multimodal phenomenon (Vigliocco, Perniss, & Vinson, 2014). As such, our models of language need to take these visual modes of communication into account and provide a unified framework for how the semiotic and expressive resources of the visual modality are recruited in both spoken and sign languages.

This issue brings together researchers who work at the interface of sign and gesture and whose research illuminates two main areas of current debate and interest: (a) How and to what extent is gesture (with or without speech) similar to or different from sign language? and (b) How can the process of conventionalization from gesture to sign be characterized, both with respect to emerging linguistic/communicative systems and in learning an established sign language? In this introduction, we first situate the debates about the relationship between sign and gesture in a historical context. We then outline the state-of-the-art on this topic related to the two guiding questions. We also provide brief descriptions of how each of the papers in this issue contributes to these areas of research before ending with some discussion as to why these questions are of theoretical interest.

2. Relationship between sign language and gesture: Historical context

Since the linguistic study of sign languages began in earnest about half a century ago, a primary concern has been to show that sign languages, while exploiting forms and constructions that are visually similar to co-speech gestures, are clearly more than sequences of gestures. Sign languages are fully fledged natural languages that exhibit linguistic structure at all levels of formal description (phonological, morphological, and syntactic), and whose organization is supported by a similar neural architecture as is found for spoken languages. Partially, as a result of the need to establish the status of sign languages as natural human languages, most of the sign language research to date, as well as much of the gesture research, has emphasized the *differences* between signs and gestures with respect to linguistic and semiotic properties and conventionalization of form (Özyürek, 2012). In parallel with this, research has emphasized the *similarities* between sign language and speech, both in terms of linguistic structure and language processing (see Emmorey, 2007 for an overview).

However, more recently, attention has shifted to an interest in understanding the extent to which affordances of the visual modality may give rise to *similar* representations by signers and gesturers. This shift aims to understand more about the role of modality in shaping communication, in general, and in shaping sign language and gesture use, in particular. In both signing and gesturing, the use of the hands allows visually motivated,

iconic representations of objects, events, and spatial relations, which can exhibit a high degree of resemblance between form and meaning. The possible similarities between signs and gestures in these types of representations have important implications for theoretical questions about shared event conceptualization and underlying mental imagery (Liddell, 2003; Schembri, Jones, & Burnham, 2005); about the involvement of sensorimotor systems in language processing (Barsalou, Simmons, Barbey, & Wilson, 2003; Hostetter & Alibali, 2008, 2010); and about the role of iconicity in language evolution, development, and processing (Perniss, Thompson, & Vigliocco, 2010; Perniss & Vigliocco, 2014).

In a second major domain of interest, the affordances of the visual modality have been studied with respect to the emergence of linguistic/communicative systems, as found in homesign systems (Goldin-Meadow, 2003) and in new sign languages like Nicaraguan Sign Language (NSL) (Senghas, Kita, & Özyürek, 2004) or Al-Sayyid Bedouin Sign Language (ABSL) (Sandler, Meir, Padden, & Aronoff, 2005). Sign languages are thought to have evolved out of non-linguistic gestural communication, and in the emergence of linguistic/communicative systems the conventionalization of gesture into sign is a documentable process.

3. How does (co-speech) gesture resemble or not sign languages?

As outlined above, gestures and sign languages use the same modality and thus share access to the possibilities of visual representation afforded by the use of the hands in a visible space. Recent research, including papers in this issue, attempts to answer broader questions regarding how access to these affordances of visual–spatial representation shapes expression—including reference to objects, actions, and the relations between them, either in single forms or in more complex constructions. However, it is also clear that sign language and gesture are produced within linguistic/communicative contexts that differ in important respects. While gestures are used in conjunction with the linguistic structure of speech, the visual signal in sign languages is the sole channel of expression and the signs themselves are part of a complex grammar. Does this difference result in mere quantitative differences in how the visual–spatial modality is used for communicative expression or in more profound qualitative differences (see e.g., Brentari, Coppola, Mazzoni, & Goldin-Meadow, 2012 on the emergence of phonological structure)? The comparison of sign language and co-speech gesture can provide important insights into the role of modality in shaping language structure in different communicative contexts and the possible shared cognitive basis for communication using the visual modality.

As originally proposed by McNeill (1992), the comparison of sign language and co-speech gesture can shed light on the interplay of gestural (imagistic) and linguistic forms in communicative expression (see McNeill, 1992 on the shared contribution of gradient/imagistic and discrete/morphological content to language). In spoken language, gestural and linguistic forms constitute a tightly integrated unit (as research on both speech–gesture production and comprehension has shown, e.g., Kelly, Özyürek, & Maris,

2010; Kita & Özyürek, 2003; McNeill, 2005), but they remain clearly separable from each other by virtue of being produced in different channels. Sign languages are similarly characterized by the use of both gestural and linguistic forms, but the fact that all of sign language expression takes place in the visual modality has consequences for how these elements may co-occur. On the one hand, signers may intersperse the stream of (linguistic) signs with gestures (Emmorey, 1999). On the other hand, many morphologically complex signs (e.g., classifier predicates, directional verbs) have been argued to combine both gestural and linguistic elements (Liddell, 2003).

However, the extent to which these kinds of comparisons between sign and gesture can be made has also been questioned. Kendon (2008) has cautioned against too readily deriving conclusions about the "gestural" nature of (certain domains of) sign language and has emphasized the need for separate evaluation of co-speech gestures and signs in their respective contexts of use (i.e., a composite system with speech in the case of gesture and a fully visual system in the case of sign). Note that this does not argue against the notion that shared cognitive systems supporting representations in the visual modality may give rise to similarities between sign and gesture, nor does it suggest that comparisons between sign and gesture should be abandoned altogether. Rather it encourages careful consideration of the different semiotic contexts in which visual representations occur in the signed and spoken language modalities (Green & Wilkins, 2014; Kendon, 2014).

A number of the papers in the current issue address the question of similarities and differences between visual representations used by signers and speakers. In the contributions by *Quinto-Pozos & Parrill* and *Perniss & Özyürek*, comparisons are made between signs and co-speech gestures in two core domains of discourse: event representation and reference tracking. *Quinto-Pozos & Parrill* find strong similarities in the use of viewpoint-taking strategies in sign and co-speech gesture in a comparison of narratives in American Sign Language (ASL) and English. They demonstrate the existence of consistent correspondences in signers and co-speech gesturers between the strategy for viewpoint encoding and the type of event encoded. The implications of their findings are discussed in terms of indicating shared conceptualization of space and shared generation of mental and motor imagery for the purposes of communication, despite the different constraints on how the visual modality is used in a sign language versus in co-speech gesture. *Perniss & Özyürek* describe features for maintaining referential cohesion in the visual modality in a comparison of narratives in German Sign Language (DGS) and German co-speech gesture. They find that both signers and co-speech gesturers use spatial modification to mark referential context by associating referents with certain locations in space. However, they show that the two systems differ markedly in the nature and type of spatial modification exhibited. The differences are discussed in terms of the different semiotic contexts of sign and co-speech gesture: Whereas gesturers can rely on speech to carry the burden of reference tracking, signers must rely fully on the visual modality and thus make more use of its spatial affordances for maintaining discourse cohesion.

The paper by *Marshall & Morgan* argues that studying the forms that hearing speakers use in the early stages of learning a sign language can reveal how a gestural repertoire can scaffold learning to use linguistic structures in a sign language. Specifically, *Marshall*

& Morgan compare spatial descriptions by hearing, English-speaking adult learners of British Sign Language (BSL) to those by Deaf adult native signers of BSL. The study examines the role of gestural representation in learning iconic classifier morphology in sign language, providing insight into the challenges of learning the conventionalized structure of these iconic forms. The aspects of the sign language that were more easily learned were those that bore similarities to gesture use, notably location representation. The aspects that were harder to learn were those that were rarely used by gesturers, specifically, the use of distinct handshapes to represent different object types. Thus, where possible, learners of a sign language recruited those aspects of spatial expression that are shared between sign and gesture.

Another approach to understanding how the visual modality shapes language structure is to compare signs with silent gesturing, that is, gestures made without speech. The term "gesture" is sometimes used in reference to either co-speech gesture or silent gesturing, but it is important to distinguish between the two, as they denote very different contexts of use and imply the engagement of different processes. Co-speech gestures are a natural accompaniment to speech, and are made by speakers unwittingly while speaking. Silent gesturing, on the other hand, removes the expressive dominance and influence of speech and has been shown to differ in its patterning of expression from co-speech gesture (Goldin-Meadow, McNeill, & Singleton, 1996). When gestures are used without speech, they take on structure that resembles that found in many sign languages, for example, in the ordering of event constituents (Goldin-Meadow, So, Özyürek, & Mylander, 2008).

In the present issue, two papers (the contributions by *Padden, Hwang, Lepic & Seegers* and *Brentari, Di Renzo, Keane, & Volterra*) compare silent gestures used by speakers across different cultures to signed expressions by signers in the same cultures. Silent gesturing allows researchers to understand the visual strategies that speakers resort to in order to convey meaning when the visual modality becomes their only expressive resource. In this way, the use of silent gesturing constitutes an approximation of an important factor that contributes to the emergence of sign languages, namely use of the visual modality as the primary means of communication.

Padden, Hwang, Lepic, & Seegers describe the use of two iconic strategies in ASL signs for man-made tools: a handling strategy, where the hands depict holding or grasping an object; and an instrument strategy, where the hands represent the shape or a dimension of an object. They show that hearing non-signers use these same iconic strategies when asked to name man-made tools using gestures only. Moreover, they show that signers and (silent) gesturers alternated between the handling and instrument strategies for describing objects displayed in pictures versus in action videos, pointing to a common cognitive basis for differentiating objects from actions. However, the signers' choice of iconic strategy was more systematic compared to gesturers, suggesting that "patterned iconicity" can be exploited for grammatical purposes, in this case, for marking the distinction between nouns and verbs.

The paper by *Brentari, Di Renzo, Keane, & Volterra* investigates handshapes used in agentive versus non-agentive event descriptions in ASL and Italian Sign Language (LIS) by adults and children as well as in the corresponding groups of gesturers in each country

using gesture without speech. The findings parallel findings by Padden et al. (this issue) in that both signers and gesturers, and across languages, exhibit the use of handling handshapes to describe agentive events (in which an agent is acting on an object), but use an object handshape to describe non-agentive events. They discuss this similar pattern in terms of shared cognition driving the conventionalization of a distinction of handshape type. They also find influences of culture: the handshape distinction is found to be more pronounced in Italian gesturers compared to American gesturers, suggesting a higher sensitivity to gestural form-meaning pairings in Italian gesturers due to the gesture-rich culture. Finally, differences between LIS signers and ASL signers in marking the distinction are explained by linguistic effects.

4. How can the process of conventionalization from gesture to sign be characterized?

The discovery of communities using emergent sign languages (differing in number of generations of signers and varying in community size) as well as of homesigning individuals in different parts of the world have provided new insights into the emergence of language (Goldin-Meadow, 2003; Sandler et al., 2005). Specifically, these cases can shed light on the conventionalization of linguistic structure in the visual modality from non-linguistic gestural origins, where gesture is a substrate for sign (Janzen & Shaffer, 2002; Wilcox, Rossini, & Pizzuto, 2010). Factors explored with respect to the process of conventionalization from gesture to sign include the age of exposure to and the amount of time spent using the visual modality as the primary modality of communication, and the influence of number and kind of communication partners (i.e., large vs. small community of users and Deaf–Deaf vs. Deaf–hearing interactions). In addition, the existence of multiple generations of language learners/users, where conventionalized structure is passed from one generation to the next, is an important factor in the emergence of a sign language.

In comparing sign and (co-speech) gesture from the perspective of conventionalization from gesture to sign, investigation of the degree of conventionalization can reveal new insights into lexicalization, linguisticization, and grammaticalization processes. Papers in this section look at the emergence and conventionalization of sign language structure from "gestural origins." *Haviland*, on a homesign community in highland Chiapas, Mexico, and *de Vos*, on a village sign language in Bali, describe how co-speech gestures—summoning and pointing gestures, respectively—used by the surrounding speaking community take on grammatical properties in the sign language. *Goldin-Meadow* provides a window into language creation by observing manual forms used to describe actions over three time spans of use of the visual modality: hearing speakers asked to use gesture only, homesigners, and signers of an established sign language.

Haviland investigates the emergence of a new sign language (Zinacantec Family Homesign) across two generations of a single family in a remote Mayan Indian village. Haviland demonstrates a grammaticalization path from a co-speech gesture meaning

"come," commonly used in the surrounding Tzotzil-speaking community, to a turn-taking marker in the emergent sign language. The data show how interactive and communicative constraints converge to drive the conventionalization of a holophrastic gesture to grammaticalized linguistic elements.

De Vos examines pointing signs in spontaneous conversations in Kata Kolok, a village sign language in Bali. She argues that pointing signs may become an intrinsic aspect of sign language grammars through two mechanisms: morphemization and syntactic integration. The analysis provides an understanding of the mechanisms of conventionalization from gesture to sign that may contribute to the emergence of village sign languages such as Kata Kolok. In addition, the analysis suggests the possibility of grammaticality in highly systematized pointing systems used in some speaking communities.

Finally, the paper by *Goldin-Meadow* draws a general and unifying picture of the topic of gesture to sign conventionalization. The paper contrasts manual forms for actions produced by silent gesturers who are asked to invent gestures on the spot; by homesigners who have created gesture systems over their life spans; and by signers who have learned a conventional sign language from other signers. She finds that properties of the predicate (particularly, the use of location to establish co-reference, the representation of path and manner components, and the use of handshape distinctions) differ across these three time spans. These findings offer unique insight into the creation of language from gestural input and argue for the importance of a community of users who provide linguistic input and enable the transfer of conventional systems over generations of users.

5. Conclusions

Taken together, by examining linguistic/communicative expression in the visual modality, the papers in this issue contribute to our understanding of how the visual modality shapes language and the emergence of linguistic structure in newly developing systems. Studying the relationship—the similarities and differences—between signs and gestures provides a window into the human ability to recruit multiple levels of representation (e.g., categorical, gradient, iconic, abstract) in the service of using or creating conventionalized communicative systems. This research clearly demonstrates that no matter which channel of transmission is dominant or preferred in different systems of communication, our human language capacity is multimodal in nature and conveys information at different semiotic and representational levels.

In further specifying the interplay of these multiple levels of representations in speakers' and signers' recruitment of the visual modality for linguistic/communicative expression, the papers in this issue demonstrate that gesture (with or without speech) and sign exhibit similarities in the visual representation of information, possibly due to shared conceptualizations of space and shared mental and motor imagery of events. The papers in this issue also show that the differences between sign and gesture, on the other hand, are attributable to use of the visual modality as the sole modality of expression carrying the full burden of communication (as in sign) or as part of a composite system together with

speech (as in gesture). The current collection of papers is notable in the range of data that are represented: from different established sign languages (including urban and rural varieties), emerging sign systems, homesign systems, different spoken languages, as well as gestures with and without speech from different communities. In addition, the papers investigate a range of core domains of communication and aspects of representation, including reference tracking, event representation, pointing, use of viewpoint, action and object representation, and turn-taking in conversational interactions.

The studies in this volume make clear that further careful research is required to understand the role that the visual modality plays in sign versus spoken languages and to further our insights into the cognitive influences on language structure and language emergence. We hope that this collection of papers will help to facilitate further fruitful exchanges between gesture and sign language researchers, taking both similar and different theoretical standpoints (see also Green, Kelly, & Schembri, 2014). Finally, it is important to note that the field of (comparative) gesture and sign language research is still in its early stages and that more research on different sign languages and on the co-speech gestures used by speakers of different spoken languages is needed to better understand the fundamental features of our capacity for language in its multimodal form.

Acknowledgments

This work was supported by a Dutch Science Foundation (NWO)-VIDI grant (no. 276-70-009) awarded to A.O. and supporting P.P. as a postdoc; by a European Research Council (ERC) Starting Grant awarded to A.O. and by the Economic and Social Research Council (ESRC) of Great Britain (grant no. RES-620-28-6002) to the Deafness, Cognition and Language Research Centre (DCAL) supporting P.P. and G.M.

References

Barsalou, L. W., Simmons, K., Barbey, A. K., & Wilson, C. D. (2003). Grounding conceptual knowledge in modality-specific systems. *Trends in Cognitive Science*, *7*(2), 84–91.

Brentari, D. (2010) (Ed.). *Sign Languages*. Cambridge language surveys. Cambridge, England: Cambridge University Press.

Brentari, D., Coppola, M., Mazzoni, L., & Goldin-Meadow, S. (2012). When does a system become phonological? Handshape production in gesturers, signers, and homesigners. *Natural Language and Linguistic Theory*, *30*(1), 1–31.

Emmorey, K. (1999). Do signers gesture? In L. Messing & R. Campbell (Eds.), *Gesture, speech, and sign*. Oxford, England: Oxford University Press.

Emmorey, K. (2002). *Language, cognition, and the brain: Insights from sign language research*. Mahwah, NJ: Lawrence Erlbaum and Associates.

Emmorey, K. (2007). The psycholinguistics of signed and spoken languages: How biology affects processing. In G. Gaskell (Ed.), *The oxford handbook of psycholinguistics* (pp. 703–721). Oxford, England: Oxford University Press.

Goldin-Meadow, S. (2003). *The resilience of language: What gesture creation in deaf children can tell us about how all children learn language*. New York: Psychology Press.

Goldin-Meadow, S., McNeill, D., & Singleton, J. (1996). Silence is liberating: Removing the handcuffs on grammatical expression in the manual modality. *Psychological Review*, *103*, 34–55.

Goldin-Meadow, S., So, W.-C., Özyürek, A., & Mylander, C. (2008). The natural order of events: How speakers of different languages represent events nonverbally. *Proceedings of the National Academy of Sciences*, *105*(27), 9163–9168.

Green, J., Kelly, B., & Schembri, A. (2014). Finding common ground: Sign language and gesture research in Australia. *Australian Journal of Linguistics*, *34*(2), 185–192.

Green, J., & Wilkins, D. P. (2014). With or without speech: Arandic sign language from central Australia. *Australian Journal of Linguistics*, *34*(2), 234–261.

Hostetter, A. B., & Alibali, M. W. (2008). Visible embodiment: Gestures as simulated action. *Psychonomic Bulletin & Review*, *15*(3), 495–514.

Hostetter, A. B., & Alibali, M. W. (2010). Language, gesture, action! A test of the gesture as simulated action framework. *Journal of Memory and Language*, *63*, 245–257.

Janzen, T., & Shaffer, B. (2002). Gesture as the substrate in the process of ASL grammaticization. In R. P. Meier, D. Quinto-Pozos, & K. Cormier (Eds.), *Modality and structure in signed and spoken languages* (pp. 199–223). Cambridge, England: Cambridge University Press.

Kelly, S. D., Özyürek, A., & Maris, E. (2010). Two sides of the same coin: Speech and gesture mutually interact to enhance comprehension. *Psychological Science*, *21*(2), 260–267.

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

Kendon, A. (2008). Some reflections on the relationship between 'gesture' and 'sign'. *Gesture*, *8*(3), 348–366.

Kendon, A. (2014). Semiotic diversity in utterance production and the concept of 'language'. *Philosophical Transactions of the Royal Society B*, *369*, 20130293.

Kita, S., & Özyürek, A. (2003). What does cross-linguistic variation in semantic coordination of speech and gesture reveal? Evidence for an interface representation of spatial thinking and speaking. *Journal of Memory and Language*, *48*(1), 16–32.

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

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

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

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

Özyürek, A. (2012). Gesture. In R. Pfau, M. Steinbach, & B. Woll (Eds.), *Sign language: An international handbook* (pp. 626–646). Berlin: Mouton.

Perniss, P., Thompson, R. L., & Vigliocco, G. (2010). Iconicity as a general property of language: Evidence from spoken and signed languages. *Frontiers in Psychology*, *1*, 1–15. doi:10.3389/fpsyg.2010.00227.

Perniss, P., & Vigliocco, G. (2014). The bridge of iconicity: From a world of experience to the experience of language. *Philosophical Transactions of the Royal Society B*, *369*, 20130300.

Sandler, W., Meir, I., Padden, C., & Aronoff, M. (2005). The emergence of grammar: Systematic structure in a new language. *Proceedings of the National Academy of Sciences*, *102*(7), 2661–2665.

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

Senghas, A., Kita, S., & Özyürek, A. (2004). Children creating core properties of language: Evidence from an emerging sign language in Nicaragua. *Science*, *305*(5691), 1779–1782.

Stokoe, W. (1960). *Sign language structure: An outline of the visual communication systems of the American Deaf*. Silver Spring, MD: Linstok Press.

Vigliocco, G., Perniss, P., & Vinson, D. (2014). Language as a multimodal phenomenon: Implications for language learning, processing, and evolution. *Philosophical Transactions of the Royal Society B*, *369*, 20130292.

Wilcox, S., Rossini, P., & Pizzuto, E. (2010). Grammaticalization in sign languages. In D. Brentari (Ed.), *Sign languages* (pp. 332–354). Cambridge Language Surveys. Cambridge, England: Cambridge University Press.