



# Experiences of silent reading

Carola Barbero<sup>1</sup> · Fabrizio Calzavarini<sup>1</sup>

Accepted: 8 February 2024  
© The Author(s) 2024

## Abstract

In *The Performance of Reading*, Peter Kivy introduces, on a purely phenomenological basis, an interesting and potentially fruitful analogy between the experience of silently reading literary texts and the experience of silently reading musical scores. In Kivy's view, both mental experiences involve a critical element of auditory mental imagery, consisting in having a performance "in the head" or the mind's ear. This analogy might have significant implications for the ontological status of literary works, as well as for the theoretical relations between music and language. Nevertheless, Kivy's hypothesis has never been investigated and discussed in its empirical merits. In the present paper, we shall claim that neuroscience data support, at least in part, Kivy's phenomenological observations about the relation between reading musical scores and reading texts. Despite being functionally and anatomically dissociated at the cognitive level, the two reading experiences both involve an auditory simulation of the content, which seems to be functionally critical for a deep and rich experience of literary texts and musical scores.

**Keywords** Silent Reading · Musical Scores · Literary Texts

## 1 Reading texts and musical scores

In the philosophical literature, the comparison between music and language has been investigated at various levels. For example, in *Music, Language, and Cognition*, Peter Kivy has famously argued that music has a linguistic nature, for it has a syntactic component that is similar –although not identical– to that of natural language: "musical 'syntax' may not be syntax in the literal sense, but syntax-like, as music is language-like, but not language" (Kivy, 2007, p. 216). On the other hand,

---

✉ Fabrizio Calzavarini  
fabrizio.calzavarini@unito.it

Carola Barbero  
carola.barbero@unito.it

<sup>1</sup> Department of Philosophy and Educational Science, University of Turin Via Sant', Ottavio 20, 10124 Turin, Italy

he has underlined that music is different from language since it does not possess a semantic component similar to linguistic meaning – although it still has emotive properties that can be recognized by competent listeners and constitute a sort of emotive “vocabulary” (Kivy, 2007, p. 220; see Di Bona, 2019).

In *The Performance of Reading* (2006; see also Kivy, 2010), Kivy discusses the comparison between music and language, particularly literary language, on a different ground. According to Kivy reading a literary text can be seen as a performance, similar to the way musicians perform music.<sup>1</sup> By defining reading as performance, he aims at defending an original ontological thesis according to which the distinction between performing artworks –or two-stage artforms (where the realization of the work requires some form of execution)– and standard artworks (or one-stage artforms), literature has to be seen as belonging to the former kind. To explain the kind of objects literary works are he uses the type/token distinction, but distancing himself from the traditional view (basically, Goodman (1968) and Wollheim (1968) consider the type as instantiated by the text tokens, i.e., the written copies of books) and considering readings seen as performances as tokens of those types (see Ribeiro, 2009 for discussion). In Kivy’s view,

You have your copy of *Pride and Prejudice* and I have mine. But, I would urge, our copies of the novel are not tokens of the type *Pride and Prejudice*, any more than our scores of Beethoven’s Fifth Symphony are tokens of the type Beethoven’s Fifth Symphony. All of the many copies of *Pride and Prejudice* are tokens of a type, but that type is not the work: it is the notation of the work (Kivy, 2006, p. 4).

It is on the strength of this ontological thesis that the analogy between literature and music can be investigated, and reading literary texts could be seen as similar to reading musical scores. Indeed, this does not mean that the two reading experiences are exactly on par in Kivy’s view, but only that they are similar in relevant respects: “it is not being argued on these pages that score-reading and novel-reading are identical; only that the former can provide for us a useful, illuminating analogy for the latter” (Kivy, 2006, p. 32).

<sup>1</sup> The idea that the silent reading is a form of performance is not new. Wolfgang Iser, in *The Act of Reading* (1978) and in *Prospecting: From Reader Response to Literary Anthropology* (1989), explores this point in depth. In Iser’s view, reading a literary text is a dynamic gap-filling activity similar to what a performer does while interpreting and enacting a script. Stanley Fish, in *Is there a text in this class?* (1987), defends a similar perspective (“Interpretation is not the art of construing but the art of constructing, interpreters do not decode poems; they make them”, p. 327). In *Experiencing Narrative Worlds: On the Psychological Activities of Reading* (1993), Richard Gerrig observes that “readers are often described as being transported by a narrative by virtue of performing that narrative” (p. 2). In Gerrig’s theory, the reader’s engagement with a narrative is considerably similar to an actor’s performance when understanding and bringing characters and stories to life. None of these authors, however, made an explicit analogy between reading literary texts and reading musical scores, which is the primary target of the present article.

## 1.1 Disanalogies

Kivy grants in several points that the analogy between literature and music, as their ontological structure and reading performance are concerned, is not perfect – some important differences between the two need to be pointed out. First, musical scores give instructions to readers for realizing a performance in their head, whereas literary texts do not give precise parameters. Maybe this depends on the fact that musical performance concerns one sense modality, whereas literary performance is a matter of two (visual and aural imagination) (Kivy, 2006, pp. 40–41). Second, it seems that the criterion of complete success in reading a musical score consists in realizing a musical performance, whereas the criterion of success in reading a literary text is understanding it (Kivy, 2006, p. 41). Text reading has to do with comprehension, which is a complex process that involves more than just decoding words. It also includes making connections, making inferences, and grasping the general meaning in its context. For sure, reading musical scores does not appear to capture all these cognitive aspects.

Another important disanalogy between silently reading a score and a novel is the one concerning multiple voices. When the score is of polyphony, the silent reading will consist in performing simultaneously “in the head” many “voices”; but when reading a novel, it seems that only one voice sounds at a time, even though it may be a different voice from that of the storyteller – for instance, it might be of one of the characters in direct quotation (Kivy, 2006, p. 39). One significant disanalogy worth highlighting is that silently reading of musical scores is a competence reserved for a few experts even within the musical domain. On the other hand, it seems that no exceptional talent or specialized training appears to be necessary for the silent reading of literary works. This means that anyone, regardless of their background or expertise, can engage in the silent reading of literary texts. (Kivy, 2006, pp. 63–4). In Kivy’s view,

There is a disanalogy here between the silent reading of novels and the silent reading of musical scores. The former is possible for any literate person, which means, potentially, every normal person, whereas the latter is possible for only the musically gifted few (Kivy, 2006, p.116).

In this respect, an important issue is the one concerning interpretation, which is considered a fundamental element of reading as performance in Kivy’s account. According to Kivy, “[a] performance is a version of the work performed. And for a performer to produce a credible performance, a credible version or “reading” of the work, she must have an interpretation of it” (2006, p. 61). This might highlight a further potential difference between reading texts and reading scores. On the one hand, according to Kivy, we have “no difficulty in thinking of score readers as hearing in the head interpretation-driven performances: Score readers are highly trained, gifted, and experienced musicians, musical performers in their own right, at the professional level, who know the nuts and bolts of the compositions they perform and of the scores that they read and realize in silent performance” (2006, p. 67). On the other hand, it is not immediately clear whether also

novel reading always involves an interpretation, since “ordinary readers of novels [...] are not “professional” novel readers (whatever that would mean). They are not writers, book reviewers, literary critics, or professors of English literature” (2006, pp. 67–68).

Nevertheless, according to Kivy, truly competent readers of literary texts must possess a significant level of literary sensibility, appreciate a diverse array of genres and forms, and demonstrate a sincere commitment to the endeavor by dedicating some of their time to reading literary criticism (2006, p. 64). Readers of this level of sophistication “are not without preconceived ideas of how what they are about to read will go” (p. 69), and always read with interpretation and *con espressione*, unless they are autistic or there is a special request for an expressionless monotone reading (Kivy, 2006, p. 46).<sup>2</sup> Thus, both the silent reading of literary and musical works comes with an interpretation, and on the very first “sounding in the head”. More importantly, according to Kivy, potential disanalogies at this level are irrelevant to the main point he wants to defend: “The point is that what happens in the head of the novel reader is the same thing that happens in the head of the score reader: a silent performance of the notated work” (2010, p. 116).

## 1.2 The fundamental analogy

In Kivy’s view, the experiences of reading musical scores and reading texts are characterized by a fundamental analogy, that is, the involvement in both cases of a critical experience of auditory mental imagery.<sup>3</sup> Exactly as happens with music, where

<sup>2</sup> Kivy also says that ancient rhapsodes in general used to criticize while acting (or singing) (Section 23) and this could be seen as part of the very act of interpretation. Therefore, reading and performing, together with reflection and criticism, are to be considered as part of the literary experience (pp. 102, 110). That’s why silent readers turn out to be “silent Ions”: because the performance in the silent reading of a novel is analogue to Ion’s performance, as described by Plato in the homonymous dialogue (p. 116). Anna Christina Ribeiro (2015, pp. 127-148), while remembering how, since poems were originally made and performed orally, the performance of literature should be seen as paradigmatic, underlines the difficulty of the relation between work and interpretation when appreciation is concerned: “Where do we draw the line of scholarly erudition appropriate to the experience of a work? The demarcation cannot but be arbitrary. Should we grant ‘experience’ of the *Iliad* only to Martin West and Gregory Nagy? Tough, since they have different and conflicting views of that epic alone. This kind of question invariably runs into the ‘paradox of distance’ long ago discussed by Edward Bullough. Homer (if he existed) was not composing for scholars; it is a question whether they can experience the work in the ‘right’ way, since they are arguably often unable to leave aside their scholarly concerns” (p. 144).

<sup>3</sup> One might object that the term “auditory imagery” is hard to understand, since the term “image” is intrinsically tied to the visual modality (as noted by Hobbes, for instance, “[a]n image, in the most strict signification of the word, is resemblance of something visible; [a]nd from hence it is manifest, that there neither is, nor can be any image made of a thing invisible”; 1651/2010, XLV, 358). Nevertheless, although a visual bias has certainly characterized research on mental imagery for a long time, it is now established that imagery can characterize all sense modalities. As observed by Bence Nanay in his recent influential book on this topic (2023), for instance, “vision is not the only sense modality. So, if we can perceive auditorily, olfactorily, and so on, we can also have auditory, olfactory, tactile, etc. mental imagery. [I]t should be clear that the word “imagery” does not here denote anything that has to do with pictures (which would usually be something visual): mental imagery exists in all sense modalities” (p. 17).

one can silently read a musical score and “hear” the music in one’s mind, also when silently reading a literary work we have a performance of it, and in both cases, we could be said having a performance “in the head” or in the mind’s ear. In text reading, the imagery component that Kivy has in mind is different from the reader’s ability to simulate the narrative world, which involves visualizing scenes, understanding spatial relations, or empathizing with characters (Kuzmičová, 2013, 2014). It is instead a kind of verbal simulation, which has to do with the auditory aspect of reading.<sup>4</sup> Thus, in Kivy’s view, both reading texts and reading musical scores involve the mental recreation of sounds, be it the voice of the narrator in a literary work or the timbre and pitch of musical notes in a score. On the one hand, according to Kivy, “one can, if one is a good enough and talented enough musician, internally realize the sound of a work by means of simply reading the score. (...) One can silently read a musical score and, through the silent reading, ‘hear’ in one’s mind the musical work: a realization of the sound of the work. One can ‘hear’ a production in the mind”. (Kivy, 2006, p. 35–36). On the other hand, it is “consistent with what we know about thought and consciousness that reading a story might be experienced as ‘hearing’ a story told in your head, in the way that reading a score, we know, is experienced, by those few who can, as ‘hearing’ music played in your head” (Kivy, 2010, pp. 111–112). In this sense, according to Kivy, “we hear stories in the head, the way Beethoven, when he read the scores of Handel, heard musical performances in the head” (Kivy, 2006, p. 63).

### 1.3 Deafness and direct quotes

An immediate objection to this position is that it makes the ability to read literary texts parasitic on the ability to hear sounds, making it impossible for congenitally deaf people to experience a true appreciation of literary works: “(i)f reading novels is hearing storytelling in the head, as reading scores is hearing music in the head, then surely, it will be argued, just as it is agreed on all hands that someone born deaf could not read a musical score with musical understanding and appreciation, he could not read a novel with literary understanding and appreciation” (Kivy, 2006, p. 119).

Kivy argues that this position is not as absurd as it may seem. From an empirical point of view, according to Kivy, there is already established evidence that reading

<sup>4</sup> As is known, the involvement of mental imagery in reading narrative texts has been intensively investigated within the embodied cognition literature and within the cognitive literary studies (see, for instance, Paivio, 1986; Sadoski & Paivio, 2001; Kuzmičová 2013, 2014), also by relying on neuroimaging techniques (e.g., Mak et al., 2023). In this context, Anezka Kuzmičová has explicitly distinguished between two forms of literary-related mental imagery, which she calls, *verbal presence* and *direct presence*: “While verbal presence refers to the reader’s vicarious perception of the voices of narrators and characters, direct presence refers to the emulated sensorimotor experience of the imaginary worlds that the narrators’ and characters’ utterances refer to” (2013, pp. 107–108). Kivy is primarily interested in the first kind of speech-imagery, which, according to Kuzmičová, “puts the reader in the position of a vicarious listener, someone who is receiving the text as if it were spoken out loud by an extraneous speaker” (2014, p. 284).

competence in deaf persons is severely defective: “congenitally deaf children have extreme difficulty in learning to read, and consistently perform below the base level of normal children in school –a condition which [...] persists after they have left school, and into adulthood” (Kivy, 2006, p. 125). Hence, their ability to have a rich and deep experience of literary works cannot be taken for granted (Kivy, 2006, pp. 125–25).

The difference between normal and deaf subjects is particularly pressing when direct quotes are involved in novels. According to Kivy, the normal silent reader experiences characters’ direct speech in a phenomenologically distinct way from the one typical of mere storytelling, since he would experience it as the voice of a character speaking “in his head” (in a similar way as a score reader would “hear in his head” the clarinet *solo* in a symphony) (Kivy, 2006, p. 119). And from this should follow that a person born deaf – even if able to distinguish, from the marks on the page, the different parts of the text, for instance, a dialogue –, being unable to know what a dialogue sounds like, because she has never heard a human voice, wouldn’t experience that dialogue/speech phenomenologically in the same way as someone with unimpaired hearing, and therefore she wouldn’t be, under this specific point of view, able to fully appreciate the literary work. (cfr. Shusterman (1978, pp. 324–325): the “appreciation (of the deaf) lacks an important element in the case of works built heavily on oral effects”).

it is clear that if the phenomenological experience we get when silently reading direct quotation in novels is anything like hearing declamation “in the head,” then readers born deaf cannot have that experience, and are deficient in that respect (Kivy, 2006, p. 120).

Kivy, therefore, needs to conclude that readers born deaf would have (in a very real sense) a defective, not full, aesthetic experience (that is, the silent reading as performance) and that this account of the silent novel-reading does not work for the congenitally deaf persons (Kivy, 2006, p. 120).

## 2 The cognitive perspective

As is known, Peter Kivy’s theory in *The Performance of Reading* has sparked a lively debate in philosophy of mind, particularly regarding the concept of performance in the silent reading, as well as the ontological implications of this idea (for discussion, see Feagin, 2008; Ribeiro, 2009; Morais, 2010, 2017). Critics like Susan Feagin (2008), for instance, argue that the concept of “performance” inherently implies an audience, a condition that is not present in the silent reading. She distinguishes between the skills needed for performance, like playing the piano in front of an audience, and the private act of reading. More generally, according to Feagin, “[w]hat it is to have the ability to read — even to read literature as art — is not the same as what it is to have the ability to perform or give a performance, by

reading a novel or playing a piece of music. Reading literature as art is not necessarily reading as performance” (p.92).<sup>5</sup> Similarly, Ribeiro argues that “performances [...] are themselves objects of aesthetic experience and appreciation, something that silent readings (or soundings) [...] cannot be” (2009, p. 190). Ribeiro also questions Kivy’s idea that readers always approach literary texts with interpretation and expectations, for “this seems to be the exception rather than the rule” (2009, p. 191).

The aim of the present paper is not to delve deeply into these philosophical issues. Instead, we will focus on a specific and more basic aspect of Kivy’s theory, exploring it independently from the broader debate about reading as a kind of performance (with interpretation) and laying the groundwork for further research in this area. We have seen that Kivy advances, on a purely phenomenological basis, an interesting and potentially fruitful analogy between the experience of reading texts and the experience of reading musical scores.<sup>6</sup> In Kivy’s view, both mental experiences involve having a performance “in the head” or in the mind’s ear. This phenomenological analogy lies at the basis of most of the philosophical observations put forward by Kivy. Based on a widely accepted methodological assumption (cfr. Drożdżowicz, 2023), however, it is not possible to determine the nature of mental experiences based solely on phenomenology. Philosophical theories regarding mental experiences should, at the very least, align with our empirical knowledge of them. Is the analogy advanced by Kivy justified on the empirical ground? In other words, are there at least minimal pieces of empirical evidence supporting the idea that the silent reading of scores involves similar cognitive structures as those involved in reading texts? The remainder of this paper will be dedicated to establishing this issue.

## 2.1 Cognitive disanalogies

Experimental evidence suggests that reading texts and reading scores are largely independent cognitive processes. Conclusions from the eye-tracking studies, for instance, suggest that the two modes of reading are differentiated due to several factors. One key

<sup>5</sup> This problem is at least partly originated by the intrinsic semantic ambiguity of the word “performance”. The *Oxford English Dictionary*, for instance, lists several meanings of this word, including “(1) The accomplishment or carrying out of something commanded or undertaken; the doing of an action or operation” and “(4) The action of performing a play, piece of music, ceremony, etc.; execution, interpretation”. According to Feagin (2008), the passage from (1) to (4) is not automatic or direct at all, also because (4) demands a kind of cognitive effort that might distract the reader from the content of the text: “I may, occasionally, silently impersonate a storyteller when I read, but I do so rarely because I find that it interferes with my efforts to understand and it slows things down. If I think about how to read the sentence — gravely, with a certain light-heartedness, with the stress on this phrase rather than that one — it takes attention away from the content of my reading, for example, the plot, characters, symbolism, and so on.” (p. 93).

<sup>6</sup> The analogy between reading literature and reading scores has been already formulated by other scholars, notably Levi Strauss in his “The Structural Study of Myth” (1955). In particular, Strauss was here reflecting on the vertical and horizontal dimensions of the reading of orchestra scores as compared to the reading of myths (p. 432; see also pp. 443–44). We thank an anonymous reviewer for this suggestion. As we have seen, however, Kivy’s comparison is based on a different aspect, which has to do with auditory imagery.

distinction is that text reading occurs in a sequential or horizontal manner, whereas music reading involves both sequential and simultaneous or vertical processing (cfr. Sloboda, 1984). There is evidence that reading musical scores is linked to extended periods of focused eye gaze, whereas reading written language, especially literary texts, is associated with a higher frequency of backward eye movements (Cara and Gómez Vera, 2016).

Neuropsychological observations –that is, observations from brain-damaged patients– also support this conclusion. In many cases, disorders in reading music and reading language are associated (cfr. Hébert and Cuddy, 2006). Nevertheless, clinical cases have been reported of musicians with an inability to read words but the preserved ability to read musical notation (e.g., Mendez, 2001). As a representative example, Cappelletti et al., (2000) reported the case of a professional musician, PKC, who suffered lesions in the left posterior temporal lobe and right occipitotemporal regions following hemorrhagic encephalitis. Due to the illness, PKC “was quite unable to read musical notes, whatever the task and whatever response was required by it”, although she “performed flawlessly or at a very high level on all tests of reading aloud letters, words, novel letter strings and sentences” (*ibidem*, p. 329). The reverse pattern of dissociation, with the preserved reading of texts and impaired reading of musical scores, has also been reported (e.g., Ayotte et al., 2000; Steinke et al., 2001). As observed by Mendez, (2001),

the finding of two reverse forms of dissociation constitutes an instance of neuropsychological double dissociation, (suggesting) a functional autonomy of text and music reading as well as structural independence of their neurobiological substrates (p. 201).

Until some years ago, there was a limited number of studies directly comparing the neural activation patterns associated with score reading and language reading (cfr. Besson and Schön, 2001). This scarcity is explained, at least in part, by the difficulty of individuating the right level of language competence that has to be compared with music processing. However, more recently the use of direct cortical stimulation by Roux et al., (2007) has shed some light on this issue, demonstrating that the brain regions involved in score reading are typically distinct from those involved in text reading. Across subjects, the activation of the intraparietal sulcus and the middle temporal gyrus was found to be selective to music reading, while only specific regions of the supramarginal gyri were associated with text reading (Roux et al., 2007). Furthermore, a study by Mongelli et al. (2017) showed that letters and musical notes elicit different patterns of activation in the occipitotemporal cortex, with music-related activations being more posterior and lateral compared to word-related activations. These intriguing findings, which align with previous patient case studies, provide further evidence for the existence of a distinct neural basis underlying text and music reading.

## 2.2 Reading scores in the mind’s ear

As we have seen, reading scores and reading texts appear to be largely autonomous processes at the cognitive level. Of course, these findings do not entirely rule out the



hypothesis that the two reading experiences are both accompanied by a vivid experience of mental auditory imagery, namely a simulation of the actual auditory perception of musical notes and voices, respectively, as suggested by Kivy in *Performance of Reading*. Is this phenomenological observation supported by empirical data?

As regards musical experts, a possible alternative hypothesis is that silently score reading is exclusively a visual-motor activity that doesn't involve any of the distinct cognitive processes required for perceiving and understanding music (cfr. Sloboda, 1984 for discussion). According to this hypothesis, visual information about written musical notation is directly transformed by the brain into motion patterns, without the involvement of any imagery-based musical mediation: "(t)hus, a particular symbol may be interpreted as a prescription for a particular movement pattern with respect to a particular instrument without any internal representations of the musical sound of the note. On this hypothesis, it is only after the notes have been produced, thereby allowing the performer to actually hear them, that any auditory perception takes place" (*ibidem*, p. 223). Overall, however, empirical data suggest that silent reading of musical scores is not a purely visuomotor task. Schön et al., (2002) have proposed the existence of at least three cognitive pathways involved when reading music: (i) a *visual-to-motor* pathway for playing on an instrument (transcoding the visual symbols into a movement); (ii) a *visual-to-verbal* pathway for naming notes (transcoding the visual symbols into verbal articulation); and (iii) a *visual-to-singing* pathway (transforming a visual symbol into auditory representations).

Critically, some form of an auditory internal representation of the written music seems also automatically involved when reading scores. This is suggested by the fact that the reading performance is sensitive to the auditory features of the scores (see Sloboda, 1984 for discussion), such as tonality coherence (Halpern & Bower, 1979), musical plausibility (Sloboda, 1977), and musical phrase structure (Sloboda, 1974). More importantly, there is evidence that the brain regions associated with auditory music perception are activated during the silent music reading of musical scores. For instance, using the neuroimaging technique (fMRI), Mongelli and colleagues have demonstrated that the regions in the primary and secondary auditory cortices involved in pitch and rhythmic processing are also activated when both novices and experts are silently reading musical notation, and their activation is correlated to musical expertise (Mongelli et al., 2017). This strongly suggests that an auditory simulation of the musical content is automatically triggered when reading musical scores, following Kivy's hypothesis.

### 2.3 Reading texts in the mind's ear

As we have seen, empirical data appear to support Kivy's hypothesis that silent reading involves hearing a musical performance in the mind's ear. Is it the same for reading texts? Although many of us share the intuitive subjective experience of inner speech, or a "voice inside our heads" while reading texts silently, as confirmed by subjective reports (Vilhauer, 2017), this phenomenon remains underspecified at the cognitive level. According to an alternative hypothesis, detailed auditory representations (i.e., information about the sound of a word) are not necessary for silent

reading, since phonological representations are coded in an abstract format (cfr. Frost, 1998).

Nevertheless, evidence suggests that when reading, the mental representations accessed by individuals may include perceptual aspects similar to actual speech (cfr. Lukatela et al., 2004). For example, Abramson and Goldinger, (1997) conducted a study where subjects were asked to read lists of visually presented words. Words were matched for length but varied in the length of the vowel sounds when spoken. Surprisingly, subjects in this study took longer to respond to words with long vowels compared to words with short vowels, despite the presentation being visual. This finding suggests that during silent reading, individuals access representations that preserve auditory temporal properties similar to those of spoken words. In another investigation (Ashby & Clifton, 2005), it was found that words containing more emphasized syllables tend to be observed for a longer period, suggesting that even subtle phonetic aspects of speech are mentally represented during the reading process.

Neuroimaging studies have provided strong empirical support for the hypothesis proposed by Kivy regarding the involvement of auditory imagery during text reading, particularly when reading direct quotes. Yao et al., (2011) used functional magnetic resonance imaging (fMRI) to examine if direct speech triggers specific activation in brain regions known to be responsive to human voices. The study revealed that certain areas of the temporal lobe, which are sensitive to voice perception, were specifically engaged when processing direct speech as opposed to indirect speech. Additionally, the researchers noted a connection between the experience of an "inner voice" during the reading of direct speech and the patterns of eye movements observed during silent reading (Yao et al., 2011). Electroencephalographic studies have additionally shown that neural activity in the voice-selective auditory areas correlates with the acoustic characteristics of utterances, providing further evidence of the close resemblance between the experience of silently reading and the perception of external voices (Petkov & Belin, 2013).

## 2.4 Deafness and direct quotes

Is auditory imagery critical for reading literary texts and musical scores with understanding and appreciation, as conjectured by Kivy? Also, in this case, the answer appears to be positive. On the one hand, although people who cannot engage in auditory imagery, that is, congenitally deaf subjects, can sometimes achieve impressive results in musical training and performance (cfr. Hash, 2003), especially with the aid of cochlear implants that provide temporal and spectral auditory cues (cfr. McDermott, 2004), their ability to understand (cfr. Trehub et al., 2009) and appreciate (PrevotEAU et al., 2018) musical pieces remain significantly below that of normal subjects.

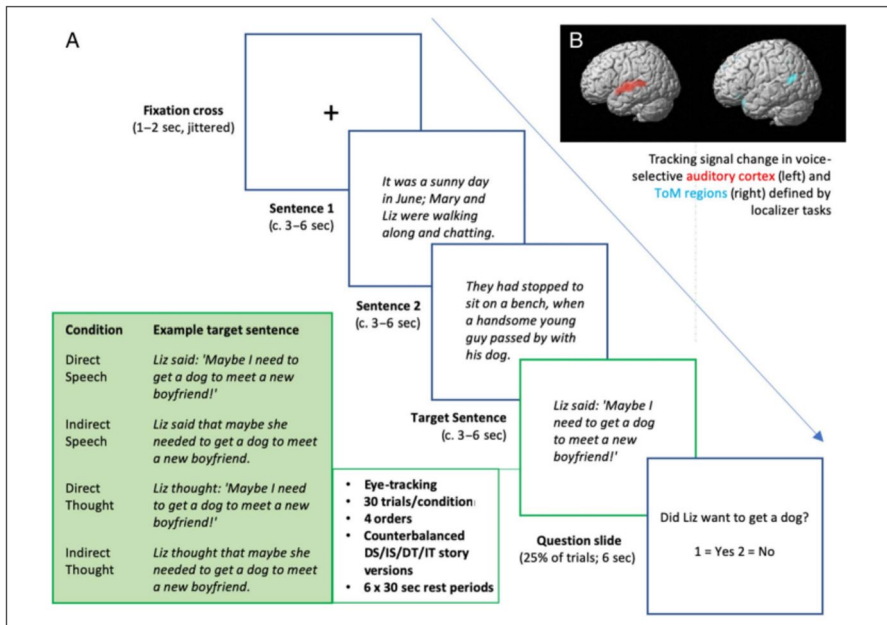
On the other hand, as observed in passing by Kivy, there is robust and established empirical evidence that deaf subjects tend to develop severe difficulties in several reading skills involving texts, such as reading fluency (Luckner and Urbach, 2012), word recognition (Luckner & Handley, 2008), or syntax identification (Traxler

et al., 2014). In a broader sense, deaf subjects consistently achieve lower scores on standardized assessments of reading comprehension (for review, see Alothman, 2021). Interestingly, recent studies have demonstrated that deaf children showed marked deficits in storytelling skills, scoring poorly on narrative structure and cohesion (Crosson & Geers, 2001). Even though continuous use of cochlear implants improves narrative abilities during development (Nikolopoulos et al., 2003), children with auditory implants still perform below the standard in narrative production tasks (Boons et al., 2013). Critically, auditorily impaired students tend to be outperformed by their peers in understanding simple fictional pieces (Gómez-Merino et al., 2022). Perhaps as a result of these difficulties, deaf students tend to remain reluctant readers during all phases of their development (Smetana et al., 2009), showing a low appreciation of narrative texts (Kim, 2021).

Based on the available data, however, it is not possible to directly verify Kivy's conjecture that deaf subjects have particular difficulties with direct quotes in novels. To our knowledge, no study has specifically assessed this aspect. Nevertheless, studies involving hearing readers seem to point in this direction. As we argued, in normal subjects silent reading of direct (vs indirect) speech in literary texts activates brain regions that have been associated to perception of human voices, following the thesis that "(c)omprehension of direct speech is [...] grounded in the perceptual experience of a vocal demonstration or dramatization of a reported speaker's utterance, and thus more likely to invoke perceptual simulations of the reported speaker's voice" (Yao et al., 2011, p. 252).<sup>7</sup> This effect is observed only when characters are represented as speaking, but not when they are represented as thinking (Alderson-Day et al., 2020, see Fig. 1). Moreover, it has been shown that neural (theta-band, 4–7 Hz) activity in the auditory cortex encodes the rhythmic dynamics of direct (vs indirect) quotes reading, suggesting "a functional role of theta-phase modulation in reading-induced inner speech" (Yao et al, 2021, p. 2).

Critically, it has been observed that phonological interference (tongue-twister effects) has a selective disruptive impact on the silent reading, specifically in the case of direct speech rather than indirect speech., pointing to "a causal link between

<sup>7</sup> Such effects are typically explained in terms of the tenets of recent "simulation" theories of language understanding (e.g., Barsalou 1999; 2009). As noted by Yao and Scheepers (2015), for instance, "[s]uch theories propose that language processing is grounded in mental simulations (or re-enactments) of sensory and motor experiences that have been acquired through individuals' interaction with the environment and their internal states. Under such a premise, accumulated experiences with how direct versus indirect speech are typically reported in spoken language could form the basis for differential mental simulations during written-language processing. In other words, silent reading of direct speech would be grounded in mental simulations of vivid vocal depictions whereas silent reading of indirect speech would be grounded in simulations of voices that are more neutral. The brain may therefore be more prone to activate "audible-speech"-like representations during silent reading of direct speech than of indirect speech" (p 291). The notion of "simulation" used in this framework is similar to that of "sensory imagery", that is, the "the re-enactment of perceptual, motor and introspective states acquired during experience with the world, body and mind" (Barsalou, 2009, p. 1281). The main difference is that perceptual simulations can be unconscious, whilst imagery experiences are typically considered to be conscious. For instance, Barsalou argues that perceptual re-enactments are potentially accessible in the form of explicit imagery (they could quickly generate conscious images), even though sometimes they do not gain phenomenal access (1999; 2009).



**Fig. 1** Experimental design of the study of Alderson-Day et al., (2020). See Alderson-Day et al., (2020) for further details

phonological simulations and silent reading of direct speech” (Yao, 2021). This in turn suggests that the inner voice is not just epiphenomenal but is a critical component of the silent reading of direct quotes in literary works, a component that is probably not at the disposal of deaf readers, making their overall aesthetic experience defective.

### 3 Conclusions and future directions

To summarize, neuroscience data support Kivy’s phenomenological observations about the relation between reading musical scores and reading texts. Despite being functionally and anatomically dissociated at the cognitive level, the two reading experiences involve an auditory simulation of the content, similar to having a performance in the mind’s ear. Interestingly, this auditory imagery experience seems to be critical for a deep understanding and appreciation of musical scores and literary texts, as suggested by studies on reading abilities in deaf subjects. Particularly noteworthy are Kivy’s intuitions about the silent reading of direct quotes in novels, which appear to be vindicated by recent neuroscience research. Overall, these considerations certainly justify further lines of investigations into the philosophical and cognitive relation between musical and linguistic abilities.

To make a few examples, it might be interesting to assess how the analogy between reading scores and reading texts, especially in terms of auditory imagery,

can be influenced by literary genres. Postmodernism and formalism, usually known as having offered “musical literary works” (take, for instance, the first lines of *Lolita* by Nabokov),<sup>8</sup> strengthen this analogy. Here, the rhythm and sound of the language can evoke a musical experience in the reader’s mind, similar to interpreting a musical score. Conversely, genres like realism, exemplified by the beginning of Flaubert’s *Madame Bovary*,<sup>9</sup> focus on descriptive precision, potentially engaging different cognitive processes less akin to musical interpretation. This suggests that the genre significantly influences how closely the reading of texts aligns with the experience of reading music scores at the cognitive level. To our knowledge, no experimental studies have been conducted so far on this topic.

Another potential line of research involves the role of interpretation and expectations, which, in Kivy’s view, accompany silent readings of both literary and musical works on the very first “sounding in the head” (see § 1.1). An initial piece of evidence in support of this view is the observation that, during silent reading of narratives, the strength of activation in imagery-related areas of the brain, including the auditory cortex, is correlated with quantitative measures of subjective reading experiences (story world absorption, story appreciation) and reading habits (Mak et al., 2023). Similarly, it has been shown using EEG that melodic experiences and expectations, quantified with statistical models trained on a musical database summarizing the music listeners have encountered, modulate the neural responses observed during musical imagery tasks (Marion et al., 2021). Future research might further explore how individuals’ previous knowledge and contextual understanding influence their reading of musical and literary texts.

A further open issue concerns the perceptual nature of the inner voice involved in silent reading, which possesses a strictly auditory character in Kivy’s view. Against this view, it might be argued that deaf individuals can experience internal speech hallucinations (for review, see Atkinson 2006). Perhaps more critically, it has been shown that the voice-selective regions of the brain tend to respond to visual sign language in congenitally deaf individuals, suggesting that these regions “may not be exclusively dedicated to processing speech sounds, but may be specialized for processing more abstract properties essential to language that can engage multiple modalities” (Petitto et al., 2000 p. 13,691; see also Lomber et al., 2010). Similarly, “auditory” cortices in deaf individuals are activated by rhythm information in the visual modality, suggesting a task-based rather than modality-specific organization of these areas (Bola et al., 2017; Heimler & Amedi, 2020). Future research with sensory-deprived subjects could profoundly reshape our understanding of inner speech processes in reading across different sensory experiences.

Finally, further experimental research could explore different kinds of verbal imagery processes involved in silent reading. For instance, it is plausible that reading

<sup>8</sup> “*Lolita*, light of my life, fire of my loins. My sin, my soul. Lo-lee-ta: the tip of the tongue taking a trip of three steps down the palate to tap, at three, on the teeth. Lo. Lee. Ta”.

<sup>9</sup> “We were in the prep-room when the Head came in, followed by a new boy in mufti and a beadle carrying a big desk. The sleepers aroused themselves, and we all stood up, putting on a startled look, as if we had been buried in our work”.

a literary text or a score also activates a speech-production component, which has mainly a kinesthetics nature (see Kuzmičová, 2014).<sup>10</sup> This component, which is not investigated by Kivy at the phenomenological level, is related to the fact that, as observed by Ribeiro (2015), the silent reading of literary texts involves two conceptually distinct activities, declaiming and listening.<sup>11</sup> Consistent with this intuition, neuroimaging studies have shown that speech imagery activates also regions outside the auditory cortex, such as the inferior frontal gyrus, associated with speech production (e.g., Kleber et al., 2007; Tian et al., 2016; Lu et al., 2021), suggesting a complex interaction of auditory and motor processes. Exploring these dynamics at the experimental level could shed new light into the neural similarities between silently reading literary texts and musical scores.

To conclude, the exploration of Kivy's analogy between reading texts and musical scores, although it touches only a circumscribed aspect of the complex experience of reading, opens several interesting lines of experimental and theoretical research. Overall, these lines promise to improve our understanding of silent reading both in its phenomenological and cognitive aspects, connecting more tightly the fields of literature and music.

**Funding** Open access funding provided by Università degli Studi di Torino within the CRUI-CARE Agreement.

**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

<sup>10</sup> This is the aspect of verbal imagery that Kuzmičová calls “rehearsal-imagery” (2014, pp. 285-86). According to Kuzmičová, “the voice imaged in rehearsal-imagery belongs to the reader, putting her in the position of a vicarious speaker. The distinctive corporeal feature of rehearsal-imagery vis-à-vis speech-imagery is that it is consciously felt to deploy the reader's vocal cords and the muscles in her mouth and throat. It is literally inner in that it originates in the reader's articulatory apparatus. Thus rehearsal-imagery is not only auditory; it is also, and necessarily, kinesthetic. There need not be much of a voice for one to experience rehearsal-imagery as long as one feels the vibrations” (2014, pp. 285-286).

<sup>11</sup> In “The Spoken and the Written. An Ontology of Poems” (2015), Ribeiro observes that, in silent reading, “[s]ometimes we perform two activities, declaiming and listening. In those cases, we are both performer and listener, ‘instantiator’ and ‘experiencer’. It is important to see that these are conceptually distinct activities, even in silent reading. Just as we would not want to say that the person in the audience at the concert hall is the one instantiating the symphony (the orchestra is), so we would not want to say that the individual as *listener* is the one instantiating the poem. It is that individual as *declaimer* (whether out loud or in silent reading) who is. It is as a rule unavoidable that a declaimer or a musician experience the work they are performing, but they could in principle instantiate the work even being deaf, in which case they would be unable concurrently to experience it” (Ribeiro 2015, p. 144).

## References

- Abramson, M., & Goldinger, S. D. (1997). What the reader's eye tells the mind's ear: Silent reading activates inner speech. *Perception & Psychophysics*, *59*(7), 1059–1068.
- Alderson-Day, B., Moffatt, J., Bernini, M., Mitrenga, K., Yao, B., & Fernyhough, C. (2020). Processing Speech and Thoughts during Silent Reading: Direct Reference Effects for Speech by Fictional Characters in Voice-Selective Auditory Cortex and a Theory-of-Mind Network. *Journal of Cognitive Neuroscience*, *32*(9), 1637–1653.
- Allothman, A. A. (2021). Language and Literacy of Deaf Children. *Psychology and Educational Journal*, *58*, 799–819.
- Ashby, J., & Clifton, C., Jr. (2005). The prosodic property of lexical stress affects eye movements during silent reading. *Cognition*, *96*(3), B89–B100.
- Atkinson, J. R. (2006). The perceptual characteristics of voice-hallucinations in deaf people: Insights into the nature of subvocal thought and sensory feedback loops. *Schizophrenia Bulletin*, *32*(4), 701–708.
- Ayotte, J., Peretz, I., Rousseau, I., Bard, C., & Bojanowski, M. (2000). Patterns of music agnosia associated with middle cerebral artery infarcts. *Brain*, *123*, 1926–1938.
- Barsalou, L. W. (1999). Perceptual symbol systems. *The Behavioral and Brain Sciences*, *22*(4), 577–609. discussion 610–660.
- Barsalou, L. W. (2009). Simulation, situated conceptualization, and prediction. *Philosophical Transactions of the Royal Society b: Biological Sciences*, *364*(1521), 1281–1289.
- Besson, M., & Schön, D. (2001). Comparison between Language and Music. *Annals of the New York Academy of Sciences*, *930*, 232–258.
- Bola, L., Zimmermann, M., Mostowski, P., Jednoróg, K., Marchewka, A., Rutkowski, P., & Szwed, M. (2017). Task-specific reorganization of the auditory cortex in deaf humans. *Proceedings of the National Academy of Sciences of the United States of America*, *114*(4), E600–E609.
- Boons, T., De Raeye, L., Langereis, M., Peeraer, L., Wouters, J., & van Wieringen, A. (2013). Expressive vocabulary, morphology, syntax and narrative skills in profoundly deaf children after early cochlear implantation. *Research in Developmental Disabilities*, *34*(6), 2008–2022.
- Cappelletti, M., Waley-Cohen, H., Butterworth, B., & Kopelman, M. (2000). A selective loss of the ability to read and to write music. *Neurocase*, *6*, 321–332.
- Cara, M. A., & Gómez Vera, G. (2016). Silent reading of music and texts; eye movements and integrative reading mechanisms. *Journal of Eye Movement Research*, *9*(7):2, 1–17. <https://doi.org/10.16910/jemr.9.7.2>
- Crosson, J., & Geers, A. (2001). Analysis of narrative ability in children with cochlear implants. *Ear and Hearing*, *22*(5), 381–394.
- Di Bona, E. (2019). Music is not even language-like: Analyzing Kivy's view on music and language. *Argumenta*, *8*(2), 71–84.
- Drożdżowicz, A. (2023). Do we hear meanings? – between perception and cognition. *Inquiry*, *66*(2), 196–228.
- Feagin, S. L. (2008). Critical Study: Reading and Performing. *The British Journal of Aesthetics*, *48*(1), 89–97.
- Fish, S. (1980). *Is there a text in this class?* Harvard University Press.
- Frost, R. (1998). Toward a strong phonological theory of visual word recognition: True issues and false trails. *Psychological Bulletin*, *123*(1), 71–99.
- Gerrig, R. J. (1993). *Experiencing Narrative Worlds: On the Psychological Activities of Reading*. Yale University Press. <https://doi.org/10.12987/9780300159240>
- Gómez-Merino, N., Fajardo, I., Ferrer, A., & Joseph, H. (2022). Eye Movements of Deaf Students in Expository Versus Narrative Texts. *American Annals of the Deaf*, *167*(3), 313–333.
- Goodman, N. (1968). *Languages of art: An approach to a theory of symbols*. Hackett Publishing Company.
- Halpern, A. R., & Bower, G. H. (1979). *Interference tasks and music reading*. Unpublished manuscript, Stanford University.
- Hash, P. M. (2003). Teaching instrumental music to deaf and hard of hearing students. *Research & Issues in Music Education*, *1*(5), 1–8.
- Hébert, S., & Cuddy, L. L. (2006). Music-reading deficiencies and the brain. *Advances in Cognitive Psychology*, *2*(2–3), 199–206.

- Heimler, B., & Amedi, A. (2020). Are critical periods reversible in the adult brain? Insights on cortical specializations based on sensory deprivation studies. *Neuroscience & Biobehavioral Reviews*, *116*, 494–507.
- Hobbes, T. (2010/1651). *Leviathan*. In A. B. Martinich & B. Battiste (Eds.). Peterborough, CA: Broadview Press.
- Iser, W. (1989). *Prospecting: From Reader Response to Literary Anthropology*. John Hopkins University Press.
- Iser, W. (1978). *The Act of Reading*. John Hopkins University Press. <https://doi.org/10.56021/9780801821011>
- Kim, D. (2021). The Impact of Graphic Novels on Reluctant Deaf Readers (Thesis). *ProQuest Dissertations Publishing*, 2021, 28418888.
- Kivy, P. (2006). *The Performance of Reading: An Essay in the Philosophy of Literature*. Blackwell.
- Kivy, P. (2007). Music, Language, and Cognition. In P. Kivy (Ed.), *Music, Language and Cognition: And Other Essays in the Aesthetics of Music* (pp. 214–232). Oxford University Press.
- Kivy, P. (2010). The experience of reading. In G. L. Hagberg & W. Jost (Eds.), *A Companion to Philosophy of Literature* (pp. 106–119). Blackwell Publishing.
- Kleber, B., Birbaumer, N., Veit, R., Trevorrow, T., & Lotze, M. (2007). Overt and imagined singing of an Italian aria. *NeuroImage*, *36*(3), 889–900.
- Kuzmičová, A. (2013). The Words and Worlds of Literary Narrative: The Trade-off between Verbal Presence and Direct Presence in the Activity of Reading. In L. Bernaerts, D. De Geest, L. Herman, & B. Vervaeck (Eds.), *Stories and Minds: Cognitive Approaches to Literary Narrative* (pp. 107–128). University of Nebraska Press.
- Kuzmičová, A. (2014). Literary Narrative and Mental Imagery: A View from Embodied Cognition. *Style*, *48*(3), 275–293.
- Lévy-Strauss, C. (1955). The Structural Study of Myth. *The Journal of American Folklore*, *68*(270), 428–444.
- Lomber, S. G., Meredith, M. A., & Kral, A. (2010). Cross-modal plasticity in specific auditory cortices underlies visual compensations in the deaf. *Nature Neuroscience*, *13*(11), 1421–1427.
- Lu, L., Sheng, J., Liu, Z., & Gao, J. H. (2021). Neural representations of imagined speech revealed by frequency-tagged magnetoencephalography responses. *NeuroImage*, *229*, 117724.
- Luckner, J. L., & Handley, C. M. (2008). A summary of the reading comprehension research undertaken with students who are deaf or hard of hearing. *American Annals of the Deaf*, *153*(1), 6–36.
- Luckner, J. L., & Urbach, J. (2012). Reading Fluency and Students Who Are Deaf or Hard of Hearing: Synthesis of the Research. *Communication Disorders Quarterly*, *33*(4), 230–241.
- Lukatela, G., Eaton, T., Sabadini, L., & Turvey, M. T. (2004). Vowel duration affects visual word identification: Evidence that the mediating phonology is phonetically informed. *Journal of Experimental Psychology: Human Perception and Performance*, *30*(1), 151–162.
- Mak, M., Faber, M., & Willems, R. M. (2023). Different kinds of simulation during literary reading: Insights from a combined fMRI and eye-tracking study. *Cortex; a Journal Devoted to the Study of the Nervous System and Behavior*, *162*, 115–135.
- Marion, G., Di Liberto, G. M., & Shamma, S. A. (2021). The Music of Silence: Part I: Responses to Musical Imagery Encode Melodic Expectations and Acoustics. *The Journal of Neuroscience*, *41*(35), 7435–7448.
- McDermott, H. J. (2004). Music perception with cochlear implants: A review. *Trends in Amplification*, *8*(2), 49–82.
- Mendez, M. (2001). Generalized auditory agnosia with spared music recognition in a lefthander: Analysis of a case with a right temporal stroke. *Cortex*, *37*, 139–150.
- Mongelli, V., Dehaene, S., Vinckier, F., Peretz, I., Bartolomeo, P., & Cohen, L. (2017). Music and words in the visual cortex: The impact of musical expertise. *Cortex; a Journal Devoted to the Study of the Nervous System and Behavior*, *86*, 260–274.
- Morais, I. (2010). Book review: Peter Kivy, the performance of reading: An essay in the philosophy of literature. *Disputatio*, *3*(28), 329–334.
- Morais, I. (2017). "Peter Kivy, The Performance of Reading: An Essay in the Philosophy of Literature." *Forma de Vida*. <https://doi.org/10.51427/ptl.fdv.2017.0054>
- Nanay, B. (2023). *Mental Imagery: Philosophy, Psychology*. Oxford University Press.
- Nikolopoulos, T. P., Lloyd, H., Starczewski, H., & Gallaway, C. (2003). Using SNAP Dragons to monitor narrative abilities in young deaf children following cochlear implantation. *International Journal of Pediatric Otorhinolaryngology*, *67*(5), 535–541.



- Paivio, A. (1986). *Mental representations: A dual coding approach*. Oxford: Oxford University Press.
- Petitto, L. A., Zatorre, R. J., Gauna, K., Nikelski, E. J., Dostie, D., & Evans, A. C. (2000). Speech-like cerebral activity in profoundly deaf people processing signed languages: Implications for the neural basis of human language. *Proceedings of the National Academy of Sciences of the United States of America*, *97*(25), 13961–13966.
- Petkov, C. I., & Belin, P. (2013). Silent reading: Does the brain “hear” both speech and voices? *Current Biology: CB*, *23*(4), R155–R156.
- PrevotEAU, C., Chen, S. Y., & Lalwani, A. K. (2018). Music enjoyment with cochlear implantation. *Auris, Nasus, Larynx*, *45*(5), 895–902.
- Prieto, E. (2002). *Listening In: Music, Mind, and the Modernist Narrative*. University of Nebraska Press.
- Ribeiro, A. C. (2009). Review: Peter Kivy: The Performance of Reading: An Essay in the Philosophy of Literature. *Mind*, *118*(469), 186–191.
- Ribeiro, C. S. (2015). The Spoken and the Written. An Ontology of Poems. In J. Gibson (Ed.), *The Philosophy of Poetry* (pp. 127–148). Oxford University Press.
- Roux, F. E., Lubrano, V., Lotterie, J. A., Giussani, C., Pierroux, C., & Démonet, J. F. (2007). When “abegg” is read and (“A, B, E, G, G”) is not: A cortical stimulation study of musical score reading. *Journal of Neurosurgery*, *106*(6), 1017–1027.
- Sadoski, M., & Paivio, A. (2001). *Imagery and text: A dual coding theory of reading and writing*. Lawrence Erlbaum Associates Publishers.
- Sadoski, M., & Paivio, A. (2013). *Imagery and Text: A Dual Coding Theory of Reading and Writing*. Routledge.
- Schön, D., Anton, J. L., Roth, M., & Besson, M. (2002). An fMRI study of music sight-reading. *NeuroReport*, *13*, 2285–2289.
- Shusterman, R. (1978). The Anomalous Nature of Literature. *British Journal of Aesthetics*, *18*. <https://doi.org/10.1093/bjaesthetics/18.4.317>
- Sloboda, J. A. (1977). Phrase units as determinants of visual processing in music reading. *British Journal of Psychology*, *68*, 117–124.
- Sloboda, J. A. (1974). The eye-hand span: An approach to the study of sight reading. *Psychology of Music*, *19*.
- Sloboda, J. A. (1984). Experimental studies of music reading: A review. *Music Perception*, *2*, 222–236.
- Smetana, L., Odelson, D., Burns, H., & Grisham, D. L. (2009). Using Graphic Novels in the High School Classroom: Engaging Deaf Students With a New Genre. *Journal of Adolescent & Adult Literacy*, *53*(3), 228–240.
- Steinke, W. R., Cuddy, L. L., & Jakobson, L. S. (2001). Dissociations among functional subsystems governing melody recognition after right-hemisphere damage. *Cognitive Neuropsychology*, *18*, 411–437.
- Tian, X., Zarate, J. M., & Poeppel, D. (2016). Mental imagery of speech implicates two mechanisms of perceptual reactivation. *Cortex; a Journal Devoted to the Study of the Nervous System and Behavior*, *77*, 1–12.
- Traxler, M. J., Corina, D. P., Morford, J. P., & Gernsbacher, M. A. (2014). Deaf readers’ response to syntactic complexity: Evidence from self-paced reading. *Memory & Cognition*, *42*, 97–111.
- Trehub, S. E., Vongpaisal, T., & Nakata, T. (2009). Music in the lives of deaf children with cochlear implants. *Annals of the New York Academy of Sciences*, *1169*, 534–542.
- Vilhauer, R. P. (2017). Characteristics of inner reading voices. *Scandinavian Journal of Psychology*, *58*, 269–274.
- Wolf, T. (1976). A cognitive model of musical sight reading. *Journal of Psycholinguistic Research*, *5*, 143–171.
- Wollheim, R. (1968). *Art and its objects*. Harper & Row.
- Yao, B. (2021). Mental Simulations of Phonological Representations Are Causally Linked to Silent Reading of Direct Versus Indirect Speech. *Journal of Cognition*, *4*(1), 6.
- Yao, B., & Scheepers, C. (2011). Contextual modulation of reading rate for direct versus indirect speech quotations. *Cognition*, *121*(3), 447–453.
- Yao, B., & Scheepers, C. (2015). *Inner voice experiences during processing of direct and indirect speech*. [https://doi.org/10.1007/978-3-319-12961-7\\_15](https://doi.org/10.1007/978-3-319-12961-7_15)
- Yao, B., Belin, P., & Scheepers, C. (2011). Silent reading of direct versus indirect speech activates voice-selective areas in the auditory cortex. *Journal of Cognitive Neuroscience*, *23*(10), 3146–3152.

Yao, B., Taylor, J. R., Banks, B., & Kotz, S. A. (2021). Reading direct speech quotes increases theta phase-locking: Evidence for cortical tracking of inner speech? *NeuroImage*, 239, 118313.

**Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.