

Kakinuma, Toshie: «Jōji Yuasa and the Apparatus Called Language», in: *Ongaku Geijutsu*, Issue 11, November 1984, pp. 105–109.

Abstract Part II

Part II explores how Yuasa's groundbreaking composition *Voices Coming* redefined the relationship between voice, sound, and meaning.

Yuasa used recorded human speech rather than newly produced sounds, emphasizing how voice inherently carries meaning. Through fragmentation and editing, he suspended meaning and broke down normal communication, turning language itself into musical material. Dismantling question–answer structures and expose the emptiness of habitual speech. Kakinuma compares Yuasa's approach to Luciano Berio and John Cage, highlighting Yuasa's distinct focus on the phenomenology of language. Over time, Yuasa expanded these ideas into live and instrumental works, maintaining speech-like rhythms even without actual words. The essay concludes that *Voices Coming* forms the philosophical and artistic foundation of Yuasa's career, revealing how fragile and yet vital communication is in the modern world.

*Voices Coming*¹ by Yūji Yuasa was a startling work that raised many questions. When one listens to Jōji Yuasa's *Voices Coming* (1969), one is immediately struck by the radicality with which the *voice* itself is treated as a sound object [*onkyō sozai*]. The recorded fragments of ordinary conversation, stripped of semantic content, become a musical material in which meaning is erased, and only the acoustic trace of the *human voice* [*kotoba = koe*] remains.

Regarding his composition *Question [Toi]*² Yuasa composer wrote as follows:

“Having pondered deeply upon the voice, I came to think that its most important point – or perhaps one of its advantages – in contrast to instruments, is precisely that it can utter

¹ Jōji Yuasa: *Voices Coming* (1969) – Tape composition for recorded voices, ca. 25 minutes, realized at the NHK Electronic Music Studio, Tokyo. The work consists of three parts:

I. *Telephonopathy* – fragments of international telephone calls

II. *Interview* – everyday Japanese speech with interjections and fillers (“ano,” “sono,” etc.)

III. *In Memory of Two Slain Peace Warriors* – overlapping speeches by Inejiro Asanuma (1898–1960 leader of the Japanese Socialist Party) and Martin Luther King Jr. (1929–1968), civil rights activist, use of non-violent resistance and civil disobedience against discrimination.

² Jōji Yuasa: *Question [Toi]*, 1971, for mixed choir, [12-14'], premiered in Tokyo, on March 9, 1971.

Question is structured in six sections, each exploring a distinct aspect of speech, intonation, and communication.

1. “*Un an Un?*” [ンとン?] – questions and replies cross each other without communication.
2. “*Madrigal*” [マドリガル] – exchanges of trivial, everyday questions and answers:
3. “*Routine*” [ルーティーン] – formulaic and emotionless replies, e.g., ritualized “Thank you.”
4. (not separately titled) – a change of speakers and tonal tension; dialogue between new male/female voices.
5. (not separately titled) – a sparse monologue by a nearly silent male speaker.
6. “*Right to Silence*” – an interrogation scene with a male voice demanding: “Name? Address? Domicile?”

meaningful words. As a result of being preoccupied with the voice, I found myself returning once again to this very obvious realization.”

That is to say, *Question [Toi]*, using words by Tanigawa Shuntarō,³ was written, in contrast to *Voices Coming*, on the premise of meaningful words. Nevertheless, in Yuasa’s case, the meaningfulness of language is in fact very subtly conditioned.

The first section – “Un and Un?” – consists of *questions and answers [toi to kotae]* that cross paths but never coincide. To a question such as “How are you?” [*gokigen ikaga*] or “Will mankind perish?” [*jinrui wa horobiru ka?*], the answer is an ambiguous murmur – “n” or “n?” – that never truly affirms or denies. In the second, “Madrigal”, this pattern persists: “What?” – “Anything.” “How many?” – “As many as you like.” Trivial exchanges that dissolve meaning through repetition and indifference. The third section, “Routine”, prescribes fixed, affectless speech, exemplified by ritualized expressions such as “Thank you.” The sixth section “Right to Silence [Mokuhiken]”, features a male voice – “forceful, high-pressure” – asking “Name? Address? Permanent domicile?” It is a one-way communication.

Thus, while the *form* of conversation remains – question and answer – the *content* produces no comprehension [*ryōkai*]. The work focuses instead on the *non-communicative* dimension of language. In this respect, *Toi* places upon the table the habitualized words of everyday life – those utterances that pass through our ears without becoming fully conscious – in just the same way that *Voices Coming* did.

– page 105 –

Toi [Question] might be called a continuation of *Voices Coming*. In *Toi*, too, most of the words are not sung but spoken, and in this respect the connection between the two works cannot be denied.

To obtain the effect of the spoken voice, in the first, third, fifth, and sixth sections, even an approximate pitch is not assigned to the words. The performance is entrusted to the acting of the performers’ voices; only the nuance of speech is indicated. For example, in the first section:

³ Shuntarō Tanikawa (1931–2024), a Japanese poet and translator, wrote the text for Jōji Yuasa’s composition *Question [Toi, 1971]*. It is not a separate literary work but was written specifically as a libretto for Yuasa’s piece, and it remains unpublished outside that musical context. Kakinuma emphasizes that Yuasa used Tanikawa’s *words* not as literary material but as triggers for examining the mechanisms of speech.

`slightly pleading,` `introspective,` `quietly` in the first movement; `gently,` `wearily,` `insistently` in the sixth and so on. In the third section, however, the instruction remains loose: “Although each *question* may have individual variation, the *answers* may be considered fixed and unemotional, as a *routine*.” In the fifth section, there is no direction concerning utterance. The performer must distinguish between “Aka?” and “Aka?” [*red?*]⁴ through *intonation* [*intonēshon*] and *gestural nuance* [*kotoba no shigusa*].

Intonation [*intonēshon*],⁵ Yuasa implies, always lies at the boundary between what is language and what is not, between what is spoken and what is unspoken. Primitive intonation may once have been the origin of incantation or prayer. One might even say that *music* [*ongaku*] is itself a kind of sublimated intonation.

By means of *intonation*, “How are you?” may acquire the same weight as “How should one live?”, and “Will humankind perish?” may sound as carefree as “How’s it going?” And *intonation*, through its gestural quality, inevitably draws the listener in. Through *intonation*, contact between the speaker and the hearer becomes possible. There, centering upon a pole formed by the magnetic force of *intonation*, a kind of pseudo-space or situation – including the composer himself – is established.

In *Voices Coming*, the `series of spaces` [*kūkan no keiretsu*] was transformed into a `sequence of times` [*jikan no keiretsu*]; here, conversely, the recreation of words themselves within the `series of spaces,` that is, within the ordinary series, is intended.

The fictive space of words produced by questions and answers that pass each other without ever reaching true understanding should, through the effect of *intonation*, become even clearer and more forcefully confront us. However, in this work, the *concretization of language* through *intonation* depends entirely upon the performer’s ability. The issue of how Japanese is

⁴ The syllable *Aka* (アカ / 赤) is ambiguous in Japanese. While *aka* most directly means “red” (赤), it also carries other resonances — as in *aka* “grime, residue” (汚れのアカ), or colloquially “Communist, radical” (“red” in a political sense). In Yuasa’s *Toi* (*Question*), Kakinuma highlights the performer’s task of distinguishing *Aka?* as mere phonetic sound from *Aka?* as a word bearing semantic weight (“red?”). The sound *a-ka* thereby oscillates between meaningful language and pure vocal material, exemplifying Yuasa’s exploration of the border between *kotoba* (word) and *koe* (voice).

⁵ Kakinuma treats intonation as a liminal phenomenon between sound and meaning – both pre-linguistic (ritual, prayer) and musical (structured sound).

pronounced and vocalized is a major task shared by all vocalists, yet in this work, owing to its heavy reliance on *intonation*, that issue must again be recognized.

Berio's *Sequenza III*⁶ owed much to the vocal virtuosity of Cathy Berberian, but what *Toi* demands is not virtuosity. Rather, it concerns how one responds to words, not the display of technique

Then, how effective was the *device* [*shikake*] of this work, designed to elicit a living response?

The mechanism that demands vivid *intonation* for utterances such as “N” – replies that must be spoken while anticipating emotionless or perfunctory answers – is rather difficult for the performer. The words, though possessing concreteness, are not sufficiently concrete to acquire lively *intonation*, nor are they abstract enough to spin a smoothly reactive sonic texture.

And probably the composer himself was already aware of this when the piece was completed. In the program note for the regular concert of the Tokyo Mixed Chorus, Yuasa wrote as follows:

“Without being limited to the characteristics of the Japanese language, I tried to approach the work from a global... – page 106 –

perspective and complete it as a composition. As a result, I found myself deeply reflecting once again on the question of the abstraction and concreteness of language.”

After *Toi* [*Question*], Yuasa composed *Utterance* [Jp. *Ataransu*; *Gen-hassei* = *Genesis of Utterance*]⁷ and *Yobikawashi* [*Calling to Each Other*].⁸ *Utterance* combines the component elements of language – vowels and consonants – and pursues the sound itself of language,⁹ while *Yobikawashi* is a kind of theater piece. If the composer immediately set about addressing the new problems posed by *Toi*, then the ‘abstractness of words’ he mentions must have been

⁶ Luciano Berio: *Sequenza III*, Solo female voice (unaccompanied), [9’], 1965.

For Kakinuma *Sequenza III* serves as a foil: where Berio foregrounds the singer’s theatricality and expressive mastery, Yuasa suppresses virtuosity to focus on the system of linguistic exchange itself.

Implicitly contrasts with Berio’s *Sequenza III* and *Visage*: whereas Berberian’s virtuosic vocal display [*meijin-gei*] foregrounds the expressive body, Yuasa’s apparatus suppresses virtuosity, focusing instead on the social mechanism of speech and non-communication.

⁷ Jōji Yuasa: *Utterance* [Jp. *Ataransu*; *Gen-hassei* = *Genesis of Utterance*], for 12 soloists (SATB), ca. 18’, 1971.

⁸ Jōji Yuasa: *Yobikawashi* [Calling to Each Other] for for mixed-choir, [15’], 1973.

⁹ The concept of ‘Genesis of Utterance’ [*原発声* = *Hassei*] – philosophically akin to Merleau-Ponty’s *le langage avant le langage* but grounded in corporeal sound: the *body’s act of uttering* as the condition of language’s emergence.

taken up in *Ataransu*, and the `concreteness of words`¹⁰ in *Yobikawashi*.

Utterance was commissioned for the *Third Min-On Contemporary Choral Music Festival*.¹¹ In contrast to *Voices Coming* or *Toi* – works whose very status as `music` was questioned – this piece provoked almost no such doubts.

Utterance contains neither sentences nor verses nor words. Without basing itself on any text whatsoever, there are only sounds. Unlike Boulez or Berio, Yuasa did not perform the operation of decomposing a text into syllables or phonemes [*onsetsu ya onse*]. This was probably not for the sake of procedural simplification. Rather, the very core of the work lies in not beginning from a text that possesses explicit meaning, but instead returning to a certain prior state of language – to a kind of chaos. Therefore, the text should be something that comes into being *after* the work, not something that exists before it.

“Before the birth of words, in the protoplasmic existence of the human being, only through being subjectively vocalized can it first exist as sound and as music. It was such a meaning that gave birth to *Ataransu*.” said the composer himself.

A language prior to language, words before words – that is, a genesis constituted by sound = voice rather than by words. Here, the concrete operations are guided by the question of how to construct and arrange sounds, how to delineate the phases of sound.

For example, the unification of sound obtained by combining each vowel with the same consonant is one of the frequently employed methods. In the opening section, the pairing of *shi* and *shu*, or the sequence *pa, pi, pu, pe, po*, strikes the ear first and foremost as a presentation of the fricative consonant /ʃ/ and the plosive consonant /p/.

¹⁰ Kakinuma hints that Yuasa, having explored the limit of *pre-linguistic abstraction*, now turned back toward *concreteness*, creating a *mechanism [shikake]* that generates real speech in performance. *Yobikawashi* as `performative apparatus`: Kakinuma interprets the work as a turning-point from the exploration of *pre-linguistic voice [hatsusei no koe]* to the creation of *linguistic situations*.

¹¹ *The Min-On Festival of Contemporary Music* (1969–1990) in conjunction with the *Tokyo International Music Competition*. Yuasa premiered *Utterance* there in 1971.

The following section consists of free combinations of several sounds and several rhythmic patterns; there is no indication of pitch. It takes the form of repetitions of the same sound on arbitrary pitches. That is to say, what is at issue here is solely the voice = timbre and rhythm.

In contrast, in the latter part, where sound groups repeatedly disperse and converge, the generated consonants acquire motional quality and aim toward *plasticity* [*kasosei*]¹². Yet this movement does not head toward the ordering of explicit language. The energies of sound that trace parabolic, oblique, or wave-like figures are abstracted expressions of the stirrings that precede linguistic formation; they do not depict the structured process that leads to linguistic articulation.

The field in which pre-linguistic language = original voice is generated – that is *Utterance* – is a composition of meaningless sounds = voices. Yet why is it that these meaningless sounds = voices, a kind of abstraction, are at the same time felt to possess a strangely real substance? If we look at this in comparison with European works, the reason should become clearer.

Take, for example, Ligeti's *Aventures* and *Nouvelles Aventures*.¹³

— page 107 —

In that *Aventures* does not rely upon any predetermined text, it shares the same condition as *Ataransu*. However, in *Aventures*, the voice, above all, consists in the articulation of phonemes, and can rather be called an *acoustic material* than a voice. Although *Aventures* is a work accompanied by flute, horn, cello, contrabass, and percussion, the instruments and voices are placed, so to speak, on an equal footing – not in a relationship of song and accompaniment. The voice utters chains of phonemes purely as a continuous acoustic flow. A single phoneme has no more concrete substance than a single note in a serial structure.

The voice as sound – that is clearly inscribed in *Aventures* through the exhaustive use of phonemes belonging to the same series and the subtle differentiation between slightly differing phonemes. By contrast, in *Ataransu*, the types of voice used are limited, and each possesses a

¹² *kasosei* [*plasticity*]: Yuasa's term for the dynamic malleability of consonantal movement.

¹³ György Ligeti: *Aventures*, for soprano, alto, baritone and seven instruments (flute, horn, trumpet, trombone, harp, celesta, percussion), [12'], 1962

György Ligeti: *Nouvelles Aventures*, for soprano, alto, baritone and seven instruments (flute, horn, trumpet, trombone, harp, celesta, percussion), [12'], 1962–65.

Kakinuma reads Yuasa's *Utterance* as dialoguing with these works. While Ligeti's phonetic vocalizations embody a European *abstract modernism*, Yuasa's approach re-roots abstraction in a spatial, gestural, almost ritual Japanese sense of *voice before meaning*.

kind of *spatial unity* [*kūkan-sei*]. The sense of substantiality that the voice has in [*Ataransu*] is closely bound to this spatiality of language – that is, the spatiality conferred by the very place where language arises, prior to any explicit signification. Within that, there lies the germ of a *gesture* [*shigusa*], and the process of thorough abstraction is avoided.

“The human voice, when once reduced to sound itself, reveals that various kinds of pronunciation and vocalization become the characteristic existences of sound itself,” the composer has said. To “reduce the human voice to sound itself” may be regarded as an abstraction of words = voice. Yet we must keep in mind that the abstraction here exists within a completely different context from that of abstraction in Europe.

The original voice [*gen hatsusei*] that precedes the segmentation of language is one extreme. Beyond that, there is nothing further – a terminus. Yet this experience was something the composer had to pass through once. And that, in turn, leads to *Yobikawashi* [*Call and Response*].

– page 108 –

The other problem generated by *Question* [*Toi*] was that of the *concreteness of words*. *Calling Each Other* [*Yobikawashi*] attempts to recover this concreteness of words without relying on any explicit directions regarding words or *intonation* [*intonēshon*]. If *Yobikawashi* may be said to be somewhat more successful than *Toi*, this lies not in the fact that it pushed concreteness further, but rather that it improved the *device* [*shikake*] so as to generate even greater concreteness.

In performance, five players draw a figure passing twice through the vertices of a pentagon¹⁴ and perform while moving according to it. Outside the pentagon, four other performers are placed: one reciting the names of constellations, one reading from a telephone directory, one performing vowels, and one performing consonants. At each vertex, the five players call out freely the names of people or animals, places, abstract nouns, and so on; at the center of the pentagon, the frozen gesture “Nn” – which also appeared in the first piece of *Toi* – is again used. Every sort of utterance becomes possible: conversations, greetings, murmurs exchanged in every scene. However, this *device* [*shikake*] is constructed so that no reply or response arises.

¹⁴ Spatial composition (pentagon movement): The performers’ geometric motion recalls happenings and early sound-theatre. Kakinuma sees it as a literalization of ‘voice in space,’ extending *Ataransu*’s ‘language space [*gengo no kūkan*]’ into performative architecture.

Speech is always one-way;¹⁵ it is never `called back.` The words are uttered toward completely different directions, flying through a space of *inner-tactile cosmos* [*naishokkakuteki uchū kūkan*]. Here again, there is dialogue that does not mesh, a fragmented space

Yet the point that differs entirely from *Toi* is that any kind of word may be used for the calling. The performers have become freer and lighter. Words of still greater concreteness acquire the *reality of intonation*. Indeed, if one seeks reality, there is nothing simpler than to rely on the recording of real conversation. But that method had already been tested in *Voices Coming*, and in *Toi* the interest had shifted to the problem of how to construct the *device* [*shikake*] that emits the voice. Persisting in not depending on raw actuality, *Yobikawashi* is a work = apparatus that seeks to obtain the reality of words = voice through the mechanism itself.

This *device* [*shikake*] is a machine equipped with the function of producing spoken language = voice. (Here, the composer, so to speak, attempts to efface himself as a *maker of machines*.) In contrast to *Voices Coming*, which was a work constructed out of spoken language as material, *Yobikawashi* became an *apparatus* that creates spoken language. Hence, here a reversal occurs: instead of *a work made of words = voice*, there emerges *words = voice produced by a work = apparatus*.

The work comes into being through the present act of utterance – through the fact that words = voice are uttered *now*, in the present. In *Voices Coming*, the voices existed prior to the work; in *Toi*, the voices were emitted as a result. Here, speech exists in the present moment, and in that present moment the work itself comes into existence.

However, the reality, presentness, and concreteness of the words exchanged in *Yobikawashi* also constituted the pitfall of this work. Depending on the performer's attitude, the chosen words might remain at a level of concreteness or might reach toward a certain abstraction. One may remain within presently existing utterance, yet it is also possible to obtain new words = voices untainted by habitual usage. How deeply those words, which have dissolved into

¹⁵ The one-way nature of speech (*ichihō tsūkō no komyunikēshon*) recalls post-war Japanese reflections on alienation (e.g., Kojin Karasawa, Shūzō Kuki). For Kakinuma, Yuasa's `apparatus' enacts the impossibility of mutual understanding within modern linguistic experience.

everyday unconsciousness, penetrate the listener depends upon the stance of both performer and listener.

Such works inevitably entail risk: they may lapse into tepid depictions of daily life, or they may turn into sharp denunciations. It is possible to communicate precisely that which cannot be communicated; yet, depending on the case, one may also remain without communicating at all.

— page 109 —

(To be continued)

Kakinuma Toshie, music critic

Editorial note:

This text is a transcription of the original Japanese article by Toshie Kakinuma and does not constitute a full translation. It was transcribed, rendered, and edited by Dorothea Schürch using Google Drive OCR, DeepL, and ChatGPT. Source: Toshie Kakinuma, “Jōji Yuasa and the Apparatus Called Language” [湯浅譲二と言語という装置], *Ongaku Geijutsu 音楽芸術*, no. 10 (Oct. 1984), pp. 105–109. Abstract and footnotes by Dorothea Schürch.