Charles Eidsvik, University of Georgia

“Voice and Gesture within the Context of Mirror Neuron Research”

In the last decade, advances in neuroscience have altered how voice and gesture can be approached as elements in cinema. At CSMI conferences, references to neuroscientists such as Jaak Panksepp, Joseph Ledoux and Antonio Damasio are common. As brain-imaging technology improves, neuroscience increasingly deals with issues involved in response to actors in films, such as responses to gesture and voice. Fifteen years ago, a discovery at the University of Parma made with primitive technology jolted the field. Researchers doing single neuron research on motor neurons accidentally found that the same neuron fired when a macaque monkey saw a lab worker grasping food and eating it as fired when the monkey did it himself (Rizzolati, Blakesly). Because the nerve’s firing was amplified and transposed to sound, a speaker “blatted” alerting the researchers. Eventually they discovered that the same cells fired whether sight or sound of the action occurred, and that parts of an understood action – e.g. showing the action’s goal-stage - excited the neurons nearly as well as the whole action (Keysers, Kohler, et al). The mirror neurons were concentrated in area F5 of the macacque brain, an area involved with hand movements and facial expressions. The world of science was skeptical about speculations that something similar occurred in human brains. However, advances in brain mapping (using several techniques, including magnetoencephalography (MEG) and functional Magnetic Resonance Imaging (fMRI)), have allowed neuroscientists to show not only that humans have mirror cells, but that they are crucial for learning (especially of language and communicative gestures with the hands and face), for response to other people's voices and gestures, and even for various kinds of motor and emotional empathy.

As with any "hot" set of discoveries in science, popular media have both publicized and, to some extent, perhaps overstated what mirror neurons do. For example, in a January 10 "Science Times" section of the *New York Times* that headlined and summarized mirror neuron research, Sandra Blakeslee called mirror neurons in people "smarter" than those in monkeys. Perhaps. But that is not what the scientific research shows so far. The mirror neuron *system* is remarkable, even "smart", and certainly makes *us* smarter. But the neurons themselves are just neurons, with one added behavior. They activate when a person performs an action ***or*** perceives it. But how and under what conditions that happens and affects us involves not just isolated neurons but complicated circuitry between parts of the brain; the circuits vary for the different actions the neurons perceive. What seems to happen is that we see and/or hear someone doing something that, to one extent or another, we know how to do. Usually this involves the face, hands, speech, or, less frequently, feet. That excites mirror neuron clusters, which copy the incoming signal and pass it to the premotor cortex, perhaps for comparison with actions it "knows how to do." The signal also is copied or relayed to other relevant areas. Facial expressions, hand movements and voice are copied to speech generating areas (which also deal with face and hand movement) and to areas dealing with emotional expression. Whether that response is strong or weak depends on the action, how it is perceived, and on prior experience with similar actions (Glaser).

Within the research community, there is considerable divergence about how broadly mirror neurons affect our perceptual response. Dr. Marco Iacoboni (UCLA) comments "when you see me perform an action – such as picking up a baseball – you automatically simulate the action in your own brain. . . . you understand my action because you have in your brain a template for that action based on your own movements." (Blakeslee) Almost every researcher would agree with him – picking up and grasping an object is the archetypal mirror neuron stimulator. But whether imitative behavior visible even in newborns, or our unconscious imitation of the people we are with, involves mirror networks, has little consensus; how (or whether) the mirror neuron system and imitative behavior are related or synonymous is unclear How far scientists go along with Byrne's motor-system theory of empathy varies (Rizzollati, et al 2002).

A major limit on what is known so far has to do with the state of fMRI technology. An fMRI machine is a two ton, hot and noisy tube in which subjects lie, their heads still, with whatever they are to react to either put in front of them, sent through earphones or shown to them on video.   Each volunteer in a group being studied is apt to undergo no more than four or five minutes of actual brain scanning at a time. The scan "slices" are four or five millimeters wide, with a millimeter between them, and take two seconds apiece. The research is pricey and computer-intensive. To create replicable results, the majority of experiments go a step at a time and only rarely play with the multimodality and interactivity of everyday perception. Whatever areas of the brain show increased oxygen use – frequently the the insula, the premotor cortex, the anterior and posterior cingulate, the posterior parietal lobe, the superior temporal sulcus, Broca's Area, and so forth – get a lot of attention.   Experiments have been cleverly designed to show that mirror neurons grasp processes, intentions, and emotional expressions directly, without recourse to "higher" brain levels (Iacaboni, Molnar-Szakacs, et al).   As an immediate response, empathy is a physiological process that does not need "top down" cognition to work. An unknown in the process is how the premotor cortex creates or references memory.

What fMRI imaging of mirror neurons that deal with speech shows is not only counter-intuitive but counter to much “standard” thinking about the different functions of speech areas in the brain. The commonplace view for decades was that Wernicke’s area took care of language understanding, while Broca’s took care of speaking; the two had little to do with one another. Functional Magnetic Resonance Imaging seems to show that all areas involved with communication are “recruited” for a wide number of tasks. Further, fMRI imaging has shown not only that there are aural speech equivalents to mirror neurons (sometimes called “echo” neurons) but that some are multi-functional: “Some neurons developed to be able both to generate sound and discharge (resonate) in response to that sound (echo-neurons).” Rizzolati marvels at “the incredibly confusing organization of Broca’s area in humans, where phonology, semantics, hand actions, ingestive actions, and syntax are all intermixed in a rather restricted neural space” (Rizzolatti and Craighero, 186). Neuroscientists such as Philip Lieberman who had been viewed as radical for arguing that speech acquisition and production are completely interdependent, now seem to have been prescient.

In many respects, of course, mirror cell research reinforces what we intuitively suspected all along. Empathy for others’ pain involves activity in affective centers, not in areas of the brain that “feel” pain (Singer, et al) . Ballet dancers respond more strongly to films of ballet than do dancers who specialize in Brazilian dance; non-dancers respond less than dancers (Glaser). We respond more to loved ones than to strangers (Järveläinens), respond more strongly “in person” than to television images, and respond more strongly to motion images than to static ones (Iacaboni cited by Blakeslee). We respond differently to observed social interactions than to images of people alone (Iacaboni, Lieberman, Knowlton et al). These and numerous small but tantalizing discoveries (such as the utter indifference of mirror neurons to Virtual Reality images) do not overthrow or undermine older forms of research such as lab work with animals and study of brain-damaged accident or disease victims. Rather, they add to the body of work by scientists such as Antony Damasio, Jaak Panksepp, and Joseph Ledoux.

For cognitive studies of motion images, however, problems to be solved now are layer-caked. Small-scale studies focusing on gestures, voice, and the like can utilize mirror neuron systems research as an explanatory tool and as a way of thinking through what to look closely at. But cognitive film studies often goes well beyond what scientists can currently ask. Further, cognitive film studies tends to ask complicated questions even when working on limited subjects. The following paragraphs condensed from a CSMI conference paper I gave in 2004, are offered merely as a utility door into what the new research tells us about gesture and voice as indicators and producers of emotion, and into how that relates to prior research using other methods.

As a film studies teacher and film festival goer, I have long been curious about the perceptual dissonance inherent in subtitled films. Reading subtitles (normally weighted toward expositional clarity, not faithfulness to what is said) while observing actors' gestures and voices in foreign languages supposedly preserves the ability to hear and feel the emotional content in the original gestures and voices. For me this raises the question of how much emotion carries purely via sound in a foreign language accompanied by gesture (which might or might not be "foreign"). Film buffs prefer subtitling. Mass audiences prefer dubbed translations, especially in Europe; to be marketable, a film needs to be dubbed. But, given the semantic problems of actually translating dialogue and "doubling" it in the local language, lip synch is lost, the actors' emotions are apt to be badly imitated and the unity of voice and visual gesture is compromised. In 2004, I put together a clip of subtitled and dubbed versions of two adjacent scenes from the Region 2 DVD of the Two Towers section of the *Lord of the Rings* film trilogy as a perceptual and cognitive “response damage” test to get preliminary data on how well voice and gesture as indicators and producers of emotion did and did not “work” and for whom. I used “classical” neuroscience and social psychology as sources of insight and vocabulary. On one hand, I was looking to the possibility of eventually doing a meaningful cross-cultural experiment with students both in the U.S. and Germany; on the other, I was just putting together teaching materials to get students to watch and listen more attentively and analytically, paying attention to their responses.

I recorded both English and German voiced versions from the DVD, with English titles on the German (dubbed) version and German on the English to neutralize preferences for reading or not reading. The characters are Eowyn and Aragorn; my focus was on Eowyn, played by Miranda Otto but voiced in the German version by Alexandra Wilcke. One scene involves Eowyn’s tight-throated expression of her greatest fear: to be caged, imprisoned, enslaved. In the second Eowyn offers Aragorn soup she has made that no one else wants to eat. He politely accepts, tastes the soup, and tries to disguise the revulsion he feels in order to be kind to Eowyn, a great warrior but execrable cook. She sees through him, but is grateful for his courtesy and all the more smitten by his attractiveness. The first scene is moving, the second, amusing, a fairly typical cooling off interlude of the sort common in thrillers. In the first scene, shot in a studio, Otto’s dialogue appears to be actual location sound; in the second, done outside, both Otto’s and Wilcke’s voices are done with ADR (that is, dubbed in later, following the guide track done on location). Both actresses use Stanislavsky techniques, based on authentically feeling the emotions of the character. (When I spoke with Wilcke about her work; she described this scene as being close to identical with Otto’s work (Interview in Athens, GA., February 2004)). Wilcke is among the top voice actors in Germany, often dubbing one or more characters for major Hollywood film releases in Germany. As dubs go, this is at least close to as good as it gets.

The first scene is two-part.   Eowyn is practicing her sword skills. Aragorn spies on her. She discovers he is watching, and they spar. He is a skilled warrior, better than she is, but fascinated by both her skills and her rage. After the swordplay they talk, with Eowyn describing her motivation for becoming a warrior. She is terrified of becoming a prisoner, enduring humiliation, becoming a victim of sadists. In the first, sword-wielding scene, the clearest difference between Otto's voice performance and Wilcke's is the tight connection between Otto's movements and posture, her facial expressions, and her voice. Her violent exertions affect her breath when she talks immediately after. In her body movements, her voice, and her facial expressions, one senses several amygdalic reactions to Aragorn's easy parrying of her sword and his inquiry about her motives: seething anger at Aragorn's intrusion into her space, jealous rage at Aragorn's skill, then, as she talks, fear so deep her throat tightens, quivers, and hoarsens; then in response to Aragorn's friendly, "mentoring" demeanor, a kind of attraction or affiliation response, carried by facial expression, not voice. The emotions are strong. Wilcke's anger and fear are softer, not quite matching the adrenaline–fueled edge-of-panic look of Otto's torso and arms and neck. Though Wilcke's throat sounds tighten in the "fear" section, the grip of the fear seems a little less, and the breathless feeling and rhythms in Otto's voice as a result of her physical work are not matched in Wilcke's version. (Though Wilcke said that she acted out the scene physically, she complained that dubbing studios have little space, and that scenes tend to be dubbed one actor at a time. This hurts performance, especially in scenes involving action and strong emotions.)

What would “classical” neuroscience say about this scene? In 2000, Philip Lieberman pointed out the tightly wired connections between the auditory cortex, speech motor areas the physical speech-delivery and control mechanism. Lieberman argues: “Speech does not just involve Broca’s area and other “motor” parts of the brain; the auditory cortex also is active (37-38).” Because we learn to use the lung, larynx, and supralarygeal vocal tract at the same time as we learn our own “native” language(s) our physical memory is specifically adapted to our native tongue(s). So English speakers could expect stronger reactions to watching and hearing Otto speak than Wilcke; German speakers might well have the opposite reaction. But as Lieberman argues: “Verbal working memory, comprised of both an executive area in the dorsolateral prefrontal cortex, and a covert “rehearsal” of verbal information (a sort of “silent speech”) allow us to remember what we hear and to form grammatical speech into trains of thought (77-81).” Lieberman saw mirroring behavior before it was proven that humans have mirror neurons.

How well would the voiced emotions “travel” to other cultures? Research on emotional identification of emotions in speech, such as Klaus Scherer’s study of voiced emotions in a symantics-free hybrid language (created by linguists out of elements of many “real” languages, and spoken by professional voice actors) indicates that people who spoke the actors’ native language (German) interpreted at least four emotions (anger, sadness, fear, and joy) with about 85% accuracy; those with other European-language backgrounds were perhaps half as accurate (Scherer). Given the visual as well as narrative contexts of the above scene, one would expect little difference between viewers of the English and German versions in “recognizing” the emotions involved, especially since the most accurate judgements of emotional content across and within cultures are for fear expressed by *both* the face and voice (de Gelder). Judging what emotion is being expressed, of course, is a very different thing from feeling that emotion.

What might mirror cell research add to a "classical" analysis of this scene? The first difference would involve the swordplay section. We respond to action via visual mirror cells, with our premotor responses dependent for intensity on our experience with the kinds of action viewed . But even those who do not dance professionally respond to dance movements, albeit at low intensities (Glaser). Though few of us sword-fight, we still can "feel" how it is done. For all of us, mirror circuits arouse premotor responses that allow "participation" in the movements and predictions of what will happen next . A good deal of the pleasure of watching action films likely comes from the physiological similarity of actions we know how to do and portrayed actions, along with our accompanying premotor activation.

Response to the conversation between Eowyn and Aragorn would involve identification with Eowyn's fear. FMRI studies not only reveal mirror/echo cell cluster activations from voices; they show that both gesture and voice activations overlap in Broca's and adjoining areas; facial expression (or gesture) and voice are interconnected if not interlocked (Kohler, et al; Wilson, Saygin, et al; Iacaboni & Wilson). Thus we "feel" Eowyn's fear, hoarseness, and trembling somewhat as if it were our own.

The ending to the scene, as Eowyn and Aragorn become friendly, has all sorts of neuroscience explanations, both old and new. Reactions to the scene would invoke the front cortical areas studied by Damasio in his work on the sense of self, but would also affect hypothalamic, insula, and other areas involved with "social" and emotional situations (Carr, Iacaboni, et al). The largest "new" discovery is that "watching social interactions produces dorsomedial prefrontal and medial parietal BOLD fMRI signal increases compared to a resting baseline."(Iacaboni, Lieberman, et al) These parts of the brain are tonically active except when one is doing concentrated, single-person tasks; then they become inactive. However, watching two people interact causes signal elevation beyond their normal state. Apparently these areas are tuned to social interaction. The experimenters speculate that these areas of the brain are involved in continual cogitation about interpersonal issues.

For the above-described scene, mirror-cell research gives us much. For the following "soup" scene, it is less useful. To be sure, our identification with the two characters is facilitated by watching their faces and gestures and listening to their voices, so our whole battery of neurological reactions to social interaction could be shown to be in play. But it is a cooling-off scene, lightly comic, and mirror-neuron research as not reached the point at which the dissipation of tension has been studied. In thrillers such as *The Lord of the Rings* trilogy, tension-relief scenes often follow intensely emotional ones, perhaps to give adrenaline and dopamine circuits time to dissipate neurotransmitters before the excitement begins again.

Only one brief part of the scene offers opportunity for mirror cell research to offer insights.). Eowyn wants his approbation and he is polite to a fault. But when Aragorn is alone briefly, he reacts to the foul smell and taste of the soup Eowyn has made with disgust. A study of response to smells, "Both of us Disgusted in *MY* Insula: The Common Neural Basis of Seeing and Feeling Disgust" finds a close connection between smelling foul odors and watching other people's facial reaction to such smells. This includes the sensation of disgust (Wicker, Keysers, et al). The film scene, however, downplays the aversive side of the situation and foregrounds the humor, as Eowyn comes back, cheerfully filling up Aragorn's bowl again. This scene, shot in a windy, outdoor location but "doubled" in a studio in both the English and German versions, shows remarkable similarity in Otto's and Wilcke's performances.

To deal with the soup scene we traditionally would have relied almost exclusively on social rather than cognitive science. As Edward Sapir pointed out a half century ago (Kendon

1990, p.34), "We respond to gestures with an extreme alertness and, one might almost say, in accordance with an elaborate and secret code that is written nowhere, known by none, and understood by all."   Getting that code—actually, codes--written down is still a work in progress, and a good deal of the coding of behavior is multimodal and cultural. But *why* these codes work universally is no longer a mystery. Our mirror systems connect us to others. We *feel* what we see and hear other people doing. Of course the experiments on mirror systems do not yet speak to more than a small fraction of the questions we have about gesture and speech, or to how we respond to mirror-neuron based kinds of empathy in social situations. It will likely be a long time before brain imaging of people in actual social situations is possible, and an even longer time before the reactions of viewers to feature-length films can be studied on fMRI machines. But if or when any of that happens, we are apt to learn more about how we relate to one another than we could by any other method.

How mirror neuron research might affect our analysis of the cognitive issues in narrative remains to be explored. Mirror neuron research so far has dealt in the short term, not in what happens in a two-hour movie. The complexity of narrative probably means that no single tool or research approach speaks to more than a small part of the field.   I doubt mirror neuron research will cause a shotgun marriage of cognitive and cultural studies, but the properties of mirror neuron circuits *do* connect us inextricably with cultural questions as well as with questions we are accustomed to working on.   What mirror neuron research has discovered increases – exponentially - what there is left to discover.

References.

Blakeslee, Sandra, “Cells That Read Minds” *The New York Times* Science Times Jan. 10, 2006, D1, D4.

Carr, Lauri; Iacoboni, Marco; Dubaru, Marie-Charlotte; Mazziotta, John C.; and Lenzi, Gian “Neural mechanisms of empathy in humans: A relay from neural systems for imitation to limbic areas” Proceedings of the National Academy of Sciences. www.pnas.org/cgi/doi/10.1073/pnas.0935845100. March 4, 2003

De Gelder, Beatrice. “Most Accurate Judgements of Emotional Content Both Within Cultures and Cross-culturally Involve Both Voice and Gesture.” PNAS Vol. 98, Number 17. 2001, p 1006-1010; see also PNAS Vol. 99, Number 6, pp. 4121-4126.

Glaser, Daniel; Grezes, J.S.; Calvo, B.; Passingham, R.E. & Haggard, P. “Functional Imaging of Motor Experience and Expertise During Action Observation” Unpublished. Available online at http://www.icn.ucl.ac.uk/dglaser/sfn2003abstract. Broadcast Jan. 2005 as “Monkey DO, Monkey SEE” on NOVA SCIENCE NOW. Interview with Glaser: See also; *NOVA: science NOW Research Update* “Daniel Glaser’s Latest Study with Ballet and Capoeira Dancers” http://www.pbs.org/wgbh/nova/sciencenow/3204/01-resp.html

Iacoboni, Marco; Molnar-Szakacs, Istvan; Gallese, Vittorio; Buccino, Giovani; Mazziotta, John C.; Rizzolatti, Giacomo “Grasping the Intentions of Others with One’s Own Mirror Neuron System” *PloS Biology* Vol. 3, Issue 3 March 2005

Iacoboni, Marco; Lieberman, Mathew D.; Knowlton, Barbara J.; Molnar-Szakacs, Istvan; Moritz, Mark; Throop, C. Jason; & Fiske Alan Page “Watching social interactions produces dorsomedial prefrontal and medial parietal BOLD fMRI signal increases compared to a resting baseline.” *NeuroImage* 21 (2004) 1167-1173. Available online at www.sciencedirect.com.

Järveläinen J.; Schürmann M.; Avikainen S.; & Hari R. “Stronger reactivity of the human primary motor cortex during observation of live rather than video motor acts”. *Neuroreport*. 2001 Nov. 16;12 (16):3493-5.

Keysers, Christian; Kohler, Evelyne; Umiltà, M.Allessandra.; Nanetti, L; Fogassi, Leonardo & Gallese, Vittorio “Audiovisual mirror neurons and action recognition” *Experimental Brain Research* (2003) 153:628-636.

Kohler, Evelyne; Keysers, Christian; Umiltà, M. Alessandra; Fogassi, Leonardo; Gallese, Vittorio; & Rizzolatti, Giacomo "Hearing Sounds, Understanding Actions: Action Representation in Mirror Neurons" *Reports* 2002

Leslie, Kenneth R.; Johnson-Frey, Scott H.; & Grafton, Scott T. "Functional Imaging of Face and Hand Imitation: Towards a Motor Theory of Empathy." *NeuroImage* xx (2004) xxx-xxx www.elsevier.com/locate/yinimg Available online at www.sciencedirect.com.

Lieberman, Philip. *Human Language and Our Reptilian Brain: The Subcortical Bases of Speech, Syntax, and Thought.* Cambridge, MA: Harvard University Press, 2000.

Rizzolatti, Giacomo & Craighero, Laila. "The Mirror-Neuron System" *Annual Review of Neuroscience* 2004.27:169-192

Rizzolatti, Giacomo; Craighero, Laila & Fadiga, Luciano. "The Mirror System in Humans," in *Mirror Neurons and the Evolution of Brain and Language*. *Advances in Consciousness Research* Vol 42. Ed Maxim I. Stamenov, Vittorio Gallese. Amsterdam: John Benjamins 2003: 37-59.

Sherer, Klaus R. & Wallbott, H.G. "Emotion Inferences from Vocal Expression Correlate Across Languages and Cultures," *Journal of Cross-Cultural Psychology* 32 (1), 76-92.

Wicker, Bruno; Keysers, Christian; Plailly, Jane; Royet, Jean-Pierre; Gallese, Vittorio & Rizzolatti, Giacomo. ""Both of Us Disgusted in *MY* Insula: The Common Neural Basis of Seeing and Feeling Disgust" *Neuron*, Vol. 40, 655-664, October 30, 2003.

Wilson, Steven M.; Saygin, Ayse, Pinar; Sereno, Martin I. & Iacoboni, Marco "Listening to seech activates motor areas involved in speech production. *Nature Neuroscience* Vol 7 Number 7 July 2004. Published online 6 June 2004 Nature Publishing Group www.nature.com/natureneuroscience.