From Gesture in Conversation to Visible Action as Utterance

From Gesture in Conversation to Visible Action as Utterance

Essays in honor of Adam Kendon

Edited by

Mandana Seyfeddinipur
University of London

Marianne Gullberg
Lund University

John Benjamins Publishing Company
Amsterdam / Philadelphia

Table of contents

A foreword

This volume was born on a warm afternoon in July on a bench in front of a fountain in Lund. The Fifth Conference of the International Society for Gesture Studies in Lund was in full swing. We were musing about what Adam Kendon had meant to us individually as well as to us as a research community.

One of us remembered a first encounter that involved an abysmal curry at a university canteen accompanied by interested and interesting questions which encouraged the (somewhat nervous) novice to continue on the gestural path. Another remembered the professor who rather spent time with students than with other professors at a summer school in San Marino, the summer school where a pact was made that resulted in Adam's book in 2004. The afternoon musings also included anecdotes about editing felines (the collaborator Oscar Gatto), Jabberwocky recitals, discussions about must-see films, Indian food (Adam Kendon is known in certain circles as Mister Vindaloo), and the art of making proper tea. But the conversation was mostly concerned with gestures, utterance visible actions, long, exhilarating data sessions where videos were played and replayed, and analyzes characterized by rigour mixed with laughter.

To many of us Adam is a generous colleague, mentor, and friend. To the research community he is the source of many of our key notions and he has laid the foundation for rigorous research on the body in communication. He has thought and written about almost all aspects of this domain. An experience many of us share is that when you decide to embark on a new topic in the area, you discover that Adam has already published a groundbreaking paper about it complete with a sound methodology. Indeed, he has often addressed major issues that the field has only later picked up on. Many of us also share the happy experience of spending hours with Adam when he, generously, has discussed our work, often finding relevant and illuminating examples in his own data and butterfly collections. In his work on the journal *Gesture* he has guided our writing and our theoretical approaches with incredible patience, leaving his mark on many of us. Indeed, for us it was time to show him what his legacy means and we hope that this volume will show him some of the ways in which his influence is visible in contemporary research.

From gesture in conversation to visible action as utterance

Mandana Seyfeddinipur and Marianne Gullberg
School of African and Oriental Studies, University of London /
Lund University, Sweden

Language use is fundamentally multimodal. Speakers use their hands to point to locations, to represent content and to comment on ongoing talk; they position their bodies showing their interactional orientation; they use facial displays commenting on what is being said; and they engage in mutual gaze establishing intersubjectivity. Over the past five decades investigations into the nature of multimodality of language use have increased dramatically in different scientific areas like linguistics, sociology, psychology and cognitive science. One person who has played a major part in the development of this area of study is Adam Kendon. Kendon's work – always many years ahead of time – has laid the theoretical and methodological foundations for the study of multimodality. This volume brings together work by junior and senior researchers on the fundamentally multimodal nature of the human capacity for language. The papers highlight how Kendon's work has provided the foundation for modern rigorous research into the multimodal nature of human language. The papers also report on contemporary research in different areas from linguistics to social psychology to conversation analysis.

Kendon's investigations and the resulting publications have become seminal in many areas of visible action in utterance. His work addresses different aspects of social interaction like the spatial configuration of speakers in interaction, the use of gaze and facial gesture in conversation, and the form and function of manual gesture with a cross-cultural twist, comparing behavior in the UK and in Southern Italy. But he has also provided descriptions of a local sign language in Papua New Guinea and an alternate sign language of the Warlpiri in Australia. He has published on language origins, and on the history of gesture research, and he has translated the work by de Jorio, a priest, archeologist and gesture scholar in 19th century Italy. In all these diverse areas his work has been ahead of its time. His work continues to be of the utmost relevance for research today. His studies on

manual gesture, for example, have shown how speech and gesture contribute different types of information ranging from semantic meaning to pragmatic structuring and speech act marking in collaboration with speech. His detailed analyses have put the study of manual gesture on a rigorous empirical and methodological footing, leading away from impressionistic descriptions and interpretations and towards verifiable analyses of the phenomena at hand.

Kendon's work unites three major theoretical strands and approaches: structuralism, interactionalism, and ethnography. Uniting these strands has led him to conduct analyses of how the body is deployed in providing different kinds of meaning in conjunction with speech in everyday interaction in different cultures. Kendon has developed a **structuralist** analysis of form and function of posture and spatial organization, gaze, and manual and facial gestures in everyday interaction. His analysis involves a minute description of the temporal unfolding of the gestural movement in relation to speech units. The temporal integration and the form-function analyses are grounded in Birdwhistell's outlook and methodology in his development of kinesics. Inspired by Birdwhistell's observations on the systematic way in which visible body motion was organized in relation to speech (see Birdwhistell 1970, for example), Kendon worked on pieces of film that Birdwhistell had made available. This is what led to the paper "*Some relationships between body motion and speech*" (Kendon 1972), a paper which was to lay the foundations for the later development of 'Gesture Studies'. To this day, it remains one of the most complete attempts at addressing the issue of the organization of the flow of bodily movement and its tight fit with the structuring of associated speech.

The second characteristic of Kendon's work from the earliest days is his focus on how speakers coordinate their activities in everyday face-to-face **interaction** as the site of human communication. He conceived of and understood the spatial organization of episodes of interaction, and the criteria for defining interaction episodes, the structuring of the 'social occasion, and has discussed different 'behavior systems' (talk, gesture, posture, orientation, jointly constructed spatial-orientational systems) in terms of how they are articulated in relation to one another. In his work on greetings, for example, he came to see that the cooperation that participants enter into to create the spatial-orientational frame for the close salutation could be viewed as a naturally bounded unit of interaction. Subsequent spatial-orientational frames created in relation to other kinds of interaction 'projects' (such as conversations of various types or conversations about different topics) provided criteria for establishing further units of interaction. This work showed how units or episodes of interaction could be defined, but also the importance of distinguishing the different levels of behavioral organization in terms of which different interactional episodes, often overlapping with one another, can

be defined. Kendon applied this context-anchored approach to the study of gaze, facial and manual gesture in natural interaction, thereby providing the analytical framework for many researchers to come.

The third strand which characterizes Kendon's work is the **ethnographic approach**. Inspired by David Efron (1941/1972) and Wilhelm Wundt (1921/1973) Kendon turned his attention to the question of which factors determine cultural differences in manual gesturing. In his research on Neapolitan gesturing he attempted to understand why a particular tradition of gesture use in communication, a gestural profile, is maintained and where it comes from. His work on the gestural profile of Neapolitans took into account historical resources such as the work by de Jorio (Kendon 1995a), as well as the actual everyday use of gesture and the local affordances of communication. He suggests that one first has to look at what the modality of gesture affords its users as a means of communication. He therefore considered in detail the circumstances of its use, its ecological circumstances of daily interaction. Second, the prevailing norms for what governs behavior in co-presence have to be taken into account in order to describe the micro-ecology of everyday interaction which in turn determines gestural conduct and the resulting cultural differences.

The combination of these three strands has further informed Kendon's most recent work in which he considers how multimodal interaction, and specifically the deep intertwining of modalities in interaction, might be at the heart of the **evolution of language**. Contrary to much current theorizing, but entirely in line with his views from the earliest days, he considers speech and gestures to have co-evolved. The conviction that face-to-face interaction is the natural habitat of communication and that modalities co-perform is again central to his research.

Gaze and face

Kendon pioneered the study of the face in interaction in his 1975 paper "*Some functions of the face in a kissing round*" (Kendon 1975b). In this study he criticized studies of the face at the time for focusing primarily on the expression of emotion at the expense of the role and function of facial patterns in social interaction. He shows in his paper how facial gestures serve as a "delicate tuning device" (1975b, p. 330) regulating the interaction. In Kendon's tradition **Bavelas, Gerwing & Healing** present a careful analysis of facial gestures in speech gesture ensembles showing how facial gestures take on different functions in dialogue parallel to the functions of manual gestures described by Kendon. The paper exemplifies that facial gestures can serve referential as well as pragmatic, interpersonal, and interactive functions in conversation. The study highlights – even after Kendon's

groundbreaking work – how the study of the face continues to focus on emotion expression and neglect the social function of facial displays in interaction. In addition, the analysis takes into account the intricate temporal integration of the modalities, again following Kendon's work.

In 1967 Kendon published a paper on gaze in interaction, *"Some functions of gaze direction in two-person conversations",* which became a seminal study on gaze in interaction despite the fact that a substantial body of literature on gaze already existed. As often, Kendon was years ahead of his time, pioneering the investigation of gaze in conversation. He overcame the technological limitations of his time by setting up a mirror next to one speaker reflecting the face of the second speaker. The camera placed opposite took one picture every second and with the resulting photographs Kendon was able to relate the gaze direction of both speakers in time. This enabled the detailed temporal coding of gaze withdrawal and gaze return showing that the patterns of gaze withdrawal and gaze return are systematically distributed between speaker and interlocutor.

Streeck's paper builds on this work and on recent work by Rossano on mutual gaze in conversation. Using a micro-ethnographic approach to the phenomenon and drawing on examples of naturally occurring interaction, Streeck shows how gaze is deployed. He proposes that mutual gaze is a primitive form of the social contract between interlocutors. Streeck shows that mutual gaze displays the act of recognition and ratification within an action sequence. He argues that gaze shifts should be looked at as components of actions and that mutual gaze should be seen as part of sequence organization of interaction, going beyond the role of gaze in turn taking.

Manual gestures – Quotable gestures and pointing

A major part of Kendon' work has investigated spontaneous manual gestures. But he has also devoted his time to what he called quotable gestures (Kendon 1992) or emblems (Efron 1942). Morris et al. (1979) conducted an areal linguistic study of 20 conventionalized gestures – such as the thumbs up gesture – looking at their meanings and use all over Europe. This work drew Kendon's attention to such conventionalized movements.

Kendon labeled these gestures *quotable gestures* since they are repeatable, listable, and reportable. Speakers use these conventionalized gestures with and without speech and can be held accountable for using them. Kendon (1992) discusses how communities share repertoires of these fully conventionalized gestures. He points out that most studies only provide word list style accounts of the gesture forms and their associated meanings. At an early stage he called for the need

to study conventionalized gestures in their context of use. He provided such an analysis in his work on pragmatic gestures (Kendon 1995b) analyzing the use of some of these gestures in Italian. He showed how these gestures mark the speech act of the utterance while others mark discourse structure.

Kendon's call for studies of use in context was taken up by the work of Brookes who studied the use of emblems by young males in South African townships (Brookes 2001; 2004). Following in Kendon's tradition she extended her study by taking into account the social relationships, cultural notions, and identity shape forms of gestural use and behavior among black urban South African males. In the present paper **Brookes** explores Kendon's concept of a communicative ecology of a community and how communicative profiles are shaped through the physical environment and cultural norms. She analyses the profile of communicative behavior in South African townships by studying gesture in natural interactions and the underlying cultural norms, the physical surroundings, and their social meanings. She then adds a comparative analysis of the communicative profiles found in Naples and in South African townships.

In his paper **McNeill** similarly addresses quotable gestures and focuses on the Neapolitan quotable gestures described by Kendon (e.g. 1995b). He provides a detailed discussion of the underlying metaphoricity of these gestures. He argues that speakers use gestures created on the fly which are based on metaphor or metonymy. These root metaphors then undergo a conventionalization process through use within particular communities, thus creating particularly stable quotable gestures compared to spoken words which undergo drastic changes over the centuries.

In contrast to studies of highly culture-specific gestures such as the quotable gestures described above, other studies focus on claims of universality. Pointing is the best example of this. Contrary to such claims, Kendon's work on pointing in Naples (Kendon & Versante 2003) has shown how the pointing form (hand shape, orientation, place of articulation and trajectory) is systematically deployed to express different semiotic functions. Kendon & Versante had observed that when people engaged in what was generally recognized as pointing to something, they did not always use the same hand shape to do so. They collected examples of pointing to compare and contrast their contexts of use in terms of the hand shapes employed. Speakers can use different hand forms to provide an interpretative 'frame' to the verbal discourse it accompanies. In these uses of different hand shapes in pointing, then, the speakers are showing something about the type of discourse act they are engaging in even as, at the same time, they are engaging in an action of pointing at or indicating something.

Mondada, in the tradition of conversation analysis, expands the analysis of pointing by examining the organization of actions in which a speaker mobilizes

pointing and establishes joint attention with co-participants towards an object. In detailed analyses she shows the complexity of the act of pointing which affords high coordination between participants. While preserving the specificity of the ecology of action in its complexity, she at the same time demonstrates that the methodical mobilization of resources in interaction can be generalized.

Manual gestures – Their nature and relationship to language

As already mentioned, a core aspect of Kendon's work on manual gestures is the formal and structuralist approach by which he examines the temporal unfolding of gestural movements in relation to speech units in detailed form-function analyses. As a part of this enterprise, Kendon has kept returning to how it is that we define and recognize gestures. He addressed these questions already in his earliest studies, *"Some relationships between body motion and speech,"* published in 1972, and *"Gesticulation and speech: Two aspects of the process of utterance"*, published in 1980. In the latter paper in particular, he explored the idea that interlocutors are able to recognize movements as being deliberately expressive even when they do not understand the spoken language that accompanies them based on formal kinetic features, rhythm, etc., combined to create the impression of deliberate expressiveness.

Müller's paper elaborates on Kendon's structural and formal focus on movements displaying articulatory "features of manifest deliberate expressiveness" (Kendon 2004, pp. 13–14) and his interactionally grounded view that interlocutors can identify gestures into an argument for how linguistic structures can emerge from bodily movements. Müller discusses the relevance of Kendon's combined focus on form, context-of-use, and meaning as reflected in his notion of gesture families, which are form-meaning clusters. Müller expands on Kendon by discussing the dynamic embodied conceptual processes through which gestural forms come to mean and the modes of representation that result from these processes. She suggests that this overall approach points towards a grammar of gesture which reveals the potential of gestures to evolve into language.

Andrén similarly delves deeper into the question of how we identify gestures and distinguish them from other forms of semiotically relevant bodily behavior. Inspired by Kendon and discussions of the upper limits of gesture trying to distinguish gestures from the signs of sign language, Andrén explores what he calls a lower limit of gestures to distinguish, for example, gestures from practical actions. Moving away from the tradition of binary distinctions and building on Kendon's comparative semiotic approach (Kendon 2008), Andrén instead suggests that

distinguishing a continuum of communicative explicitness from a continuum of representational complexity can help us investigate complex interactions that help us define gestures in terms of family resemblance. He also suggests a continuum of conventionalization as being relevant.

Language evolution

The role of gesture in the origin of language and language evolution has occupied researchers over many centuries (Kendon 1991). As often before, Kendon was ahead of his time and discussed these issues already in the 1970s in the paper "*Gesticulation, speech, and the gesture theory of language origins*" (Kendon 1975a). Over the years he has criticized theories of language evolution that advocate a 'gesture first' explanation (Kendon 1975a; 1991; 1993; 2010) and propose that human language evolved through communicating through manual gesture first and then switched to the oral and auditory modality to facilitate communication over long distances. Kendon has questioned these theories on the grounds that it remains unclear why there should have been a switch of modality rather than a continued parallel use of hand and mouth.

In his paper, **Corballis** takes up this discussion and presents a 'gesture-first' position – albeit one that allows for a gradual shift of balance. He outlines arguments in its support drawing on comparisons with nonhuman primates, focusing on vocal and manual asymmetries where the greater degree of intentional and flexible use of manual actions suggest a more primary mode of expression which may have evolved into pantomime with gradual conventionalization into arbitrary symbols. Corballis's position contrasts in interesting ways with Kendon's and the engaged argumentation across the positions is clear and enlightening.

In contrast, **Goodwin** argues along the same lines as Kendon in his paper proposing that gesture is not sufficient as co-operative action but that the core of human language use requires the full multimodal power of speech and gesture. Goodwin draws upon interactions of an aphasic man to demonstrate how communicating for action moves from ambiguous gestures to speech through the development of arbitrary signs. Using examples from interactions between scientists he also shows how subsequent action is accumulatively built by performing structure-preserving transformations of the materials provided by a prior action. The complexity of the expressions speakers create by exploiting the available modalities forms the core of human communication.

Sign systems

Kendon was planning to study courtroom interactions in the Enga province of Papua New Guinea when he met Imanoli, a young deaf woman who was using a local sign language. The sign system sparked his interest and he embarked on the investigation of Enga sign language. He provided a detailed description of all of the signs in the repertoire of Imanoli, a detailed exploration of the "iconic devices" employed in Imanoli's signs, and the way in which discourse was constructed in this sign language – effectively a kind of syntactic study (Kendon 1980).

This work directed Kendon's attention to sign languages in general and he came to focus on the alternate sign language used by hearing Warlpiri speakers in Yuendumu, a Warlpiri community in north central Australia. His work resulted in the only book-length work on the topic available to this day, *Sign languages of Aboriginal Australia: Cultural, semiotic and communicative perspectives* (Kendon 1988). It presents a history of the study of sign languages in Australia, extensive ethnographic background to their use in the north central desert region of Australia, detailed discussions of the relationship between the structure of these sign languages and the structure of the associated spoken languages, and comparative analyses of the sign languages of six different Aboriginal groups. He also discusses kinship and sign language, and the relationship between alternate sign languages and primary sign languages. Finally, he also provides an analysis of the social and ecological circumstances that appear to favor the use of sign languages among Australian aborigines.

Green has continued this unique line of research. Her paper focuses on Kendon's question of how speakers utilize different modalities as a semiotic resource for expression in communication. Green analyses Arandic sand stories, a traditional form of verbal art uniquely mastered especially among Arrernte women in Central Australia. In this form of verbal art speakers draw in the sand, speak, gesture and sign. In a detailed analysis Green illustrates the temporal and semantic integration of the modalities, exemplifying the mastery of this Aboriginal art form.

Kendon's description of the hierarchical organization of body movements with respect to discourse units (1972) showed the temporal coordination of all bodily actions, ranging from body posture to head movement to manual gesture. His work provided a first detailed account of the syntagmatic organization of manual gestures through a functional analysis of manual movements. He showed that they could be distinguished into different movement phases with the stroke being the semantic nucleus of the gesture. He characterized manual gestural movements into hierarchically organized units characterizing the form features of each phase. This seminal work laid the foundation for studies of the temporal coordination of speech and gesture. **Haviland** uses Kendon's description of

phrasal organization of gestural movement and applies it to a first generation of sign language created spontaneously by three deaf siblings and their hearing age mates in an indigenous community in Mexico. His study demonstrates how this formal approach is required to derive appropriate analytical categories from the empirical materials.

Goldin-Meadow further draws inspiration from Kendon's work on signs and the relationship between signs, gestures and linguistic systems more generally. She presents an overview of a series of studies which have explored the emergent linguistic properties of the gesture systems developed by deaf children born into hearing families, so-called home signs; the linguistic properties of gestures produced by hearing people asked to rely on gestures only to communicate; and the properties of gestures that accompany speech. The careful juxtaposition of manual movements across these different contexts of use and populations reveals in great detail what characterizes movements that are more like signs and therefore like language from movements that are more like gestures.

Child language development

Kendon's work encompasses a great many domains and areas. One of the few that he has *not* worked on is development. Yet, in his thinking about the nature of gestures and their multifaceted deployment in interaction he has occasionally touched on the question of how children come to be competent interlocutors deploying speech and gestures in culturally appropriate ways in a given culture and language community. It is clear that children mobilize speech and gestures in different ways from adults, but it remains largely unknown how children become sophisticated multimodal agents and whether and if so how, the nature of the relationship between speech and gestures change over the course of development. Kendon may not have studied child language himself, but his approach on to how to study bodily communicative behaviors has predictably inspired researchers working on child development.

Graziano focuses on a Kendonian specialty, namely pragmatic gestures which do not express referential content, but rather comment on the production itself. Graziano discusses what Kendon (2004) calls the *Palm Presentation* or PP and *Palm with a Lateral Movement* or PL gestures and investigates the way in which Italian children aged between 4 and 10 produce these in narratives. The results indicate an important developmental shift in the use of such pragmatic gestures. Whereas the oldest children use them in similar ways to adults, the youngest children do not. Graziano suggests that the deployment of pragmatic gestures is a late development connected to children's growing capacity for structuring narratives

and rhetorical control over their own discourse. Her analyses shed light on a domain that has received relatively little attention in child development studies.

Cristilli examines a different aspect of how speech and gesture come together in development. Her paper focuses on how Italian children aged 6 and 10 deploy gestures in narratives as part of the effort to construct cohesive discourse and track who and what is being talked about over the course of discourse, something known as reference tracking. The analyses reveal that the relationship between spoken referential expressions (e.g., *the girl, she*) and gesture changes over development with younger children's gestures predominantly disambiguating spoken expressions, and older children's gestures instead supplementing speech to achieve redundancy.

Guidetti et al. take a more general view and discuss some fundamental questions concerning child language acquisition and gestures based on an overview of a series of studies that have examined children's speech and gestures across several ages groups, types of speech, and tasks. Building on discussions of speech and gesture evolution, they probe, for example, the continuity hypothesis suggesting a link between earlier and later forms of expression. Importantly, they also ask what it is that actually develops in the visible shift in children's gestural and multimodal behaviors documented in their previous work. This discussion lays the ground for a number of further studies that will be needed for us to better understand what children do that may or may not be different from adults' bodily communicative behavior.

And end to the beginning

In sum, the papers in this volume all testify to the enduring importance of Adam Kendon's work to multiple strands of contemporary work, and they eloquently reflect the influence of his thinking on gestures or utterance visible action, his current term for these behaviors, in a range of sub-fields. Many studies whose inspiration is clearly Kendonian did not make it into this volume. However, we hope that the selection that did clearly demonstrates Adam's influence on the vibrant new research field that is Gesture Studies. His work will continue to inspire new studies for years to come.

References

Birdwhistell, R. L. 1970. *Kinesics and Context. Essays on Body Motion Communication*. Philadelphia: University of Pennsylvania Press.

Brookes, H. J. 2001. "*O clever* 'He's streetwise.' When gestures become quotable." *Gesture* 1 (2): 167–184. DOI: 10.1075/gest. 1.2.05bro

Brookes, H. J. 2004. "A repertoire of South African quotable gestures." *Journal of Linguistic Anthropology* 14 (2): 186–224. DOI: 10.1525/jlin.2004.14.2.186

Efron, D. 1941/1972. *Gestures, Race and Culture* (First edition 1941 as Gestures and Environment. New York: King's Crown Press. ed.). The Hague: Mouton.

Kendon, A. 1967. "Some functions of gaze direction in two-person conversations." *Acta Psychologica* 26: 22–63. DOI: 10.1016/0001-6918(67)90005-4

Kendon, A. 1972. "Some relationships between body motion and speech: An analysis of an example." In *Studies in Dyadic Communication*, A. W. Siegman & B. Pope (eds), 177–210. New York: Pergamon. DOI: 10.1016/B978-0-08-015867-9.50013-7

Kendon, A. 1975a. "Gesticulation, speech, and the gesture theory of language origins." *Sign Language Studies* 9: 349–373. DOI: 10.1353/sls.1975.0016

Kendon, A. 1975b. "Some functions of the face in a kissing round." *Semiotica* 15: 299–334. DOI: 10.1515/semi.1975.15.4.299

Kendon, A. 1980. "Gesticulation and speech: Two aspects of the process of utterance." In *The Relationship of Verbal and Nonverbal Communication*, M. R. Key (ed.), 207–227. The Hague: Mouton.

Kendon, A. 1988. *Sign Languages of Aboriginal Australia: Cultural, Semiotic and Communicative Perspectives*. Cambridge: Cambridge University Press.

Kendon, A. 1991. "Some considerations for a theory of language origins." *Man* 26: 199–221. DOI: 10.2307/2803829

Kendon, A. 1992. "Some recent work from Italy on quotable gestures ('emblems')." *Journal of Linguistic Anthropology* 2 (1): 92–108. DOI: 10.1525/jlin.1992.2.1.92

Kendon, A. 1993. "Human gesture." In *Tools, Language and Cognition in Human Evolution*, K. R. Gibson and T. Ingold (eds), 43–62. Cambridge: Cambridge University Press.

Kendon, A. 1995a. "Andrea De Jorio – The first ethnographer of gesture?" *Visual Anthropology* 7: 375–394. DOI: 10.1080/08949468.1995.9966661

Kendon, A. 1995b. "Gestures as illocutionary and discourse structure markers in Southern Italian conversation." *Journal of Pragmatics* 23 (3): 247–279. DOI: 10.1016/0378-2166(94)00037-F

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. DOI: 10.1075/gest. 8.3.05ken

Kendon, A. 2010. "Some modern considerations for thinking about language evolution: A discussion of The Evolution of Language by Tecumseh Fitch." *SemiotiX* 2010 (3). http://www.semioticon.com/semiotix/files/2010/11/Fitch-Essay.pdf

Kendon, A., and Versante, L. 2003. "Pointing by hand in "Neapolitan"." In *Pointing: Where Language, Culture, and Cognition Meet*, S. Kita (ed.), 109–169. Mahwah, NJ: Erlbaum.

Morris, D., Collett, P., Marsh, P., and O'Shaughnessy, M. 1979. *Gestures, their Origins and Distribution*. London: Cape.

Wundt, W. 1921/1973. *The Language of Gestures* (J. S. Thayer, C. M. Greenlauf and M. D. Silberman, Trans. From *Völkerpsychologie: eine Untersuchung der Entwicklungsgesetze von Sprache, Mythus und Sitte*. Stuttgart: Alfred Kröner Verlag, 1921. Vol. 1, First part, Chapter 2. ed.). The Hague: Mouton.

Gaze and face

Including facial gestures
in gesture–speech ensembles

Janet Bavelas, Jennifer Gerwing and Sara Healing
Department of Psychology, University of Victoria

Conversational facial gestures fit Kendon's (2004) specifications of the functions of hand gestures. We illustrate how facial gestures in dialogue, like hand gestures, convey referential content as well as serving pragmatic, interpersonal and interactive functions. Hand and facial gestures often occur together, creating an integrated visual image in gesture–speech ensembles. A semantic features analysis demonstrates how speakers adjust their use of these visible versus audible expressive resources according to context. Speakers who were interacting face to face (compared to speakers who could not see their addressee) were significantly more likely to rely on their hand and facial gestures than on their words when describing key semantic features, and their gestures were more likely to convey information that was not in their words.

Part 1: "Ensembles of gestures and speech"

In this chapter, all Kendon quotations (including those used as section headings) are from his 2004 book.

> We shall see that speakers create *ensembles* of gesture and speech, by means of which a semantic coherence between the two modalities is attained. This is not to say that speech and gesture express the same meanings. They are often different. Nevertheless, the meanings expressed by these two components *interact* in the utterance and, through a reciprocal process, a more complex unit of meaning is the result. (pp. 108–109)

We begin by extending Kendon's theory of gesture–speech ensembles to include more than hand gestures in these ensembles. An example from a video-recorded medical consultation between physician and patient (Healing, unpublished data) will illustrate the nature of these expanded ensembles. The physician had asked about the patient's symptoms since they last met.

Patient: "Usually first thing in the morning, I'm kind of, 'Oh!' Trying to get things going, 'n yesterday I just – "
Physician: "So that's good."

Nothing the patient said seemed to answer the physician's question. He did not say how he "usually [is] first thing in the morning" or what it meant for him to be "trying to get things going," much less what he was like "yesterday." Yet the physician's reply ("So that's good") indicated that the patient's answer was both informative and positive. As shown in Table 1, it was the patient's animated and precisely timed gestures with hands, arms, shoulders, upper torso, face, head and eyes that conveyed all of this specific information. Thus, although gesture and speech expressed different meanings, they interacted in his utterances to create more complex units of meaning that were still semantically coherent.

Table 1. Patient telling his physician how his symptoms have been since his last visit

The patient began by depicting himself as he usually was when he woke up:

"Usually first thing in the morning, I'm kind of," [Physician nods]
[*hunching his shoulders forward, sitting stiffly, not moving– as if bracing himself; his face is frozen and tense*]

Then he enacted more about his usual morning:

Moaning "Ohhhh."
[*leaning over to one side, grimacing*]

He followed by showing how he usually had to try to begin moving:

"Trying to get things going," [Physician says, "m-hm"]
[*kneading his affected leg, as if trying to get it going*]

Then he quickly contrasted this description with how he felt yesterday:

"'n yesterday, I just –"
[*sits up straight and begins to mime a vigorus marching motion with his arms and legs; head held high with a determined face*]

The physician started to nod and to say "So that's good," and the patient simultaneously ceased these actions and smiled at the physician.

"Facial gestures"

While writing about the use of the kinesic medium in sign language, Kendon mentioned "facial gestures, such as eyebrow movements or positionings, movements of the mouth, head postures and sustainments and changes in gaze direction" (p. 310). His list includes most of what we consider facial gestures. More formally, *conversational facial gestures* are any configuration or movement of the face or of

the head (including the eyes) that is synchronized with speech in both timing and meaning. For example, the patient's grimace, leaning over, and the long "Ohhhh" all occurred together and all contributed to the picture of how he felt. Similarly, precisely with "'n yesterday I just – ," he held his head high with a determined look and marched with his arms and legs.

"Closing off further investigation"

Anyone who watches the faces of interlocutors in a dialogue will see a virtually constant succession of rapid and diverse facial gestures by both speaker and addressee. Yet conversational facial displays in dialogues are the subject of only three descriptive studies (Brunner 1979; Chovil 1989, 1991/1992; Ekman 1979) and two experiments (Chovil 1991; Bavelas, Gerwing & Healing in press). Given the abundance of these co-speech gestures in face-to-face dialogues, it is initially difficult to understand why there is so little investigation of them in the literature.

We propose that the answer is an unquestioned assumption that facial expression equals emotional expression, which has dominated interpretations of the face for centuries. Kendon (p. 31) illustrated this historical bias with sketches of fixed and stereotypic facial expressions of "contempt," published by Lebrun in 1734 as lessons on how to portray "the passions". These pictures are typical of the literature of the time, in which certain facial expressions were equated with specific emotions. Almost 300 years later, Ekman and Friesen's (1969) typology of non-verbal behavior included many functions of hand gestures but only one for facial expression: "the face is the primary site of affect displays" (p. 71). They narrowed affect displays even further, proposing that there is a fixed set of muscle configurations that correspond to innate, universally recognized emotions.

Although Birdwhistell (1970) almost simultaneously proposed that facial displays serve linguistic functions (and Ekman 1979, 1997) occasionally wrote about communicative facial expressions), conversational facial gestures have remained largely unrecognized, while the non-linguistic emotion approach has continued to motivate contemporary research. Even researchers who propose social ("audience") effects on facial expressions (e.g. Kraut & Johnston 1979) have limited themselves to emotional expressions such as smiling, fear or anger (see review in Chovil 1997). We affirm Kendon's observation that "the typology Ekman and Friesen presented... might almost be said to have had the effect of closing off further investigation" (p. 72) – even more so for facial than for hand gestures.

Although Ekman and Friesen's focus on emotional expressions may have hampered investigation of other possible functions of facial gestures, it is interesting to consider the results of their examination of nearly 6000 facial actions

of patients with affective disorders. These were depressed or bipolar individuals who were talking about their feelings, which could range from depression to mania. Yet *fewer than a third* of their facial actions were classifiable as emotional expressions (Fridlund, Ekman, & Oster 1987, pp. 160–161). Similarly, in the only systematic analysis of facial actions in dialogues in a non-clinical setting, Chovil (1989, 1991/1992; see also Bavelas & Chovil 1997) found that personal reactions of any kind were only about a quarter of the 720 meaningful facial gestures identified in the analysis; the remainder were semantically or syntactically related to the wide range of topics in the dialogues. Here we focus on this neglected majority, specifically on how conversational facial gestures share many of the characteristics and functions that Kendon outlined for hand gestures. The examples are from data gathered for Bavelas, Gerwing and Healing (in press).

A note on terminology

The literature is populated with a variety of terms for what the face does, each of which has underlying theoretical assumptions. The term *facial action* focuses on the formal musculature of the face instead of on the functions these actions might serve in interaction. *Facial* (or *emotional*) *expression* focuses on what the face reveals about an internal emotional state. Kraut and Johnston (1979) borrowed the ethological term *display* in order to distinguish between a social facial *display* and an emotional facial *expression.* Chovil (1989, 1991, 1991/1992) used "display" for the same reason. Although the above terms could convey important theoretical distinctions, they appear to be used interchangeably in the literature. Bavelas, Gerwing, and Healing (2014) have proposed *conversational facial gestures* as a term that emphasizes the close functional similarities to conversational hand gestures. The next sections document these similarities by mapping Kendon's (2004: Chapter 9) outline of the features of hand gestures onto the features of facial gestures.

"The gestured component of an utterance"

"Referential content"

"The gestured component of an utterance can be a part of its *referential* content" (p. 158), either by pointing or representing. Facial gestures can also serve a pointing function: the simplest deictic is a quick sideways head motion toward something ("It's over there"). More subtly, one person can simply shift his or her eyes to the side to indicate that the interlocutor should look at someone in that direction. The deictic function of gaze direction is apparent even in infants: by 12 months, human infants followed the direction of the experimenter's eyes, whereas even adult great apes were less likely to do so (Tomasello, Hare, Lehmann & Call 2007).

Sherzer (1973) and Enfield (2001) described the more complex "lip-pointing" deictic, which has meanings that depend on the immediate conversational context.

The more common and varied way in which facial gestures convey referential content is by various techniques of *representation*, and it is here that the capacities of hand and facial gestures differ the most. Hand gestures often use the techniques of modeling (e.g. forming a shape) or depicting (e.g. sketching in the air), but the face is not well suited to either of these. Facial gestures excel at *enacting* any imaginable face, that is, demonstrating anything that any face can look like. In contrast to a small set of stereotypic affect displays, conversational facial gestures are virtually unlimited in number and kind. The rapidity and flexibility of many muscle groups enables the face to enact the way the speaker looked or might have looked in a past situation – or might look in a future or even a hypothetical situation.

Nor is the face limited to enacting oneself; it can just as easily represent someone else's reaction in the past, present or future. Such facial gestures can represent a real person, a character in a story (human or not), or someone entirely hypothetical or generic. For example, in Figure 1, while retelling excerpts from the movie *Shrek 2*, the speaker described a scene in which Shrek had captured the cat who had attacked him. Frame 1 shows her own (non-representational) animated story-telling face; frame 2 shows her version of Shrek's slightly fiendish triumph; and frame 3 depicts the suddenly concerned and apologizing cat. (The experiment described below focuses on this capacity of facial gestures to enact the face of a movie character.)

"And Shrek, um picks him up by, the"   "nape of his neck and then he was like--"   "the cat was like, Oh I'm so sorry! "

Figure 1. While describing a scene from the movie *Shrek 2*, the speaker made the hand gesture in frame 2 and the facial gestures in frames 2 and 3. The three frame shots cover a 6.75 second period. (The face of the addressee is inset at the upper right)

"Pragmatic" and "interactive or interpersonal functions"

Even further from the world of emotion are the pragmatic and interactive functions of facial gestures. These non-referential functions in dialogue (also called collateral communication or meta-communication) are about the dialogue itself rather than about its topic.

"Modal functions" are the first of Kendon's three pragmatic functions of hand gestures. They alter the way in which "the utterance is to be interpreted" (pp. 158–159). An example of a facial gesture that serves a modal function is the *facial shrug,* which is analogous to a shoulder shrug. It typically involves a quick eyebrow flash and the retraction of a corner of the mouth; see Figure 2. Just as the shoulders can "shrug something off," a facial shrug can convey that something does not matter (e.g. that enough has been said or that it has been said well enough). Both Ekman (1985) and Chovil (1989, 1991/1992) observed facial shrugs. Smiles can also serve a modal function; Coates (1991) found that smiles played a role in marking ironic humor.
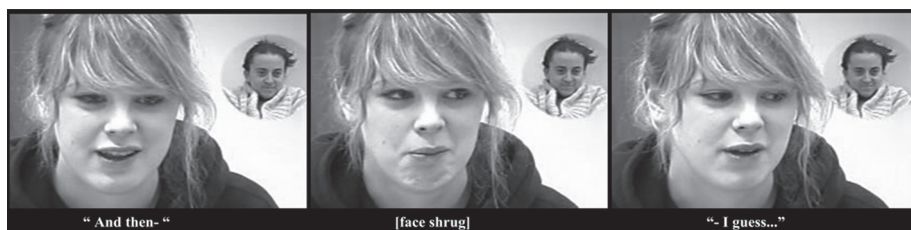
Figure 2. The speaker was completing her description of a scene from *Shrek 2* when, instead of continuing on from frame 1, she made a *facial shrug* (in frame 2). Then she said "I guess" (in frame 3) and went on. The three frame shots cover a 1.3 second period

"Performative functions... indicate the kind of speech act or interactional move a person is engaging in" (p. 159). For example, speakers often raise their eyebrows to indicate a question, even if the syntax was not interrogative (Ekman 1979: 185). Brunner's analysis (1979) showed how addressees' smiles can function as back-channels, with the same timing, placement and function as verbal back-channels. Their performative function is to indicate that the addressee is following what the speaker is saying. Nods serve the same function so often that they are often treated as verbal rather than gestural.

"Parsing functions" are useful for "punctuating the spoken discourse or... marking out its different logical components" (p. 159). The most common group in Chovil's (1989; 1991/1992) data were what she called *syntactic* displays, especially eyebrow movements that either emphasized a single word by a quick flash or underlined a whole phrase by staying up for the duration. Chovil also found that speakers who were relating a story or anecdote could use their smiles, not to show happiness, but to punctuate their narratives. For example, the patient in Table 1 smiled to mark a shift from telling about his symptoms to attending to the physician's comment. Similarly, the speaker in Figure 1 smiled immediately after frame 3 to close off that part of her narrative.

"*Interactive or interpersonal functions*" include "the use of gestures as a way of indicating to whom a current utterance is addressed" (p. 159), for example by gaze and head direction. Another easily recognized facial gesture with an interactive function is the *thinking face*. As shown in Figure 3, this facial gesture usually involves shifting the gaze away from the addressee and looking thoughtful or searching, which indicates "that a current speaker, though not actually speaking, is nevertheless still claiming a role as speaker (still 'holding the floor')" (p. 159).

Figure 3. While recalling a scene from *Shrek 2*, the speaker made an extended (1.7 second) *thinking face* in frames 2 and 3

"Two different kinds of expressive resource"
Kendon emphasized that:

> The gesture–speech relationship... is best understood in terms of a point of view that sees gesture and speech as two different kinds of *expressive resource* available to speakers, and that the gestures employed within an utterance, like the words that are employed, are components of the speaker's *final product*.   (p. 111)

By illustrating how facial gestures share Kendon's characteristics of hand gestures, we have proposed that facial gestures are also a visible expressive resource that is a part of the speaker's final product. The rest of this chapter focuses on flexibility between speech and gestures (of either kind), as a function of the "expressive resources … existing within the context of the given moment of interaction" (p. 111).

Part 2: "There is flexibility in the gesture–speech relationship"

> It appears that there is flexibility in the gesture-speech relationship.... Both the gestures and the verbal expressions used are to be accounted for in terms of such factors as... the speaker's knowledge of various expressive resources... as well as the constraints and possibilities existing within the context of the given moment of interaction.   (p. 111)

To assess and test this flexibility, we applied a *semantic features analysis* (Beattie & Shovelton 1999, 2002; Gerwing & Allison 2009, 2011; Holler & Beattie 2002, 2003, 2004; Holler & Stevens 2007; Holler & Wilkin 2009) to speakers' retelling of a video excerpt that could be described with speech, hand gestures and facial gestures. This method starts by identifying a set of semantic features that are specific to the material the speaker is describing. That is, the analysts stipulate in advance certain specific information in the stimulus material, then they assess whether words, gestures or both contribute information about each of these features.

Gerwing and Allison (2009) compared the semantic features method to two other ways of studying the relationships between gestures and speech (i.e. deictic references and redundancy) and found that the semantic features analysis had the advantage of identifying precisely how and when speakers distribute information between the two modes. The studies cited above have shown the utility of semantic features analysis for understanding the relationship between speech and hand gestures, but to the best of our knowledge, this method has not previously been applied to facial gestures.

The data are a subset of Experiment 1 in Bavelas et al. (in press), in which speakers retold several scenes from the movie *Shrek 2* in either a face-to-face dialogue, a telephone dialogue or a monologue. This new analysis focused on the face-to-face and telephone dialogues[1] in order to assess how visibility would affect the use of audible versus visible expressive resources. One scene was particularly suitable because it included features that could be described with words, hand gestures or facial gestures. In this scene, Puss in Boots (a cat) wants to join Shrek on his journey. The cat, who suddenly appears very small, clutches his hat under

---

1.   Two experiments (Bavelas et al. 2008, in press) have shown that holding a phone does not significantly decrease the overall rate of hand gestures. These two studies found no difference between face-to-face and telephone conditions, which replicated the results of five similar experiments that compared face-to-face versus partition conditions: Rimé (1982), Bavelas, Chovil, Lawrie and Wade (1992: Exp. 2), Pine, Burney and Fletcher (2010), Holler, Tutton and Wilkin (2011), de Ruiter, Bangerter and Dings (2012). (Note that the de Ruiter et al. data on overall gesture rate were obtained from J. P. de Ruiter, personal communication, July 13, 2012.) See Bavelas and Healing's (2013) review of visibility effects on hand gestures.

his chin and, with enormous eyes, looks up at the much bigger Shrek, silently using his cute face and huge eyes to manipulate Shrek into taking him along (see Figure 4). The comic elements of the scene are the cat's pose and especially his eyes, which he makes impossibly large and endearing – a strategy that any cat owner will recognize.

Figure 4. Puss in Boots silently pleading to be taken along in *Shrek 2*

The semantic feature analysis focused on how speakers conveyed three key pieces of information in the scene: the cat's clutching the hat, making very big eyes, and looking up at Shrek. We compared the proportions of visible versus audible means of conveying these three features when speakers were in two different conversational contexts: a face-to-face dialogue versus a telephone dialogue.

The hand or facial gestures that could describe the three features were all *enactments*, that is, first-person portrayals in which the speaker momentarily presented herself in the way the cat looked, clutching the hat or looking up with big eyes. First-person enactments of oneself or someone else are a common kind of hand or facial gesture. For example, the patient in the initial example used hand and facial gestures to portray himself, first as he was on a bad day, then on a recent good day. Recall that very little of this information was in his words; the visible information was not redundant with the audible information.

We predicted that the experimental condition would change the relative distribution of the information that speakers presented using visible versus audible means: in face-to-face dialogues, speakers would convey more information by visible means (hand and/or facial gestures), and in telephone dialogues, speakers would convey more information by audible means. This difference would demonstrate flexibility in the relationship between speech and gestures. We further predicted that the information conveyed by visible means would be redundant with speech less often in the face-to-face condition and more often in the telephone condition. In de Ruiter, Bangerter and Ding's (2012) terms, the visible information in the face-to-face condition would be *obligatory*, while in the telephone condition, it would be *non-obligatory*.

Method

Participants
Initially, 40 female undergraduate psychology students participated for course credit: 20 in the face-to-face condition (10 dyads) and 20 in the telephone condition (10 dyads). We randomly assigned both the experimental conditions and the participants' roles as speaker or addressee. We analyzed the 16 speakers who included the scene described above, 9 who were face to face and 7 who were on the phone.

Materials
The experimental stimulus was a 2 min., 45 s. video containing two excerpts from *Shrek 2*. The scene chosen for this analysis, in which the cat is silently appealing to Shrek, was 10 s. long.

Equipment
The experiment was held in the University of Victoria Psychology Department's Human Interaction Laboratory suite, using three Panasonic WV–CP474 color cameras to capture a large front view and narrow side view of the speaker, plus a smaller inset of the addressee. The speakers viewed the movie excerpts on a small color TV/VCR in an adjacent room within the suite. Speakers in the telephone condition used a handheld phone with the dialogue tapped into the audio track of the video recording. We digitized the videos with Broadway ProDVD (www.b-way.com) and analyzed them on an 18-inch ViewSonic G90fb color monitor using ELAN (www.lat-mpi.eu›Tools›Elan (Brugman & Russel 2004; Wittenburg, Brugman, Russel, Klassmann & Sloetjes 2006).

Procedure
Before recording began, the participants met in the main recording room and provided written consent. In the telephone condition, the addressee then moved to a nearby office while the speaker remained in the main recording room, and they did the experimental tasks over the telephone. In both conditions, the participants spent a few minutes getting acquainted, then did two unrelated pilot tasks followed by the main task of viewing and describing the movie excerpts. The speaker watched these excerpts twice in an adjacent room, then returned to the main recording room to describe the scenes to the addressee either face to face or by telephone. Afterward, the experimenters debriefed the participants, answered questions and showed them the video of their participation. Each participant then signed a form indicating the permissible uses of their video (e.g. permission to view for analysis only, permission for viewing by professional audiences, permission to include a still photo in an academic journal).

Semantic features analysis

The first step was to identify the speakers who included this particular scene in their descriptions. In the face-to-face condition, nine speakers described it. In the telephone condition, eight did, but one participant was not analyzable because she had moved her head off-screen while describing the scene, leaving seven participants for analysis in that condition.

The analysis focused on three key semantic features of the cat's silent persuasion of Shrek during this scene: (1) the cat clutched his hat under his chin; (2) he made his eyes disproportionately large; and (3) he was looking up at Shrek, who was much taller. (Detailed operational definitions for these features are available from the authors.)

First, using only the audio, the analysts located when each speaker used words to refer to one of these features:

For the cat clutching his hat, the words could be any reference to "clutching," including synonyms (e.g. "holds his hat"). Verbal references to the hat alone, which did not include how the cat was holding it, were not sufficient to count for this feature.

For the cat's big eyes, verbal references counted only if they conveyed that the cat's eyes were unusually or disproportionately big (e.g. "huge" or "really, really big"). References to "big eyes" alone were not considered sufficient to count as a reference to this feature.

For the cat looking up, speakers had to convey the upwards direction of the cat's gaze, so the speaker's words counted only if they combined "looking" and "up". "Looking" alone was not sufficient because a key component of this feature was the height difference between the cat and Shrek.

The analysts then used the video to decide whether the participants used hand or facial gestures to depict any of the three features:

For clutching the hat, a reference counted as visible if the speaker's hands were in a clutching position (i.e. in fists) and held somewhere between under the chin and in front of the chest.

For the big eyes, visible references were widening the eyes, making them look bigger than the speaker's own baseline position, based on examining the speaker's eyes before and after the scene. Speakers could also use hand gestures to demonstrate big eyes by suggesting two large circles with curled thumbs and index fingers, then holding these up in front of their eyes.

For looking up at Shrek, the speaker had to look away from the addressee in an upwards direction (e.g. gazing upward and sometimes also tilting the head upward). The precise timing of the gaze was important. The analysts had to watch more than just this scene in order to ensure that this was in fact a portrayal of the

cat looking up and not the speaker raising her eyes to show that she was searching for a word (i.e. not a "thinking face" as in Figure 3).

Two analysts (JG and SH) conducted all of the analysis together. For reliability, a third analyst (JB) trained on a randomly selected 30% of the excerpts, then worked independently on a new randomly selected 30%. Agreement required identifying exactly the same word, hand gesture or facial gesture, and they agreed on 93% of their decisions.

Qualitative results

The words that speakers used for clutching the hat included "holding his hat," "clutches, like, his hat," or "holds his little hat in his hands." Their most common hand gesture was to hold one or both fists close together, right under the chin (see Figure 5).

"like --"
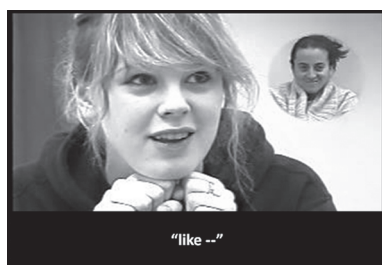
Figure 5. The speaker gestures the cat *clutching the hat* with her hands (.23 s) as well as the cat *looking up* by gazing upward (.43 s)

Speakers described the cat's big eyes in a wide variety of ways: "big big black eyes," "makes his eyes all big," "he does this whole, like, big big eye thing," "like big pussy cat eyes, they are huge," "huge like puppy dog's eyes." They could also use a facial gesture to make their own eyes appear unusually big, for example, by opening their eyes more widely and looking far to the side, which combined to show much more white, as in Figure 6. Several speakers used a hand gesture that projected huge eyes in front of their face, as in Figure 7a.

For the cat looking up at Shrek, speakers said, for example, "peers up at Shrek" or "looks up to Shrek". They gestured this feature by casting their own eyes up, as in Figure 5, or tilting their head up, as in Figure 7b.

These figures also illustrate two other patterns of the visible enactments. First, the speakers often demonstrated one or more features without any accompanying speech, just introducing the enactment with the discourse marker "like" (Figures 5
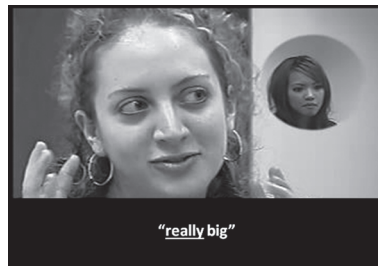
Figure 6. The speaker gestures the cat's *big eyes* by showing more whites of her own eyes (.40 s)

Figures 7a and 7b. The speaker gestures the cat's *big eyes* using a hand gesture in the first frame and then gestures the cat *looking up* with a facial gesture in the second frame. (Total duration =.81 s)

and 7b). Second, and perhaps most important, speakers frequently used their hand and facial gestures simultaneously (as in Figure 5) or in quick succession (as in Figures 7a and 7b). The result was to provide a more complete image of how the cat looked at that moment.

Quantitative results

Creating proportions
Table 2 shows the results for each speaker, aggregated across the three features. For each speaker, we first summed the number of references that this speaker made to any of the features using words, using hand gestures and using facial gestures. The sum of hand and facial gestures was the total number of visible references. The sum of visible and audible (verbal) references was the total number that the speaker made in any modality. The two key measures were the proportions of total references that were audible versus visible.

Table 2. Individual speakers' use of visible versus audible references to semantic features as a function of experimental condition

	Face-to-face condition						Telephone condition				
	Visible references						Visible references				
Dyad #	Hand gestures	Facial gestures	Total visible[a]	Total audible[a]	Total references	Dyad #	Hand gestures	Facial gestures	Total visible[a]	Total audible[a]	Total references
2	2	1	3 (.75)	1 (.25)	4	3	0	4	4 (.80)	1 (.20)	5
5	3	2	5 (1.0)	0 (0)	5	7	1	0	1 (.5)	1 (.5)	2
9	3	4	7 (.875)	1 (.125)	8	16	0	0	0	0	0
12	3	2	5 (.71)	2 (.29)	7	17	0	0	0 (0)	1 (1.0)	1
18	2	0	2 (.67)	1 (.33)	3	19	2	2	4 (.80)	1 (.20)	5
21	2	0	2 (1.0)	0 (0)	2	28	1	0	1 (.5)	1 (.5)	2
26	0	0	0	0	0	34	0	0	0	0	0
27	1	3	4 (.57)	3 (.43)	7						
29	2	2	4 (.80)	1 (.20)	5						
Mean proportions			.71	.18		Mean proportions			.37	.34	
Standard deviations			.30	.16		Standard deviations			.37	.36	

[a] The numbers in italics are the proportions of total references.

Note that we aggregated hand and facial gestures for two reasons. First, the theoretical focus was on visible versus audible enactments as a function of visible and not-visible experimental conditions. Second, as illustrated in Figures 5 and 6, the hand, head, gaze and facial gestures often occurred together, portraying an integrated picture of the cat at a particular moment. Separating them would obscure this integration.

Use of speech versus gesture within experimental conditions
We first examined (a) whether speakers were more likely to use words or gestures to convey information about the three semantic features and (b) whether these distributions varied as a function of the experimental condition. As shown in Table 2, every speaker in the face-to-face condition used a higher proportion of gestures than words. The difference in the mean audible and visible proportions was statistically significant: within-subjects t (8) = 4.530, one-tailed $p < 0.002$. In contrast, speakers in the telephone condition were more varied, and the mean audible and visible proportions were identical to each other.

Use of speech versus gesture between experimental conditions
We also compared how the conversational context (i.e. speaking face to face or on the telephone) affected how much information speakers conveyed in each modality. Comparing across conditions in Table 2 shows that the mean proportion of visible references to the three semantic features in the face-to-face condition was almost double the proportion in the telephone condition, a difference that was statistically significant: between-subjects t (14) = −.015, one-tailed $p < 0.032$). That is, speakers who were talking face-to-face made visible references to the features significantly more often than speakers on the telephone made visible references to the same features. Conversely, the mean proportion of audible references was higher in the telephone condition than in the face-to-face condition. However, this difference was not significant, which led to our third analysis.

Redundancy between speech and gesture as a function
of experimental condition
We hypothesized that, although the proportional use of speech to convey information about the semantic features did not differ between experimental conditions, the relationship between the speakers' words and gestures would differ in the two conditions. Specifically, the hand and facial gestures in the telephone condition would tend to convey the *same* information as the words and would therefore be redundant (i.e. not obligatory). As a result, information would be available to the addressee via the speakers' words, even though the gestures were not visible. In contrast, the hand and facial gestures in the face-to-face condition would convey

different information than was conveyed in words and would therefore be non-redundant (i.e. obligatory). The addressee would have to see the speaker's visual enactment to get all of the information the speaker was providing about the three features.

To test this hypothesis, we returned to each speaker's raw scores for each semantic feature and recorded whether the speaker's visible contribution was obligatory (i.e. the speaker did not convey the same information in words) or not obligatory (i.e. the speaker also conveyed the same information in words). We then collapsed the three semantic categories and recorded whether or not that speaker conveyed at least one semantic feature using an obligatory hand or facial gesture. Three speakers were excluded because, as shown in Table 2, they did not refer to any of the three features.

As shown in Table 3, seven of the nine speakers in the face-to-face dialogues conveyed information in hand or facial gestures that was obligatory, that is, not conveyed at all in the words. One speaker conveyed information about the semantic features in both words and hand/facial gestures. In telephone dialogues, four of the seven speakers conveyed information in hand or facial gestures that was redundant with the words or was non-obligatory. One participant conveyed information in at least one category using obligatory hand or facial gestures.

Table 3. Effect of experimental condition on obligatory vs. non-obligatory gestures

	Experimental condition	
Relation of gestures to words	Face-to-face	Telephone
At least one hand or facial gesture that was obligatory (not redundant with words)	7	1
No hand or facial gestures that were obligatory (all were redundant with words)	1	4

Note: One speaker in the face-to-face condition and two in the telephone condition did not describe any of the semantic features analyzed.

*$\chi^2_{(1, N = 13)}$ = 5.923; $p < 0.05$. Note that two of the expected frequencies are lower than conservative practice recommends. However, the pattern is clear in the observed frequencies themselves.

These results are consistent with the only other semantic features analysis that has compared how speakers distributed information in speech and hand gestures in different conversational contexts. Gerwing and Allison (2011) found that speakers who were describing the shape of the skirt on an unusual dress conveyed significantly more of this information in their gestures than in their words when speaking to an addressee in a face-to-face dialogue. When the addressee was on the telephone, speakers conveyed significantly more of the information in their words.

Discussion

This chapter has provided two related illustrations of the lasting influence and relevance of Kendon's work, especially his magnum opus (2004). Part 1 used his extensive and detailed specifications of the characteristics, contributions and functions of conversational hand gestures as a framework and set of standards for including *conversational facial gestures* as another instance of "visible action as utterance". This framework made it possible to articulate and document the extensive similarities of facial to hand gestures, which offer an alternative to approaches that see the face as stereotypic configurations related to a few emotional expressions. Facial gestures include anything the face, head and eyes can do to convey any meaning related to the talk in progress: they can convey referential content either deictically or by direct representation. They can serve pragmatic functions, such as indicating the mode or frame of an utterance (e.g. the facial shrug), indicating the kind of speech act (e.g. eyebrows marking a question), parsing the utterance (e.g. a smile closing a narrative), or indicating the status of turns (e.g. a speaker's thinking face). In all of these functions, facial gestures are part of the speech-gesture ensembles that constitute language in dialogues.

Hand and facial gestures, as well as other bodily movements of the torso or legs, often act in concert with speech and each other. The gestural components of these ensembles present detailed and nuanced *images* of how someone looked (e.g. the cat) or acted (e.g. the patient). Therefore, in drawing attention to facial gestures, we are not suggesting that they should be studied in isolation from gestures made by other parts of the body. The challenge is to appreciate the parts while still keeping them in the context of the whole ensemble, in which the gestures "serve to create an image of the object that is the topic of the spoken component" (p. 161).

Part 2 added facial gestures to an experimental demonstration of Kendon's insights on the flexibility of the relationship between parts of the speech-gesture ensembles in different interactional contexts. A semantic features analysis showed how speakers represent the same material differently depending on changes in transmission conditions. When conversing face-to-face, speakers conveyed information about semantic features more in facial and hand gestures than in words. In some instances, speakers conveyed information about a feature entirely in gestures, making these references obligatory or necessary for accurate comprehension of the description. When conversing on the telephone, speakers conveyed information about the three features equally in words and gesture. However, the gestures that they used were also more likely to be redundant with their words and were therefore not necessary for accurate comprehension. In other words, "Speakers... can control these two components and can orchestrate them differently, according to the occasion" (p. 127).