



---

**BETWEEN THEATRICALITY AND SLOWNESS**  
**Phonetic, Prosodic and Pausal Strategies in the Readings of Yoshihara Sachiko and Arai Takako**

FEDERICO LO IACONO<sup>1</sup> AND EDOARDO OCCHIONERO<sup>2</sup>

<sup>1</sup> Università di Torino/Università di Genova – <sup>2</sup>Università di Torino/Università di Genova

<sup>1</sup>[federico.loiacono@unito.it](mailto:federico.loiacono@unito.it) – <sup>2</sup>[edoardo.occhionero@unito.it](mailto:edoardo.occhionero@unito.it)

**Abstract**

**(EN)** Applying the VIP methodology to Japanese poetry, this study analyses recitations by Yoshihara Sachiko and Arai Takako. We propose a “theatricalizing rhetoric” beyond the traditional speed/slowness binary. Quantitative analysis of acoustic and rhythmic indices reveals that, despite comparable speech rates, the poets diverge structurally: Yoshihara embodies slowness through extended silence, while Arai displays a dense, melodically complex theatricality distinct from simple acceleration.

**KEYWORDS:** Yoshihara Sachiko; Arai Takako; Prosody; Japanese poetic speech; Digital Humanities.

**(ITA)** Applicando la metodologia VIP alla poesia giapponese, lo studio analizza le recitazioni di Yoshihara Sachiko e Arai Takako, proponendo una “retorica teatralizzante” oltre al binomio velocità/lentezza. L'analisi quantitativa degli indici acustici rivela che, a parità di speech rate, le poetesse divergono strutturalmente: Yoshihara incarna la lentezza tramite silenzi estesi; Arai esibisce una teatralità densa e melodicamente complessa, distinta dalla mera accelerazione.

**PAROLE CHIAVE:** Yoshihara Sachiko; Arai Takako; prosodia; Parlato poetico giapponese; Digital Humanities.

**1. Introduction**

In the field of Japanese poetic recitation, Wakui Takashi (1994) identifies two principal rhetorical modes of *modus legendi*: slowness and speed. Since these categories stem from a theoretical and empirical reflection that is not based on a systematic collection of data but rather on an approach closer to literary criticism, they should be understood as conceptual tools functional to the present analysis, which instead adopts a more distinctly multidisciplinary perspective, as the following sections will demonstrate.

The rhetoric of slowness is rooted in the long tradition of oral performance, dating back to public recitations of *waka* and *renga*, as well as *Nō*, where a deliberately slowed rhythm fostered reflection and intensified emotional resonance. This mode of reading was once considered necessary for the text to maintain its generic integrity and to distinguish itself from other expressive forms; reading more slowly than ordinary speech was, at one time, an effective way to underscore its “poeticity” (Wakui 1994: 49).

The rhetoric of speed – often associated with “free verse with speed” (Wakui 1994: 49) – emerged more recently, in conjunction with the neo-avant-garde experiments of the 1920s. Within this context, Hagiwara Kyōjirō (1899–1938) is frequently cited as a precursor, while the performative



practices that developed from the 1960s onward with Yoshimasu Gōzō (b. 1939) offer more fully realized examples of this mode, characterized by a rapid, linear delivery that follows the natural flow of speech up to the physiological limit of breath.

One of the most recognizable manifestations of this rhetoric is the abandonment of traditional lineation in favor of a constant pace, with pauses determined solely by the exhaustion of breath (Wakui 1994: 49–50). This strategy, central to certain readings by Nejime Shōichi (b. 1948), also appears in later performances such as Itō Hiromi’s “*Imi no gyakutai*”<sup>1</sup> (“The Maltreatment of Meaning”) (b. 1955). In light of the above, Yoshihara Sachiko’s (1932–2002) reading practices clearly fall within the rhetoric of slowness.

The primary aim of this study is to propose a third, hybrid rhetorical category: while it may overlap with the two previously identified modes, it distinguishes itself through the integration of additional expressive dimensions. This category, which we term *theatricalizing rhetoric*, combines the rapidity and semantic concentration characteristic of poetic reading with an expansion toward more complex performative parameters. From this perspective, theatricalizing rhetoric encompasses the analysis of vocal features that are markedly expressive and performative, including timbral modulation, tonal variation, and, more broadly, the ensemble of vocal strategies that contribute to the staging of the poetic text. It thus functions as an interpretive framework capable of accounting for a wider spectrum of performative modalities than those captured by the earlier categories.

The need to formulate this third category is justified by the emergence, in contemporary practices of poetic reading, of modes that cannot be fully accommodated within traditional rhetorical models. In this sense, Arai Takako (b. 1966) offers an exemplary case: her vocal practice paradigmatically embodies the dynamics that this new category seeks to describe.

## 2. Description of the Corpus

The oral corpus analyzed in this study was defined primarily on the basis of the availability of online multimedia materials, with particular reliance on publicly accessible audio recordings – especially those circulated through YouTube. This choice, while methodologically pragmatic, necessarily introduces limitations: the recordings are heterogeneous in quality, often lack detailed contextual information, and reflect only a portion of the poets’ broader performance practices. These constraints are acknowledged as inherent to research on contemporary poetic performance, where systematic archival resources remain unevenly developed.

In the case of Yoshihara Sachiko, the corpus draws not only on freely available digital sources but also on materials preserved by her son, Yoshihara Jun. His dispersed and still-expanding archive<sup>2</sup> includes typed manuscripts, journal publications, and, crucially, analog recordings of the poet, which significantly contribute to safeguarding her vocal legacy. Among the four selected texts – “*Ame na no ni*” (“Despite the Rain”), “*Nichibotsu*” (“Sunset”), “*Inori*” (“Prayer”), and “*Jikai*” (“Self-Discipline”) – only the latter two could be precisely contextualized: both were performed during a lecture delivered on 25 January 1984 at Notre Dame Seishin Women’s University in Okayama. For the remaining pieces, the absence of detailed temporal or situational metadata suggests that they were likely recorded during small-scale *rōdokukai* (public readings) or *kinenkai* (annual commemorative gatherings).

The situation differs markedly for Arai Takako. The recordings analyzed here document two events of clear international scope: “*Asa o kudasai*” (“Give Us Morning”) and “*Irogarasu*” (“Coloured Glass”), both included in *Tamashii dansu* (Soul Dance, 2007), were performed respectively at the 2022 Poetry International Festival in Rotterdam and at the second Tokyo Poetry

<sup>1</sup> “The Maltreatment of Meaning”, Poetry International: [https://www.poetryinternational.com/en/poets-poems/poems/poem/103-7849\\_THE-MALTREATMENT-OF-MEANING#lang-org](https://www.poetryinternational.com/en/poets-poems/poems/poem/103-7849_THE-MALTREATMENT-OF-MEANING#lang-org) (20.04.2026).

<sup>2</sup> Yoshihara Sachiko *Ākaibu*: [https://www.facebook.com/SachikoYoshihara1932/?locale=ja\\_JP](https://www.facebook.com/SachikoYoshihara1932/?locale=ja_JP) (20.04.2026).

Festival, held alongside the sixth World Haiku Association Conference in September 2011. Their well-documented context provides a more stable basis for performance analysis.

The selection of four texts by Yoshihara and two by Arai aims to ensure a balanced comparison in terms of total reading time. The numerical disparity reflects both the differing length of their compositions – Yoshihara generally writes within forty lines, whereas Arai often favors long-poem forms – and the duration of the available recordings, which in both cases amount to approximately seven minutes. While necessarily partial, this corpus offers a representative sample of each poet’s vocal practice and provides a workable foundation for the comparative analysis developed in this study.

## 2.1 Yoshihara Sachiko

Yoshihara Sachiko made her debut in 1964 with the collection *Yōnen rentō* (*Childhood Litanies*), a work that bears witness to the hardships she endured during the Second World War. Her early biography is marked by events of profound emotional and experiential impact, including evacuation to Kōmyōji Temple in Yamagata, the news of her father’s death, and her direct exposure to the air raids on Tokyo (Kuninaka 2000).

In 1972 she published *Ondīnu* (*Ondine*), the collection that earned her the Takami Jun Prize and firmly established her within the contemporary poetic landscape. Over the following decade, in collaboration with Shinkawa Kazue (1929–2024), she helped found the journal *Gendaishi – La Mer* (1983–1993), an editorial project that not only played a decisive role in fostering the next generation of women poets but also became central to the redefinition of *joseishi* (women’s poetry).

Yoshihara’s stylistic and textual signature lies in her preference for *kyū-kana zukai*<sup>3</sup> (the historical orthography of Japanese), through which she signals her intention to reconnect with early twentieth-century poetic traditions. Although this archaic script does not differ phonetically or semantically from modern orthography, it produces a striking visual effect that generates a complex layering of expression (Hayes *et al.* 2017). Yoshihara draws on the full linguistic resources of Japanese to articulate multiple levels of meaning, using *kanji* and *furigana* strategically in a metatextual capacity.

From a thematic standpoint, her poetics can be effectively understood through the lens of lyricism, which, as Mizuta (2012) argues, intertwines with a distinctly confessional tone. The result is a poetry sustained by a deep existential polarity: life is set against death; desire against its refusal; love against its negation; and solitude against the inescapable presence of the Other. This dialectical tension manifests in a poetic language that combines conceptual abstraction with rhetorical density, and it is reflected in her versification – often syntactically complete and structured around recurring anaphoric patterns – that lends each stanza a sense of autonomy and formal closure, as can be seen, for example, in “*Nichibotsu*”:

雲が沈む

*Kumo ga shizumu*

そばにゐてほしい

*Soba ni ite hoshii*

鳥が燃える

*Tori ga moeru*

---

<sup>3</sup> The new orthographic transcription system introduced in 1946 aimed to eliminate obsolete *kana* such as ゐ (*wi*), ゑ (*we*), and を (*wo*), even in cases where their pronunciation coincided with い (*i*), え (*e*), and お (*o*). It also sought to abolish the distinction between ぢ (*ji*) and じ (*ji*), as well as づ (*zu*) and ず (*zu*), which are now homophonous in standard Japanese. For historiographical discussions of the Japanese language, see Bowring (2005) and Irwin (2011).

そばにゐてほしい

*Soba ni ite hoshii*

海が逃げる

*Umi ga nigeru*

そばにゐてほしい

*Soba ni ite hoshii*

もうぢき

何もかもがひとつになる

*Mou jiki*

*Nani mo kamo ga hitotsu ni naru*

指がなぞる

匂はない時間のなかで

*Yubi ga nazoru*

死がふるへる

*Nioi wa nai jikan no naka de*

*Shi ga furueru*

蟻が眠る

そばにゐてほしい

*Ari ga nemuru*

風がつまづく

*Soba ni ite hoshii*

そばにゐてほしい

もうぢき

*Kaze ga tsumazuku*

夢が終る

*Soba ni ite hoshii*

何もかもが

黙る

*Mou jiki*

*Yume ga owaru*

*Nani mo ka mo ga*

*damaru<sup>4</sup>*

---

<sup>4</sup> Sunset // The clouds descend / I wish you stood beside me // The birds blaze / I wish you stood beside me // The sea vanishes / I wish you stood beside me // Soon all things / Will become one // The finger retraces / In scentless time / Death trembles // The ants sleep / I wish you stood beside me // The wind stumbles // I wish you stood beside me // Soon / The dream will end // All things / Will fall silent (Unless stated otherwise, all translations from Japanese are the author's own).

(Yoshihara 2021: 64)

Another recurring feature is the specular dynamic that emerges most clearly in the relationship between the poetic “I” and a fictive “you.” The latter, devoid of any discursive autonomy, functions as a reflected projection of the self, made possible by the fundamentally monological structure of the texts. The reader is thus encouraged to perceive an overlap between the two figures, with the “you” operating as a mirror that refracts the author’s inner drama (Ōoka 2021). As a result, the poetic “I,” positioned as the spokesperson for a form of narcissism, develops an acute sensitivity to suffering that drives a gradual assimilation of the other into the self. In this process, the identities of “I” and “you” interpenetrate to the point of erasing any distinction between them, a dynamic sustained by a solipsistic framework: «there exists only one consciousness, and it is one’s own consciousness; therefore, if the other feels pain, the one who feels pain is the speaking “I”» (Azuma 2003, 29).

#### Invocation

Let me go, I beg you  
I contemplate from a stained-glass window  
a tiny *stain* tinged by the shadow on the floor

Love frightens me, kindness frightens me  
The regret of a nauseating thought frightens me  
Do not invite me to life, I beg you  
I will go nowhere, I will not laugh  
Let me go, I beg you  
Here, in this one unmoving place

Do not lay your hand upon my shoulder  
Do not force me to turn around  
Do not rub with your large, benevolent palm  
The *stain* I am contemplating  
(Yoshihara 2021: 53-54)

## 2.2 Arai Takako

With the publication of *Tamashii dansu* (*Soul Dance*, 2007), *Betto to shokki* (*Beds and Looms*, 2013), and *Oshirako-sama kibun* (*The Chronicles of Oshirako-sama*, 2024), Arai Takako moves increasingly toward a poetry that is both «social» (Angles 2016) and «political» (Morton 2023), shaped by an ethnographic and ‘micro-epic’ sensibility. This shift unfolds through the development of a poetic idiolect – referred to by Arai herself as *arataka ben* – that blends three linguistic varieties: *jōshū ben* (Gunma), *kesengo* (Iwate), and selected elements of *tsugaru ben* (Hokkaidō). Yet within this linguistic cosmogony, the hyperrealistic quality of these local speech forms confronts its own status as a constructed or deliberately simulated register: the idiolect takes on imagined contours (Hayes *et al.* 2016; Hayes 2016), distancing itself from the linguistic practices actually used by the communities it evokes. In this way, the ethnographic impulse comes into tension with a dimension of pure poetic invention, understood as an individual and aesthetic process.

*Tamashii dansu*, the work that chronologically inaugurates the process that can now be read as a trilogy – and one of the central objects of analysis in this study – includes a section devoted to Arai’s father’s textile factory in Kiryū, her hometown. This part of the book seeks to reclaim the centrality of women’s labor: the conversations of female workers and the transformation of the surrounding landscape, recalled through childhood memory, converge into a reservoir of chronicles and oral histories that Arai rewrites. Through this process, she constructs a radically innovative poetic framework, sharply at odds with the dominant tendencies of Japanese poetry, which remain anchored

in the expression of an elegiac self. Moreover, through her use of unconventional grammar and punctuation, she subverts standard Japanese linguistic norms (Arai 2025), generating a sense of displacement that becomes a defining feature of her poetic style.

Although certain performative characteristics allow Arai's delivery to be associated with the rhetoric of speed – particularly in passages marked by sudden intensifications and tighter rhythmic patterns than ordinary speech – the linearity of breath is not a defining characteristic of her vocality. A decisive influence, especially in the oralization of her texts, can be traced to the *jōkyō gekijō* (Situation Theater) of Kara Jūrō (1940–2024), which may be understood as an implicit statement of poetics.

In a lecture given at Kokugakuin University, Arai stated:

I was particularly influenced by the comedies and plays of Kara Jūrō. I reflected on how different their content was from modern literature, on their unprecedented unfolding, and on how they offered a concrete representation of the thoughts of artisans and laborers who appeared on stage, as if they were a group of people deeply rooted in the earth. [...] In this kind of theatre, there are moments when people, stepping outside the everyday world, enter another realm, another elsewhere. [...] Instead, acceleration is given to something already underway, and existence and words step outside themselves, overstepping the planned framework. This kind of mysterious moment of transience is the illusion [...]. I go to see Kara Jūrō's works to encounter these things. I feel that I myself write poetry in search of such an 'overstepping'. (Arai 2021: 7)

The foregoing discussion makes clear that Arai's poetic project is not simply a matter of stylistic experimentation, but rather an effort to reach a performative beyond – one in which word and body push past the boundaries of the written text and open onto an ephemeral, unrepeatably elsewhere. From this standpoint, a theatricalizing rhetoric becomes essential to understanding her relationship to poetic speech. Anticipating the analysis that follows, which incorporates quantitative data according to the parameters outlined above, this rhetoric can be observed in an excerpt from “*Asa o kudasai*”:

ズット、浮カンドイル、アタシデス	Zutto, ukande iru, atashi desu
踞マツタ、ママノ、オイラサ	Uzukumatta, mama no, oira sa
ボクデス、右腕ガ、モゲタノハ	Boku desu, migi ude ga, mogeta no wa
会イタイ、ヨ、会イタイ、ヨ	Aitai, yo, aitai, yo
コメカミニ、一発ノタマ	Komekami ni, ippatsu no tama
搔イテ、喉ヲ、苦シンデ	Kaite, nodo o, kurushinde
今ハモウ、沈ンデ、イクダケ	Ima wa mou, shizunde, iku dake
ドーシテ、ナンデ、オレナンダ	Dōshite, nande, ore nan da
吹ッ飛バサレル、爆風ニ	Futtobasareru, bakufū ni
アットイウ間ノ、火ノ、手ガ、クル	Atto iu ma no, hi no, te ga, kuru
モガイテモ土砂、モガイテモ土砂	Mogaite mo dosha, mogaite mo dosha <sup>5</sup>

(Arai 2007: 35)

---

<sup>5</sup> It's me, the girl floating here this whole time / It's me, Mama's boy crouched down / It's me, the boy with the right arm wrenched off / I want to see you again, I want to see you again / A bullet to the temple / I scratch my throat, it hurts / Now I'm sinking as far as I can go / Why? Why was I the boy / Blown aside by the bomb blast? / The fingers of flame came in no time / I struggle but there's only sand, I struggle but there's only sand (Angles 2019: 41).

As noted in the translators' commentary to the English collection *Factory Girl* (Angles 2019), the poem was composed against the backdrop of the Iraq War and the 2004 Indian Ocean tsunami. Its sob-like voice, which conveys the fragmentation experienced by victims, finds a visual counterpart in the use of *katakana* and italics, accompanied by punctuation that interrupts and destabilizes the normative flow of syntax. Within this framework, the methodology of the VIP project – previously applied to Italian poetry – was extended to the analysis of Japanese poetry, bringing its differences and specificities into focus. Rather than concentrating on poetic declaratives, attention was directed toward the rhetorical device known as *tōchihō* 倒置法 (inversion): an expressive technique that deliberately alters word order within the sentence to heighten emphasis, as in the line “*Boku desu, mi-gi ude ga, mogeta no wa,*” whose neutral order would be “*Mi-gi ude ga mogeta no wa boku desu.*”

Another significant element, emerging from the analysis of the second selected text, is Arai's use of *ioriten* 庵点 (ゝ), the grapheme for theatrical reading. In manuscript tradition, this conventional sign functioned as a paratextual marker indicating the beginning of a recited or lyrical section. In the premodern period, it found specific application in *Nō* theatre manuscripts and in *renga*, where it served as a guide for vocal performance.

ゝかいぐり	かいぐり	とつとの目	<i>Kaiguri</i>	<i>kaiguri</i>	<i>totto no me</i>
かいぐり	かいぐり	とつとの目	<i>Kaiguri</i>	<i>kaiguri</i>	<i>totto no me</i>
かいぐり	かいぐり	とつてよ目	<i>Kaiguri</i>	<i>kaiguri</i>	<i>totte yo me</i> <sup>6</sup>

(Arai 2007: 168)

### 3. Problem statement

The experimental study of poetry read aloud has already shown that it can offer significant avenues of inquiry and generate new insights into the prosodic organization of a language. This has been demonstrated not only by the interest of many phoneticians, who have used this register to test intuitions and hypotheses, but also by the need of phonologists to evaluate metrical and textual structures in light of the expected modes of enunciative scanning associated with those forms (Byers 1979; Nord *et al.* 1989; 1990; Lehiste 1990; 1994; Fant *et al.* 1991; Kruckenberg *et al.* 1993; Fant *et al.* 1996; Barney 1999; Bröggelwirth 2005; Madureira 2008; Wagner 2010; 2012; Colonna 2022; 2024; 2025; Barbosa 2023; Wagner *et al.* 2023; Lo Iacono *et al.* 2026a; 2026b; Lo Iacono *et al.* *Forthcoming*). What remains far from settled, however, is the debate on the potential contribution that an experimental description of poetic performance can make to our understanding of authors' stylistic practices and their processes of textual composition (Colonna 2022; Lo Iacono 2024; Lo Iacono *et al.* 2026). Our theoretical position aligns with that of several scholars who have laid the foundations for this field (Tomaševskij 1927). We argue that the study of poetry read aloud can serve both linguists and literary critics. Experimental work of this kind can reveal how far the prosody of a language can extend, where its limits lie, where stylistic variation is most pronounced, which tendencies stabilize into tradition, and what degrees of symmetry or asymmetry emerge between enunciation and composition – and what these can tell us about an author's poetics.

Our primary goal is therefore to capture these metrical and stylistic elements and interpret them through an extensive linguistic and metrical analysis that remains attentive to the authors' poetics. This scientific need has driven the development of an analytical methodology that brings together multidisciplinary expertise and tools. Such an approach allows us to account for the many

---

<sup>6</sup> Sing: Roll your hands round and round pull your eyes flat / Roll your hands round and round pull your eyes flat / Roll your hands round and round pull your eyes out (Angles 2019: 33).

parameters involved in the study of poetry readings in a holistic way, avoiding the fragmentation of structural elements that collectively shape poetic rhythm and language (Beccaria 1975).

#### 4. Methodology

As noted earlier, we aim to identify – and subsequently describe – points of convergence and divergence between composition and enunciation in the writing and reading styles of the selected poets. To do so, we adopted the methodology developed within the *Voices of Italian Poets* (VIP) project (Colonna 2022; Lo Iacono *et al.* 2024), given its versatility, replicability, and adaptability to different linguistic and cultural contexts. This methodology has already been used to conduct experimental analyses on several European languages (Colonna *et al.* 2023a; Colonna 2024; 2025), as well as on invented languages (Colonna *et al.* 2022), and has proven to be an effective tool for capturing convergences and divergences across assorted styles, languages, and cultural settings. This contribution, however, is the first to focus entirely on the reading of poetry aloud in Japanese. It is also the first to apply the VIP methodology to a language that is typologically and phylogenetically distant from those of the Romance and broader European area.

Following a protocol adopted in several previous studies of the VIP group, the six poetic recordings were first automatically aligned with the online tool WebMAUS (Kisler *et al.* 2017), and subsequently manually annotated and corrected across four prosodic tiers using *Praat* (Boersma *et al.* 2024). On the first tier, the rhythmic word (PR), all accentual groups in the recordings were manually identified. The identification of these groups follows exclusively phonetic criteria and is sensitive to the distribution of accentual prominences. These prominences can be interpreted as *accent commands*<sup>7</sup> within superpositional models of intonation – among them the Fujisaki model, which was developed first based on descriptions of Japanese prosody (Fujisaki *et al.* 1969; Fujisaki *et al.* 1984; Hirose *et al.* 1984). The second annotation tier is that of prosodic curves (CP), that is, the interpausal units. Because these units form internally coherent segments of speech, they are considered the central units of the methodology and serve as the basis for extracting all prosodic, acoustic, agogic, and rhythmic indices. The concatenation of multiple prosodic curves linked by the same overarching pragmatic–informational function is termed a *poetic utterance* (EN). This constitutes the third annotation tier. The hierarchy is completed by the tier of verses (VS), where the graphic–textual lines of the poems are annotated. This level is essential in a study that simultaneously addresses both enunciative and compositional aspects.

In addition to these four tiers, all pauses occurring in the performances were also annotated. The distribution of pauses is central to cross-linguistic descriptions of the poetic speech register, and in this study, it proves to be a highly discriminating factor for identifying divergences among the poets' different reading styles. Following the protocol, we used the following labels: short pause <pb> (0.1 s < t ≤ 0.4 s); medium pause <pm> (0.4 s < t ≤ 0.6 s); long pause <pl> (0.6 s < t ≤ 1 s); very long pause <pll> (t > 1 s); and, at stanza boundaries, reset pause <P>.

Table A provides a schematic overview of all the indices extracted for the analysis of the poems read aloud. These indices – reflecting a strongly multidisciplinary foundation – can be viewed collectively and in context through the VIP-Radar (Colonna *et al.* 2023b). This visualization offers not only a highly representative snapshot of each poet's reading style but also enables the immediate comparison of acoustic indices across multiple voices. Compared with the methodology used in the project dedicated to Italian poetry, we replaced one index – the measure indicating the number of declaratives and poetic declaratives – with an index specifically designed to describe and capture the realization of *tōchihō*. This structural feature of Japanese poetry refers to a marked word order deliberately employed to emphasize the message.

---

<sup>7</sup> In cases of *accent sandhi*, we treated as rhythmic words those portions of speech governed by the most prominent command accent.

vs(CP)	CPs that coincide perfectly with a single verse line
CP(vs)	CPs that encompass only a portion of a single verse
CP(vs)CP	CPs that bridge the boundary between two verses, including a portion of the preceding verse and a portion of the following one
vs(CP)vs	CPs that encompass two or more complete verses within a single breath group
Pitchspan	The total extent of fundamental frequency variation across the entire reading, calculated in semitones (st)
Rel_meanpitch	The average fundamental frequency (f0) normalized relative to the speaker's baseline, measured in Hertz (Hz)
Rel_meanI	The average vocal intensity normalized relative to the recording baseline, measured in decibels (dB)
Voice Setting Changes	The frequency of perceptible shifts in vocal register or tonal center, normalized against the total number of CPs. This metric quantifies changes in the speaker's vocal setting
Speech rate	The ratio of total phonetic units to the total duration of CP. Given Japanese phonology, this is calculated both as Syllabic Rate (syllables/sec) and Moraic Rate (morae/sec)
Accelerando	Segments characterized by a perceptually distinct increase in articulation rate relative to the global mean
Trattenuto	Segments characterized by a perceptually distinct decrease in articulation rate (lengthening)
Plenus	The ratio of vocal activity to total duration (Speech / [Speech + Silence]), measuring the density of the speech signal versus the pause component
Focus	The presence of focal prominences employed to highlight specific informational or prosodic units
Inversione	Instances where the canonical word order is deliberately inverted for emphatic or stylistic effect (Japanese hyperbaton)
Interrupt	A measure of disfluency or deliberate fragmentation, manifested as perceptible breaks within a prosodic curve or an irregular/stilted use of pauses
Appoggiato	Borrowed from musical terminology; this index measures the density of rhythmical words within a Prosodic Curve. Calculated as the ratio of PR to CP
Articolato	Borrowed from Baroque musical practice; this index quantifies the segmentation of EN into CP, reflecting the respiratory planning of phrases. Calculated as the ratio of CP to total EN
Synonymia & palilogia	Rhetorical-musical figures indicating repetitive intonation contours. <i>Palilogia</i> refers to repetition on the same pitch level; <i>Synonymia</i> refers to repetition on different pitch levels. Also accounts for melodic <i>variatio</i>
Enjambement	The count of enjambments that are acoustically realized through a silent pause (breaking the syntactic flow at the line end)
Plan	A metric evaluating the structural organization of the poetic reading, calculated as the ratio of total EN to the total number of verses

Table 1  
Acoustic, Agogic, and Rhythmic Indices of the VIP Methodology.

### 5. Phonetic, acoustic, and prosodic analysis

At the core of the experimental analysis, as noted earlier, is the aim of determining whether the rhetoric of speed and the rhetoric of slowness described by Wakui (1994) structurally characterize the enunciative modes of the two poets under examination, in addition to shaping their textual poetic style. Furthermore, we sought to validate the proposed theatrical category by determining whether it exhibits distinct phonetic and prosodic characteristics.

When we look at the indices displayed in the *VJP-Radar*<sup>8</sup> (Figures 1 and 2), it becomes immediately clear that *Speech Rate* alone (see also Figure 3) – the index measuring speaking tempo, and therefore central to identifying timing categories such as those under consideration – does not sharply distinguish the two styles. All readings fall within a narrow range of values, both in terms of syllables per second and morae per second, as shown in Table B. When viewed against the broader landscape of Japanese read speech – specifically the benchmarks established by Ohno *et al.* (1998), who classify rates of 6.6, 7.6, and 8.8 words per second as slow, average, and fast, respectively – the poetry readings in this study are clearly characterized by a reduced tempo (Table B). This reduction indicates a significant temporal dilation of rhythmic segments typical of poetic delivery.<sup>9</sup> However, the most salient distinctions emerge from the *Fluency Rate*, the *Plenus* (the ratio of phonation to pause), and the complex interplay of associated melodic and acoustic indices.

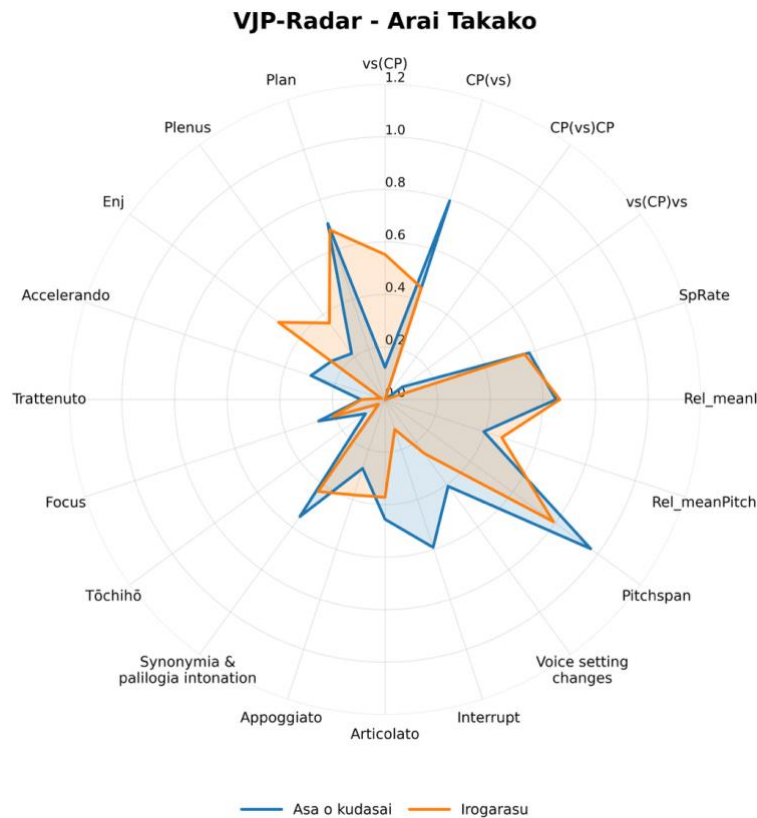


Figure 1. Comprehensive VJP-Radar showing the phonetic profile of Arai Takako’s readings

<sup>8</sup> In this paper, we have adapted the name of the chart from *Voices of Italian Poets* to *Voices of Japanese Poets* for this application to poetry read aloud in Japanese.

<sup>9</sup> This finding aligns with established cross-linguistic and cross-cultural research, which consistently characterizes the poetic register by increased syllabic duration, reduced speech rates, extended pauses, and a marked decrease in phonetic and syllabic reduction phenomena (Lo Iacono *et al.* *Forthcoming*; Lo Iacono *et al.* 2026).

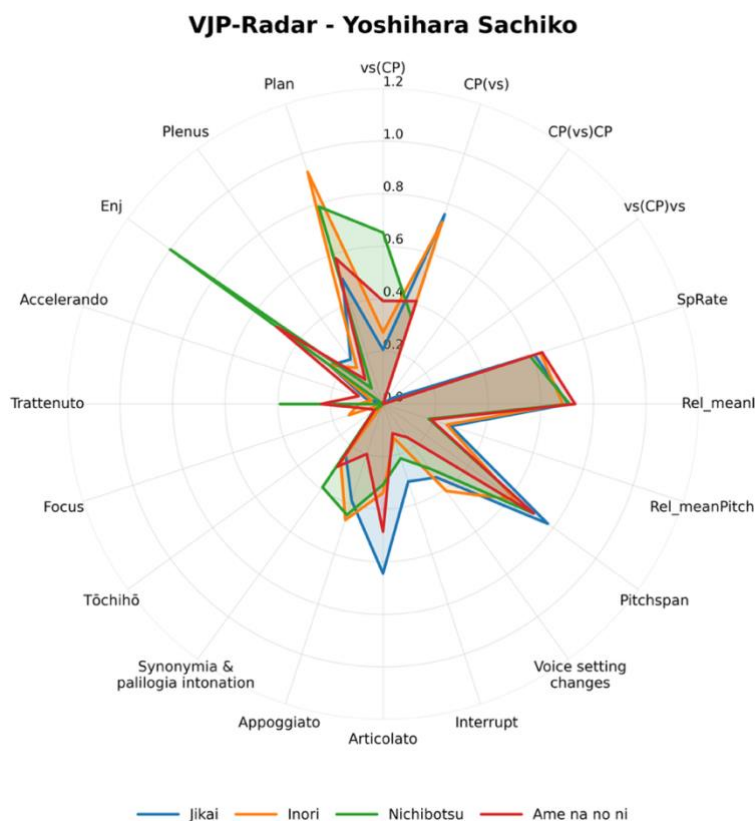


Figure 2.

Comprehensive VJP-Radar showing the phonetic profile of Yoshihara Sachiko’s readings.

Poet	Poem	Syll. Rate	Mor. Rate	Ohno <i>et al.</i> (1998) classification
Arai Takako	<i>Asa o kudasai</i>	5,51 syll/s	6,61 mr/s	Slow rate reading
Arai Takako	<i>Irogarasu</i>	5,36 syll/s	6,54 mr/s	Slow rate reading
Yoshihara Sachiko	<i>Jikai</i>	6,05 syll/s	6,64 mr/s	Slow rate reading
Yoshihara Sachiko	<i>Inori</i>	6,20 syll/s	7,41 mr/s	Medium rate reading
Yoshihara Sachiko	<i>Nichibotsu</i>	6,18 syll/s	6,75 mr/s	Slow rate reading
Yoshihara Sachiko	<i>Ame na no ni</i>	6,40 syll/s	7,39 mr/s	Medium rate reading

Table 2.

Comparison between Syllabic Rate and Moraic Rate across the analyzed recordings.

Broadly speaking, Arai’s readings are distinguished by a considerable melodic-intonational variability and a wide array of vocal registers. Within the *VJP-Radar* (Figure 1), this versatility is quantified by a broad *Pitchspan* (0.97 in “*Asa o kudasai*” and 0.79 in “*Irogarasu*”), correlated with elevated *Rel\_meanPitch* values (0.40 and 0.47, respectively). Furthermore, the substantial indices for *Voice setting changes* (0.41 and 0.25) provide a measurable correlate for the perceptually salient shifts in register and phonation modes. Consequently, when voice quality parameters are factored in, the interpretations reveal a highly complex and variegated vocal profile. Both “*Asa o kudasai*” and “*Irogarasu*” exhibit a markedly dynamic, almost theatrical contour, traversing a broad vocal register

continuum that suggests a flexible manipulation of strategies and a substantial excursion between laryngeal tension and relaxation.<sup>10</sup>

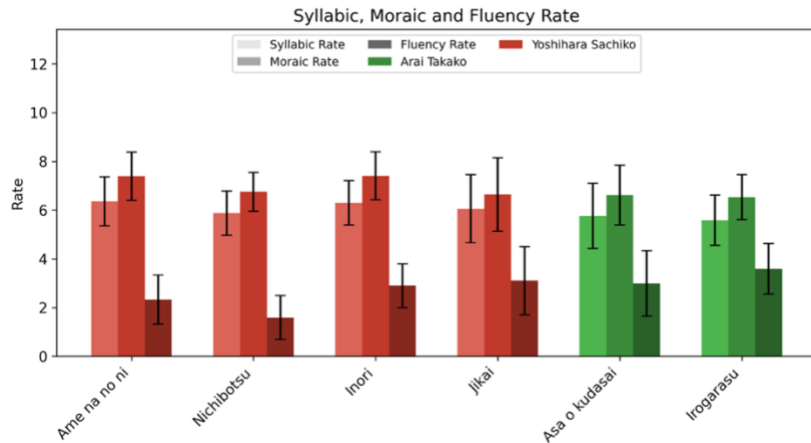


Figure 3. Syllabic, Moraic, and Fluency Rates, sorted by chromatic intensity.

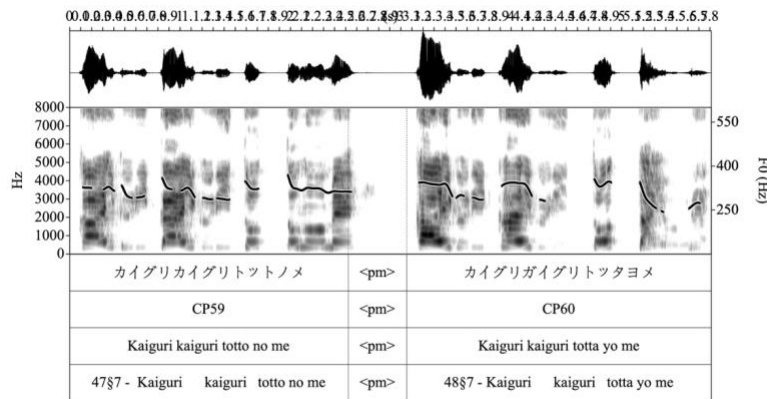


Figure 4.

Praat analysis window displaying an instance of *synonymia* between CP 59 and CP 60 in “Irogarasu”.<sup>11</sup>

The *Synonymia & Palilogia* indices are the highest in the sample (0.55 in *Asa o kudasai*; 0.43 in *Irogarasu*), confirming a marked use of chant-like cadences characterized by repetitive yet modulated intonation structures. These include tripartite structures with *variatio* (vv. 33–35) and bipartite forms (vv. 47 and 48, see Figure 4), which generate rhetorical redundancy and strong intra-textual correspondences. The theatricality of Arai’s performance is substantiated by the stylistic prolongation of plosive closures; a technique reflected in a high *Interrupt* value (0.59 in “*Asa o kudasai*”). This articulatory strategy generates a palpable sense of emphasis and suspension. Simultaneously, the *Accelerando* index (0.30 in “*Asa o kudasai*”) suggests Prosodic Curves characterized by urgent, driving momentum, resulting in a performance that is taut, vibrant, and marked by a distinct crescendo.

Regarding the metrical-rhythmic structure, segmentation remains relatively respectful of the verse structure (CP(vs) at 0.80 and 0.45; vs(CP) at 0.12 and 0.55). However, the *Appoggiato* and

<sup>10</sup>Voice Quality indices were extracted using a custom Praat script developed by Chen Gafni (available at: <https://github.com/chengafni/praat?tab=readme-ov-file>, last accessed: 30/01/2026). Due to the suboptimal acoustic quality of the source recordings, quantitative data were analyzed qualitatively. Validity was ensured through a perceptual assessment conducted by two expert phoneticians to corroborate the extracted metrics with the poets’ actual voice quality.

<sup>11</sup>All analysis windows (spectrograms/waveforms) were generated using the script developed by Wendy Elvira-Garcia, available at: [https://github.com/wendylviragarcia/create\\_pictures](https://github.com/wendylviragarcia/create_pictures) (last accessed: 16/10/2025).

*Articolato* indices (ranging 0.28–0.46; see Figure 5) signal a heavily marked scansion of rhythmic units and enunciations within the phrase. The use of focal prominence (*Focus*) is notably more frequent here (0.27; 0.21) than in Yoshihara’s readings, underscoring an emphatic use of melodic excursion. Ultimately, while maintaining a *Speech Rate* comparable to Yoshihara’s, Arai presents higher *Plenus* values (0.22; 0.36; see also Figure 6). This indicates a denser signal ratio: despite occasional stiltedness or fragmentation, the overall delivery is rapid, compact, and structurally cohesive.

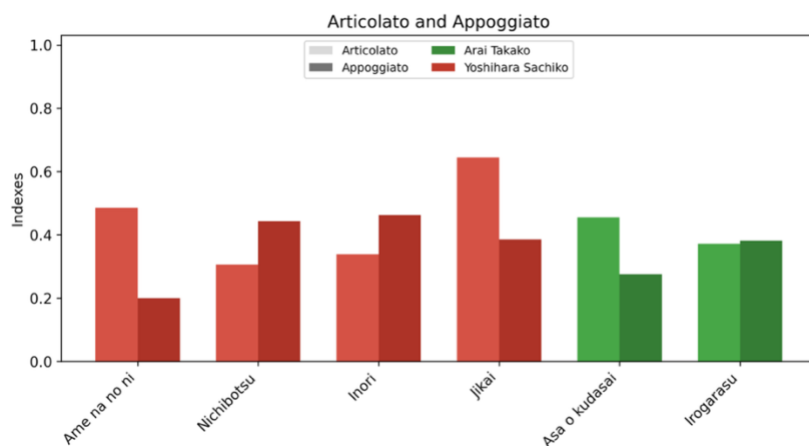


Figure 5.  
Comparative visualization of Articolato (light shading) and Appoggiato (dark shading) indices.

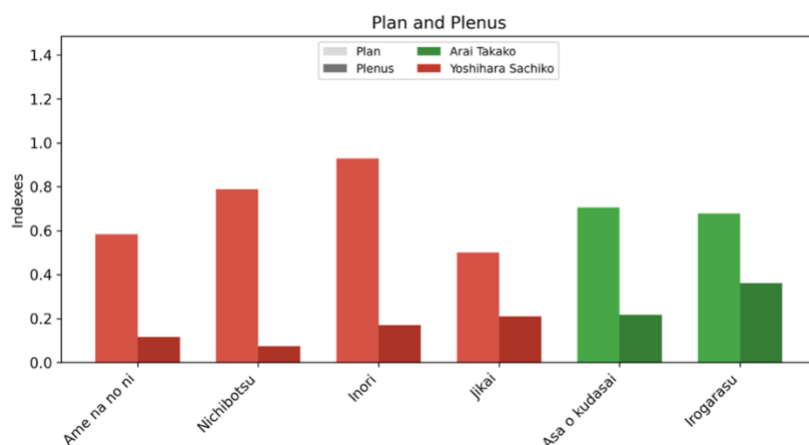


Figure 6.  
Comparative visualization of Plan (light shading) and Plenus (dark shading) indices.

In stark contrast, Yoshihara’s readings are profoundly shaped by the dimension of silence: pause durations frequently exceed those of the prosodic curves themselves, resulting in a markedly slow delivery that operates as a veritable ‘aesthetic of emptiness’. Within the *VJP-Radar* (Figure 2), this prosodic organization is mirrored by systematically lower *Plenus* values (0.21 in “*Jikai*”, 0.17 in “*Inori*”, 0.07 in “*Nichibotsu*”, 0.12 in “*Ame na no ni*”), which quantify the rarefaction of the speech signal. Furthermore, elevated indices of *Trattenuto* – particularly in “*Nichibotsu*” (0.39) and “*Ame na no ni*” (0.24) – signal segments where elocutionary velocity is deliberately suspended. Notably, while Yoshihara’s global *Speech Rate* remains within the same range as Arai’s, her significantly lower *Fluency Rate* indicates that syllables and morae are distributed into short units (CP) punctuated by extensive silence. This fragmentation is corroborated by the *Interrupt* indices (0.31 in “*Jikai*”, 0.22 in “*Nichibotsu*”, 0.12 in “*Ame na no ni*”).

From a distributional perspective, pause duration analysis reveals a heterogeneous, multimodal structure, effectively modeled by three distinct Gaussian components (extracted via

unsupervised probabilistic clustering using *Gaussian Mixture Models*<sup>12</sup>) (Figure 7). The majority of pauses (Cluster 1;  $w = 0.57$ ) cluster around 0.66 s, while the remaining volume is distributed between longer pauses (mean = 1.61 s) and a less frequent tail of extremely long silences (mean: 3.51 s). Consequently, the pause structure is heavily skewed toward extended durations, exceeding norms even for read speech. Disaggregating the data by reader exposes a clear stylistic dichotomy: Arai Takako (gray bars) exhibits a more agitated management of time, dominating the short-pause component with an overall mean of 0.95 s; conversely, Yoshihara Sachiko (green bars) displays a more dilated distribution with a markedly high mean (1.71 s), contributing significantly to the components of extreme duration.

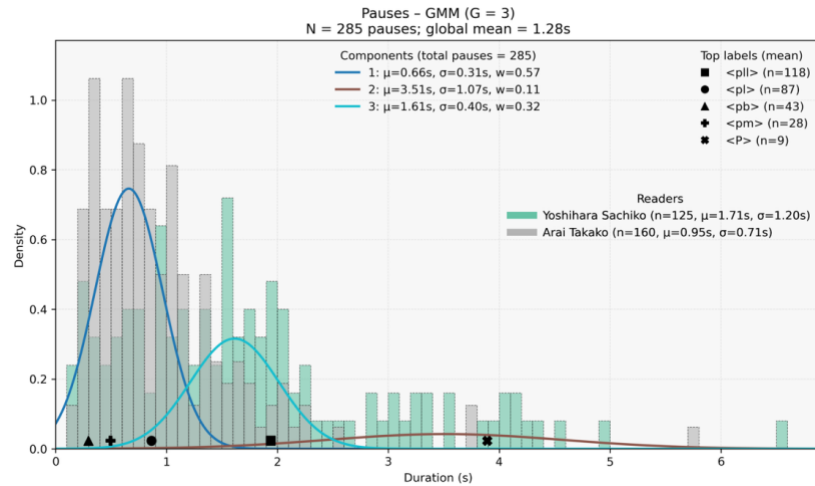


Figure 7.

Pause distribution across the six readings, including occurrences, mean duration, standard deviation, and pause weight. Data is presented both globally and disaggregated by poet. Top right: Manually assigned pause labels according to the VIP methodology.

Furthermore, an analysis of Yoshihara’s pause management reveals significant asymmetries between the textual substrate and the enunciative activity. The realization of acoustic silence does not maintain a mechanical correspondence with the graphic spaces of the text, despite their visual prominence as potential rhythmic cues. Representative examples from the poem “*Ame na no ni*” illustrate this non-isomorphic relationship: while Yoshihara often aligns pauses with graphic spaces (e.g., v. 9; Figure 8), she frequently deviates from the visual template. This is evidenced by the insertion of pauses in the absence of graphic spacing (v. 17; Figure 9) and, conversely, by the suppression of silence despite the presence of a graphic gap (v. 29, between “*ga*” and “*hon no*”; Figure 10). These discrepancies signal a strategy of prosodic planning that operates independently of the *mise en page*.

Regarding metrical organization, Yoshihara exhibits consistent values for both CP(vs) (0.76 in “*Jikai*”, 0.73 in “*Inori*”, 0.35 in “*Nichibotsu*”, 0.41 in “*Ame na no ni*”) and vs(CP) (0.21, 0.27, 0.65, and 0.39, respectively). Notably, the *Plan* index is particularly elevated in “*Inori*” (0.93) and “*Nichibotsu*” (0.79), indicating a prosodic planning strategy that is tightly coordinated with the versification structure (Figure 6). When analyzing the specific indices of *Appoggiato* and *Articolato*, a comparison with Arai reveals generally higher *Articolato* values in Yoshihara’s performance; this suggests a more distended and relaxed management of the relationship between prosodic construction and enunciative delivery (Figure 5). While the *Synonymia & Palilogia* indices (0.24–0.39 range) confirm the presence of intonational repetition figures, they are less systematic than in Arai’s readings. Conversely, the extensive use of enjambment achieved through pause (reaching 1.00 in

<sup>12</sup>Gaussian Mixture Model (GMM) analysis was implemented in a Python environment utilizing the following libraries: Matplotlib (Hunter, 2007) for data visualization; scikit-learn (Pedregosa *et al.* 2011) for the statistical computation of the GMMs; and SciPy (Virtanen *et al.* 2020) for generating the Gaussian curves superimposed onto the histograms.

“*Nichibotsu*”), combined with relatively low *Focus* values (0–0.14), produces a style that aligns with Tada Chimako’s concept of candid lyricism (Tada 2021).

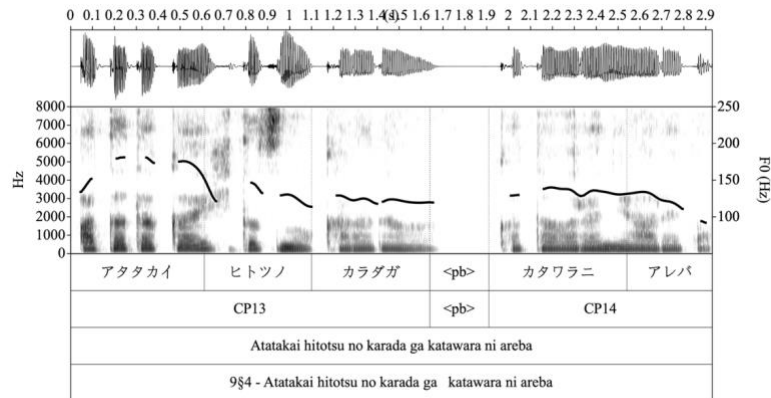


Figure 8.  
Praat analysis window demonstrating alignment between graphic space and pausation in “Ame na no ni”.

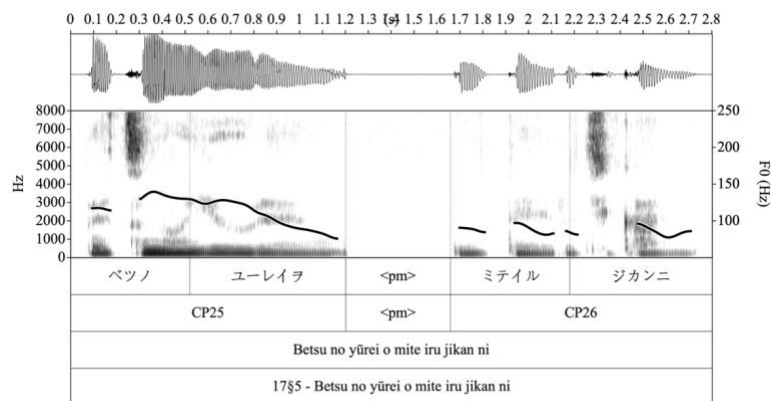


Figure 9.  
Praat analysis window demonstrating a textual-prosodic asymmetry: realization of a pause despite the absence of graphic space (from “Ame na no ni”).

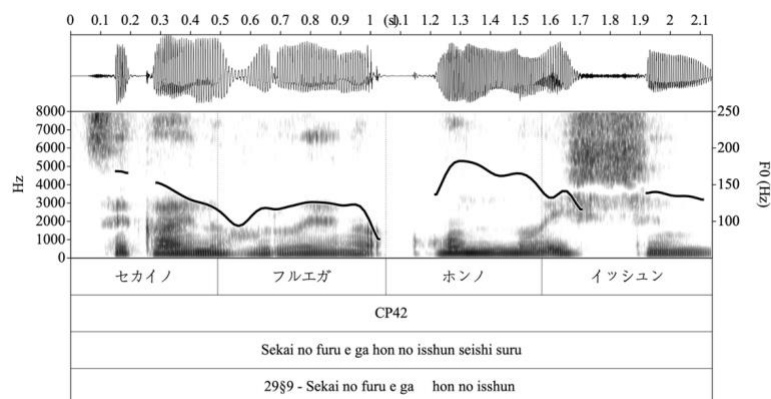


Figure 10.  
Praat analysis window demonstrating a textual-prosodic asymmetry: absence of a pause despite the presence of graphic space (from “Ame na no ni”).

This approach emphasizes a strong respect for the text's implicit prosody without emphatic distortion, while the fragmented use of inter-pausal units (CP) highlights the 'audacity of concise affirmation'. In terms of voice quality, Yoshihara's four readings demonstrate a more stable and contained timbral palette. Her interpretations – particularly *Ame na no ni* and *Nichibotsu* – privilege a 'softened' profile (Gafni *et al.* 2017), acoustically characterized by steep spectral tilt and reduced internal variability. Where Arai relies on contrasts in register and phonation modes, Yoshihara adopts an expressive strategy that is intimate, compressed, and acoustically homogeneous.

## 6. Conclusions

In conclusion, when speech rate is held constant, the combination of *Fluency Rate*, *Plenus*, pause management, and acoustic indices shows that Arai tends to move toward a rhetoric of speed. However, this tendency does not result in a simple acceleration of the enunciative flow. The prosodic density and the breadth of melodic expansion that characterize her readings reveal a more complex expressive control, which allows us to place her within the theatricalizing rhetoric. In other words, her rapidity is never neutral; it becomes a performative device that intensifies the presence of the voice and the construction of her poetic imaginary. Yoshihara, by contrast, more faithfully embodies a rhetoric of slowness, grounded in emptiness, shorter prosodic curves, and structural silences that redefine the temporality of poetic reading. In her case, the rarefaction of the vocal gesture is not a mere slowing down, but an organizing principle that reshapes the perception of the utterance, modulating expectation, suspension, and the informational value of pauses.

The comparative analysis of the two poets thus highlights the limits of the two traditional rhetorical categories, which prove insufficient to account for the complexity of contemporary performative practices. It is precisely within this gap that the theatricalizing rhetoric finds its theoretical justification: it enables us to grasp the ways in which the poetic voice expands beyond the slowness/speed dichotomy, integrating corporeal, timbral, and scenic dimensions that redefine the very nature of poetic reading. The theatricalizing rhetoric does not necessarily entail a spectacularization of the text; rather, it denotes an intensification of vocal parameters that transforms reading into a complex performative act capable of generating a stratified perceptual field. From this perspective, Arai represents a paradigmatic case: her vocality does not merely modulate rhythm or speed but constructs a genuine performative space in which the voice functions as a vector of meaning, affect, and presence. Yoshihara, conversely, demonstrates how slowness can be reinterpreted not as a residual trace of tradition but as a conscious aesthetic choice that reshapes the temporality of listening. Together, the two poets delineate a performative landscape that demands more articulated analytical tools capable of capturing the complexity of contemporary vocal practices.

More broadly, the results of this study have theoretical implications that extend beyond the description of two individual reading styles. The analysis of poetic speech can contribute not only to a more precise understanding of poetry as an oral and performative practice, but also to the linguistic study of read speech in general. Because poetic reading often pushes prosodic organization toward its expressive limits (Lehiste 1990, 1994), it offers a privileged domain for observing how far style can modify the temporal, melodic, pausal, and phonatory dimensions of language. In this sense, the study of poetic performance does not concern a marginal or exceptional register alone; rather, it provides evidence for the plasticity of spoken language and for how culturally situated practices can reshape the phonetic and prosodic realization of written texts. This point is particularly relevant in a cross-linguistic perspective. The present work constitutes the first application of the VIP methodology to Japanese, a language typologically and historically distant from the Romance languages for which the method was originally developed. The results suggest that several analytical categories – such as prosodic curves, *Plenus*, pause typology, melodic span, and the relationship between verse and prosodic segmentation – can be fruitfully extended beyond their initial linguistic domain. At the same time, the Japanese data also show that such models cannot simply be transferred

mechanically from one language to another. They must be tested, refined, and, where necessary, adapted to the prosodic, metrical, graphic, and performative specificities of each language and poetic tradition.

The study also has several limitations. The corpus is necessarily small and partly constrained by the availability and quality of archival or online recordings. The selected materials differ in recording context, technical conditions, and degree of documentation, especially in the case of Yoshihara, whose vocal archive remains dispersed and only partially contextualized. Moreover, the analysis focuses on two poets and six readings and therefore does not aim to provide a general model of Japanese poetic recitation. Rather, it offers an exploratory and methodologically controlled case study, designed to test the descriptive potential of a multidimensional phonetic and prosodic approach. Future research should therefore expand the corpus in several directions: by including a larger number of Japanese poets, by comparing multiple readings of the same text, by integrating more homogeneous recording conditions where possible, and by applying the same analytical framework to other poetic traditions. A broader multilingual dataset would make it possible to evaluate which prosodic indices are language-specific, which are genre-specific, and which may instead capture more general tendencies of poetic speech as a cross-cultural and cross-media vocal practice. Such work would also make it possible to investigate more systematically the relationship between textual composition, metrical organization, graphic layout, and oral performance.

Finally, the accumulation of larger and more diverse datasets may have important applications beyond literary and phonetic analysis. More robust descriptions of poetic and stylistically marked speech could contribute to the development of computational models capable of automatically recognizing performative styles, detecting prosodic signatures, and distinguishing between different modes of read speech. In the longer term, such findings may also inform artificial intelligence applications, including more refined Text-to-Speech systems able to model not only linguistic content but also stylistic, rhetorical, and expressive variation (Delmonte 2015; Wagner *et al.* 2023). From this perspective, the study of poetic speech offers a particularly rich testing ground: it allows us to observe language at the intersection of structure, voice, performance, and cultural practice, and therefore to better understand both the limits and the expressive possibilities of spoken language.

## References

- Angles, J., 2016, “Poetry in an Era of Nuclear Power. Three Poetic Responses to Fukushima”, in Geilhorn B., Iwata-Weickgenannt K. (eds.), *Fukushima and the Arts. Negotiating Nuclear Disaster*, Routledge, London: 144–160.
- Angles, J. (ed.), 2019, *Factory Girls. Selected Poems of Arai Takako* Action Books, Notre Dame.
- Arai, T., 2007, *Tamashii dansu*, Michitani, Tokyo.
- Arai, T., 2022, “Kotoba wa odoru” [Kōenroku], *Kokugakuin zasshi*, 123(10): 1–23.
- Arai, T.; Asabuki, R., 2025, “Dairokukai Ōoka Makoto shō o jushō. Shijin, Saitama daigaku kyōju, shishi ‘Mite’ henshūnin Arai Takako”, *Mita hyōron*, 1299: 62–67.
- Azuma, N., 2003, “Yoshihara Sachiko-ron. Hikaru kizu”, *Otani Joshi Daigaku Kokubun*, 33: 23–33.
- Barbosa P. A., 2023, “The Dance of Pauses in Poetry Declamation”, *Languages*, 8: 76.
- Barney T. H., 1999, “Readers as Text Processors and Performers: A New Formula for Poetic Intonation”, *Discourse Processes*, 28(2): 155–167.
- Beccaria G. L., 1975, *L'autonomia del significante*, Einaudi, Torino.
- Boersma, P., Weenink, D. (2024). *Praat: doing phonetics by computer*. Version 6.4.06.
- Bröggelwirth J., 2005, “A Rhythmic-Prosodic Model of Poetic Speech”, in *Proceedings of Interspeech 2005*, Trancoso I. (ed.), 2397–2400.
- Browning, C., 2005, “Toward a New Word Order: Early Twentieth Century Orthographic Reform in Japan”, *Monumenta Nipponica*, 60(1): 1–35.

- Byers P. P., 1979, “A Formula for Poetic Intonation”, *Poetics*, 8(4): 367–380.
- Colonna V., 2022, *Voices of Italian Poets. Storia e analisi fonetica della lettura della poesia italiana del Novecento*, Edizioni dell’Orso, Alessandria.
- Colonna V., Romano A., 2022, “Reading Poetry in Invented Languages: Prosodic-Syntactic Analysis of Some Cases”, *RiCOGNIZIONI. Rivista di Lingue e Letterature Straniere e Culture Moderne*, 9(18).
- Colonna V., Romano A., 2023a, “The prosody of Seamus Heaney: A phonetic study on some original readings”, in *Proceedings of the 13th Nordic Prosody Conference: Applied and Multimodal Prosody Research*, 238–244, Sciendo/de Gruyter.
- Colonna V., Romano A., 2023b, “VIP – RADAR: A model for the phonetic study of poetry reading”, in Skarnitzl R., Volín J. (eds.), *Proceedings of the 20th International Congress of Phonetic Sciences (ICPhS)*, 594–598, Guarant International.
- Colonna V., Pamies Bertrán A., Damato S., 2024, “Towards a Phonetic History of the Voices of Spanish Poets”, *Journal of Experimental Phonetics*, 33: 7–34.
- Colonna V., 2024, *Voices of Spanish Poets*, risorsa web: <https://voicesofspanishpoets.ugr.es/>.
- Colonna V., 2025, “How Has Poets’ Reading Style Changed? A Phonetic Analysis of the Effects of Historical Phases and Gender on 20th Century Spanish Poetry Reading”, *Languages*, 10(10): 255.
- Colonna V., in press, “Le Modèle «Voices of Italian Poets» Adapté Aux Poètes Français: Yves Bonnefoy, Jean-Pierre Lemaire, Anne Portugal et Jacqueline Risset”, in Lang A., Murat M., Pardo C. (eds.), *Archives sonores de la poésie II*, Les Presses du réel.
- Delmonte R. (2015). “Visualizing poetry with SPARSAR–visual maps from poetic content”, in *Proceedings of the Fourth Workshop on Computational Linguistics for Literature*, 68–78.
- Fant G., Kruckenberg A., Nord L., 1991, “Durational correlates of stress in Swedish, French, and English”, *Journal of Phonetics*, 19(3–4): 351–365.
- Fant G., Kruckenberg A., 1996, “On the Quantal Nature of Speech Timing”, *Proceedings of ICSLP 1996*, 2044–2047.
- Fujisaki H., Nagashima S., 1969, “A Model for the Synthesis of Pitch Contours of Connected Speech”, *Annual Report of the Engineering Research Institute, Faculty of Engineering, University of Tokyo*, 28: 52–60.
- Fujisaki H., Hirose K., 1984, “Analysis of voice fundamental frequency contours for declarative sentences of Japanese”, *Journal of the Acoustical Society of Japan (E)*, 5(4): 233–242.
- Gafni C., Tsur R., 2017, “Softened” Voice Quality in Poetry Reading: An Acoustic Study”, *Style*, 51(4): 456–481.
- Hayes, C., 2016, “Translating Contemporary Japanese Poetry”, *Australian Poetry Journal*, 6(2): 38–57.
- Hayes, C., Kikuchi, R., 2016, “Flared Skirt and The Healds by Arai Takako”, *Transference*, 3(1): 68–76.
- Hayes, C., Kikuchi, R., 2017, “Untitled Nonsense, She, and Contradictions by Yoshihara Sachiko”, *Transference*, 5(1): 59.
- Hirose K., Fujisaki H., Yamaguchi M., 1984, “Synthesis by Rule of Voice Fundamental Frequency Contours of Spoken Japanese from Linguistic Information”, in *Proc. IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, San Diego, CA, USA, 1984, 597–600.
- Hunter J. D., “Matplotlib: A 2D Graphics Environment”, in *Computing in Science & Engineering*, vol. 9, no. 3, pp. 90-95.
- Irwin, M., 2011, “Japanese Loanword Orthography from 1955”, *Yamagata University Faculty of Humanities Annual Report*, 39–57.
- Kisler, T., Reichel U. D., Schiel, F. (2017): Multilingual processing of speech via web services, *Computer Speech & Language*, 45, 2017, 326–347.

- Kruckenbergh A., Fant G., 1993, “Iambic versus Trochaic Patterns in Poetry Reading”, in *Nordic Prosody VI*, 123–135.
- Kuninaka, O., 2000, “Yoshihara Sachiko kaisetsu”, in *Shinkenkyū shiryō gendai Nihon bungaku 7: Shi*, Meiji Shoin, Tokyo.
- Lehiste I., 1990, “Phonetic Investigation of Metrical Structure Orally Produced Poetry”, *Journal of Phonetics*, 18(2): 123–133.
- Lehiste I., 1994, “Poetic Metre, Prominence, and the Perception of Prosody: A Case of Intersection of Art and Science of Spoken Language”, *Proc. 3rd International Conference on Spoken Language Processing (ICSLP 1994)*, 2237–2244.
- Lo Iacono F., 2024, “«Quadrati a profondità ritmiche»: la dimensione orale come chiave interpretativa degli «Spazi metrici» di Amelia Rosselli”, *Finzioni*, 4(8): 49–92.
- Lo Iacono F., Colonna V., Romano A., forthcoming, “Stylistic Vowel Lengthening in Italian Poetic Speech. A Comparative Analysis of 14 Contemporary Italian Poetry Readings”, in *Italo-Romance, Romance and Beyond. Linguistik Aktuell / Linguistics Today*, John Benjamins, Amsterdam.
- Lo Iacono F., Colonna V., Romano A. 2024, “Preservation, conservation and phonetic study of the voices of Italian poets. A study on the seven years of the VIP archive”, *Proceedings of Interspeech 2024*, pp. 3664-3668.
- Lo Iacono F., Romano A., 2026, “A Multi-Dimensional Pipeline for Holistic Prosodic Visualization: Integrating *f<sub>0</sub>*, Intensity, and Syllabic Rate”, *Proceedings of Speech Prosody 2026*, University of Pennsylvania, Philadelphia.
- Lo Iacono F., Romano A., 2026, “Di Montale e del leggere: analisi ritmico-fonetica dei ‘lineamenti fissi’ e delle ‘sillabe storte’ nella prosodia di *Ossi di seppia*”, *Luziana*, Istituti editoriali e poligrafici internazionali, Pisa-Roma.
- Madureira S., 2008, “Reciting a Sonnet: Production Strategies and Perceptual Effects”, *Proc. Speech Prosody 2008*, 697–700, doi: 10.21437/SpeechProsody.2008-154
- Mizuta, N., 2012, *Modanizumu to “sengo joseishi” no tenkai*, Shichōsha, Tokyo.
- Morton, L., 2023, “Contemporary Japanese Poetry and Politics: An Overview”, *Internationale Zeitschrift für Kulturkomparatistik*, 10: 83–109.
- Nord L., Kruckenbergh A., Fant G., 1989, “Timing Studies of Read Prose and Poetry with Parallels in Music: A Research Proposal”, *STL-QPSR*, 30(1): 1–6.
- Nord L., Kruckenbergh A., Fant G., 1990, “Some Timing Studies of Prose, Poetry and Music”, *Speech Communication*, 9: 477–483.
- Ohno S., Fujisaki H., Hara Y., 1998, “On the Effects of Speech Rate upon Parameters of the Command-Response Model for the Fundamental Frequency Contours of Speech”, *Proc. 5th International Conference on Spoken Language Processing (ICSLP 1998)*, paper 0936.
- Ōoka, M., 2021, “Yoshihara Sachiko no shi”, in Yoshihara, Sachiko, *Shishū*, Shichōsha, Tokyo: 143.
- Pedregosa F., Varoquaux G., Gramfort A., et al., 2011. “Scikit-learn: Machine Learning in Python”, in *J. Mach. Learn. Res.* 12, 2825–2830.
- Tada, C., 2021, “Chōritsu sareta nikusei”, in Yoshihara, S., 2021, *Shishū*, vol. 1, Shichōsha, Tokyo.
- Todorov T. (a cura di), 1968, *I formalisti russi. Teoria della letteratura e metodo critico*, Einaudi, Torino.
- Virtanen, P., Gommers, R., Oliphant, T. E., Haberland, M., Reddy, T., Cournapeau, D., Burovski, E., Peterson, P., Weckesser, W., Bright, J., van der Walt, S. J., Brett, M., Wilson, J., Millman, K. J., Mayorov, N., Nelson, A. R. J., Jones, E., Kern, R., Larson, E., Carey, C. J., ... SciPy 1.0 Contributors, 2020. SciPy 1.0: fundamental algorithms for scientific computing in Python”. *Nature methods*, 17(3), 261–272.
- Wagner P., 2010, “Two Sides of the Same Coin? Investigating Iambic and Trochaic Timing and Prominence in German Poetry”, *Proceedings of Speech Prosody 2010*, Hasegawa-Johnson M. (ed.), paper 324: 1–4.

- Wagner P., 2012, “Meter-Specific Timing and Prominence in German Poetry and Prose”, in Niebuhr O. (ed.), *Understanding Prosodies: The Role of Context, Function and Communication*, De Gruyter, Berlin/New York: 219–236.
- Wagner P., Betz S., 2023, “Effects of Meter, Genre and Experience on Pausing, Lengthening and Prosodic Phrasing in German Poetry Reading”, *Proceedings of Interspeech 2023*, Harte N., Carson-Berndsen J., Jones G. (eds.), 2538–2542.
- Wakui, T., 1994, *Prosody, Diction, and Lyricism in Modern Japanese Poetry (shi)*, Ph.D. dissertation, Columbia University.
- Yoshihara, S., 2021, *Shishū*, vol. 1, Shichōsha, Tokyo.