

Stefan Blohm, Maria Kraxenberger, Christine A. Knoop,
Mathias Scharinger

Sound Shape and Sound Effects of Literary Texts

Abstract: The sound of language comprises all articulatory, acoustic, and perceptual aspects of speech, including the phonological and phonetic recoding of orthographic symbols. The sound of casual speech is widely considered a mere vehicle of meaning; in literary genres, however, such as proverbs, poetry, or even the novel, the sound shape of language serves an aesthetic function and constitutes an integral component of the literary work of art, resulting in a pronounced “palpability” of form (Jakobson, 1960). This chapter selectively reviews the growing body of empirical research that is concerned with sound-related aspects of literary texts; particular attention is paid to prevalent concepts, theories, and methods, concluding with suggestions and recommendations for future investigation.

Introduction

A large body of research on literature, especially in literary studies, focuses nearly exclusively on questions of explicit and implicit meaning. However, this focus disregards the sound shape of literary texts, which, while serving as a vehicle of meaning, often involves a distinct aesthetic function. The sound shape may highlight the text’s form and create a variety of effects, ranging from mere euphony to the enhancing or impeding of cognitive processes. In some genres, like poetry, such sound effects occur across the board, whereas they tend to be more subtle and less regular in prose genres. This chapter provides some background information about linguistic sound structure and phonological systems before we selectively review empirical investigations into sound-related aspects of literary texts, identify dominant approaches and underlying assumptions, and describe key theoretical positions, pivotal methods, and major findings. It concludes with suggestions for future research into the sound shape and sound effects of literary texts, providing concrete recommendations for research design and identifying evidential gaps.

The Sound of Language

The “sound of language” refers – in its narrowest sense – to acoustic signals shaped to encode information according to linguistic semiotic systems shared by sender and receiver. Linguistic theory distinguishes the actual acoustic speech signal from the sound conceptions stored in the minds of speakers, hearers, and readers. This crucial conceptual distinction gives rise to two sub-disciplines concerned with speech sounds: *Phonetics* investigates the production, structure, transmission, and perception of speech signals; *phonology* studies the systems of abstract knowledge acquired by competent speakers and hearers, including inventories of basic categories and combinatorial rules and principles (Hayes, 2009, p. 19). During perception, continuous acoustic information is mapped onto discrete phonological categories, so that many stylistic effects of sound structures in literature depend on the phonological system of the language at hand. However, since sound sequences carry lexical meaning, stylistic effects of sound structures may also be indirect and mediated by the activation of lexical-conceptual knowledge, both denotational and connotational, in the mind of the reader/listener.

Sound Structure and Phonological Systems

A simple way to think of the structure of speech is as a series of distinct sounds. Although this conception is clearly inaccurate from the viewpoint of phonetics, it is a fair approximation of what we do when we speak or comprehend, which is to assemble complex sound structures from smaller, more basic, and abstract building blocks, or to segment continuous speech streams into such discrete smaller units and components (see Figure 1).

Phonemes are the elementary speech sounds that any given language uses to distinguish lexical meaning, as in *hit* – *pit*. All languages have both vowels and consonants, but the inventories of these phoneme classes may differ dramatically across languages, ranging from 6 to 122 consonants and from 2 to 14 basic vowels (Maddieson, 2013a, 2013c). Although usually considered basic units of speech, phonemes can be described as bundles of *distinctive features* that reflect articulatory gestures and states, and that are more or less transparently related to modulations of the speech signal. Shared distinctive features define classes of speech sounds; and phonological rules apply to such classes, or feature bundles, rather than to distinct phonemes. The descriptive level of distinctive features and their phonetic correlates is relevant and necessary for

the adequate formal description of literary phenomena such as assonance and imperfect rhyme. Some radically economic theories of spoken word recognition (e. g., Lahiri & Reetz, 2002) maintain that distinctive features are sufficient to distinguish and activate lexical meaning and that the phoneme level can be dispensed with altogether.

Syllables are complex structures that arise from the combination of phonemes; they are consciously perceived as speech units and as building blocks of words. The internal structure of syllables is usually seen as tripartite, consisting of an optional consonantal *onset*, a usually vocalic *nucleus*, and (optional) postvocalic consonants in the *coda* constituent; the complex formed by the nucleus and coda is referred to as *rime* or *rhyme*. Although languages differ in terms of the phoneme combinations they permit in these syllabic constituents (Maddieson, 2013b), many cross-linguistic (and language-specific) generalizations about the internal structure of syllables can be captured by the so-called *sonority principle* – the tendency to combine speech sounds in such a way that, roughly, overall acoustic energy and perceptual prominence is greatest at the syllable nucleus (its most sonorous constituent) and decreases towards the edges. On this view, a syllable corresponds to a stretch of a speech stream spanning from one local sonority minimum to the next. In many languages, syllable sequences exhibit prominence contrasts, e. g., levels of syllable stress as in English, or tone as in Mandarin Chinese. Syllable prominence and phoneme inventories constitute the basis of tradition-specific sound patterning rules.

Syllables constitute the elementary units of *prosody*, which refers to structural phenomena above and beyond single speech sounds or phonemes, such as stress, accent, and rhythm, tone and intonation, etc. (see Cutler et al., 1997; Nespor & Vogel, 1986). The sound structure above the syllable is captured in the *prosodic hierarchy* (Nespor & Vogel, 1986), wherein larger units dominate and constrain smaller units. Not all languages exhibit all levels of prosodic structure, i. e., descriptively, some prosodic systems are flatter than others. Syllables constitute the lowest level of this hierarchy, receiving syllable weight (either weak=w or strong=s). Depending on the language, strong syllables have, e. g., either a long vowel nucleus or a coda and can bear stress. The patterning of strong and weak syllables defines a *prosodic foot*, with possible binary sequences of strong-weak (trochee) and weak-strong (iamb). Feet are dominated by the level of *prosodic words*, whose edges may be marked by boundary elements such as lengthening or pauses; prosodic words provide an important interface between prosody and syntax. Individual prosodic words are subsumed in *phonological phrases*, which also align with syntactic structure. Phonological phrases then define *intonational phrases*, themselves dominated by the utter-

ance level. Figure 1 depicts the phonological hierarchy from the subsegmental feature level up to the utterance level.

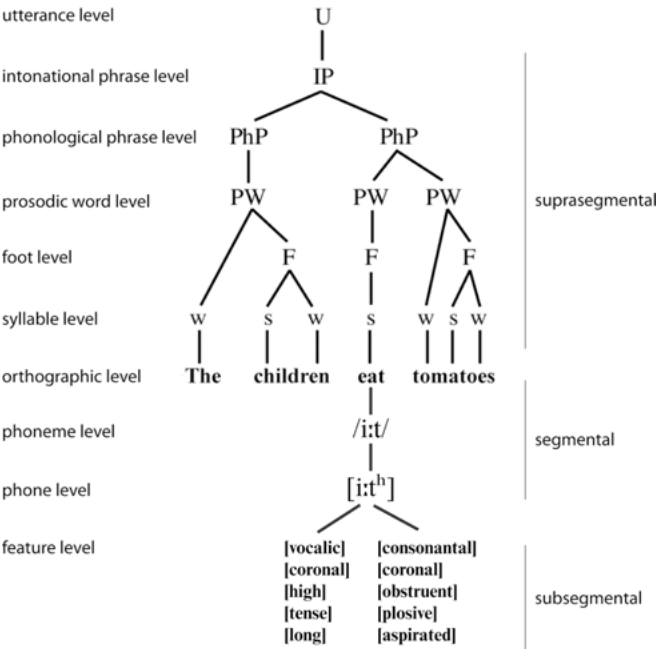


Figure 1: The hierarchical sound structure of language, combining the prosodic hierarchy (Nespor & Vogel, 1986) with assumptions from feature theory (Lahiri & Reetz, 2002).

The Sound of Written Text

Writing systems differ with respect to the “unit of linguistic structure that is represented most directly” (Comrie, 2013), i. e., the phoneme, the syllable, or the word – and thus in the amount of sound-related information they convey. Any information that is not conveyed explicitly has to be supplied by the reader, which creates instructional gaps and interpretive freedom. For instance, written Chinese texts convey pitch information via lexical units, whereas the pitch contour of English texts is underspecified and supplied by the reader.

Orthographic symbols (written words) are tightly connected with their corresponding phonological representations in the minds of experienced readers, regardless of whether or how well the writing system represents sound. Thus,

reading a word automatically activates its abstract sound representation, which is referred to as *phonological recoding* and comprises both phonemic and syllabic representations (see, e. g., Braun et al., 2009; Conrad et al., 2009; McCutchen & Perfetti, 1982; Ziegler et al., 2000). But what readers experience as an “inner voice” (Huey, 1908) during silent reading is the *phonetic recoding* of written text. This sub-vocalization comprises not only individual speech sounds and syllables but, like overt articulation, also sentence intonation, phrasing, stress, and rhythm, partly guided by punctuation (Chafe, 1988; Steinhauer & Friederici, 2001; Stolterfoht et al., 2007). In line with Fodor’s (1998, 2002) *implicit prosody hypothesis*, these rhythmic and melodic aspects may affect how readers perceive a text, co-determining gaze positions (Ashby & Clifton, 2005), as well as phonological and syntactic parsing (Bader, 1998; Kadota, 1987; Kentner & Vasishth, 2016; see the reviews in Breen, 2014, 2015). Crucially, the stream of inner speech is the actual realization of the sound shape of the written literary work of art; and it is usually these realizations that we study when we investigate sound effects of written literary texts.

Although phonological and prosodic recoding generally occur during silent reading, they may be of particular importance when reading poetry (cf. Schrott & Jacobs, 2011), as this literary genre usually displays the greatest degree of phonological constraint. Highlighting the importance of genre-specific processing strategies, De Beaugrande (1978, p. 24) argued that poetry readers “have as part of the text-type frame the instructions to attend to sound recurrences in the assumption that these are not random.” Hence, poetry-appropriate reading strategies may lead to an increase in the quantity and quality of phonological and prosodic recoding during poetry reading (cf. Kraxenberger, 2017). However, rather than being the result of genre-appropriate a priori adjustments, as argued by De Beaugrande, both increased attention to sound and deeper phonological recoding can be triggered by the sound recurrences themselves or by other sound features of poetry (Blohm et al., 2017; Blohm et al., in press). In their recent eye tracking study of Shakespeare’s sonnets, for instance, Xue et al. (2019) have argued that more intensive phonological recoding during poetry reading may be related to sonority, which not only plays a role in silent reading (Maionchi-Pino et al., 2008; Berent, 2013) but also influences the subjective beauty of words (Jacobs, 2017).

Sound Shape and Sound Effects of Literary Texts

Research into the sound shape and sound effects of literary texts falls into two major categories. First, there is continuing interest in how sound conveys non-denotational meaning, particularly in the phenomenon of sound iconicity. The second major research tradition examines non-semantic aspects of literary texts and literary comprehension – and focuses on how systematic and sporadic recurrences of sounds and sound sequences affect the sound shape and the reception of literary texts.

Sound Iconicity

The term *sound iconicity* – also referred to as *phonological iconicity*, *sound symbolism* (Hinton et al., 1994), *phonetic symbolism* (Sapir, 1929), or *hypo-iconicity* (Nöth, 2000) – is commonly understood as “an inmost, natural similarity association between sound and meaning” (Jakobson & Waugh, 1979/2002, p. 182). Whether there is indeed an inherent, natural association between a concept and its name, or whether this relation is purely conventional, is an ancient question (see Plato’s *Cratylus*). De Saussure is arguably the most prominent proponent of the *conventionalist* position; his claim (1916/1983) that the relation between the signifier and the signified is arbitrary has – not least because of the striking variation in the phonological and lexical systems of the world’s languages – been the dominant view in the language sciences in the past century.

With respect to an inherent semantics of isolated speech sounds, strict conventionalism further maintains that “the sounds of a language are intrinsically meaningless” and that “their only purpose is to form the building blocks of which words are made” (Hayes, 2009, p. 19). The *naturalist* position has been convincingly argued by Sapir (1929), who – acknowledging a high degree of apparent arbitrariness in sound-meaning mappings – demonstrated that some associations of concepts and sounds are more natural than others and that people intuitively agree on such associations. More recently, non-arbitrary sound-meaning relations have received increased attention in the language sciences (e.g., Aryani et al., 2019; Dingemanse et al., 2015; Perniss et al., 2010; Schmidtke et al., 2014), and the emerging picture is that the development of lexical and grammatical forms is driven by several competing motivations, particularly practicality and ease of articulation, but also iconicity and intuitive comprehensibility.

Sound-iconic relations usually link portions of conceptual space to phonetic and phonological features, as well as to the phonemes and phoneme classes they define. For instance, studies in several languages have examined the potential link between vowel quality and the perception of brightness/darkness. Fechner's (1876, p. 318, our translation) claim that "*a, e, i* appear as brighter and *o, u* as darker" has received cross-linguistic support (Moos et al., 2014; Tsur, 1992, 1997; Wrembel, 2009; but see Kraxenberger & Menninghaus, 2016b). Furthermore, there is evidence indicating that some acoustic and phonological features iconically express the size of denoted objects (Huang et al., 1969; Sapir, 1929; Thompson & Estes, 2011) and their (e. g., angular vs. round) shape (Köhler, 1929; Ramachandran & Hubbard, 2001; Westbury, 2005). Of course, these sound-symbolic biases of language may be more fully developed and more extensively exploited in carefully constructed literary texts.

Poetry, being firmly rooted in oral traditions, is the literary genre in which such sound-iconic relations have been studied most extensively (Fónagy, 1961; Jakobson & Waugh, 1979/2002; Jespersen, 1933; Tsur, 1992), following the notion that its sound "must seem an echo to the sense" (Pope, 1711/2010). In modern literary theory, the school of (Russian) structuralism was the first to highlight the phonological and phonetic aspects of literary texts, especially of poetry, which was claimed to be "a province where the internal nexus between sound and meaning changes from latent into patent and manifests itself most palpably and intensely" (Jakobson, 1960, p. 373). Interestingly, phono-semantic iconicity in verse appears not to be restricted to individual sound-iconic words; rather, the meaning of one word may also be sound-iconically reflected in the phonetic properties of adjacent words in the same line (Auracher et al., 2019).

The relation between the sound of a poem and its emotional meaning has been frequently highlighted in theoretical reflections as well as empirical studies (e. g., Jakobson & Waugh, 1979/2002; Tsur, 1992). Indeed, most empirical studies on (phonological) iconicity in poetry have focused on readers' perceptions and assessments of a poem's key emotional tonality (e. g., Aryani, et al., 2016; Auracher et al., 2010; Kraxenberger & Menninghaus, 2016a, 2017; Kraxenberger, 2017; Whissell, 2002, 2011). Hevner (1937), for instance, examined the connotations of vowels (high front vowels vs. round back vowels) and basic rhythmic patterns in poetry (iambes and anapests), asking listeners to describe carefully constructed nonsense verse by choosing appropriate adjectives from a 64-item list. She observed that "meter is very effective in determining the happiness or sadness of the poetry, while vowel sounds are quite ineffectual in this capacity" (p. 431); iambes were perceived as being solemn, serious, earnest, sad, heavy, dignified, etc., whereas anapests were described as merry, humorous,

playful, joyous, light, gay, etc. (for a further investigation of prosodic aspects of language in relation to emotion perception, see Kraxenberger et al., 2018).

The idea that phonemic contrasts systematically support and signal thematic contrasts (e. g., Tsur, 1997) has been pursued by Miall (2001), who compared phoneme distributions in selected texts, e. g., in passages of Milton's *Paradise Lost* featuring depictions of Hell or Eden. Contrary to the proposed mapping of vowel qualities onto the brightness/darkness dimension, he observed that passages about Hell contained considerably more front vowels than passages about Eden, which contained more medial back vowels than the Hell passages (Eden=dark, Hell=bright). Phonemic contrast has, further, been shown to effectively signal narrative shifts in a short story by Katherine Mansfield (Miall & Kuiken, 2002).

The hypothesis that nasal sounds (e. g., /m/, /n/) are naturally associated with negative affective states, whereas plosive sounds (e. g., /p/, /d/) are naturally associated with positive ones, constitutes an example of a manner-of-articulation-emotion association. This link seems to coincide with the descriptive observation that there are strikingly persistent associations between negation and nasal sounds in both Indo-European (/n/) and Sino-Tibetan (/m/) languages. It remains unclear, however, whether these are indeed natural, “pre-linguistic” sound-meaning associations, or rather acquired, language- and culture-specific links between certain sound properties and negative affective states. In any case, this association has been examined in more detail in literary texts: Plosives are more frequent in Old Egyptian hymns than in Old Egyptian lamentations and in Goethe's hymns than in his ballads; by contrast, nasals are more frequent in the ancient lamentations and in the Goethean ballads (Albers, 2008). Corroborating evidence stems from a study reporting that German, Russian, and Ukrainian (all Indo-European), as well as Chinese (Sino-Tibetan), native speakers perceive poems with a relatively high frequency of nasals as sad and poems with a high frequency of plosives as happy/joyful (Auracher et al., 2010). However, a more recent study could not replicate this finding for German poems (Kraxenberger & Menninghaus, 2016a). The authors suggest that the assumed poetry-specific sound-emotion nexus – if it exists – may in fact be driven by the acoustic features of emotional prosody during poetry reading rather than by the frequency of specific phoneme classes, based on the observation that emotional prosody is an important factor in both the recitation and the perception of German poems by both native speakers and non-native listeners who have no access to the semantic content of the poems. Semantic and thematic text aspects of poetry – if available to recipients – are nevertheless far better predictors of perceived emotions than emotional prosody (Kraxenberger et al., 2018).

A number of studies have investigated sound-iconic relations in prose texts (for an overview of diverse aspects of iconicity in spoken language, see Hinton et al., 1994). For example, Perlman et al. (2015) used a story reading task to investigate iconic prosody in terms of articulation rate and pitch. Results show different temporal patterns of iconic prosody between concrete and abstract dimensions of speed (e. g., *a sluggish walk*, *a fast track to success*) and size (e. g., *a small grasshopper*, *a really big deal*) as described in short stories. These findings corroborate previous findings showing that speakers spontaneously iconically modulate prosody to fit the meaning they are expressing, for instance when describing short video clips showing fast or slow-paced events (Perlman & Benitez, 2010; for evidence on markers of emotional prosody when reading sad or joyful poetry, see Kraxenberger et al., 2018).

Despite the existing research, a comprehensive picture of sound iconicity remains elusive, and results on the topic often contradict one another. A possible reason for this, apart from the broadness of this issue and the methodological diversity, might lie in the often insufficient scope of the underlying operational definitions of the sound component (cf. Aryani et al., 2016; Kraxenberger, 2017).

Sound Recurrences and Phonological Parallelism

We distinguish the more general term *sound recurrence*, which denotes all kinds of phonemic and suprasegmental sound repetitions, from *phonological parallelism*, which occurs if identical phonological units (e. g., phonemes) appear in identical positions of higher order units of phonological structure (e. g., a syllable) or poetic structure (e. g., a line of verse). Higher order units usually contain contrasting elements in their other positions (Fabb, 1997; Frog & Tarkka, 2017; Jakobson, 1960, 1966). For instance, alliteration refers to the (re-)occurrence of identical consonants in onset positions of neighboring words or syllables; end rhyme refers to the (re-)occurrence of identical phoneme sequences in final metrical positions of (half-)lines of verse or other types of regulated speech. By contrast, the isolated (re-)occurrence of consonantal sounds in onset and coda positions of neighboring syllables counts as non-systematic sound recurrence. We further distinguish four broad aspects of literary comprehension that are affected by sound recurrences and phonological parallelism: the sound shape of the literary work of art, text processing, text evaluation, and text memory.

The *sound shape* of a literary text corresponds to the phonetic and phonological properties of an actual realization of the text's surface form, whether acoustically and explicitly during performance or oral reading, or internally during silent reading. Sound shapes of written literary texts are often empirically investigated by analyzing the acoustic properties of spoken renditions of literature, including the study of (a) material recorded independently of the study, such as recorded performances and audiobooks, and (b) material created and recorded expressly for the purpose of a study, often in different variations or including comparisons between lay and professional speakers. One approach frequently applied to written texts involves calculating text statistics by (a) counting certain phonemic and prosodic phenomena in a text based on the phonological information provided by a lexicon, grapheme-to-phoneme conversion tools, or other means of approximating actual speech and (b) calculating relations (ratios, intervals) between frequently occurring phenomena in order to derive patterns. While most such work has been done computationally in recent years, earlier research employed manual annotations, and some scholars of Classical languages and literatures continue to do so today.

Text processing comprises all cognitive mechanisms involved in converting acoustic or visual input into the linguistic or conceptual representations that enter the consciousness of the recipient. The effects of sound on processing are usually examined by measuring recipients' behavior or physiological responses during reading/listening; these so-called *online methods* – usually adopted from psychology and psycholinguistics – tap into the comprehension process as it happens. In the *self-paced reading/listening* paradigm (Just et al., 1982), for instance, experimenters split texts into smaller regions of interest and record processing times per region while participants navigate through the text at their own pace. Observed processing times are indicative of cognitive effort associated with the respective text section. *Eye tracking* refers to the monitoring of recipients' gaze position and/or pupil dilation while they process a visual or auditory stimulus. Allowing for more natural reading and providing a better signal-to-noise ratio than self-paced reading, eye tracking not only yields total processing times but also allows for the computation of further dependent variables indicative of cognitive processing, e. g., the duration of the first fixation on a particular word, or the length, frequency, start, and landing sites of regressive eye movements (see Rayner, 1998). *Electroencephalography* (EEG) refers to the noninvasive recording of stimulus-related electrical brain activity via electrodes placed on the surface of the scalp; this technique has a high temporal resolution (~1 msec), which allows one to study the time-course of cognitive processes. EEG yields multidimensional (temporal, spatial, voltage, oscillatory activity) data that can be interpreted qualitatively, i. e., certain com-

ponents in the EEG signal can – on the basis of previous research – be associated with specific cognitive processes, such as the phonological recoding of orthographic symbols. Differences in such signal components are interpreted as evidence for the presence and/or depth of the respective cognitive processes (Luck, 2014).

Evaluation taps into recipients' experience and comprises their conscious judgment about (aspects of) the text; note that recipients are usually aware of the judgment itself but often unconscious of the processes and factors that bring it about. The effects of sound on recipients' text evaluation are usually assessed by means of self-report and intuitive judgments. These techniques differ primarily in terms of (a) the conceptual dimensions they aim to assess, (b) the kinds of stimuli they use, and (c) the format of recipients' response. For instance, the *semantic differential* technique (Osgood & Snider, 1969) taps into the connotative meaning of objects, words, concepts, and – in the empirical study of literature – text passages or entire texts. This technique requires recipients to rate literary stimuli on a number of bipolar conceptual dimensions (e. g., *dark-bright*, *simple-complex*); the overlap and divergence of these ratings reveal the connotative dimension and directionality associated with the contrast of interest, e. g., between groups of texts or readers.

Memory refers to the mental representation that recipients maintain of a text after they have processed it. In text comprehension, surface information is usually rapidly forgotten, whereas semantic and particularly conceptual information is retained over longer periods (Kintsch et al., 1990). The effects of sound on memory are usually examined by testing recipients' text representation after reading/listening. In *recall* tasks, recipients are prompted to reproduce (parts of) the text, whereas *recognition* tasks require participants to identify parts of the text and discriminate them from a set of alternatives; in both tasks, the accuracy of recipients' responses serves as the dependent variable of primary interest.

The combination of (a) multiple phonemic and prosodic parallelisms (meter, (b) rhyme, and (c) several local sound patterns renders lyrical poems more beautiful, moving, and melodious, and may increase – depending on the theme of the poem – either the sadness or joy that is conveyed (Menninghaus et al., 2017). In addition to these aesthetic and stylistic effects, multiple sound recurrences and parallelisms facilitate memorizing and recalling poetry and other genres of phonologically regulated speech (Rubin, 1995). Studying children's memory for poetry, Ballard (1913) observed a phenomenon called *hypermnnesia*, i. e., that (verbatim) recall improves for a few days without re-reading the text, as evidenced by better performance upon repeated testing. However, it appears that hypermnnesia also occurs for the semantic content of narrative prose if the

intervals between tests are short enough (Wheeler & Roediger, 1992). More recently, Tillmann and Dowling (2007) have reported further evidence in support of the idea that the combination of multiple sound recurrences – analogous to harmonic recurrences in musical structures – facilitate the memorization and verbatim recall of poetry. They compared memory for (auditorily presented) poetry and prose in immediate and delayed (<1 minute) recognition and observed that memory for linguistic surface form decreased rapidly for prose but not for poetry. Rubin et al. (1997) collected corpora of children’s counting-out rhymes in Romanian and English, comparing principles of composition and changes in wording that occur during oral transmission. They noted that counting-out rhymes are structurally very similar in both languages, their form being subject to multiple phonological and lexical constraints (alliteration, rhyme, word repetition, etc.). Verbatim recall was very high even for children who performed less well on standardized memory tasks such as list learning. The observed changes in the wording of counting-out rhymes revealed that lexical changes (e.g., rhyme word substitutions) usually preserved the constraints on phonological structure, suggesting that it is the schematic representation of the sound-based poetic structure that improves recall and that constrains variation in reproduction.

Phonemic Parallelism

Phonemic parallelism refers to the systematic patterning of phonemes and syllabic constituents and comprises *alliteration*, *assonance*, *consonance*, and *rhyme*. It is governed by meter and verse constituency in many poetic systems (see Fabb, 1999; Žirmunskij, 1966), but during literary comprehension it is word constituency that matters most, i.e., the fact that a sound repetition involves, for example, the onsets of meaningful units.

The Musical Qualities of Phonemic Recurrence. The poetry-specific formulation of Birkhoff’s (1933) aesthetic measure was an early attempt to quantify the musical qualities of regular and sporadic recurrence at the level of phonemes and syllabic constituents. Here, the aesthetic measure M reflects the ratio of order O and complexity C of an aesthetic stimulus ($M = O/C$). Applied to the musical qualities of verse, the complexity C corresponds to the number of phonemes and word boundaries incapable of liaison, whereas the order O corresponds to the number of “musical vowels” and harmonic recurrences of phonemes and syllabic constituents; recurrences are considered harmonic if they are not too excessive and repetitive. There is some empirical support for the validity of Birkhoff’s proposal and for the predictive value of his own calculations. Davis

(1936) had readers rank lines of verse according to their poetic quality, paying “attention only to the sound and structure of verse”; the obtained ranking showed a moderate correlation with Birkhoff’s values. Davis also observed that other linguistic and non-linguistic measures of complexity – syllable count and line length in cm – work nearly as well as the phoneme count proposed by Birkhoff. Beebe-Center and Pratt (1937) reported three experiments in which participants, instructed “to judge on musical value disregarding meaning,” assessed lines of poetry they read and listened to. In one experiment, participants were presented with constructed lines of meaningless pseudoword verse. The observed preferences for actual verse showed considerable inter-individual variation, and they did not correlate with Birkhoff’s measure – regardless of whether preferences were assessed by means of ratings (7-point scale) or two-alternative forced choices. By contrast, the preference pattern for meaningless pseudoword verse showed considerable inter-individual agreement and strongly correlated with Birkhoff’s musicality measure.

Effects of Phonemic Parallelism

During comprehension, both alliteration and rhyme can prime cohorts of phonological neighbors in the mental lexicon, but it seems that phonological overlap in word onsets – as in alliteration – is a more potent prime than the word-offset overlap of rhyme (Marslen-Wilson & Zwitserlood, 1989). Combined with the descriptive generalization that systematic alliteration is local whereas systematic rhyme is distal (Fabb, 1999), this and related findings suggest that – depending on the experimental task and context – alliteration locally facilitates or disrupts word processing. How exactly this affects the processing of literary texts is unclear at present; similarly, little is known about the aesthetic and stylistic effects of alliteration, at least independent of other types of parallelism. What seems fairly well-established, however, is that alliteration is an effective mnemonic device. For instance, Lea et al. (2008) investigated memory effects of alliteration in a series of experiments in which participants orally and silently read excerpts of free verse (thus avoiding confounds of systematic rhyme and meter) or prose and performed a subsequent sentence recognition task. Alliteration increased recognition accuracy in silent and oral reading (showing that overt articulation was not a prerequisite for its mnemonic effect) and for poetry as well as prose (indicating that the memory effect was not a result of genre-appropriate processing strategies triggered by prior text categorization. Corroborating evidence was reported by Atchley and Hare (2013), who demonstrated that systematic alliteration enhances immediate and delayed memory for regu-

lated verse. Their participants repeatedly read a set of verse lines aloud during a learning phase, and recognition tests were administered immediately afterwards and after a delay of twelve hours, which either contained a sleep period or not. Verbatim memory was less accurate after twelve hours, but the alliterative schema was largely retained, with the sleep group performing better during delayed recognition.

Alliteration constrains the occurrence of consonants in syllable onsets, but other consonantal patterns may be more complex and involve both onset and coda segments as well as other phonological variables. The traditional Welsh poetic form *cynghanedd* features complex consonantal and stress patterns involving several words within a line of verse, e. g., “A **daeth** i **ben** | **deithio** **byd**.” Vaughan-Evans and colleagues (2016) studied the processing and evaluation of this poetic form, combining EEG during sentence reading with intuitive judgments about how “good” the sentences sounded. They observed a distinctive brain response associated with attentional re-orientation in the *cynghanedd*, but not in control conditions that violated one or both of the constraints on stress and consonantal re-occurrence; readers’ conscious judgments, by contrast, did not distinguish between sentences that conformed to the constraints of the form and those that did not.

Rhyme occurs if two words share the same sounds from their last prominent syllable onwards (see, e. g., Fabb, 1997); slight deviations from the strict identity constraint are conventionally licensed in many poetic traditions, and a few studies have examined the perception of such imperfect or slant rhymes (e. g., Knoop et al., 2019; Stausland Johnsen, 2011). The bulk of rhyme-related empirical research so far has examined *end rhyme*, i. e., the systematic use of rhyme to mark the closure of poetic structures such as (half-)lines of verse. This regular recurrence of identical vowel sounds makes rhyme both a rhythmic (Breen, 2018; Menninghaus et al., 2014) and a distinctly melodic element of verse, comparable to the return to the tonic in music (Lanz, 1926; Schramm, 1935a, 1935b). During text processing, systematic end rhyme allows one to predict upcoming input and facilitates word processing if the prediction is fulfilled (Menninghaus et al., 2014; Obermeier et al., 2016). Even during reading, the non-fulfillment of a prediction modulates brain responses as early as 150 ms after the onset of a non-rhyming word. Although the available ERP evidence on rhyme processing in poetry is not consistent, it appears to converge in the modulation of three distinct phases of word processing: (a) an early phase related to phonological processing proper (including phonological recoding during reading); (b) an intermediate phase related to lexical processing and the semantic integration of word meaning; and (c) a later phase related to the (semi-)conscious evaluation and the controlled resolution of processing conflicts that arise from earlier proc-

essing steps (Chinese: Chen et al., 2016; Dutch: Hoorn, 1996; German: Obermeier et al., 2016). Further physiological evidence for rhyme predictability is provided by a study demonstrating that the pupillary responses of listeners to limericks are sensitive to violated rhyme expectations, but not to violations of metrical requirements or of syntactic and semantic constraints (Scheepers et al., 2013). Carminati and colleagues (2006) used self-paced reading to examine how transitions in sub-genre (narrative vs. lyric poetry) and/or changes in rhyme scheme affect poetry reading; whereas readers slowed down after transitions between sub-genres, their reading times were unaffected by rhyme scheme changes. The aesthetic and stylistic effects of rhyme seem to be strongly genre- and context-dependent: Rhymed aphorisms appear more convincing and accurate (McGlone & Tofighbakhsh, 2000); rhymed proverbs more beautiful, succinct, and persuasive (Menninghaus et al., 2015); and rhymed lyrical poems more emotionally intense (Obermeier et al., 2013). Unconventional and imperfect rhymes may have a comical effect and render humorous verse funnier (Menninghaus et al., 2014).

The mnemonic effects of rhyme are well-known. In a series of experiments, Bower and Bolton (1969) used list learning and recall to study why rhymes are easy to memorize. They observed that rhyme words are recalled more accurately than non-rhymes, but that other phonemic parallelisms (*assonance*) have comparable effects if used systematically. Moreover, they found that rhyme – if overgeneralized – may also interfere with task performance. Taken together, their results suggest that the mnemonic effect of end rhyme depends on the restriction of the set of response alternatives, “practically converting recall into a recognition test” (Bower & Bolton, 1969, p. 453). A similar conclusion was reached by Rubin and Wallace (1989), who demonstrated that combined constraints (semantics *and* rhyme) provided far better cues to memory retrieval than would be expected from the combination of their individual cue strengths. Rubin et al. (1993) examined the emerging expertise for English ballads in literary novices who (repeatedly) heard and recalled several ballads over a period of five weeks. Repeated exposure to the same ballad increased verbatim recall, as did repeated exposure to the ballad structure. More importantly, as participants became familiar with the form and the typical ABCB rhyme scheme, they better recalled rhyme words than non-rhymes even after first exposure to a text, suggesting that they developed a schematic representation of the formal (and thematic) poetic structure of ballads and ballad stanzas, on which they relied during listening and recall as well as during an additional ballad composition task. Recall rates were conflated across the two rhyme words of the ABCB ballad stanza, leaving open the question as to whether both rhyme words were recalled more accurately. Blohm and colleagues (2021) combined self-paced reading with a probe recognition task to study how end rhyme in quatrains affects word

processing and memory. They observed that recognition is more accurate for the second but not for the first word of a rhyme pair and that only recognition of the first word suffers interference from rhyming alternatives. In keeping with the idea of multiple constraints that incrementally accrue to restrict the set of possible continuations (Rubin, 1995), this suggests that the memory effect of rhyme is unidirectional.

Prosodic Parallelism and Musical Qualities

Prosodic parallelism refers to the systematic patterning of syllables and higher order prosodic units and comprises both rhythmic and melodic patterns. Rhythmic patterns are usually based on the relative prominence of prosodic units, e. g., duration and quantity or stress and accent. Note that poetic systems frequently neutralize prominence distinctions in the underlying phonological systems, for instance by reducing four-level systems of accent or tone to binary oppositions (Fabb, 1997).

The Rhythm of Prose. In the empirical study of literature, the concept of prose rhythm does not pertain simply to rhythmic structures preferred in particular languages, but rather aims to explore and measure author-, genre-, or time-specific rhythmical properties of literary language that are assumed to have been deliberately realized by an artful selection and combination of lexical elements. Strictly speaking, the vast majority of these rhythmic patterns are sporadic recurrences rather than systematic parallelisms (if one takes the entire text into account), but they may – in certain passages of rhythmic prose – approximate the strict rhythmicity of metered verse.

Depending on the language studied, the analysis of prose rhythm usually focuses on syllable prominence or syllable duration, that is, on patterns of and intervals between stressed and unstressed or long and short syllables, often depending on their position in larger entities. With regard to methodology, this means that most of the studies in question syllabify sentences or phrases to identify their quantitative or accentual patterns and to determine average lengths and length distributions of rhythmical units and intervals between these units. In addition, some studies include larger-scale linguistic phenomena, such as regular or partly regular patterns of word, phrase, sentence, paragraph, or even chapter length.

Empirical studies on prose rhythm usually resort to manual or computational corpus annotations based on the lexicon; studies focusing on or including the description of acoustic properties as realized by readers/speakers remain rare (but see, e. g., Esser, 1988). In terms of knowledge production, most of this

research aims (a) to study the occurrence of rhythmical units in particular text positions, (b) to quantitatively distinguish where verse ends and prose begins, or (c) to identify lengths and patterns of rhythmical units, partly in combination with other textual variables.

Inspired by Cicero's (55BC/2001) programmatic claims that rhythmic units tend to occur in specific text positions, several studies have investigated rhythmical patterns in the so-called *clausulae*, sentence- and passage-final cadences in antique prose. In these studies, rhythmical patterns conforming to known antique metrical grids are the primary field of interest. The underlying method consists in quantifying the occurrence of metrical feet and strings of metrical feet in terminal cadence positions and assigning different levels of significance to the pauses following them (e. g., Aumont, 1996; Bornecque, 1907; de Groot, 1921; Primmer, 1968; Zielinski, 1904, 1914). In particular, Aili (1979) was able to show distinctive differences between authors using this method. While all of these studies share similar aims and methodology, in the context of Classical Studies the differences between them remain subject to much disagreement. In particular, they differ on the principles underlying the assignment of clausula length; primary, secondary, and sentence accents; and caesura prominence, as well as the relative significance of stress and syllable quantity in Latin, about which definitive historical evidence is lacking.

Remarkably, most research on antique prose rhythm has been conducted without the help of computers despite involving, in part, very large corpora (one of the largest being Zielinski's, who analyzed nearly 125,000 clausulae). However, one of the exceptions, a computational study by Knapp (2015) using the software *Numerator*, was able to largely replicate Primmer's (1968) and Aili's (1979) results regarding the occurrence, distribution, and combination of specific metrical feet.

Taking a similar approach, while analyzing a vast corpus by hand and extending the method to a modern language, Saintsbury (1912) attempted to identify metrical feet and their positions in English prose. He explored phenomena as diverse as Old and Middle English literatures, ornate and plain styles, nineteenth-century novelists, and lyrical prose. His calculations revealed a bias for quadrisyllabic feet (paeons, diiamb, ditrochees, etc.) in much English prose, especially in passage-final positions, as well as a trend for rhythmic groups that are marked by slight variations in length due to additional or omitted monosyllables, for polymetrical groups, and for various foot combinations. Overall, however, Saintsbury arrived at the conclusion that literary prose, while exploring rhythmic regularity, always stops "short...of admitting the recurrent combinations proper to metre" (p. 344). In a diachronic computational study comparing prose and verse, Borgeson and colleagues (2018) corroborate this result,

showing that English prose was less metrical in periods when verse was particularly popular and more metrical when verse was less popular. In addition, they develop individual measures for the identification of accidental metrical feet, for positions where a text only partly matches a metrical grid, and for instances where different acoustic realizations of the same text parts are equally likely.

Studies focusing on metrical similarities and differences between prose and poetry have a considerable tradition in Russia (see, e.g., Yarkho, 1925; Shengeli, 1921). Yarkho (1925), whose work greatly promoted text statistics, claimed that verse and prose are not distinguished by categorical differences, but represent gradations on a scale, with verse defined as language use in which the same type of interval between rhythmical units occurs in no more than 50% of the cases. In addition, Yarkho (1969) assumed that the distribution of rhythmical units in non-literary language should conform to a Gaussian distribution and that literary language, by virtue of being art and thus a deviation from the norm, should be recognizable by not fitting into this distribution. However, this claim could not be substantiated empirically (Grzybek, 2013). A recent computational study by Anttila and Heuser (2015) used the software *Prosodic* (Heuser, 2011) to examine differences between prose and verse in both English and Finnish. These differences include the choice and linearization of words that strategically weaken or support particular metrical positions, as well as a genre-specific propensity for stress clashes (in verse) vs. lapses (in prose).

A different approach relies on exploratively assigning stress to individual syllables to determine the average length of, intervals between, and patterns resulting from, rhythmical units in literary prose (Kagarov, 1928; Lipsky, 1907). This method was made popular by Marbe (1904), who, assuming a binary system of stressed and unstressed syllables, calculated average intervals between stressed syllables and mean variation in a corpus of text extracts by Goethe and Heine. Although interindividual differences between coders turned out to be relatively high, an analysis by Kagarov (1928) for a Russian corpus arrived at similar results. More recent studies that have partly used manual annotation and partly relied on computational methods, have been able to refine those early distributional principles by annotating whole texts instead of merely extracts, and have proposed different models for different languages (e.g., Best, 2002; Kaßel, 2002; Knaus, 2008). Also drawing on Marbe's results (1904), a number of studies have explored interactions between the distribution of rhythmical units and word length, genre, and affective content (e.g., Unser, 1906; Kullmann, 1909; Gropp, 1915).

Building on Karagov's (1928) results and drawing on probability theory, Tomashevsky (1929) systematically analyzed the frequencies with which certain

intervals between rhythmical units occurred and used them to formulate predictions for general, theoretically expectable interval frequencies. Shengeli (1921) manually analyzed the accent structures in ten Russian prose passages of 5,000 words each from the nineteenth and twentieth centuries, extracting word length frequencies and distributions of accent positions in relation to word length.

Aiming to differentiate between the stress or quantity distributions of the lexicon and the acoustic prose rhythm actually realized by readers, whether they read silently or aloud, Esser (1988) evaluated reading performances and found significant differences between readers based on their levels of experience and proficiency. In addition, rhythm performance has been shown to be marked by the complexity of interactions between syllable rhythm and accent rhythm, the knowledge and foresight a reader needs to recognize and represent thematic focus, as well as the level of text interpretation through performance (Esser, 2011).

Meter and Verse Rhythm. Like phonemic parallelism, the prosodic regularity imposed by meter affects the sound shape of literary texts, text processing, text evaluation, and text memory. Research into the sound shape of metered and non-metered texts has a long-standing tradition in the empirical study of verbal art (Metcalf [1938] reviews a number of early studies of literary form). Readers tend to emphasize the rhythm of verse by various acoustic means, many of which are captured in the “formula for poetic intonation” proposed by Byers (1979); see also Barney (1999). Poetic intonation involves reduced speech rates, which reflect the systematic lengthening of speech units in the service of metrical requirements (e. g., Swedish: Fant et al., 1991; German: Wagner, 2012) as well as the number and duration of silent speech pauses (e. g., Kowal et al., 1975). Recent investigations of the acoustic correlates of the metrical hierarchy and the rhyme scheme in Dr. Seuss’s *The Cat in the Hat* (Breen, 2018; Fitzroy & Breen, 2019) have shown that readers modulate syllable durations, intensity, and intersyllable intervals to realize different levels of the metrical hierarchy.

Research into processing effects of meter has relied on various techniques to measure processing effort. Menninghaus and colleagues (2014) used self-paced reading to examine how meter affects global reading fluency; their participants read original and modified couplets of humorous German verse that fully crossed meter and rhyme. Readers slowed down if the implicit rhythm did not conform to the regularity of the metrical scheme, and they rated these couplets lower in rhythmicity, providing strong evidence that meter increases global reading fluency and the perception of rhythmical qualities. Breen and Clifton

(2011) used eye tracking to examine prosodic expectations while their participants silently read limericks containing rhyme words that were either consistent with metrical requirements or not. They observed that readers systematically slowed down at metrically inconsistent rhyme words, providing strong evidence for concrete metrical expectations during silent reading (for a review of the role of implicit prosody in sentence processing, see Breen, 2015). Recording EEG while recipients listened to original and modified quatrains of German lyrical poetry, Obermeier et al. (2016) found that only the combination of continuous meter and systematic rhyme decreased amplitudes of ERP components related to predictive lexical-semantic processing, whereas meter and rhyme individually had no such effect, indicating that meter was a prerequisite for the facilitative effect of rhyme on word processing. van Peer (1990) used a recognition task and semantic differentials to examine the effects of meter on memory and aesthetic evaluation, presenting readers with original and modified stanzas of Dutch poetry. He found that meter enhanced memory for verse and increased its perceived “smoothness.” Taken together, these results are in line with the view that meter, typically seen as a global property of a poem, generally increases reading fluency and plays a crucial role in supporting the stylistic effects of systematic end rhyme by providing the prosodic grid necessary to predict the position (i. e., the “when”) of this phonemic parallelism. While the impression of smoothness appears to reflect processing fluency – most likely a result of enhancing the reliability of predictive prosodic cues – other stylistic effects of meter appear to be more context-dependent, such as the observation that meter may render humorous verse funnier (Menninghaus et al., 2014).

Speech Melody. The term *speech melody* describes arranged sequences and contours of audible speech tones (Hart et al., 1990; Patel et al., 2006; Sievers, 1912) and implies strong affinities between perceived qualities of music and (poetic) language. To date, most elaborate melody definitions originate in the fields of music theory and music cognition; as they relate predominantly to differences between musical traditions, they are less applicable to melody in speech. Rousseau (1768) more broadly defines melody as an arranged succession of sounds in time that follows the laws of rhythm and modulation and thus creates a sensation pleasant to the ear.

As perceived, incremental distributions of tone heights and tone durations are sufficient for melodic impressions; but they are by no means the only melody-defining features. However, they are the ones shared by music and speech and will, therefore, be highlighted here. Studies in music cognition suggest that melodic percepts are mainly shaped by tone contours and tone intervals. By contrast, speech is not marked by fixed tone heights and intervals of musical melodies, but by approximate pitches (Sievers, 1912, p. 57). In addition, the dis-

crete character of tones with regard to their pitch and duration in musical melodies, is opposed to the continuous changes of tone in speech over phrases and sentences.

If considered in relation to syllables, speech tones are also rather variable in terms of duration. Furthermore, speech melodies do not seem to have tonal centers (Krumhansl, 1990), although the median pitch of speakers (depending on age and gender) may be seen as an approximation to such tonal stability. Finally, the range of intervals in musical melodies is larger than the range of intervals in speech. This holds true even when comparing a singing voice (with a range of two octaves) with a speaking voice (with a range of one octave; see Fant, 1956). With regard to poetic language, Lanz (1926) highlights rhyme as the melodically most relevant feature, conceptualizing its occurrence as the return to the tonic in language.

Methodologically, the basis for prosodic analyses of speech, including analyses of speech melody, is the digitized speech signal, represented as amplitudes over time. Speech intonation analyses require the extraction of the pitch track, i. e., the varying fundamental frequency based on vocal fold oscillation during vowels and sonorous consonants. This pitch track is based on the repetition rate of the vocal folds and is estimated by autocorrelation analyses (de Cheveigné & Kawahara, 2002; Paliwal & Rao, 1981). The lines in the pitch track illustrate the continuous changes in speech intonation (speech melody). In linguistics, pitch tracks are the basis of prosodic analyses, involving a notational system that allows the characterization of different tone types in relation to phonological and syntactic phrases (e. g., Pierrehumbert, 1980).

Studies of music and speech cognition suggest that speech melody can also be modelled more discretely by taking mean pitch values per syllable together with (absolute) syllable durations (e. g., Patel et al., 2006). When this is done, speech intonation can be expressed in musical notation systems. The advantage of these representations is that the discrete notes can then undergo statistical analyses commonly applied for studying the (more abstract) Gestalt of melodies (see, e. g., Müllensiefen & Frieler, 2007). Musical approaches to speech melodies have a long history (Steele, 1775) and are still relevant to describing commonalities of speech and music (Chow & Brown, 2018), potentially linking to evolutionary perspectives (Fenk-Oczlon, 2017).

Two early studies by Schramm (1935a, 1935b) corroborate Lanz's (1926) hypothesis that rhyme is the primary melodically relevant element in poetic language by providing empirical evidence for recurrent syllable pitch. In addition, Menninghaus et al. (2018) have found evidence for the recurrence of entire pitch sequences, underscoring Rousseau's (1768) understanding of melodies as incremental arrangements of tones. In this study, discrete speech tones were sub-

jected to autocorrelation analyses on the level of the stanza, revealing recurrences in the pitch track that the linguistic notational system could not capture, given that it was restricted to local phenomena. The autocorrelation approach, on the other hand, was able to quantify longer range dependencies and patterns, especially the recurrence of similar pitches across stanzas. Importantly, the size of the autocorrelation value at the stanza-lag was shown not only to predict melodiousness ratings of spoken poems, but also to correlate with the likelihood that a particular poem had been set to music. Speech melody, therefore, despite well-attested differences to musical melodies on local levels, seems to be perceived on similarly Gestalt-like grounds, providing empirical evidence for the century-old assumption that poetic language (speech) and music are intrinsically linked by melodic properties.

Suggestions and Recommendations for Future Research

Based on our review, we provide some recommendations for researchers interested in studying the sound effects of literary texts and suggest areas for future research. With regard to the experimental study of literary comprehension more generally, we have observed that, in many studies, samples of recipients and particularly of texts – the actual object of inquiry – were surprisingly small, reducing the external validity of the results. Including sufficiently large samples of both participants and texts permits the generalization of results to recipients and literary texts not investigated in a given experiment. It also avoids confounds with inter-individual differences and features of individual texts that often yield stronger effects than sound structure does, i. e., semantic and thematic ones.

Regarding internal validity, we note considerable differences in the specification and operationalization of theoretical constructs and in the control of potential confounds. We recommend paying particular attention to the control of linguistic complexity and, again, to semantic and thematic text features whenever testing claims about sound effects. Further, a thorough understanding of a given phonological system should guide researchers as they operationalize their research questions, create appropriate and sufficient contrasts in the materials, draw conclusions and recognize the limits of applicability, as well as formulate further hypotheses.

Regarding the modality (spoken/written) of stimulus material or raw data in investigations of sound-related literary comprehension, we recommend following a sound-over-letter principle whenever possible. Spoken words provide richer sound information than printed ones. Speech signals thus constitute (a) explicit and more effective stimulus materials for investigating sound effects during literary comprehension and (b) more informative data than approximations based on written texts (e. g., via grapheme-to-phoneme conversion).

We have further observed that most of the existing empirical research on sound and sound effects in literary texts has been conducted in only a small set of languages. A deeper understanding of these phenomena will be achieved by taking more diverse languages and literatures into account and by systematically comparing sound-related literary comprehension across languages and literatures. However, evidential gaps also exist within better researched literatures and well-established research areas, such as the study of phonemic parallelism and its effects. Little is known, for instance, about how alliteration affects the processing and evaluation of poems and prose passages. Likewise, even apparently well-established insights, such as the context-dependence of end rhyme's stylistic effects, leave ample room for empirical refinement. Finally, a large number of empirical studies of sound shape and sound effects in literary texts focus either on large corpora, as is the case for most research into prose rhythm, or on the way literary texts are processed by the reader. However, few engage the combination of both (but see Xue et al., 2019). A tighter alliance between researchers digitally analyzing large corpora and researchers studying readers' responses to selected stimuli would allow for more specific and nuanced insights into the effects and the development of literary forms.

References

- Aili, H. (1979). *The prose rhythm of Sallust and Livy*. Almqvist & Wiksell.
- Albers, S. (2008). *Lautsymbolik in Ägyptischen Texten*. Philipp Von Zabern.
- Anttila, A., & Heuser, R. (2015). Phonological and metrical variation across genres. *Proceedings of the 2015 annual meeting on phonology*, 1–12. <https://dx.doi.org/10.3765/amp.v3i0.3679>
- Aryani, A., Kraxenberger, M., Ullrich, S., Jacobs, A. M., & Conrad, M. (2016). Measuring the basic affective tone of poems via phonological saliency and iconicity. *Psychology of Aesthetics, Creativity, and the Arts*, 10(2), 191–204. <https://dx.doi.org/10.1037/aca0000033>
- Aryani, A., Hsu, C. T., & Jacobs, A. M. (2019). Affective iconic words benefit from additional sound-meaning integration in the left amygdala. *Human Brain Mapping*, 40(18), 5289–5300. <https://dx.doi.org/10.1002/hbm.24772>

- Ashby, J., & Clifton, C., Jr. (2005). The prosodic property of lexical stress affects eye movements during silent reading. *Cognition*, 96(3), B89-B100. <https://dx.doi.org/10.1016/j.cognition.2004.12.006>
- Atchley, R. M., & Hare, M. L. (2013). Memory for poetry: More than meaning? *International Journal of Cognitive Linguistics*, 4(1), 35–50.
- Aumont, J. (1996). *Métrique et stylistique des clausules dans la prose latine: De Cicéron à Plinie le Jeune et de César à Florus* [Metrics and stylistics of clausulae in Latin prose: From Cicero to Pliny the Younger and from Caesar to Florus]. Champion.
- Auracher, J., Albers, S., Zhai, Y., Gareeva, G., & Stavniychuk, T. (2010). P is for happiness, N is for sadness: Universals in sound iconicity to detect emotions in poetry. *Discourse Processes*, 48(1), 1–25. <https://dx.doi.org/10.1080/01638531003674894>
- Auracher, J., Scharinger, M., & Menninghaus, W. (2019). Contiguity-based sound iconicity: The meaning of words resonates with phonetic properties of their immediate verbal contexts. *PLoS ONE*, 14(5), e0216930. <https://dx.doi.org/10.1371/journal.pone.0216930>
- Bader, M. (1998). Prosodic influences on reading syntactically ambiguous sentences. In J. D. Fodor & F. Ferreira (Eds.), *Reanalysis in sentence processing* (pp. 1–46). Springer. https://dx.doi.org/10.1007/978-94-015-9070-9_1
- Ballard, P. B. (1913). Oblivescence and reminiscence. *British Journal of Psychology*. Monograph Supplements, 1, 1–82.
- Barney, T. (1999). Readers as text processors and performers: A new formula for poetic intonation. *Discourse Processes*, 28(2), 155–167. <https://dx.doi.org/10.1080/01638539909545078>
- Beebe-Center, J. G., & Pratt, C. C. (1937). A test of Birkhoff's aesthetic measure. *The Journal of General Psychology*, 17(2), 339–353. <https://dx.doi.org/10.1080/00221309.1937.9918004>
- Berent, I. (2013). The phonological mind. *Trends in Cognitive Sciences*, 17(7), 319–327. <https://dx.doi.org/10.1016/j.tics.2013.05.004>
- Best, K.-H. (2002). The distribution of rhythmic units in German short prose. *Glottometrics*, 3, 136–142. <https://www.ram-verlag.eu/wp-content/uploads/2018/08/glo3abstract.pdf>
- Birkhoff, G. D. (1933). *Aesthetic measure*. Harvard UP.
- Blohm, S., Menninghaus, W., & Schlesewsky, M. (2017). Sentence-level effects of literary genre: Behavioral and electrophysiological evidence. *Frontiers in Psychology*, 8, 1887. <https://dx.doi.org/10.3389/fpsyg.2017.01887>
- Blohm, S., Versace, S., Methner, S., Wagner, V., Schlesewsky, M., & Menninghaus, W. (in press). Reading poetry and prose: Eye movements and acoustic evidence. *Discourse Processes*.
- Borgeson, S., Anttila, A., Heuser, R., & Kiparsky, P. (2018). *The rise and fall of antimetricity*. [Unpublished manuscript]. Stanford University, Palo Alto, USA. <https://web.stanford.edu/~kiparsky/Papers/Antimetricity-Draft-2018-04-20.pdf>
- Bornecque, H. (1907). *Les clausules métriques latines* [Metrical clausulae in Latin]. Au siège de l'Université de Lille.
- Bower, G. H., & Bolton, L. S. (1969). Why are rhymes easy to learn? *Journal of Experimental Psychology*, 82(3), 453–461. <https://dx.doi.org/10.1037/h0028365>
- Braun, M., Hutzler, F., Ziegler, J. C., Dambacher, M., & Jacobs, A. M. (2009). Pseudohomophone effects provide evidence of early lexico-phonological processing in visual word recognition. *Human Brain Mapping*, 30(7), 1977–1989. <https://dx.doi.org/10.1002/hbm.20643>

- Breen, M. (2014). Empirical investigations of the role of implicit prosody in sentence processing. *Language and Linguistics Compass*, 8(2), 37–50. <https://dx.doi.org/10.1111/lnc3.12061>
- Breen, M. (2015). Empirical investigations of implicit prosody. In L. Frazier & E. Gibson (Eds.), *Explicit and implicit prosody in sentence processing: Studies in honor of Janet Dean Fodor* (pp. 177–192). Springer. https://dx.doi.org/10.1007/978-3-319-12961-7_10
- Breen, M. (2018). Effects of metric hierarchy and rhyme predictability on word duration in *The Cat in the Hat*. *Cognition*, 174, 71–81. <https://dx.doi.org/10.1016/j.cognition.2018.01.014>
- Breen, M., & Clifton, C., Jr. (2011). Stress matters: Effects of anticipated lexical stress on silent reading. *Journal of Memory and Language*, 64(2), 153–170. <https://dx.doi.org/10.1016/j.jml.2010.11.001>
- Byers, P. P. (1979). A formula for poetic intonation. *Poetics*, 8(4), 367–380. [https://dx.doi.org/10.1016/0304-422X\(79\)90007-X](https://dx.doi.org/10.1016/0304-422X(79)90007-X)
- Carminati, M. N., Stabler, J., Roberts, A. M., & Fischer, M. H. (2006). Readers' responses to sub-genre and rhyme scheme in poetry. *Poetics*, 34(3), 204–218. <https://dx.doi.org/10.1016/j.poetic.2006.05.001>
- Chafe, W. (1988). Punctuation and the prosody of written language. *Written Communication*, 5(4), 395–426. <https://dx.doi.org/10.1177/0741088388005004001>
- Chen, Q., Zhang, J., Xu, X., Scheepers, C., Yang, Y., & Tanenhaus, M. K. (2016). Prosodic expectations in silent reading: ERP evidence from rhyme scheme and semantic congruence in classic Chinese poems. *Cognition*, 154, 11–21. <https://dx.doi.org/10.1016/j.cognition.2016.05.007>
- Chow, I., & Brown, S. (2018). A musical approach to speech melody. *Frontiers in Psychology*, 9, 247. <https://dx.doi.org/10.3389/fpsyg.2018.00247>
- Cicero, M. T. (2001). *De oratore* (E. W. Sutton & H. Rackham, Eds. and Trans.). Loeb Classical Library. Harvard UP.
- Clifton, C., Jr. (2015). The roles of phonology in silent reading: A selective review. In L. Frazier & E. Gibson (Eds.), *Explicit and implicit prosody in sentence processing: Studies in honor of Janet Dean Fodor* (pp. 161–176). Springer. https://dx.doi.org/10.1007/978-3-319-12961-7_10
- Comrie, B. (2013). Writing systems. In M. Dryer & M. Haspelmath (Eds.), *The world atlas of language structures (WALS) online*. Leipzig: Max Planck Institute for Evolutionary Anthropology. <https://wals.info/chapter/141>
- Conrad, M., Carreiras, M., Tamm, S., & Jacobs, A. M. (2009). Syllables and bigrams: Orthographic redundancy and syllabic units affect visual word recognition at different processing levels. *Journal of Experimental Psychology: Human Perception and Performance*, 35(2), 461–479. <https://dx.doi.org/10.1037/a0013480>
- Cutler, A., Dahan, D., & van Donselaar, W. (1997). Prosody in the comprehension of spoken language: A literature review. *Language and Speech*, 40(2), 141–201. <https://dx.doi.org/10.1177/002383099704000203>
- Davis, R. C. (1936). An evaluation and test of Birkhoff's aesthetic measure formula. *The Journal of General Psychology*, 15(2), 231–240. <https://dx.doi.org/10.1080/00221309.1936.9917921>
- De Beaugrande, R.-A. (1978). Information, expectation, and processing: On classifying poetic texts. *Poetics*, 7(1), 3–44. [https://dx.doi.org/10.1016/0304-422X\(78\)90003-7](https://dx.doi.org/10.1016/0304-422X(78)90003-7)

- de Cheveigné, A., & Kawahara, H. (2002). YIN, a fundamental frequency estimator for speech and music. *The Journal of the Acoustical Society of America*, 111(4), 1917–1930. <https://doi.org/10.1121/1.1458024>
- de Groot, A. W. (1921). *Der antike Prosarhythmus* [Prose rhythm in antiquity]. J. B. Wolters.
- de Saussure, F. (1983). *Course in general linguistics* (R. Harris, Trans.). Duckworth.
- Dingemanse, M., Blasi, D. E., Lupyan, G., Christiansen, M. H., & Monaghan, P. (2015). Arbitrariness, iconicity, and systematicity in language. *Trends in Cognitive Sciences*, 19(10), 603–615. <https://dx.doi.org/10.1016/j.tics.2015.07.013>
- Esser, J. (1988). *Comparing reading and speaking intonation*. Rodopi.
- Esser, J. (2011). *Rhythm in speech, prose, and verse: A linguistic description*. Logos.
- Fabb, N. (1997). *Linguistics and literature: Language in the verbal arts of the world*. Blackwell.
- Fabb, N. (1999). Verse constituency and the locality of alliteration. *Lingua*, 108(4), 223–245. [https://dx.doi.org/10.1016/S0024-3841\(98\)00054-0](https://dx.doi.org/10.1016/S0024-3841(98)00054-0)
- Fant, G. (1956). On the predictability of formant levels and spectrum envelopes from formant frequencies. In M. Halle, H. Lunt, & H. Maclean (Eds.), *For Roman Jakobson* (pp. 109–120). Mouton.
- Fant, G., Kruckenberg, A., & Nord, L. (1991). Stress patterns and rhythm in the reading of prose and poetry with analogies to music performance. In J. Sundberg, L. Nord, & R. Carlson (Eds.), *Music, language, speech and brain: Proceedings of an international symposium at the Wenner-Gren Center, Stockholm, 5–8 September 1990* (pp. 380–407). Macmillan.
- Fechner, G. T. (1876). *Vorschule der Ästhetik* [Primary school of aesthetics]. Breitkopf & Härtel.
- Fenk-Oczlon, G. (2017). What vowels can tell us about the evolution of music. *Frontiers in Psychology*, 8, 1581. <https://dx.doi.org/10.3389/fpsyg.2017.01581>
- Fitzroy, A. B., & Breen, M. (2019). Metric structure and rhyme predictability modulate speech intensity during child-directed and read-alone productions of children's literature. *Language and Speech*, 63(2), 292–305. <https://dx.doi.org/10.1177/0023830919843158>
- Fodor, J. D. (1998). Learning to parse? *Journal of Psycholinguistic Research*, 27(2), 285–319. <https://dx.doi.org/10.1023/A:1023258301588>
- Fodor, J. D. (2002). Prosodic disambiguation in silent reading. *Proceedings of NELS*, 32(1), 113–132.
- Fónagy, I. (1961). Communication in poetry. *Word*, 17(2), 194–218. <https://dx.doi.org/10.1080/00437956.1961.11659754>
- Frog, M., & Tarkka, L. (Eds.) (2017). Parallelism in verbal art and performance [Special issue]. *Oral Tradition*, 31(2). Center for Studies in Oral Tradition.
- Gropp, F. (1915). *Zur Ästhetik und statistischen Beschreibung des Prosarhythmus* [On the aesthetics and statistical description of prose rhythm]. Königliche Universitätsdruckerei H. Stütz.
- Grzybek, P. (2013). Empirische Textwissenschaft: Prosarhythmus im ersten Drittel des 20. Jahrhunderts als historisch-systematische Fallstudie [Empirical textual science: Prose rhythm in the first third of the twentieth century as a historical-systematic case study]. In A. Hansen-Löve, B. Obermayr, & G. Witte (Eds.), *Form und Wirkung: Phänomenologische und empirische Kunstwissenschaft in der Sowjetunion der 1920er Jahre* (pp. 427–455). Fink.
- Hayes, B. (2009). *Introductory phonology*. Wiley-Blackwell.
- Heuser, R., Falk, J., & Anttila, A. (2011). *Prosodic (software)*. <https://github.com/quadrismegistus/prosodic>

- Hevner, K. (1937). An experimental study of the affective value of sounds in poetry. *The American Journal of Psychology*, 49(3), 419–434. <https://dx.doi.org/10.2307/1415776>
- Hinton, L., Nichols, J., & Ohala, J. J. (Eds.) (1994). *Sound symbolism*. Cambridge UP.
- Hoorn, J. F. (1996). Psychophysiology and literary processing: ERPs to semantic and phonological deviations in reading small verses. In R. J. Kreuz & M. S. MacNealy (Eds.), *Empirical approaches to literature and aesthetics* (pp. 339–358). Ablex.
- Huang, Y., Pratoomraj, S., & Johnson, R. C. (1969). Universal magnitude symbolism. *Journal of Verbal Learning & Verbal Behavior*, 8(1), 155–156. [https://doi.org/10.1016/S0022-5371\(69\)80028-9](https://doi.org/10.1016/S0022-5371(69)80028-9)
- Huey, E. B. (1908). *The psychology and pedagogy of reading*. Macmillan.
- Jacobs, A. M. (2017). Quantifying the beauty of words: A neurocognitive poetics perspective. *Frontiers in Human Neuroscience*, 11, 622. <https://dx.doi.org/10.3389/fnhum.2017.00622>
- Jacobs, A. M. (2018). (Neuro-)cognitive poetics and computational stylistics. *Scientific Study of Literature*, 8(1), 165–208. <https://dx.doi.org/10.1075/ssol.18002.jac>
- Jacobs, A. M. & Grainger, J. (1994). Models of visual word recognition: Sampling the state of the art. *Journal of Experimental Psychology: Human Perception and Performance*, 20(6), 1311–1334. <https://dx.doi.org/10.1037/0096-1523.20.6.1311>
- Jakobson, R. (1960). Closing statement: Linguistics and poetics. In T. A. Sebeok (Ed.), *Style in language* (pp. 350–377). MIT Press.
- Jakobson, R. (1966). Grammatical parallelism and its Russian facet. *Language*, 42(2), 399–429. <https://dx.doi.org/10.2307/411699>
- Jakobson, R. & Waugh, L. R. (2002). *The sound shape of language* (3rd ed.). Mouton de Gruyter.
- Jespersen, O. (1933). Notes on metre. *Linguistica* (pp. 249–274). Allen & Unwin.
- Just, M. A., Carpenter, P. A., & Woolley, J. D. (1982). Paradigms and processes in reading comprehension. *Journal of Experimental Psychology: General*, 111(2), 228–238. <https://dx.doi.org/10.1037/0096-3445.111.2.228>
- Kagarov, E. G. (1928). O ritme prozaicheskoi rechi [On rhythm in prose language]. *Doklady Akademii Nauk SSSR, Seria B*, 44–51.
- Kaßel, A. (2002). *Zur Verteilung rhythmischer Einheiten in deutschen und englischen Texten* [On the distribution of rhythmic units in German and English texts] Unpublished master's thesis. University of Göttingen, Göttingen.
- Kadota, S. (1987). The role of prosody in silent reading. *Language Sciences*, 9(2), 185–206. [https://dx.doi.org/10.1016/S0388-0001\(87\)80019-0](https://dx.doi.org/10.1016/S0388-0001(87)80019-0)
- Kentner, G., & Vasishth, S. (2016). Prosodic focus marking in silent reading: Effects of discourse context and rhythm. *Frontiers in Psychology*, 7, 319. <https://dx.doi.org/10.3389/fpsyg.2016.00319>
- Kintsch, W., Welsch, D., Schmalhofer, F., & Zimny, S. (1990). Sentence memory: A theoretical analysis. *Journal of Memory and Language*, 29(2), 133–159. [https://dx.doi.org/10.1016/0749-596X\(90\)90069-C](https://dx.doi.org/10.1016/0749-596X(90)90069-C)
- Knapp, R. (2013). *Der Prosarhythmus als Stilmittel: Eine Untersuchung an Ciceros De natura deorum mit statistischer Auswertung und Softwareentwicklung* [Prose rhythm as stylistic medium: A study of Cicero's De natura deorum with a statistical evaluation and software development] (Unpublished master's thesis). <https://eplus.uni-salzburg.at/obvusbhs/content/titleinfo/4375771>

- Knaus, M. (2008). Zur Verteilung rhythmischer Einheiten in russischer Prosa [On the distribution of rhythmic units in Russian prose]. *Glottometrics*, 16, 57–62. <https://www.ram-verlag.eu/wp-content/uploads/2018/08/g16zeit.pdf>
- Knoop, C. A., Blohm, S., Kraxenberger, M. (2019). How perfect are imperfect rhymes? Effects of phonological similarity and verse context on rhyme perception. *Psychology of Aesthetics, Creativity, and the Arts*. (Advance online publication). <https://dx.doi.org/10.1037/aca0000277>
- Köhler, W. (1929). *Gestalt psychology: An introduction to new concepts in modern psychology*. Liveright.
- Kowal, S., O'Connell, D. C., O'Brien, E. A., & Bryant, E. T. (1975). Temporal aspects of reading aloud and speaking: Three experiments. *The American Journal of Psychology*, 88(4), 549–569. <https://dx.doi.org/10.2307/1421893>
- Kraxenberger, M. (2017). *On sound-emotion associations in poetry*. Unpublished doctoral dissertation. Freie Universität, Berlin.
- Kraxenberger, M., & Menninghaus, W. (2016a). Emotional effects of poetic phonology, word positioning, and dominant stress peaks in poetry reading. *Scientific Study of Literature*, 6 (2), 298–313. <https://dx.doi.org/10.1075/ssol.6.2.06kra>
- Kraxenberger, M., & Menninghaus, W. (2016b). Mimological reveries? Disconfirming the hypothesis of phono-emotional iconicity in poetry. *Frontiers in Psychology*, 7, 1779. <https://dx.doi.org/10.3389/fpsyg.2016.01779>
- Kraxenberger, M., & Menninghaus, W. (2017). Affinity for poetry and aesthetic appreciation of joyful and sad poems. *Frontiers in Psychology*, 7, 2051. <https://dx.doi.org/10.3389/fpsyg.2016.02051>
- Kraxenberger, M., Menninghaus, W., Roth, A., & Scharinger, M. (2018). Prosody-based sound-emotion associations in poetry. *Frontiers in Psychology*, 9, 1284. <https://dx.doi.org/10.3389/fpsyg.2018.01284>
- Krumhansl, C. L. (1990). *Cognitive Foundations of Musical Pitch*. Oxford UP.
- Kullmann, P. (1909). Statistische Untersuchungen zur Sprachpsychologie. *Zeitschrift für Psychologie und Physiologie der Sinnesorgane*, 54, 290–310.
- Lahiri, A., & Reetz, H. (2002). Underspecified recognition. In C. Gussenhoven & N. Warner (Eds.), *Laboratory Phonology VII* (pp. 637–677). Mouton de Gruyter.
- Lanz, H. (1926). The physical basis of rime. *PMLA*, 41(4), 1011–1023. <https://dx.doi.org/10.2307/457460>
- Lea, R. B., Rapp, D. N., Elfenbein, A., Mitchel, A. D., & Romine, R. S. (2008). Sweet silent thought: Alliteration and resonance in poetry comprehension. *Psychological Science*, 19 (7), 709–716. <https://dx.doi.org/10.1111/j.1467-9280.2008.02146.x>
- Lipsky, A. (1907). *Rhythm as a distinguishing characteristic of prose style*. The Science Press.
- Luck, S. J. (2014). *An introduction to the event-related potential technique* (2nd ed.). MIT Press.
- Maddieson, I. (2013a). Consonant inventories. In M. Dryer & M. Haspelmath (Eds.), *The world atlas of language structures (WALS) online*. Leipzig: Max Planck Institute for Evolutionary Anthropology. <https://wals.info/chapter/1>
- Maddieson, I. (2013b). Syllable structure. In M. Dryer & M. Haspelmath (Eds.), *The world atlas of language structures (WALS) online*. Leipzig: Max Planck Institute for Evolutionary Anthropology. <https://wals.info/chapter/12>
- Maddieson, I. (2013c). Vowel quality inventories. In M. Dryer & M. Haspelmath (Eds.), *The world atlas of language structures (WALS) online*. Leipzig: Max Planck Institute for Evolutionary Anthropology. <https://wals.info/chapter/2>

- Maïonchi-Pino, N., Cara, B. D., Magnan, A., & Ecalles, J. (2008). Roles of consonant status and sonority in printed syllable processing: Evidence from illusory conjunction and audio-visual recognition tasks in French adults. *Current Psychology Letters: Behaviour, Brain & Cognition*, 24(2). <https://journals.openedition.org/cpl/4033>
- Marbe, K. (1904). *Über den Rhythmus der Prosa* [On prose rhythm]. J. Ricker'sche Verlagsbuchhandlung.
- Marslen-Wilson, W., & Zwitserlood, P. (1989). Accessing spoken words: The importance of word onsets. *Journal of Experimental Psychology: Human Perception and Performance*, 15(3), 576–585. <https://dx.doi.org/10.1037/0096-1523.15.3.576>
- McCutchen, D., & Perfetti, C. A. (1982). The visual tongue-twister effect: Phonological activation in silent reading. *Journal of Verbal Learning and Verbal Behavior*, 21(6), 672–687. [https://dx.doi.org/10.1016/S0022-5371\(82\)90870-2](https://dx.doi.org/10.1016/S0022-5371(82)90870-2)
- McGlone, M. S., & Tofighbakhsh, J. (2000). Birds of a feather flock conjointly (?): Rhyme as reason in aphorisms. *Psychological Science*, 11(5), 424–428. <https://dx.doi.org/10.1111/1467-9280.00282>
- Menninghaus, W., Bohrn, I. C., Altmann, U., Lubrich, O., & Jacobs, A. M. (2014). Sounds funny? Humor effects of phonological and prosodic figures of speech. *Psychology of Aesthetics, Creativity, and the Arts*, 8(1), 71–76. <https://dx.doi.org/10.1037/a0035309>
- Menninghaus, W., Bohrn, I. C., Knoop, C. A., Kotz, S. A., Schlotz, W., & Jacobs, A. M. (2015). Rhetorical features facilitate prosodic processing while handicapping ease of semantic comprehension. *Cognition*, 143, 48–60. <https://dx.doi.org/10.1016/j.cognition.2015.05.026>
- Menninghaus, W., Wagner, V., Knoop, C. A., & Scharinger, M. (2018). Poetic speech melody: A crucial link between music and language. *PLoS ONE*, 13(11), e0205980. <https://dx.doi.org/10.1371/journal.pone.0205980>
- Menninghaus, W., Wagner, V., Wassiliwizky, E., Jacobsen, T., & Knoop, C. A. (2017). The emotional and aesthetic powers of parallelistic diction. *Poetics*, 63, 47–59. <https://dx.doi.org/10.1016/j.poetic.2016.12.001>
- Metcalfe, J. T. (1938). Psychological studies of literary form. *Psychological Bulletin*, 35(6), 337–357. <https://dx.doi.org/10.1037/h0055922>
- Miall, D. S. (2001). Sounds of contrast: An empirical approach to phonemic iconicity. *Poetics*, 29(1), 55–70. [https://dx.doi.org/10.1016/S0304-422X\(00\)00025-5](https://dx.doi.org/10.1016/S0304-422X(00)00025-5)
- Miall, D. S., & Kuiken, D. (2002). The effects of local phonetic contrasts in readers' responses to a short story. *Empirical Studies of the Arts*, 20(2), 157–175. <https://dx.doi.org/10.2190/M9RC-WBP5-4NDQ-2EJD>
- Moos, A., Smith, R., Miller, S. R., & Simmons, D. R. (2014). Cross-modal associations in synaesthesia: Vowel colours in the ear of the beholder. *i-Perception*, 5(2), 132–142. <https://dx.doi.org/10.1068/i0626>
- Müllensiefen, D., & Frieler, K. (2007). Modelling experts notions of melodic similarity. *Musicae Scientiae*, Discussion Forum 4A, 183–210.
- Nespor, M., & Vogel, I. (1986). *Prosodic Phonology*. Foris.
- Nöth, W. (2000). *Handbuch der Semiotik* [Handbook of semiotics] (2nd ed.). J. B. Metzler.
- Obermeier, C., Kotz, S. A., Jessen, S., Raettig, T., von Koppenfels, M., & Menninghaus, W. (2016). Aesthetic appreciation of poetry correlates with ease of processing in event-related potentials. *Cognitive Affective & Behavioral Neuroscience*, 16(2), 362–373. <https://dx.doi.org/10.3758/s13415-015-0396-x>

- Obermeier, C., Menninghaus, W., von Koppenfels, M., Raettig, T., Schmidt-Kassow, M., Otterbein, S., & Kotz, S. A. (2013). Aesthetic and emotional effects of meter and rhyme in poetry. *Frontiers in Psychology*, 4, 10. <https://dx.doi.org/10.3389/fpsyg.2013.00010>
- Osgood, C. E., & Snider, J. G. (1969). *Semantic differential technique: A sourcebook*. Aldine.
- Paliwal, K. K., & Rao, P. V. S. (1981). A modified autocorrelation method of linear prediction for pitch-synchronous analysis of voiced speech. *Signal Processing*, 3(2), 181–185. [https://dx.doi.org/10.1016/0165-1684\(81\)90082-7](https://dx.doi.org/10.1016/0165-1684(81)90082-7)
- Patel, A. D., Iversen, J. R., & Rosenberg, J. C. (2006). Comparing the rhythm and melody of speech and music: The case of British English and French. *The Journal of the Acoustical Society of America*, 119(5), 3034–3047. <https://dx.doi.org/10.1121/1.2179657>
- Perlman, M., & Benitez, N. J. (2010). Talking fast: The use of speech rate as iconic gesture. In F. Perrill, V. Tobin, & M. Turner (Eds.), *Meaning, form, and body* (pp. 245–262). CSLI Publications.
- Perlman, M., Clark, N., & Johansson Falck, M. (2015). Iconic prosody in story reading. *Cognitive Science*, 39(6), 1348–1368. <https://dx.doi.org/10.1111/cogs.12190>
- Perniss, P., Thompson, R. L., & Vigliocco, G. (2010). Iconicity as a general property of language: Evidence from spoken and signed languages. *Frontiers in Psychology*, 1, 227. <https://dx.doi.org/10.3389/fpsyg.2010.00227>
- Pierrehumbert, J. B. (1980). *The phonology and phonetics of English intonation*. Unpublished doctoral dissertation. Massachusetts Institute of Technology. <https://dspace.mit.edu/handle/1721.1/16065>
- Plato (1892). Cratylus. In *The dialogues of Plato* (B. Jowett, Trans.) (pp. 253–289). Clarendon.
- Pope, A. (2010). *An essay on criticism: With introductory and explanatory notes*. The Floating Press.
- Primmer, A. (1968). *Cicero numerosus: Studien zum antiken Prosarhythmus* [Cicero numerosus: Studies of antique prose rhythm]. Böhlau.
- Ramachandran, V. S., & Hubbard, E. M. (2001). Synaesthesia – A window into perception, thought and language. *Journal of Consciousness Studies*, 8(12), 3–34.
- Rayner, K. (1998). Eye movements in reading and information processing: 20 years of research. *Psychological Bulletin*, 124(3), 372–422. <https://dx.doi.org/10.1037/0033-2909.124.3.372>
- Rousseau, J.-J. (1768). *Dictionnaire de musique* [Dictionary of music]. Veuve Duchesne. <https://archive.org/details/dictionnairedem00rous>
- Rubin, D. C. (1995). *Memory in oral traditions: The cognitive psychology of epic, ballads, and counting-out rhymes*. Oxford UP.
- Rubin, D. C., & Wallace, W. T. (1989). Rhyme and reason: Analyses of dual retrieval cues. *Journal of Experimental Psychology: Learning Memory and Cognition*, 15(4), 698–709. <https://dx.doi.org/10.1037/0278-7393.15.4.698>
- Rubin, D. C., Ciobanu, V., & Langston, W. (1997). Children's memory for counting-out rhymes: A cross-language comparison. *Psychonomic Bulletin & Review*, 4(3), 421–424. <https://dx.doi.org/10.3758/bf03210804>
- Rubin, D. C., Wallace, W. T., & Houston, B. C. (1993). The beginnings of expertise for ballads. *Cognitive Science*, 17(3), 435–462. https://dx.doi.org/10.1207/s15516709cog1703_4
- Saintsbury, G. (1912). *A history of English prose rhythm*. Macmillan. <https://archive.org/details/historyofenglis00sainuoft>
- Sapir, E. (1929). A study in phonetic symbolism. *Journal of Experimental Psychology*, 12(3), 225–239. <https://dx.doi.org/10.1037/h0070931>

- Scheepers, C., Mohr, S., Fischer, M. H., & Roberts, A. M. (2013). Listening to limericks: A pupillometry investigation of perceivers' expectancy. *PLoS ONE*, 8(9), e74986. <https://dx.doi.org/10.1371/journal.pone.0074986>
- Schmidtke, D. S., Conrad, M. & Jacobs, A. M. (2014). Phonological iconicity. *Frontiers in Psychology*, 5, 80. <https://dx.doi.org/10.3389/fpsyg.2014.00080>
- Schramm, W. L. (1935a). A characteristic of rime. *PMLA*, 50(4), 1223–1227. <https://dx.doi.org/10.2307/458118>
- Schramm, W. L. (1935b). The melodies of verse. *Science*, 82(2116), 61–62. <https://dx.doi.org/10.1126/science.82.2116.61>
- Schrott, R., & Jacobs, A. M. (2011). *Gehirn und Gedicht: Wie wir unsere Wirklichkeiten konstruieren (Brain and Poetry: How We Construct Our Realities)*. Hanser.
- Shengeli, G. A. (1921). *Traktat o russkom stikhe* [Treatise on Russian verse]. Gos. izd.vo.
- Sievers, E. (1912). *Rhythmisch-melodische Studien: Vorträge und Aufsätze* [Studies of rhythm and melody: Lectures and essays]. C. Winter.
- Stausland Johnsen, S. (2012). *Rhyme acceptability determined by perceived similarity*. Unpublished manuscript, University of Oslo. https://folk.uio.no/sverrej/talk_files/2012/april/nctu/stausland.johnsen.april.2012.nctu.pdf
- Steele, J. (1775). *An essay towards establishing the melody and measure of speech to be expressed and perpetuated by peculiar symbols*. Bowyer and Nichols. <https://archive.org/details/essaytowardsesta00stee>
- Steinhauer, K., & Friederici, A. D. (2001). Prosodic boundaries, comma rules, and brain responses: The closure positive shift in ERPs as a universal marker for prosodic phrasing in listeners and readers. *Journal of Psycholinguistic Research*, 30(3), 267–295. <https://dx.doi.org/10.1023/a:1010443001646>
- Stolterfoht, B., Friederici, A. D., Alter, K., & Steube, A. (2007). Processing focus structure and implicit prosody during reading: Differential ERP effects. *Cognition*, 104(3), 565–590. <https://dx.doi.org/10.1016/j.cognition.2006.08.001>
- Hart, J., Collier, R., & Cohen, A. (1990). *A perceptual study of intonation: An experimental-phonetic approach to speech melody*. Cambridge UP.
- Thompson, P. D., & Estes, Z. (2011). Sound symbolic naming of novel objects is a graded function. *Quarterly Journal of Experimental Psychology*, 64(12), 2392–2404. <https://dx.doi.org/10.1080/17470218.2011.605898>
- Tillmann, B., & Dowling, W. J. (2007). Memory decreases for prose, but not for poetry. *Memory and Cognition*, 35(4), 628–639. <https://dx.doi.org/10.3758/BF03193301>
- Tomashevsky, B. (1929). Ritm prozy: Pikovaya dama [The rhythm of prose: The Queen of Spades]. In *O stikhe* (pp. 254–318). Priboy.
- Tsur, R. (1992). *What makes sound patterns expressive? The poetic mode of speech perception*. Duke UP.
- Tsur, R. (1997). Sound affects of poetry: Critical impressionism, reductionism and cognitive poetics. *Pragmatics and Cognition*, 5(2), 283–304. <https://dx.doi.org/10.1075/pc.5.2.05tsu>
- Unser, H. (1906). *Über den Rhythmus der deutschen Prosa* [On the rhythm of German prose]. Universitäts-Buchdruckerei J. Hörning.
- van Peer, W. (1990). The measurement of metre: Its cognitive and affective functions. *Poetics*, 19(3), 259–275. [https://dx.doi.org/10.1016/0304-422X\(90\)90023-X](https://dx.doi.org/10.1016/0304-422X(90)90023-X)

- Vaughan-Evans, A., Trefor, R., Jones, L., Lynch, P., Jones, M. W., & Thierry, G. (2016). Implicit detection of poetic harmony by the naïve brain. *Frontiers in Psychology*, 7, 1859. <https://dx.doi.org/10.3389/fpsyg.2016.01859>
- Wagner, P. (2012). Meter specific timing and prominence in German poetry and prose. In O. Niebuhr (Ed.), *Understanding prosody: The role of context, function, and communication* (pp. 219–236). De Gruyter.
- Westbury, C. (2005). Implicit sound symbolism in lexical access: Evidence from an interference task. *Brain and Language*, 93(1), 10–19. <https://dx.doi.org/10.1016/j.bandl.2004.07.006>
- Wheeler, M. A., & Roediger, H. L. (1992). Disparate effects of repeated testing: Reconciling Ballard's (1913) and Bartlett's (1932) results. *Psychological Science*, 3(4), 240–245. <https://dx.doi.org/10.1111/j.1467-9280.1992.tb00036.x>
- Whissell, C. (2002). Emotion conveyed by sound in the poetry of Alfred, Lord Tennyson. *Empirical Studies of the Arts*, 20(2), 137–155. <https://dx.doi.org/10.2190/6K4G-LWPQ-RAY8-67QG>
- Whissell, C. (2003). Readers' opinions of romantic poetry are consistent with emotional measures based on the *Dictionary of Affect in Language*. *Perceptual and Motor Skills*, 96(3), 990–992. <https://dx.doi.org/10.2466/pms.2003.96.3.990>
- Whissell, C. (2011). “To those who feel rather than to those who think”: Sound and emotion in Poe's poetry. *International Journal of English and Literature*, 2(6), 149–156. <https://academicjournals.org/journal/IJEL/article-full-text-pdf/4FAB1671263>
- Wrembel, M. (2009). On hearing colours: Cross-modal associations in vowel perception in a non-synaesthetic population. *Poznań Studies in Contemporary Linguistics*, 45(4), 595–612. <https://dx.doi.org/10.2478/v10010-009-0028-0>
- Xue, S., Lüdtke, J., Sylvester, T., & Jacobs, A. M. (2019). Reading Shakespeare sonnets: Combining quantitative narrative analysis and predictive modeling – an eye tracking study. *Journal of Eye Movement Research*, 12(5). <https://dx.doi.org/10.16910/jemr.12.5.2>
- Yarkho, B. I. (1925). Granitsy nauchnogo literaturovedeniia [The boundaries of the scientific study of literature]. *Iskusstvo*, 2, 45–60.
- Yarkho, B. I. (1969). Metodologiia tochnogo literaturovedeniia: nabrosok plana [Methodology of a precise study of literature: An outline]. *Trudy po Znakovym Sistemam*, 4, 515–526.
- Ziegler, J. C., Tan, L. H., Perry, C., & Montant, M. (2000). Phonology matters: The phonological frequency effect in written Chinese. *Psychological Science*, 11(3), 234–238. <https://dx.doi.org/10.1111/1467-9280.00247>
- Zielinski, T. (1904). *Das Clauselgesetz in Ciceros Reden: Grundzüge einer oratorischen Rhythmik* [The clausula law in Cicero's speeches: Rudiments of a mode of oratory rhythm]. Dieterich'sche Verlagsbuchhandlung.
- Zielinski, T. (1914). *Der constructive Rhythmus in Ciceros Reden: Der oratorischen Rhythmik zweiter Teil* [Constructive rhythm in Cicero's speeches: Of oratory rhythm, part two]. Dieterich'sche Verlagsbuchhandlung.