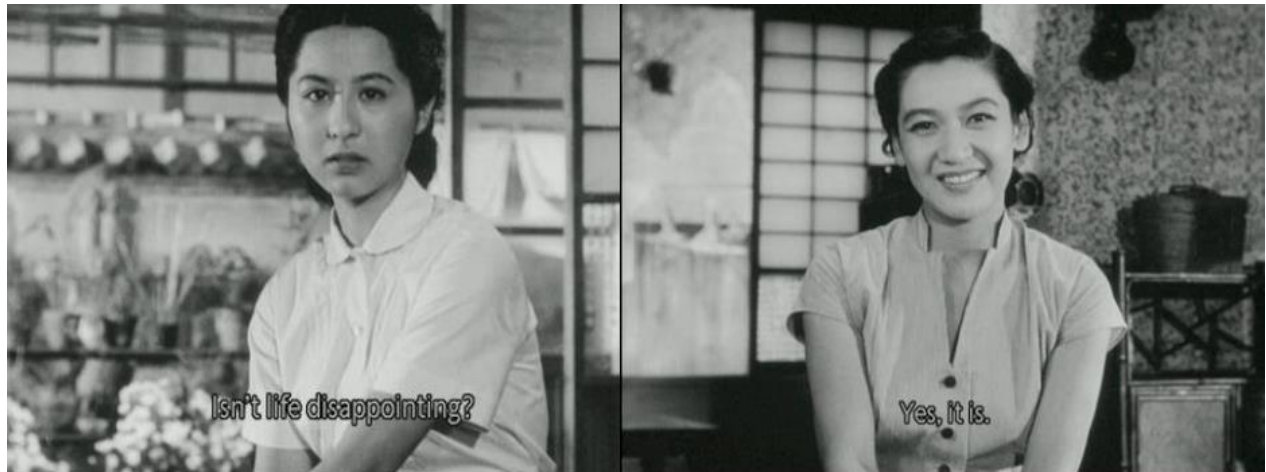


Multimodal Reading and Layered Consciousness

Robert Hanna



One of the greatest moments in the history of cinema occurs in Yasujiro Ozu's 1953 *Tokyo Story*, when, near the end of the movie, the character played by Kyōko Kagawa turns to the character played by Setsuko Hara, and asks unhappily, "Isn't life disappointing?" and Hara replies, with a wistful smile and hard-earned experiential wisdom, "Yes, it is." Both frames are shot in characteristic Ozu style, in black and white, with each actor fully facing the camera and kneeling on tatami mats, and the camera set up at the same level, with reverse shots of the same domestic setting behind them. The dialogue is in Japanese, and—in the version used in the two frames displayed directly above—obviously the subtitles are in English.

Of course, there's a great deal to be said about this particular two-frame sequence and about *Tokyo Story* itself, from the standpoint of film criticism and film theory (see, e.g., Richie, 1974). But for the purposes of this essay, I'm principally interested in the implications of this particular two-frame sequence for the philosophy of reading and for what it tells us about the nature of rational human consciousness

In an earlier essay called "How Reading Shines a Bright Light on Consciousness," I wrote this:

My proposal ... is that rational human consciousness is not only inherently schematic in nature, but also *inherently presented as inner speech whenever either scanning + parsing or comprehending occurs in reading*. In short, for those who *can* read, then *as they read*, rational human consciousness is also *inherently the subjective experience of hearing one's own voice*.

Moreover, the phenomenon of inner speech is also present in many or even most acts or processes of *silent thinking*, even when it's *not* conscious reading, since, for those who *can* read, many or even most acts or processes of silent thinking are expressed by means of sequentially generating mental imagery of *legible texts* (Hanna, 2006: ch. 4)... Ulric Neisser aptly observed that silent reading is “externally guided thinking” (Neisser, 1967, as quoted in Rayner et al., 2012: ch. 7). But by the very same token, silent reading is also rational human consciousness *externalized* onto the legible text.

This means that for those like us who *can* read, our own consciousness is characteristically directly presented to ourselves schematically *on-&-via* legible texts *as* we read them, especially when we're engaged in the highly self-conscious enterprises of formal-&-natural science or philosophy. The legible text, as read by us, is literally *the shape* of that form of our rational human consciousness. (Hanna, 2024a: p. 12)

Let's grant, for the purposes of argument, that this analysis of rational human consciousness in reading is cogent and correct.

Granting that, it's nevertheless crucial to recognize that my analysis applies to rational consciousness *in the act or process of normal silent reading of ordinary legible texts superimposed on a uniform (usually white) background*. But in the case of this particular two-frame sequence from *Tokyo Story*, and in foreign movies more generally, rational human consciousness in reading is essentially *complex*, with at least three distinct layers: (i) visual consciousness of the flow of visual images in the movie, (ii) auditory consciousness of background sound effects and the actors' voices in the original language of the movie, and (iii) inner speech consciousness of reading the superimposed subtitles as English translations of the actors' lines. All three layers are simultaneously present and also unified into a single complex visual, auditory, and inner-speech consciousness, as you watch the movie unfolding in the temporal sequence of distinct frames and as you read the subtitles from left to right in each frame and across the two frames. Let's call this *multimodal reading*.

But which, if any, of the layers of consciousness in multimodal reading is *basic*, in the sense that it unifies the single complex visual, auditory, and inner-speech consciousness of the entire phenomenological event? It seems clearly and distinctly true that the layer-(ii) auditory consciousness of the sound effects and the actor's voices in the original language of the movie, is subordinate to one or another of the other two layers of consciousness. But does the layer-(i) visual consciousness of the flow of visual images in the movie dominate over the layer-(ii) auditory consciousness of the background sound effects and the actors' voices and also over the layer-(iii) inner speech consciousness of reading the subtitles? Or, conversely, does the layer-(iii) inner speech consciousness of reading the subtitles dominate over the other two layers?

It initially seems plausible that when watching foreign movies with subtitles, the layer-(i) visual consciousness of the flow of visual images in the movie will dominate over the layer-(ii) auditory consciousness of the background sound effects and actors' voices and also over the layer-(iii) inner speech consciousness of reading the subtitles. After all, we naturally say that we *watch* foreign films with subtitles, as opposed to saying that we *listen to* them or that we *read* them.

But in this particular case, which I'm assuming to be a paradigmatic case of multimodal reading, *what is being said* by the two actors is what's basic, and therefore, for someone who speaks and reads English, but neither speaks nor reads Japanese (for convenience, let's call them *monolingual*, overlooking the possibility that they might speak and/or read languages other than English or Japanese), the layer-(iii) inner speech consciousness of reading the subtitles dominates over the other two layers. For if only the layer-(i) visual consciousness of the sequence of visual images and the layer-(ii) auditory consciousness of the background sound effects and actors' voices had been present, then for the monolingual English-speaking cognizer, the sequence could just as easily have been about some ordinary domestic event, for example, with Kyōko Kagawa asking Setsuko Hara whether she'd like a cup of tea, and Hara replying, "No, but thank you just the same." What makes the sequence so dramatically moving and existentially significant as a unified phenomenological event, is the *meanings* of the two lines of dialogue, presented to us via the layer-(iii) inner speech consciousness of reading the subtitles. Moreover, had the two lines of subtitled dialogue been merely "Would you like a cup of tea?" and "No, but thank you just the same," then our layer-(iii) inner speech consciousness of reading the subtitles would *still* have dominated over the other two layers of consciousness and determined our semantic interpretation of the two-frame sequence as only a classic Ozu-style exchange representing the finegrained texture of middle-class Japanese domestic life in the 1950s, that for most viewers would be completely forgettable in the larger context of the whole movie.

Assuming that this case from *Tokyo Story* is indeed a paradigmatic case of multimodal reading, I conclude that in multimodal reading, contrary to what initially seems plausible, our layer-(iii) inner speech consciousness of reading always dominates over the other two layers of consciousness, precisely because it determines our semantic interpretation of the other two layers and thereby brings about the unification of the whole phenomenological event, *provided that* the three layers are all coherent and concordant with one another. The mutual coherence-&-concordance provision is important. If any of the layers is incoherent or discordant with any of the others, then the layer-(iii) inner speech consciousness might not be dominant. For example, suppose that our layer-(i) visual consciousness of the flow of visual images in the movie represents an automobile speeding down a road in a big city on a sunny day, that the layer-(ii) auditory

consciousness of the background sound effects and actors' voices represents the sounds of barnyard animals quacking like ducks, snorting like pigs, and mooing like cows, with no human voices to be heard, and that our layer-(iii) inner speech consciousness of reading the subtitles represents the words "It is a truth universally acknowledged, that a single man in possession of a good fortune, must be in want of a wife" —i.e., the famous first sentence of Jane Austen's *Pride and Prejudice*. In this case, none of the layers of consciousness would clearly dominate over the others, and the phenomenological event as a whole wouldn't be unified.

With the mutual coherence and concordance provision in place, then the thesis that our layer-(iii) inner speech consciousness of reading always dominates over the other two layers of consciousness, precisely because it determines our semantic interpretation of the other two layers and thereby brings about the unification of the whole phenomenological event, also generalizes to the slightly differing cases of *closed captioning*, in which the voices are speaking in the same language as the subtitles, and *silent movies*, in which the soundtrack is exclusively musical and the legible text is presented as *intertitles* between the flow of visual images in frames or sequences. That all being so, then we can rightly say that monolingual rational human animals basically *read* (i) foreign films with subtitles, (ii) closed captioned TV, movies, and other video presentations, and (iii) silent films, and only in a derivative way either *watch* any of them or *listen to* any of them. More synoptically, *consciousness is a many-layered thing*, and the inner speech consciousness of reading is arguably the cognitively and philosophically *most* basic layer of consciousness for rational human animals (Hanna, 2023a, 2023b, 2024a, 2024b).¹

¹ In the context of this essay, I've defined basicness as the "top-down" feature whereby one mode of consciousness controls and unifies other modes of consciousness. But if we focus instead on the "bottom-up" feature whereby consciousness is necessarily and completely embodied in a suitably complex living organismic body —i.e., the feature whereby consciousness is *essentially embodied* — then pre-reflective desire-based emotional feeling and primitive bodily awareness jointly constitute the *fundamental* layer of rational human consciousness (Hanna and Maiese, 2009: esp. section 1.2).

REFERENCES

(Hanna, 2006). Hanna, R. *Rationality and Logic*. Cambridge: MIT Press. Available online in preview at URL = <[https://www.academia.edu/21202624/Rationality and Logic](https://www.academia.edu/21202624/Rationality_and_Logic)>.

(Hanna, 2023a). Hanna, R. "The Philosophy of Reading as First Philosophy." Unpublished MS. Available online at URL = <[https://www.academia.edu/107390679/The Philosophy of Reading as First Philosophy September 2023 version](https://www.academia.edu/107390679/The_Philosophy_of_Reading_as_First_Philosophy_September_2023_version)>.

(Hanna, 2023b). Hanna, R. "Caveat Lector: From Wittgenstein to The Philosophy of Reading." Unpublished MS. Available online at URL = <[https://www.academia.edu/103950659/Caveat Lector From Wittgenstein to The Philosophy of Reading October 2023 version](https://www.academia.edu/103950659/Caveat_Lector_From_Wittgenstein_to_The_Philosophy_of_Reading_October_2023_version)>.

(Hanna, 2024a). Hanna, R. "How Reading Shines a Bright Light on Consciousness: The Science of Reading." Unpublished MS, Available online at URL = <[https://www.academia.edu/108993104/How Reading Shines a Bright Light on Consciousness The Science of Reading January 2024 version](https://www.academia.edu/108993104/How_Reading_Shines_a_Bright_Light_on_Consciousness_The_Science_of_Reading_January_2024_version)>.

(Hanna, 2024b). Hanna, R. "The Internal Structure of Reading and the Internal Structure of Philosophizing." Unpublished MS. Available online at URL = <[https://www.academia.edu/108765998/The Internal Structure of Reading and the Internal Structure of Philosophizing January 2024 version](https://www.academia.edu/108765998/The_Internal_Structure_of_Reading_and_the_Internal_Structure_of_Philosophizing_January_2024_version)>.

(Hanna and Maiese, 2009). Hanna, R. and Maiese, M., *Embodied Minds in Action*. Oxford: Oxford Univ. Press. Available online in preview at URL = <[https://www.academia.edu/21620839/Embodied Minds in Action](https://www.academia.edu/21620839/Embodied_Minds_in_Action)>.

(Neisser, 1967). Neisser, U. *Cognitive Psychology*. New York: Appleton Century Crofts.

(Rayner et al., 2012). Rayner, K., Pollatsek, A., Ashby, J., and Clifton Jr, C. *Psychology of Reading*. 2nd edn., New York: Psychology Press/Routledge.

(Richie, 1974). Richie, D. *Ozu: His Life and Films*. Berkeley CA: Univ. of California Press.