V. New Music and Beyond: Music-Historical and Cultural Entanglements

This chapter focuses on three intercultural contexts in rather loose sequence (though arranged roughly chronologically) that may share some overarching themes and lines of development. Beside the inevitable confrontation with cultural essentialism, questions of political and social function and communicability of new music are foregrounded along with aspects that have not yet been discussed sufficiently or in detail, including the relevance of voice, body, and movement as well as questions of perception and metaphorical hearing. Deviating from the emphasis placed elsewhere on the East Asian space, problems of the conception and compositional reception of African rhythmic structures against the background of systematic and historical aspects of polymeter are outlined in part 2 of the chapter, proceeding from György Ligeti's much-discussed Piano Concerto. All three parts might also be understood as case studies for an entangled transnational music history as discussed in Chapter II.1.

1. The Rediscovery of Presence: Intercultural Passages Through Vocal Spaces Between Speech and Song

The Voice in Twentieth- and Twenty-First Century Music

Owing to its mythical and magical implications and its close relation to speech and language, the human voice has always been a crucial hinge between musical and social, spiritual, religious experience. It has also been a medium of sound that from the beginning was considered highly capable of communicating meanings, ideas, and ideologies. Researchers of evolutionary musicology and linguistic anthropologists have suggested that human language originated from symbolic patterns of vocalization and singing. A naive genealogy that attributes speech and singing to a common origin has become questionable at least since Jacques Derrida's deconstruction of Jean-Jacques Rousseau's *Essai sur l'origine des langues*, arguing that speaking and singing have always already deviated from each other.[1]

1 Derrida, *Of Grammatology*, 199. Rousseau's *Essai* (written in 1755, published posthumously in 1781) argues that speech and song would have been nothing more than language itself in a prehistoric era (see Rousseau, "Essay on the Origin of Languages"). Rousseau intended to prove that melody, not harmony, should claim priority in terms

If one considers anthropological theses about the origin of language through the formation of semiotic patterns in archaic vocal music, however,[2] it also becomes apparent that, on the one hand, a complete "desemanticization" of the human voice, as was sometimes rigorously sought in twentieth-century music, soon reaches its limits: speaking and singing, even non-verbal forms, generate meaning. On the other hand, it is also clear that the apparently "most neutral" forms of organization of language, such as the standardized voices of newsreaders, contain a subversive, sonically autonomous element that cannot be grasped using structural-syntactic language models. Nevertheless, in musical and linguistic environments – which often tend to stress this autonomy of vocal sonority – the voice is embedded in structural and cultural frameworks or codes that may both restrict and expand this sonorous presence and thus influence or multiply its connotations. These meanings are usually more ambiguous in performed language and vocal musical structures than in everyday language, and often intentionally so.

If the level of musical globalization comes into view regarding such ambiguity, understood as a constantly procedurally shifting constellation of homogenizing and differentiating movements (→ I.2–3), it becomes clear that the voice may reinforce both movements – cultural convergences or hybridizations as well as the articulation of local, regional, or national identities.[3] Furthermore, it has become clear in the extensive studies on the aesthetics and cultural history of the voice in recent years[4] that collective models of voice use and articulation are reflected in a highly complex manner in individual voices through personal appropriation or musical training, so that it seems impossible to separate or isolate them. Indeed, it is precisely the pluralistic variety of vocal characters and techniques that enables storytellers, singers of traditional narrative genres, opera singers, avant-garde vocalists, or theater actors to switch between different moods and characters in the shortest possible time. In contrast to such vocal "role play," deep psychological approaches aim at a "liberation" of the voice from the formalized "masks" and conventions of the daily use of voice, an approach pursued by the vocal pioneers Alfred Wolfsohn and Roy Hart[5] as well as in the musical avant-garde since the 1960s by Dieter Schnebel, Peter Maxwell Davies, or Jani Christou.

Maybe it was precisely this tendency toward more or less subconscious articulation of a not only culturally but also historically "charged" sonority of human singing voices that, in twentieth-century music, temporarily led to a widespread eschewal of the vocal element – a point Dieter Schnebel made rather clear:

---

of music and music theory, since it imitates the accents of language and, in its own linguistic similarity, would be superior to mere speech (see also Berger, "'Harmonie' und 'mélodie'").

2 "Vocal music might have been the evolutionary laboratory in which early humans developed complex syntactic patterns and a system of multifaceted discrete contrasts that allowed them to attach to these patterns their preexisting symbolic abilities and thus to establish a stable speechlike system." (Richmann, "Did Human Speech Originate in Coordinated Vocal Music?," 244.) This argument has also been taken up in more recent research: "Performative functions associated with oral sound-making provided initial pressures for vocal communication by promoting rank and relationships. These benefits, I suggest, facilitated conflict avoidance and resolution, collaboration, and sharing of needed resources." (Locke, "Rank and Relationships in the Evolution of Spoken Language," 37, Abstract.)

3 See in particular Utz and Lau, *Vocal Music and Contemporary Identities*.

4 See among others Potter, *Vocal Authority*, Meyer-Kalkus, *Stimme und Sprechkünste im 20. Jahrhundert*, Kolesch and Krämer, *Stimme*, Kittler, Macho, and Weigel, *Zwischen Rauschen und Offenbarung*, Felderer, *Phonorama*, Dolar, *A Voice and Nothing More*, Cavarero, *For More than One Voice*, and Anhalt, *Alternative Voices*.

5 See Peters, *Wege zur Stimme*.

> Serial music was conceived as an abstract music, like abstract painting, nonrepresentational painting. We [the composers] also worked toward this. There were constructions and certain forms that we simply rejected. We didn't want to write symphonies. We didn't want to write operas. Vocal music also was almost taboo, because vocal music always has a text, and the text provides an emotional plan. The text has its own progression of sound, but we wanted to compose the sounds themselves. For this reason we turned to the phonetic script, because we had the possibility to notate minute differentiations within the [vocal] sounds.[6]

Where the voice did appear, it was sometimes approached radically in an instrumental idiom, perhaps oriented toward some of Anton Webern's vocal works. This did not preclude, in works such as Pierre Boulez's *Le marteau sans maître* (1953–57) or Luigi Nono's *Il canto sospeso* (1955–56), a specific expressivity resulting from an unresolved friction between the musical structure and the vocal-physical sonorities. On the subject of his magnum opus *Das Mädchen mit den Schwefelhölzern* (The Little Match Girl, 1990–96, → IV.2), Helmut Lachenmann (referring to his own œuvre, but also making a general point) (self-)critically discussed this tendency toward an "exclusion" of the voice:

> I have always paid special attention to familiar musical genres – in a paradigmatically modified context, so to speak. Of course, this also applies to answering the question: "How do you feel about singing?" To this day, this question has remained traumatic for me. There is something wrong with a notion of music that avoids the voice, or even shuts out singing. […] [My opera *The Little Match Girl*] can not least be attributed to an – unfinished – examination of the singing voice […].[7]

Lachenmann has explained further that his compositional practice, developed in the late 1950s and early 1960s, by definition excluded the option of simply "recuperating" established modes of vocal and instrumental expression by employing "recognizable emotional gestures." For Lachenmann, compared to instrumental music, vocal music particularly relies on such gestures since "singing without preconceived emphasis is impossible;" a "singer cannot […] deliver their tone from a distance, but must give themselves, identify themselves. That is why the sung tone today, like 500 years ago, is a magical event."[8] In his few vocal works, Lachenmann consistently

---

6 Schnebel, personal communication, 25/05/2006, quoted in Gee, "The Notation and Use of the Voice in Non-semantic Contexts," 180 ("Die serielle Musik sollte eine abstrakte Musik sein, wie abstrakte Malerei, gegenstandslose Malerei. Und das haben wir auch angestrebt. Es gab Konstruktionen und bestimmte Formen, die wir einfach abgelehnt haben. Wir wollten keine Symphonien schreiben. Wir wollten keine Opern schreiben. Vokalmusik war auch fast verpönt. Weil Vokalmusik ist ja immer mit Text, und der Text liefert einen emotionalen Ablauf. Der Text hat auch selber einen Klangablauf, aber wir wollten die Klänge selber komponieren. Deswegen sind wir auch auf die phonetische Schrift gekommen, weil wir da die Möglichkeit hatten, die Klänge sehr differenziert zu notieren." Gee, "The Relationship of Non-Semantic Vocal Music to the International Phonetic Alphabet," 281.)

7 Lachenmann, "'Klänge sind Naturereignisse,'" 33 ("So ging es mir immer wieder um die Auseinandersetzung mit den vertrauten Werkgattungen – sozusagen in paradigmatisch verwandeltem Zusammenhang. Zum Beispiel natürlich auch um die Beantwortung der Frage: 'Wie hältst Du's mit dem Gesang?' Bis heute ist für mich diese Frage traumatisch geblieben. Ein Musikbegriff, der der Stimme ausweicht, gar den Gesang aussperrt, bei dem stimmt irgendetwas nicht. […] [*Das Mädchen mit den Schwefelhölzern* verdankt] sich nicht zuletzt der noch nicht beendeten Auseinandersetzung mit der singenden Stimme […].").

8 "Helmut Lachenmann im Gespräch," SFB 2005, 69, quoted in Meyer-Kalkus, "Stimme und Atemsyntax," 105. ("vorab eingeführte Ausdrucksregister der Vokal- und der Instrumentalmusik mit wiedererkennbaren emotionalen

understood the voice as physical materiality that is made to sound on the basis of an anti-rhetorical "breathing syntax"[9] whereby the instrumental approach of the *musique concrète instrumentale* is transferred to the vocal medium.[10]

Whether such developments are traced back to the music-aesthetic primacy of instrumental music in the nineteenth century or are thought of in combination with the radical criticism of conventional ideas of similarity between linguistic and musical elements in the period after the Second World War (→ I.2), it is clear that the voice in works such as Arnold Schoenberg's *Erwartung* (1909) or Dieter Schnebel's *glossolalie* (1959–60, a concept work for which the composer made a fully-notated performance score, *Glossolalie 61*, 1961–65) breaks into situations that threaten the credibility of the language-character of music, and with it the presence of the human voice with a kind of atavistic, eruptive force. Last but not least, the incommensurability of these works can be identified in the way the vocal protagonists assume a new responsibility that significantly encroaches on the monological authority of the composer.[11] In *Erwartung* or *glossolalie*, neither the performer's nor the composer's voice communicates a stable identity any longer; rather, they articulate (deliberately) futile attempts at constructing coherence, which is clearly shown by their non-linear, anti-narrative formal design.

If the voice in European music after 1945 and until Lachenmann was deployed at all, it did not necessarily act as a medium for the articulation of linear and meaningful expressive content that had generally become questionable: an intact musical-vocal representation of texts seemed to be meaningless. Also in music composed outside of Western contexts, the vocal element was often used with great caution from the postwar period on – a turn surely indebted to Cold War aesthetics, in which the enthusiastic endorsement of vocal music during the prewar and wartime periods, for example in Japan (→ II.4), appeared problematic.[12] Moreover, instrumental works seemed easier to convey to a global audience (an audience comprising a majority accustomed to Western music) than vocal music, not least because it allowed the composer to avoid primary language and phonetic problems. Against this background, it is again no coincidence that – contrary to this tendency – in what was perhaps the most important period of Asian music in the twentieth century – the emergence of young Chinese composers since the early 1980s – the vocal element played a key role (→ II.3., III.5).[13] In the space between the vocal

Gesten einfach abzurufen;" "Singen [ist] ohne […] vorgelagerte Emphase unmöglich;" "Ein Sänger kann nicht […] seinen Ton aus der Distanz abliefern, sondern er muß sich selbst dabei geben, muß sich identifizieren. Darum ist der gesungene Ton, heute wie vor 500 Jahren, ein magisches Ereignis.")

9 See ibid.

10 See Hiekel, "Escaped from Paradise?," 164–172.

11 See Abbate, *Unsung Voices*, 11–14.

12 This thesis only applies with certain qualifications, of course, and above all to the composers or works that have become known in the West. An indicator of this may be seen in the fact that the œuvres of the Asian composers best-known to Western audiences hardly include any vocal works at all (as with Tōru Takemitsu), or references to Asian vocal music have been largely suppressed (as in Isang Yun's operas). Yun's Asian-inspired vocal works *Gagok* for voice, guitar, and percussion (1972) and *Memory* for three voices and percussion (1974) are exceptions that prove the rule (see Eikemeier, "Zum Text von Yun Isangs *Memory*," Kabisch, "Klang, Ton, Geräusch, Exkurs," and Utz, *Neue Musik und Interkulturalität*, 243–244). A distance to the vocal medium can also more generally be found in new Korean music over longer periods of time (see Lee, "Reconsidering Traditional Vocal Practices in Contemporary Korean Music"). However, it must be emphasized that Jōji Yuasa (b. 1929) worked at about the same time as Schnebel (1930–2018) on comparable body-oriented and sound-related forms of vocal music, for example in *Voices Coming* (1971), *Toi* (1971), and *Utterance* (1971). See Galliano, *Yōgaku*, 220–221.

13 See, for more detail, Utz, *Neue Musik und Interkulturalität*, 315–316.

modernism represented by Schoenberg's *Pierrot lunaire* (1912) and the almost inexhaustible reservoir of Chinese local and national vocal genres, a "third space" with great potential opened up (see below).

A special characteristic of vocal music has always been the difficulty of notating it. The voice certainly reflects the tension between aural, performative and written, denotative layers of musical practice in a particularly fundamental manner. This tension is equally familiar from ethnomusicology[14] and contemporary music. The minimal variations in pitch, vocal coloring, dynamics, and register changes that constitute an individual vocal sound can hardly ever be recorded in notation without omissions. In new music, this resulted on the one hand in overdetermined forms of notation, which, as in the aesthetics of Brian Ferneyhough, were sometimes linked to the idea of the vocal organ as a "resonator" to help filter out and individualize the notational complexity.[15] This contrasts tendencies of works that have been created in close collaboration between composers and vocal soloists relying on their very specific vocal skills and colors – including vocal "composer-performers." In such collaborations, notation often has a merely supplementary role as a mnemonic device, establishing a kind of "oral tradition" comparable to many vocal genres in traditional music. Even though some of these developments can also be observed in the field of instrumental composition, the highly individual and peculiar vocal colors and possibilities of the singers led in many cases to particularly strongly performer-centered concepts. These include Peter Maxwell Davies's *Eight Songs for a Mad King* (1969), written for Roy Hart's "polyphonic" vocal artistry, a work reflecting the anti-psychiatric debate in the UK during the late 1960s,[16] the large number of works created for, with, and by Cathy Berberian, including John Cage's *Aria* (1958), Luciano Berio's *Sequenza III* (1966), and Berberian's *Stripsody* (1966),[17] as well as Giacinto Scelsi's works for solo voice, especially the *Canti del Capricorno* (1962–72); Michiko Hirayama played a crucial part in the creative process of the *Canti*, also adding connections to traditional Japanese vocal practices.[18] In several cases, such collaborations touch on questions of authorship, since the performers often took on the role of a "co-composer."[19] In the globalized context, such close collaborations between composer and vocal soloist reflect a growing importance of local contexts, which often involve new forms of global connectivity.[20]

14 See Nettl, *The Study of Ethnomusicology*, 74–91.

15 "[P]erformers are no longer expected to function solely as optimally efficient reproducers of imagined sounds; they are also themselves 'resonators' in and through which the initial impetus provided by the score is amplified and modulated in the most varied ways imaginable." (Ferneyhough, "Unity Capsule," 100.)

16 Williams, "Madness in the Music Theatre Works of Peter Maxwell Davies," 83–84.

17 See Herzfeld-Schild, "Studien zu Cathy Berberians 'New Vocality.'"

18 See, among others, Kirchert, "Between Worlds," Hirayama, "'Mach's wie Du es fühlst,'" Tortora, *Voce come soffio, Voce come gesto*, and Hiekel, "Escaped from Paradise?," 161–164. Hirayama, however, emphasized that she had no sound knowledge of traditional Japanese vocal music (Hirayama, "[Michiko Hirayama in Conversation with Jürgen Kanold]," 16–17).

19 See especially Meehan, "Not Just a Pretty Voice."

20 This means, for example, that these collaborations often develop over a longer period of time at a certain location – often privately – but at the same time integrate the experiences of all those involved on the global "stage." See also Tomlinson, *Globalization and Culture*, 1–31.

Articulation and Codification of the Voice: Thresholds and Interactions Between Speech and Song

Analogously to the limitations of vocal notation, it was shown that even a tool designed as universalistically as the International Phonetic Alphabet (IPA), despite its high degree of differentiation, ultimately cannot capture many phonetic structures, for example, where a voice gradually changes between different phonetic types or categories.[21] An interculturally accentuated historiography (→ II.1) that critically addresses the criterion of writing must see a central task precisely in the clarification of such tensions between notation and performance, writing and sound. Indeed, it is the transitional areas between established vocal categories and vocal styles, as reflected in the problem of adequate notation, that have repeatedly become a focal point of particular importance for vocal music of the most varied historical and cultural provenance, especially in the twentieth and twenty-first century. Such passages[22] between speaking and singing voices are traced below using selected case examples from Japanese, Chinese, and European contexts. The exploration of the areas between speaking and singing seeks to question the relationship between voice and musical meaning, whereby structural and sociocultural dimensions of vocal music are interrelated. To this end, the interaction between "articulation" and "codification" is placed at the center of attention. "Articulation" is here understood not only in the technical or phonetic sense, but above all also metaphorically: the human voice "articulates" itself as an independent "agent" in a musical context where the vocal sound forms an autonomous psychophysical presence, while it might at the same time appear as one "voice" among several in the structure of the composition. Similarly, "codification" is understood not only in the obvious sense as a notation of sounding voice(s), but also as a term for those sound-immanent and language-discursive means for the constitution of cultural and historical "codes" in which a vocal sound is received and understood.

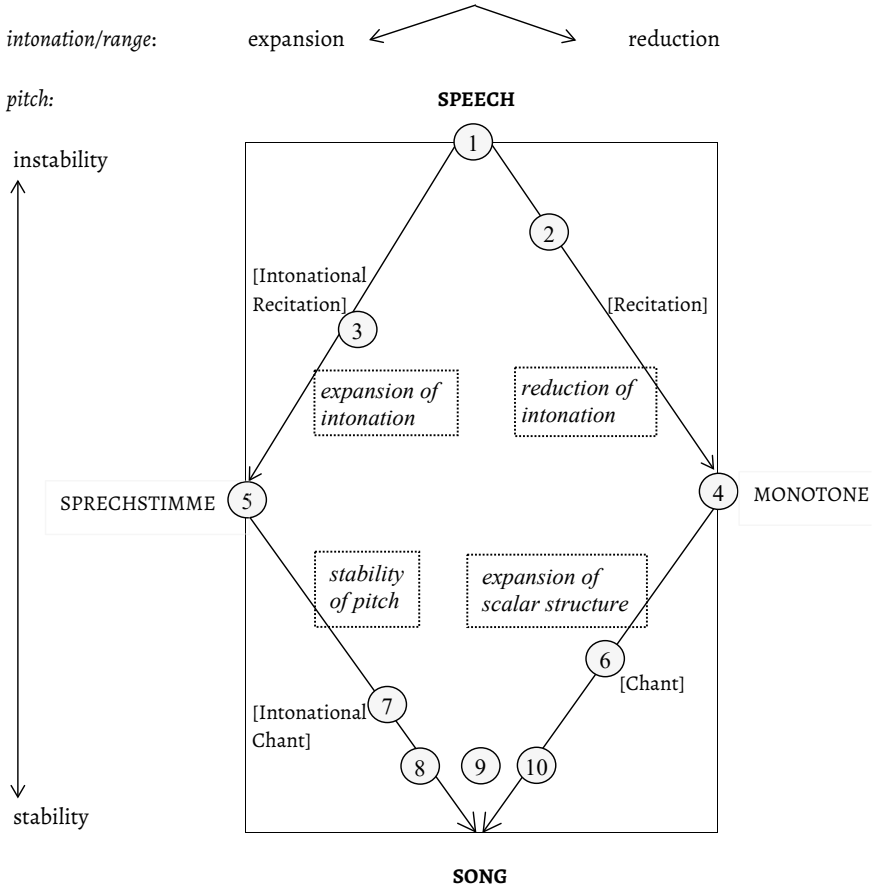
The passages between speaking and singing can be technically described in a variety of ways. Two culturally different "mappings" of this field will be discussed here, some parts of which will serve as a reference in the following analyses. As early as 1963, still somewhat indebted to the tradition of "Comparative Musicology," George List created a map conceived in an unambiguously universalist manner, based on a very wide number of traditional and new singing and speaking styles (Table 5.1).[23] The examples include Western Australian Nyangumata *wangka* speech, Hopi Native American's *laváyi* (speech), *táwi* (song), and *tí:ngava* ("announcing," a chant on two reference pitches), and the New Zealand Maori's *koorero* (speech), *karakia* (ritual chant), *waiata* (song), and *haka* (narrative form with mixed styles). In addition, List discusses forms of heightened speech in advertising speakers, Western spoken drama, and jumping rope rhymes,

21 See Gee, "The Relationship of Non-Semantic Vocal Music to the International Phonetic Alphabet," 265.

22 In a wide-ranging essay, Martin Zenck presented the potential of the term "passage" as it is used here, historically prominently introduced by Walter Benjamin in *The Arcades Project* (*Das Passagen-Werk*). Zenck thematizes topographical-cultural fields (passages between geographic cultures or regions), intercultural fields (passages as interactions between cultures), philosophical fields (passages between physical and metaphysical levels and concepts), and aesthetic fields (passages as crossings and transgressions of boundaries). These fields are also associated with music-specific meanings of the passage concept (transitions as part of musical syntax and form, continuous transition as a principle in new music, etc.). According to Zenck, passages that "lead over thresholds" sometimes approximate transgressions in which borders are violently torn down "because this is the only way to cross them" (Zenck, "Zu einer kulturwissenschaftlichen Theorie der 'Passage,'" 30, "über Schwellen führen," "weil nur so ihre Überschreitung möglich erscheint").

23 List, "The Boundaries of Speech and Song," 9.

Table 5.1: Gradation of the areas between spoken voice and singing by George List with placement of the examples discussed by List. 1: heightened speech (Western drama); 2: jump rope rhymes (USA); 3: Palau women song (Micronesia); 4: type of Thai Buddhist chant; 5: Schoenberg, Pierrot lunaire; Maori: haka; Chinese opera recitative (not specified); 6: tobacco "auctioneering" (USA); 7: Hopi Native Americans (Arizona): announcement; 8: aboriginal Australians (songs); 9: songs of black Americans; 10: bi-, tri-, tetratonic chants (Vedda: children's songs, folk songs)



forms of recitation from Palau (Micronesia), Thailand, songs of black Americans, Schoenberg's *Sprechstimme*, and Chinese opera "recitative." The presupposed assumption of comparability within this sociologically and geographically exceptionally broad spectrum naturally leads to the question of how far the "technical" aspects of the voices gathered here can be isolated from the sociocultural conditions of their production and their emic concepts – a familiar argument from the broad criticism of comparative ethnomusicology since the 1960s. The meaningfulness of this map must certainly be limited to the "technical" dimension, which should not prevent us from using it as an aid to the following analyses. The pitches, which are less profiled in the everyday speaking voice than in the singing voice, gain clearer contours in more stylized declamation or in "half-sung" forms (List places Schoenberg's *Sprechstimme* in this area, center left). In contrast, pitch organization in the form of monotone recitation is increasingly narrowed

Table 5.2: Gradations between speech and song in the four most important vocal genres of traditional Japanese music using Hirano's classification of ***ginshō*****,** ***rōshō*****,** *and* ***eishō***

	speech ⟵	⟶	*song*
	ginshō 吟誦	*rōshō* 朗誦	*eishō* 詠唱
heikyoku **平曲** (*heikebiwa* 平家琵琶)	*yomimono* 読物 ***shirakoe*** **白声 / 素声**	***kudoki*** **口説** *sage* 下ゲ *kōnokoe* 強ノ声 ***hiroi*** **拾** *shojū* 初重 *chūon* 中音	*origoe* 折声 ***sanjū*** **三重** [*kō* 甲-*jō* 上- *kudari* 下り] *sashigoe* 指声 *sage'uta* 下歌 *age'uta (kami'uta)* 上歌
gidayū-bushi **義太夫節** (*jōruri* 浄瑠璃 – style of *bunraku* 文楽)	***kotoba*** **詞** (→ first person) (*sugatari* 素語り)	*kotoba-nori* 詞ノリ ***iro*** **色** *ji iro* 地色 *naka* 中 *kakari* カカリ	***ji*** **地** (→ third person) / *jiai* 地合 *fushi* フシ *jo no kotoba* 序詞
nō **能**	***kotoba*** **詞** *nanori* 名のり, *katari* 語り *tsuki-zerifu* つき台詞 *mondō* 問答/ *kakeai* 掛合	***sashi*** **指** / *kakaru* 繋る *kudoki* 口説 *fumi* 文 *notto* のっと	***tsuyogin*** **強吟** ***yowagin*** **弱吟** *sage-uta* 下歌, *age-uta* 上歌, *kuse* 曲 [introductory songs: *issei* 一声 etc. closing songs: *waka* 和歌, *kiri* 切]
shōmyō **声明**	***sashigoe*** **サシ声**	***kirigoe*** **切り声**	***hikigoe*** **引き声**

down to a single note (center right), often in ritual contexts. Recitative styles are often characterized by repetitions of sound, such as the *parlando* of opera buffa. An increasing stabilization of the pitch or an expansion of scalar structures then leads to forms that are generally referred to as "singing."

Even on a purely technical level, List's map obviously only provides a relatively rough grid. More detailed categories are required to describe specific forms of vocal delivery. In traditional Japanese vocal music, passages between speaking and singing have received special attention and have been differentiated into complex categories within the practice of the individual genres. Table 5.2 seeks to show these categories for the vocal styles from *heikyoku* (epic narrative singing accompanied by the lute *biwa*), *gidayū-bushi* (vocal style accompanying the puppet theater *bunraku*), *nō* (archaic music theater employing male voices exclusively in both solo and choir formats), and *shōmyō* (ritual Buddhist chant) in a representation using an overarching three-level typology based on research by Kenji Hirano on *katarimono*, an umbrella term for language-oriented traditional Japanese vocal genres.[24] Hirano distinguishes between *ginshō* (declamation that moves close to the spoken language), *rōshō* (syllabic recitation at regular time intervals with pitch repetitions on one or a few pitch centers), and *eishō* (recitation tending toward singing, employing changing time intervals, with sustained notes alternating with faster notes of different pitches).

The vocal delivery techniques listed in Table 5.2 are by no means complete, however, and their assignment to the three categories is somewhat arbitrary, especially in *heikyoku* and *gidayū-bushi*, which have an extraordinarily large number of different gradations, although their exact relevance in practice often remains ambivalent and strongly dependent on additional factors such as the aural conventions of the respective schools.[25] After all, in the three transcriptions shown in Example 5.1 (*shirakoe*, *kudoki*, *sanjū*) from the *heikyoku* piece *Suzuki*, one can recognize the three categories *ginshō*, *rōshō*, and *eishō* in a relatively pure form.

The vocal delivery techniques in *heikyoku* are generally referred to as *kyokusetsu*[26] (alteration, change, melodic formula), and in *gidayū-bushi* as *senritsukei*[27] (melody type). Depending on the context, these terms may designate vocal techniques in the narrower sense but also melodic contours, lines, or even certain individual pitches.[28] *Gidayū-bushi* includes particularly numerous transitory styles that are used to switch between the main styles *kotoba* (speech), *iro* (recitation), and *ji* (chant), such as *kotoba-nori*, *ji iro*, or *kakari*. Cadential figures, melodic patterns, and quotations of specific melodies are juxtaposed to these main styles, resulting in a layered complexity that is inwardly connected to the process of oral transmission.[29]

---

24 See Hirano, "Beziehungen zwischen Sprache und Musik im *katarimono*."

25 Updated introductions into the two genres can be found in Komoda, "The Musical Narrative of *The Tale of the Heike*" and Yamada, "*Gidayū-bushi*: Music of the *bunraku* Puppet Theatre."

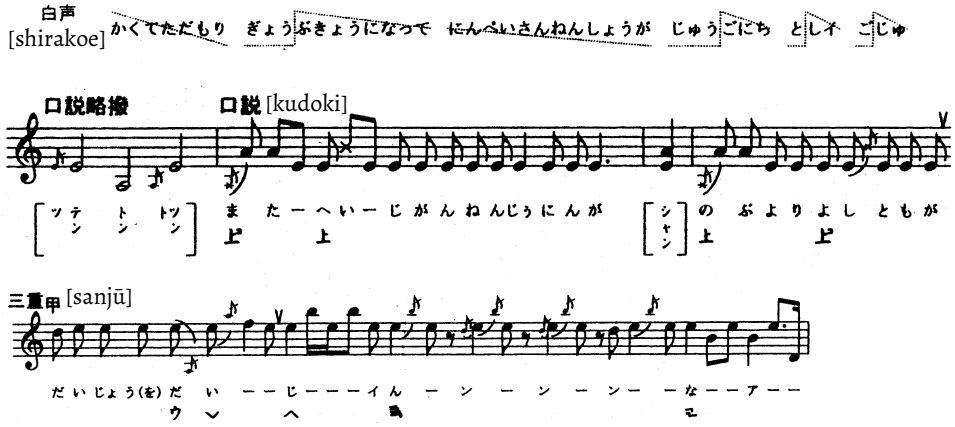
26 Komoda, "The Musical Narrative of *The Tale of the Heike*," 84. Sometimes the term *onsetsu* (syllable) is used instead of *kyokusetsu* (see Mayeda, "Über die *sanjū*-Abschnitte im *heikyoku*," 174).

27 Yamada, "*Gidayū-bushi*: Music of the *bunraku* Puppet Theatre," 205.

28 The use of these terms is not uniform in research and is not limited to the two genres: "[*kyokusetsu*] tends to indicate longer passages of section length, while [*senritsukei*] […] refers to shorter phrases" (Tokita and Hughes, "Context and Change in Japanese Music," 24).

29 "All *katarimono* of pre-modern provenance stem from oral practices in which text and music were produced by techniques of oral composition. Evidence for oral composition in such traditions is demonstrable through analyses developed from the principles of oral-formulaic theory." (Ferranti, "Transmission and Textuality in the Narrative Traditions of Blind *Biwa* Players," 141, note 22.)

Example 5.1: The three vocal delivery techniques shirakoe, kudoki, and sanjū in transcriptions of the heikyoku piece ***Suzuki*** *(performer: Tsutomu Imai; transcription: Komoda, Heike no ongaku, 404–417)*



Hirano particularly highlights the intersections and transitions between the three main categories, which show a tendency toward hybrid formation. For example, a *kotoba* recitation in *gidayū-bushi* can tend to stabilize pitches in the manner of *rōshō* practice, motivated by dramatic, phonetic, or symbolic contexts. The *hiroi* technique in *heikyoku*, in turn, can be extended both to the regions of *eishō* and *ginshō*. Furthermore, as both traditions are transmitted orally, performance styles specific to a certain school and/or individual performers may lead to substantially different interpretations of the same technical terms. We can thus infer that even the most refined typologies do not provide sufficient tools to grasp the specific manner and criteria of transformation between speech and song inherent to many elaborate vocal styles. This obvious shortcoming implies that the interaction and transformation between speech and song must be analyzed comparatively in a variety of specific contexts.

Fragmentation and Montage (*gidayū-bushi*, Tan Dun, Schoenberg)

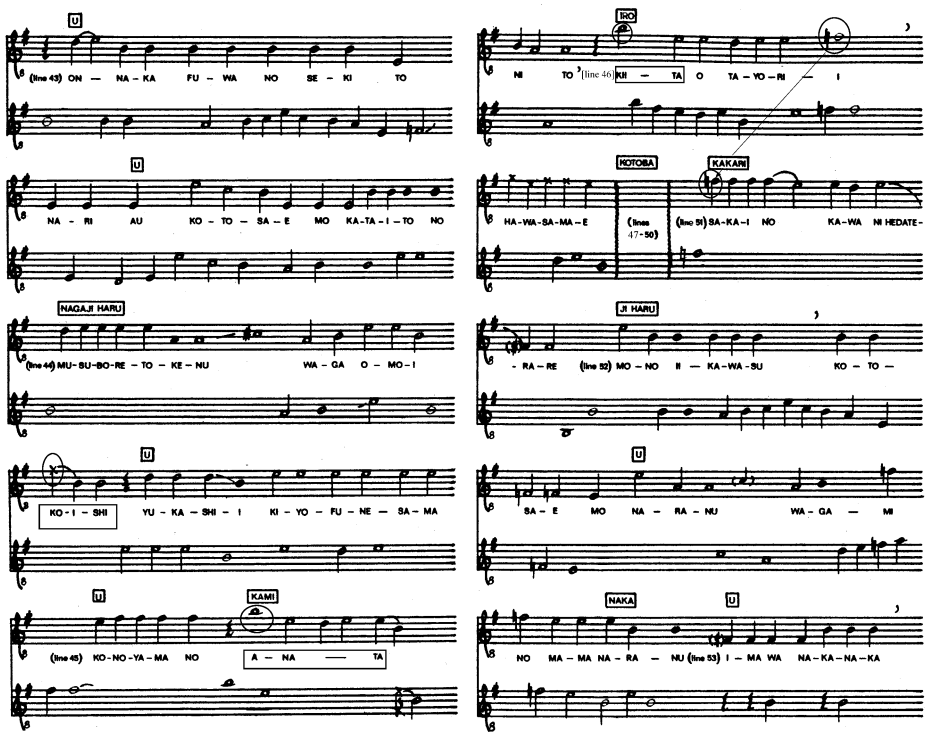
A detailed analysis of a section from *gidayū-bushi* will initially show how its "hybrid formation" of vocalizations is achieved practically. *Gidayū-bushi* is a specific vocal style of the narrative genre *jōruri* used in Japanese *bunraku* puppet theater and developed by Gidayū Takemoto (1651–1714) in Osaka around 1684. The ambivalences and complexity of this style seem to me to allow particularly vivid passages to other types of vocal articulation. *Gidayū-bushi* is performed by a reciter (*tayū*, narrator) and a player of the three-string long-necked *futozao shamisen* (a bass *shamisen*, usually tuned in *honchōshi* with the [relative] pitches B3–E4–B4). The *tayū* acts as the narrator of the story and simultaneously lends his voice to all puppet characters acting on the stage. His vocal timbre is based on a strong chest voice and diaphragmatic breathing and is of a rough and throaty quality, which is guided by the aesthetic ideal *ibushi* (oxidation)[30] and closely linked to the buzzing *sawari* timbre of the *futozao shamisen*. The *shamisen* contributes to the

30 Reese, "*Bunraku*," 148.

expression of the text by playing finely spun heterophonic lines around the reciter's part and bridging gaps between sections with solo melodies (*meriyasu*). The continuing changes in the reciter's vocal character can only inadequately be inferred from the sparse original notation, which merely consists of a few instructions noted down in the *tayū*'s book to the right of the column with the lyrics.[31]

Example 5.2 shows the transcription of a section from the scene *Yama no dan* (Mountain Scene) from *Imoseyama onna teikin* (Mount Imo and Mount Se: A Parable of Female Virtue, 1771) a *bunraku* drama written by Hanji Chikamatsu (1725–1783). This section gives a good understanding of the mosaic-like microstructure that is so characteristic of *gidayū-bushi*, in that the vocal delivery changes with every line of text, or often even within the same line.[32] The lyrical pas-

Example 5.2: Hanji Chikamatsu, ***Imoseyama onna teikin*** *(Mount Imo and Mount Se. A Parable of Female Virtue). Section from the scene* ***Yama no dan*** *(Mountain Scene), based on a performance by Sumitayū Takemoto, recitation, and Kizaemon Nozawa, futozao shamisen (transcription: Malm, "A Musical Analysis of 'The Mountains Scene,'" 78–79; the transcription follows the convention of notating the open strings of the futozao shamisen with the pitches B3–E4–B4; in the recording of the scene attached to Malm's analysis, the pitches are about a minor third lower, so the open strings of the futozao shamisen are at G#3–C#4–G#4*



31 See Yamada, "*Gidayū-bushi*: Music of the *bunraku* Puppet Theatre," 204–205.

32 See Malm, "A Musical Analysis of 'The Mountains Scene'" and Inobe, "Reading 'The Mountains Scene' as Music."

sage belonging to the overarching category *ji* at the beginning of text line 44 (third system on the left) is designated with the term *nagaji*, which symbolizes elegance, and is subdivided by the beginning of the direct speech of the young Hinadori, marked by the word "koishi" ("dearest [missed Koganosuke]"), which is clearly distinguished from the previous song by speaking. The melody then reaches a melodic and emotional climax with the B4 (*kami*) on the word "anata" ("you"). The top note, B4, returns at the beginning of the following line on the word "kiita" ("I asked [my mother to let me go]"), but with slightly less intensity and in a semi-spoken voice, marked by the vocal delivery technique *iro*. The "open end" of the sung phrase on F4, the upper secondary note of the pitch center E4, facilitates the transition to an initially half-spoken, then purely spoken passage (category: *kotoba*); another transitional passage employing the recitation technique *kakari*, in which the F4 is taken up again, finally leads back to the singing type *ji*. Hinadori's direct speech continues during all these changes in the vocal articulation; the forms of vocal delivery are thus not synchronized with the narrative-dramatic sectional structure of the text.

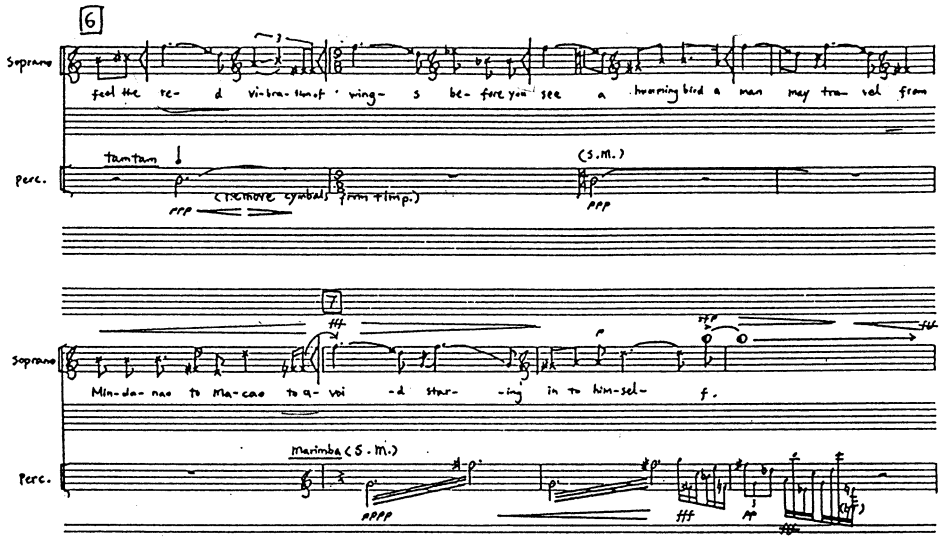
The transitions between sung and spoken passages are organized down to the last detail and at the same time with great variability, whereby pitch, timbre, and the interaction between voice and instrument continuously influence one another, "blurring" the vocal techniques. Owing to a high number of repeated pitches, most lyrical passages (*ji*) contain references to stylized speech (in Hirano's terms, it is a mixture of *rōsho* and *eishō*), while conversely spoken passages (*kotoba*) retain clearly present melodic forms (i.e., they mediate between *ginshō* and *rōsho*); references to gender, class, or age of the characters, or to the dramatic situation are encoded in these gradations. We therefore cannot reduce the complex relationships between melodic form, melodic patterns, the meaning of the words, and the phonetic structure to simple schemata or rules. The resulting high level of fragmentation is analogous to the flexibility with which the performers translate the traditional performance instructions into sound; the information in the notation is formed by the reciter and the singular timbre of his voice into a strong vocal presence whose fragmentary microstructure is shaped into an ever-changing continuum.

Comparable forms of vocalization that continually change between different articulation types can be found in numerous Asian dramatic and narrative genres, though only a few show a degree of fragmentation in the microstructure similar to *gidayū-bushi*. The vocal articulation in Beijing Opera (*jingju*), which is usually summarized with the overarching categories of *sangzi* (voice) or *changbai* (singing and speaking), encompasses a large number of formalized gradations, even if these often vary imperceptibly in practice: "Because the same basic techniques of vocal production are used for all types of vocal performance, there is no feeling that a character suddenly stops talking and starts singing, or stops singing and begins talking; 'a very smooth transition from speech to song and vice versa [is achieved], contributing to the unity of a whole play.'"[33]

In addition to the Chinese folk song type *shan'ge* ("mountain songs"), Daoist ritual music, and Schoenberg's "speech voice," the singing styles of Beijing Opera constitute a key element of the vocal style developed by the Chinese American composer Tan Dun. Comparable to many other Chinese composers of his generation such as Qu Xiaosong (b. 1952), Wen Deqing (b. 1958), or Mo Wuping (1958–1993), Tan Dun created this distinctive style primarily on the basis of his own voice and during the 1980s consequently often acted as a vocal soloist of his own works. Starting with the early *Fu, Fu, Fu* (1982) to *On Taoism* (1985) and *Nine Songs* (1989), the metaphorical saying of "finding one's own voice" can be understood literally in Tan's case. Establishing

33 Wichmann, *Listening to Theatre*, 177, quoting Hwang, "Peking Opera," 220.

Example 5.3: Tan Dun, ***Silk Road*** *for soprano and percussion, 3, systems 1+2*



an "oral tradition" of its own, Tan subsequently passed on this style to other singers who (with the exception of Beijing Opera singer and director Chen Shi-Zheng) were usually trained in European vocal technique. The American soprano Susan Botti (b. 1962), for example, worked with Tan Dun on a number of projects in the 1980s and 90s. An important step in this cooperation was *Silk Road* for soprano and percussion (1989).[34] Example 5.3 shows a section of this work in which the voice alternates regularly on almost every beat between Tan's "register notation," marked by a "register key," which distinguishes five registers, and a speech voice, marked by cross-shaped noteheads and notated in chromatic pitch degrees.

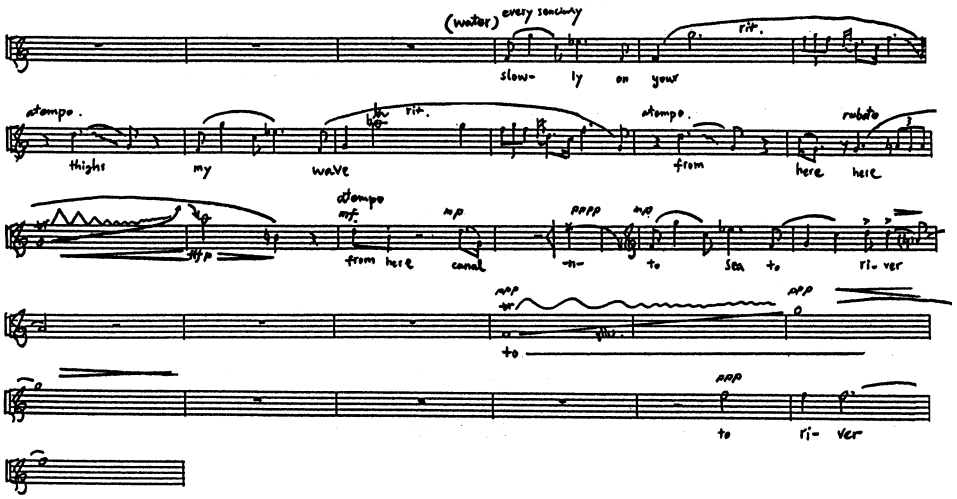
The sections in register notation connect the individual syllables by means of sliding lines whose shapes are precisely prescribed, the falsetto and breast voice constantly merging into one another, as is characteristic of the *yunbai* style of declamation in Beijing Opera, in which stylized Chinese speech tones are connected to a melodic contour through glissandi. The relative pitches in *yunbai* follow the conventions of *Zhongzhou yun* (Zhongzhou rhyme, based on a dialect of the district of Zhongzhou, Henan province in the tenth to fourteenth centuries), a style distinct from "Beijing dialect" (Mandarin) which is also used in Beijing Opera.[35] In Tan Dun's adoption, both types of vocal articulation, the "chromatic" chanting in the Schoenbergian speech voice style and the glissando-laden *yunbai* style, which is only roughly designed melodically according to registers, are combined in a way that minimizes their conceptual difference, not least by assigning them to the same performer and thus by linking them through a constant timbre.

34 A first version of *Silk Road* was presented by Joan La Barbara on 01/04/1989, but subsequently the work was performed by Susan Botti and recorded on CD by her. See Utz, *Neue Musik und Interkulturalität*, 397–398 and 423–429 (a detailed description of Tan Dun's compositions with voice up to the years 1996 to 1998).

35 Wichmann, *Listening to Theatre*, 204–211.

In his opera *Marco Polo* (1991–95), Tan Dun then referred to the *yunbai* style much more extensively and, in this context, expanded the concept of stylistically hybrid voices suggested in *Silk Road*. The type of voice treatment for the role Water, for example (Ex. 5.4), which emerged from the long-standing cooperation with Susan Botti (the soloist of the premiere), shows such a hybrid character in a particularly sustainable manner. The vocal structure aptly mirrors the symbolism of the Water-role in the opera as a mediator between geographical areas (water as a transport medium) and between the opera's spiritual and physical layers (water as a transformative medium). Although Tan Dun's elaboration for all voices in *Marco Polo* refers to the traditional singing technique of the Beijing Opera, differentiating between three phases of sound production – head (*tou*), belly (*fu*), and tail (*wei*) – according to the initial, middle, and final sounds of the Chinese syllable,[36] this characteristic and the resulting inflection techniques summarized in Tan's term "vocal calligraphy" are particularly formative for the Water-part.[37] In addition, this part is characterized by a montage-like, paratactic structure that constantly changes between allusions to Beijing Opera vocal qualities, including the hightened speech (*yunbai*) and colloquial speech (*nianbai*), as well as to the singing style of Chinese mountain songs (*shan'ge*) and a (Western music based) lyrical high soprano register with controlled vibrato. Here too, the timbre of the vocal performer lends this fragmented mosaic-like structure a certain timbral unity. In fact, Susan Botti's performance as recorded on CD does not seem to be essentially informed by Asian singing traditions,[38] so that the structural analytical impression of hybridity to some degree appears questionable in the result.

Example 5.4: Tan Dun, ***Marco Polo. An Opera within an Opera****, Scene* ***Sea****; Part "Water," mm. 35–70*



36 For more details see ibid., 194–199.

37 See Utz, *Neue Musik und Interkulturalität*, 423–427.

38 Vgl. Tan Dun, *Marco Polo: An Opera within an Opera*, Sony Classical S2K 62912 (1997), CD 1, Track 11.

It becomes clear from the examples discussed so far that "vocal styles" that are linked in such hybrid contexts are almost infinitely differentiable and "fragmentable" in themselves. And this perspective can be further condensed if one includes a historical dimension that, for example in the case of the Beijing Opera, would have to take into account all those local and national singing styles from which a specific singing style emerged.[39] The same applies, of course, to what is probably the best known form of vocal hybridization in the twentieth century – Arnold Schoenberg's *Sprechstimme*, as introduced by the composer in *Pierrot lunaire* op. 21 (1912). An interpretation of Schoenberg's *Sprechstimme* as hybrid, as a montage from fragments of different voice concepts, could initially point to the wealth of its actual and potential sources,[40] including post-Wagnerian styles of vocal declamation, for example in Hugo Wolf's songs[41] or in Engelbert Humperdinck's short experiment with *Sprechgesang*, as well as contemporary forms of heightened speech in theater and literature, prominently embodied in Karl Kraus's readings or in the speech art of the legendary *Burgtheater* actor Josef Kainz, in which the melodic ambitus was stretched far beyond what is common today.[42] Last but not least, the Sprechstimme concept in *Pierrot lunaire* is linked to the aesthetics of a "tonal freedom" of vocal articulation, as demanded in a 1920 treatise by Albertine Zehme,[43] the soloist in the premiere of *Pierrot lunaire* in 1912 and the commissioner of the work – even if one has to assume that there was ultimately no consensus between Schoenberg and Zehme regarding the degree of this "freedom." It is clear, however, that Schoenberg wanted to go beyond a simple imitation of the fin-de-siècle trend of "pathetic speaking" when, in his foreword written in 1914, he stated that the vocal performer "must always be on guard against falling into a 'singing' manner of speech,"[44] and that he disagreed with Zehme's increasing "tonal freedom" in the performances following the premiere.[45]

The microstructural fragmentation of the vocal part in *Pierrot* can be traced not only in terms of cultural history, but also in terms of the compositional structure. The peculiarity was pointed out in the literature several times that in no. 3 of *Pierrot* (*Der Dandy*) in addition to the general *Sprechstimme* as employed in all movements of the cycle, no less than six other different types of vocal articulation are explicitly distinguished (Ex. 5.5): *tonlos geflüstert* (whispered tonelessly), *tonlos* (toneless), *gesprochen* (spoken), *mit Ton gesprochen* (spoken with tone), *halb gesungen* (half sung), and *gesungen* (sung). One can find comparable instructions for the

39 The main styles converging in the Beijing Opera are *xipi* and *erhuang*, the styles of local operas in the Chinese provinces of Anhui and Hubei. See Li, *The Soul of Beijing Opera*, 18–24, Schönfelder, *Die Musik der Peking-Oper*, and Mackerras, *The Rise of the Peking Opera 1770–1870*.

40 See, in addition to countless other contributions, Brinkmann, "Kritischer Bericht," Stephan, "Sprechgesang," Meyer-Kalkus, *Stimme und Sprechkünste im 20. Jahrhundert*, 299–318, Cerha, "Zur Interpretation der Sprechstimme in Schönbergs *Pierrot lunaire*," and Byron, "The Test Pressings of Schoenberg Conducting *Pierrot lunaire*."

41 See Kravitt, "The Influence of Theatrical Declamation upon Composers of the Late Romantic Lied."

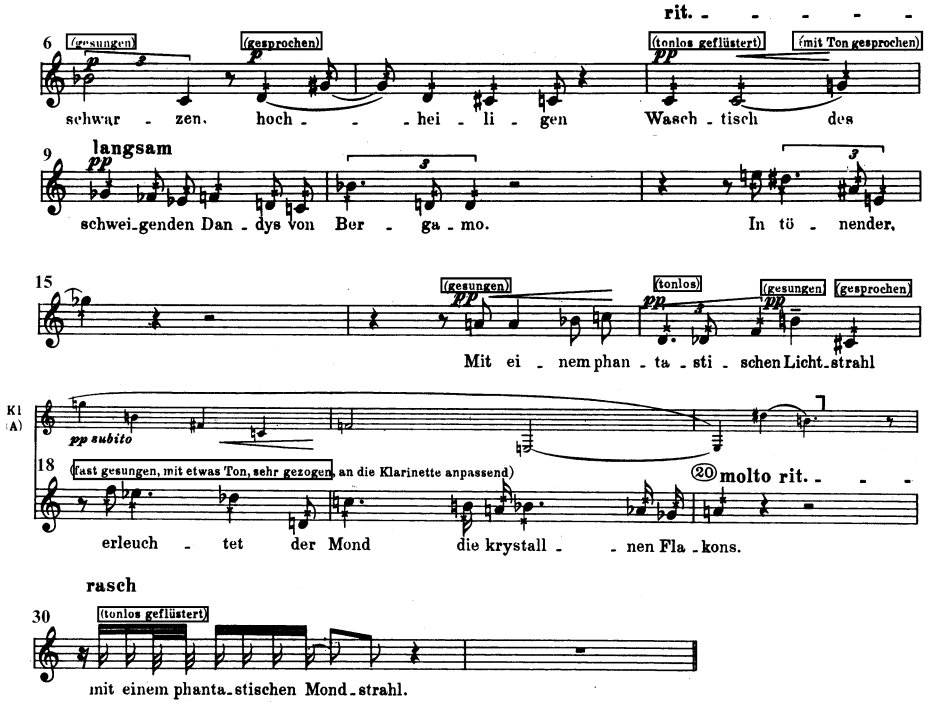
42 See Cerha, "Zur Interpretation der Sprechstimme in Schönbergs *Pierrot lunaire*" and Nöther, *Als Bürger leben, als Halbgott sprechen*.

43 See Zehme, *Die Grundlagen künstlerischen Sprechens und Singens*.

44 Schönberg, "Vorwort" ("sich sehr davor hüten [muss], in eine 'singende' Sprechweise zu verfallen"), translation after Byron, "The Test Pressings of Schoenberg Conducting *Pierrot lunaire*," 2.6; see Nöther, *Als Bürger leben, als Halbgott sprechen*, 131.

45 Meyer-Kalkus, *Stimme und Sprechkünste im 20. Jahrhundert*, 308–309.

Example 5.5: Different instructions for voice articulation in Arnold Schoenberg's ***Pierrot lunaire*** *op. 21, no. 3:* ***Der Dandy****, mm. 6–11, 15–20, 30–31*



voice only in five other movements of the cycle, only in one place each.[46] While this on the one hand surely proves that the fanciness of this movement (and the portrayed character) was of particular importance to Schoenberg, I would suggest that the continuous change between different nuances of *Sprechstimme* here only prescribes a performance practice in particular detail that can (cautiously) be transferred to the other movements as well.

This viewpoint perhaps provides a new perspective on the much-discussed problem of the interpretation of the *Sprechstimme* in *Pierrot lunaire*: Schoenberg's ambiguous information on the performance of the vocal part in the score, to which he added further, no less contradictory, restrictions and clarifications, unintentionally opened a space, a *passage* for the collaboration between the composer and the vocal soloist. He consistently notated clear chromatic levels for the *Sprechstimme* but at the same time – also in performances he directed himself – apparently allowed interpretations (or even preferred ones) in which the sung-spoken pitches differed sig-

46 Four times it is a short "sung" phrase (*gesungen*), the end of which is marked by the instruction *gesprochen* (spoken), see no. 4 (*Eine blasse Wäscherin*), mm. 14/15, no. 8 (*Die Nacht*), mm. 10/11, no. 9 (*Gebet an Pierrot*), mm. 13/14, no. 11 (*Rote Messe*), mm. 24/25. In no. 10 (*Raub*), mm. 9/10, the voice alternates between *tonlos* (toneless), here indicated by special noteheads that occasionally appear without comment in later movements, and *ton* ([with] tone).

nificantly from those specified in the score.[47] The artistic process initiated by Schoenberg can therefore neither be described as a license for "self-assumed musical reciters" nor as an unambiguous product of a "radical aesthetic of the work."[48] With this passage between speaking and singing Schoenberg also opened one between writing and orality. His concept of *Sprechstimme* "resists the view of music as solely the composer's sound which needs to be reproduced by passive performers"[49] and thus opens a space of negotiation between notation and orality, between compositional intentionality and performative presence.

There are key moments in *Pierrot lunaire* in which the voice moves to extreme registers, such as in some phrases of *Nacht* (no. 8), where an extremely low register symbolizes blackness, darkness, or in *Die Kreuze* (no. 14), where ascending leaps of over two octaves are contained within a structural cross symbolism. In these moments, the ecstatic potential of the voice can be experienced abruptly, it transcends the meaning of the word just as it goes beyond a simple illustration of rhythmized text declamation.[50] Indeed, it is probably these moments when the fragmentary microstructure is most sustainably absorbed in the presence of the voice and can be filtered by the performer into a new unity – similarly to what has been described above for the voice in *gidayū-bushi*. In all the examples discussed so far, there is this tension between fragmentation, particularity, a tendency to disintegrate into individual articulations and their connection in the act of an overarching vocal representation by the performer.

Incantation, Prosody, Ritual (Shimoyama, Takahashi, Peri, Sciarrino)

Such moments of vocal presence, immediacy, even ecstasy, were exploited most extensively in the twentieth century as a key subject of vocal composition. Some Western composers such as Dieter Schnebel, Giacinto Scelsi, or Jani Christou, encountered this field mainly through "psychoanalytic" and ritualistic approaches, and it is hardly coincidental that the attraction of the voice's presence resulted from the extreme fragmentation or parameterization in serial and postserial music.

47 See Byron, "The Test Pressings of Schoenberg Conducting *Pierrot lunaire*," 4.1–4.12. On the problems of whether the vocal part in *Pierrot* should respect the notated pitches or not, a broad discussion has continued for decades. Recently, Lukas Dorfegger made detailed transcriptions of the vocal part in three recordings (Erika Stiedry-Wagner/Arnold Schoenberg 1940, Anne-Lise Berntsen/Christian Eggen 1994, Christine Schäfer/Pierre Boulez 1998) and summarized that while the more recent recordings surely respected the pitches more than the earlier ones, "the demand to sing each pitch exactly is basically hardly feasible and thus is cannot become the most important criterion for a good performance of *Pierrot lunaire*." (Dorfegger, "Praktische und theoretische Überlegungen zur Ausführung der Sprechstimme in *Pierrot lunaire* op. 21 von Arnold Schönberg," 32; "die Forderung […], jeden einzelnen Ton genau anzusingen, [ist] praktisch kaum ausführbar ist und [kann] daher auch nicht zum wichtigsten Kriterium für eine gute Aufführung von *Pierrot lunaire* werden […].")

48 Meyer-Kalkus, *Stimme und Sprechkünste im 20. Jahrhundert*, 309 ("musikalisch sich dünkende Rezitatoren," "radikale[] Werkästhetik").

49 Byron, "The Test Pressings of Schoenberg Conducting *Pierrot lunaire*," 4.6.

50 Some research (e.g., Rapoport, "On the Origins of Schoenberg's *Sprechgesang*") argues that the melodic contours of Schoenberg's *Sprechstimme* basically follow the model of intonation in spoken language in an almost literal manner, and the musical result is "alienated" through structural procedures such as reversal or octave transposition. Even though I doubt that the "melody" in *Pierrot* actually came about in such an analogical way, it is by no means unreasonable to assume that spoken intonation could have been essential for the melodic and rhythmic design of the vocal part.

Against a comparable background, vocal works by Hifumi Shimoyama (b. 1930)[51] and Yūji Takahashi can be cited from the Japanese context, in which a countercurrent to the cases discussed so far becomes apparent: the starting point is less a variety of fragmented vocal techniques than a single reference articulation that is eroded in several ways. Since the 1970s, Japanese composers have responded to the simplistic combinations of "Western" structure and "Japanese" color, as sought in some parts of Japanese postwar music, by turning to ritualistic simplicity (→ III.1).

Shimoyama's vocal works generally use text material from Buddhist ritual music as it can be found in *shōmyō* or *nō*, specifically relying on the Japanese transliteration of archaic Sanskrit, in which the textual meaning dissolves into auratically conceptualized sequences of sounds.[52] Shimoyama explains that the meaning of these texts, if they can be reconstructed at all, has no relevance for his compositions.[53] Rather, he focuses on the magical implications of individual words or phrases. This attitude is also reflected in his approach to archaic Japanese words, which are no longer part of everyday language today but can still be found in the local Tsugaru dialect of his home prefecture Aomori and also in the vocal aesthetics of the narrative genre *tsugaru-jamisen*,[54] which is little known in or outside of Japan. Shimoyama hears these magical dimensions of words and phonemes in this genre and in the ancient Tsugaru language, in *gidayū-bushi*, as well as in *nō* singing.

His *Monolog* (1991) for *shamisen*-playing vocalist, written for Kazuko Takada, and *Breath* for three voices, two drummers, and piano (1971/77), written for Michiko Hirayama (who, as mentioned above, is noted for her performance and co-creation of Scelsi's *Canti del Capricorno*), vividly illustrate the type of vocal presence in Shimoyama's vocal works. The reference technique in both works is an almost pure variant of the *rōshō* type, which suggests Buddhist recitation; it consists of rapid repetitions of a single, not precisely defined pitch. In *Monolog* (Ex. 5.6) fragments are arranged around this reference model, which deepen into the auratic qualities of individual phonemes and describe interrelations between voice and instrument as familiar from numerous traditional genres. The Buddhist text is taken from the *nō* play *Funa Benkei* (Benkei on the Ship), where it is presented as a prayer to guardian deities.

Breath introduces a "multiplication" of Hirayama's voice into three parts, two of which were pre-recorded by the soloist.[55] The text material is compiled from the three different *nō* pieces *Okina*, *Aoi no ue* [The Court Lady Aoi], and *Atumori* (from which a prayer of the Nenbutsu sect is used). All vocal textures are based on the same *rōshō* model, which occurs in at least five different forms of articulation (Ex. 5.7), including a microtonally layered recitation tone with

51 Shimoyama belongs to the large group of Japanese composers who are hardly known in the West. He was born in the same year as Tōru Takemitsu and, after graduating from Hirosaki University (Aomori Prefecture), studied with Yoritsune Matsudaira in Tokyo from 1955. In an international context, his works were presented primarily at the festivals of the International Society for Contemporary Music (ISCM).

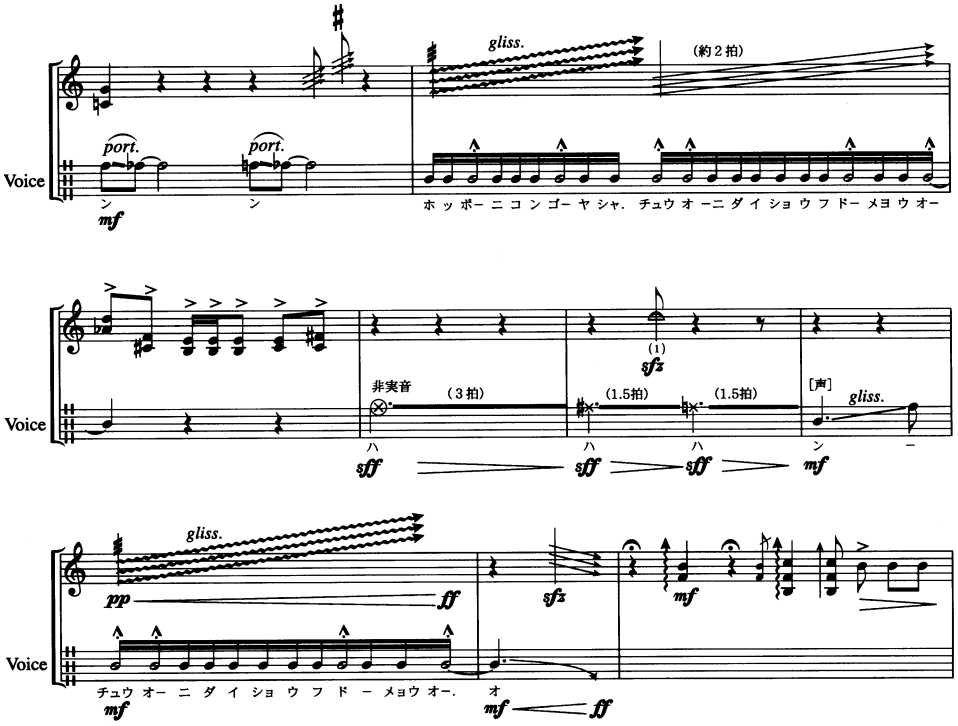
52 This transliteration tries to imitate the pronunciation of Sanskrit using the syllables or morphemes available in Japanese, but without linguistic meaning; it is a purely onomatopoetic procedure. See Nelson, "Court and Religious Music (2)," 67–68.

53 Personal communication with the author, Tokyo, 26/2/2008.

54 See Groemer, *The Spirit of Tsugaru*.

55 *Breath* was premiered by Hirayama in Rome in 1974 and revised by the composer in summer 1977. In 1978 the revised version was premiered at the ISCM festival in Helsinki. The live recording of this performance is published on CD. In addition to a performance by a live vocal soloist and two pre-recorded vocal parts, which are played on tape, a performance by three live soloists is also possible.

Example 5.6: Hifumi Shimoyama, ***Monolog*** *for shamisen and vocal soloist, p. 4, systems 1–3*



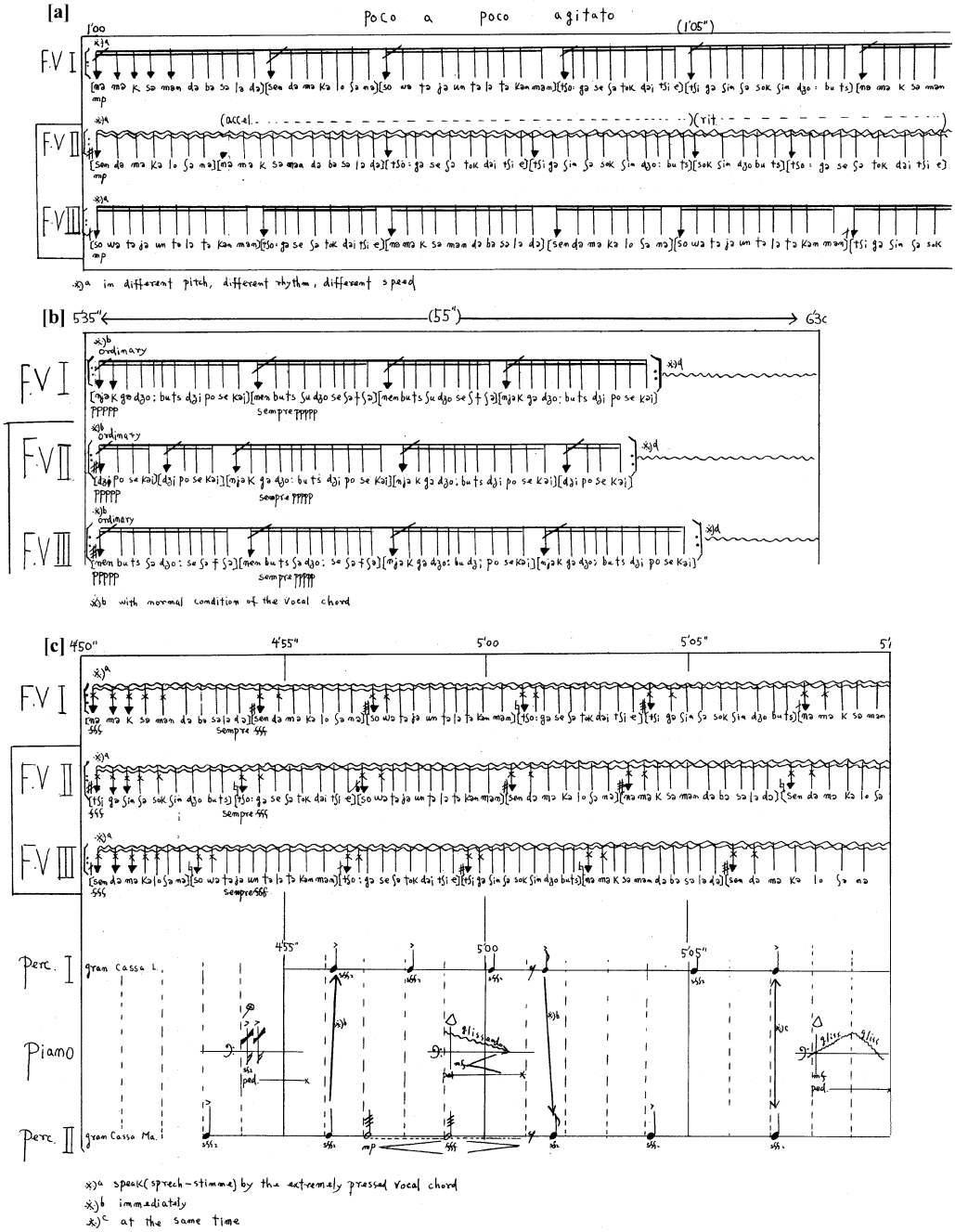
crescendo [a], also occurring in a whispered variant [b] and in a variant with "pressed" voice [c], which Shimoyama associates with *gidayū-bushi*[56] but in the score refers to with the (original German) word "sprech-stimme." In addition, there is a texture of sustained notes performed with "larynx articulation" [d], which can be seen as a much slower variant of the *rōshō* model, as well as mixed polyphonic textures [e] that also contain vocal inserts by the instrumentalists, comparable to the *kakegoe*-calls of the drummers in *nō* theater.[57]

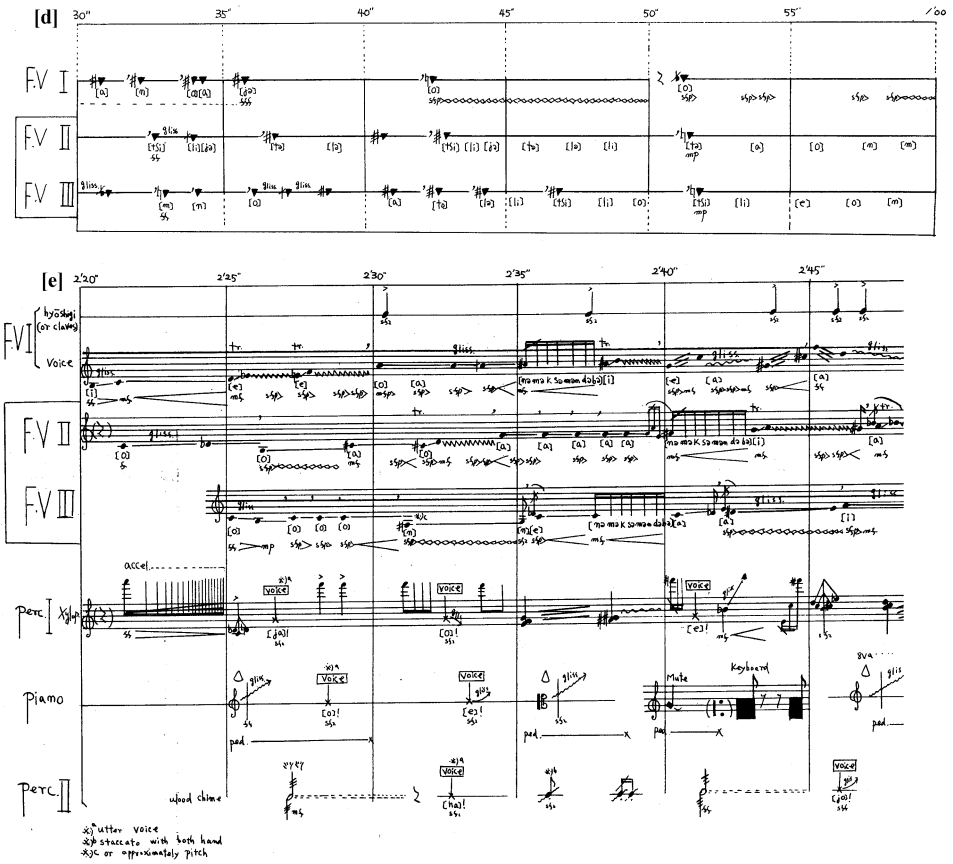
Yūji Takahashi's works since the late 1980s were introduced in Chapter III as remarkable examples of an anti-essentialist approach to traditional Asian musical practices, informed by long-term composer-performer collaborations (→ III.4). The vocal style in Takahashi's *Unebiyama* (1992) for a reconstructed archaic five-string zither and chant is mainly derived from vocal practices in shamanistic genres. Although there are still a few female shamans (*fujo*) practicing in Japan's northern Aomori Prefecture, Takahashi's work primarily refers to shamanistic music from Southeast Asian regions, particularly to "hidden traditions" that are hardly documented and which Takahashi encountered through field recordings. More specifically, the

56 Personal communication with the author, Tokyo, 26/2/2008.

57 A detailed description of the function and content of these calls, which can be associated with the cries (*katsu*) of masters of Zen Buddhism during meditation, is provided in Tamba, *The Musical Structure of Nō*, 228–229.

Example 5.7: Hifumi Shimoyama, ***Breath*** *for three voices, two percussionists, and piano, five vocal articulations (a–e) from pages 3, 13, 11, 2, 7 of the score*





Copyright © 1982 by Hifumi Shimoyama, Tokyo

composer identified music by shamans from the Temiar ethnic group in Western Malaysia as one of his main sources.[58]

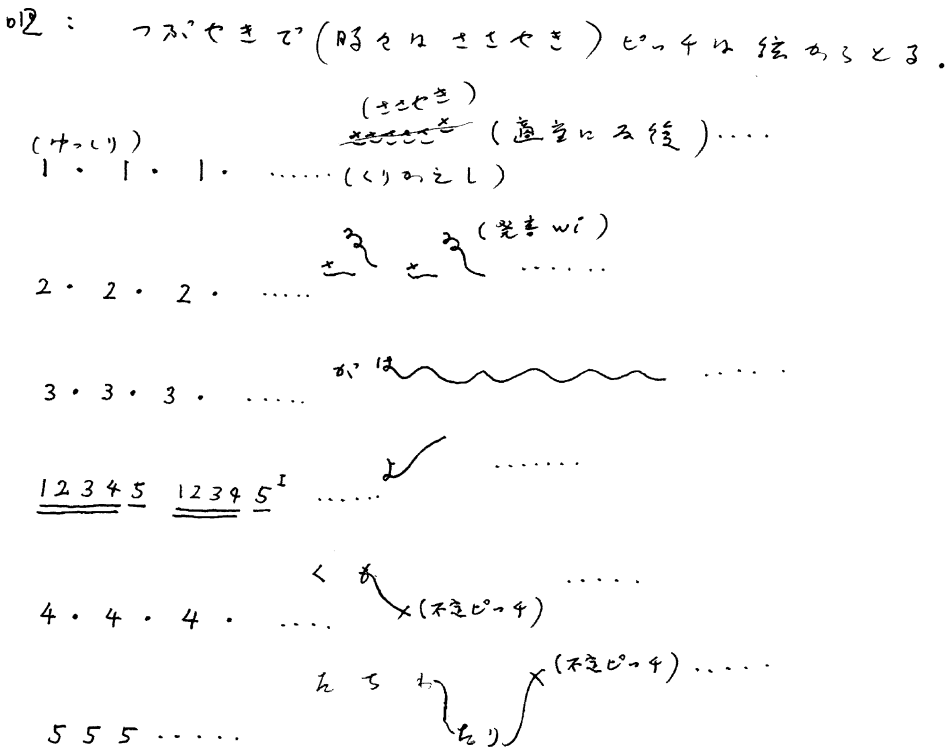
The organology and construction of the five-stringed zither (length: 115 cm) and ethnological evidence form another important background of this work. The instrument dating from the fifth century BC was discovered 1978 in a tomb of the Marquis Yi of Zeng in Hubei Province, China.[59] Not much is known about the original name or usage of this instrument. Takahashi argues that, based on iconographic evidence, it may have been used in shamanistic rituals.[60] He tunes the instrument in a modified, "oblique" pentatonic system (D2, lowered E2, A2, D3, raised C3) that bears some resemblance to the popular Japanese mode *miyakobushi* (→ II.4). The strings are plucked with a long pick made from the horn of a water buffalo. Since it is almost impossible to depress the strings, Takahashi's composition consists almost exclusively of open strings, with the repeated patterns enhancing the impression of ritual music.

58 Personal communication with the author, Tokyo, 22/2/2008. The Temiar belong to the Senoi group of the indigenous *orang asli* people. See Roseman et al., "The Indigenous Peoples (*Orang Asli*) of the Malay Peninsula."

59 Kido, *Reconstructed Music Instruments of Ancient East Asia*, 126–133.

60 Takahashi, *Astray Among Decayed Strings*, 9.

Example 5.8: Yūji Takahashi, ***Unebiyama*** *for five-string zither and incantation, p. 1*



Copyright © 1992 by Yūji Takahashi, Tokyo

The voice takes up the pitches of the five open strings (Ex. 5.8). The instructions in the score call for a voice that combines "muttering" (*tsubuyaki*) and "whispering" (*sasayaki*). The text is based on the Japanese creation story *Kojiki* (compiled around 712 CE by Ō no Yasumaro) and can be linked to the dragon iconography of the original instrument via the symbolism of water and clouds. The selected poem belongs to the category of *wakauta*, songs that handed down oracles sung by young virgins who warned of political crises.[61] It is attributed to Isukeyorihime, the wife of the mythical emperor Jinmu.[62]

61 Ibid., 11.

62 There is, however, notably no explicit reference in Takahashi's work to the archaic recitation style that is usually associated with the *Kojiki*. This style was practiced by singers called *kataribe* and is considered the origin of the Japanese *katarimono* (dramatic narrative chants, see above). Hirano connects this style to genres practiced by the Ainu and Okinawa peoples (*yūkara*, *omoro*; see Hirano, "Beziehungen zwischen Sprache und Musik im *katarimono*," 57.)

佐韋河	Saigawa yo	From Sai river
雲立ち渡り	kumo tachiwatari	Clouds rise and spread across the sky;
畝傍山	Unebiyama	Unebi Mountain
木の葉さやぎぬ	konoha sayaginu	Rustles with the leaves of trees;
風吹かむとす	kaze fukamutosu	A wind is about to blow.[63]

The notation developed here is based on the composer's reflections on the difference between the notation methods for vocal and instrumental music in the traditional music of Japan: while music for instruments is usually recorded in a form of tablature that indicates the position of the fingers, the string, etc., and is permeated with instrumental patterns learned by mnemonics, the notation of vocal music is based on gestural analogies between writing and sound as in the notation forms *hakase* (e.g., in the Buddhist chant *shōmyō*) or *gomaten* (in *nō* theater).[64] For Takahashi, these different paradigms convey the more temporal quality of vocal music, in which the sounds are continuously "stretched" over time, compared to the more spatial, punctuating characteristics of the (often plucked or struck) instrumental impulses.

One method by which Shimoyama and Takahashi circumvent the tendency toward fragmentation is to consciously refrain from specifying pitches or their constant inflection and fluctuation in an often vague (registral) area. As shown, this method can be closely related to linguistic articulation or corresponding forms of notation in traditional Japanese music.

In a completely different historical-cultural context, related methods for deriving different degrees of vocal pitch determination from linguistic particularities can be observed. In his preface to *Le musiche, sopra l'Euridice* (1601), Jacopo Peri (1561–1633), singer, composer and a key figure in the *nuove musiche* around 1600, refers to the ancient distinction between "diastematic" and "continuous" vocal articulation. This distinction can be traced back to around 500 CE, when Boethius adopted it from Aristoxenus's *Elementa Harmonica*, while describing a mixed type of vocal articulation for the presentation of "heroic poems."[65] Peri developed a theory of Italian speech and intonation patterns from these categories, conveyed by Girolamo Mei and Vincenzo Galilei, whereby clearly intoned speech sounds were assigned to a consonance or a change of harmony, whereas the syllables in between could be free and independent of the bass part.

An intercultural passage can therefore by opened in relating to the spatial characteristic of the instrumental sounds accentuated by Takahashi (Peri's "diastematic" vocal style coupled to change of harmony) and temporal vocal articulation (Peri's harmony-independent "continuous" vocal style). Peri, who, like his rival Giulio Caccini, was trained in strict counterpoint, must have been aware of the provocative effect of such a musical structure. The notable dissonant passages from Daphne's speech in Peri's *Euridice* are not merely constructed according to the text's prosody, of course, but driven primarily by the (negative) affect of the words ("Angue crudo e spietato," "con si maligno dente").[66] The freedom in the articulation of the unstressed syllables prompts a novel musical form: the recitative.

A flexible articulation of unstressed syllables was also fundamental to Giulio Caccini's theory of vocal style, as documented in the dedication of his *L'Euridice* (1600) and his collection

63 After Takahashi, *Astray Among Decayed Strings*, 11 (English translation based on *A Waka Anthology*, vol. 1: *The Gem-Glistening Cup*. Translated by Edwin A. Cranston. Stanford: Stanford University Press, 1993).

64 Personal communication with the author 22/2/2008. For a good introduction to the notation forms of traditional Japanese music see Malm, *Traditional Japanese Music*.

65 See Palisca, *Studies in the History of Italian Music and Music Theory*, 456–457.

66 See ibid., 460.

Nuove musiche (1602). Caccini used a key concept of the high Renaissance to describe the lightness of these phrases, *sprezzatura* ([noble] negligence): "In this manner of song, I used a certain negligence that I valued as having a noble quality, for it seemed to me that with it I approached natural speech that much more."[67]

In the years after 1600, there was a paradoxical development. The reference to spoken language helped to free song from the rhythmic rigor of speech-based declamation. Thus virtuoso "fioraturi" and "passagi" increasingly replaced the ascetic early syllabic style, so that in 1614 Caccini consistently applied the term "sprezzatura" to the execution of these "passagi" and no longer primarily to a vocal style close to speech:

> *Sprezzatura* is that gracefulness conferred upon singing by passing through a number of eighth notes and sixteenths over various notes [of the bass], through which, when done in time, removing from the singing a certain bare poverty and dryness, it is rendered pleasing, unhindered, and airy…[68]

The Italian composer Salvatore Sciarrino (b. 1947) has described his vocal style using the term "sillabazione scivolata" ("sliding syllabic articulation"), and one can recognize in it something akin to a shadow of Peri's recitative, interspersed with some elements of Caccini's "passagi." Sciarrino refers to spoken language in a way closely related to Peri's theory, emphasizing the "scioltezza" ("lightness") of everyday language as a model for the vocal performers. And by combining sustained, "diastematic," stressed syllables/notes with several extremely short, "continuous," unstressed ones, Sciarrino's works also use a musical approach analogous to Peri's (Ex. 5.9). Sciarrino notes that the intonation of the closing figures, consisting of short notes, should not be rendered precisely and need not be equally tempered, meaning that "the nuancing of non-tempered intervals, as is customary when speaking," is possible.[69]

Example 5.9 shows variants of the "sillabazione scivolata" from Sciarrino's opera *Luci mie traditrici* (1996–98) in which the stylization and variation of the interval structure [a, b], the breaking-up of the lines with rests (symbols for hesitation [b], fainting [c], death [e]), and their dissolution into descending micro-glissandi [b, c, e] or monotonously spoken phrases [d] are clearly recognizable. In fact, the formal process of the entire opera can be seen as an increasing reduction of intonation, culminating in the dark final scene, in which all articulation types extending to extreme interval jumps [d] recur in a (nightmarish) reminiscence.[70]

Articulation – Codification – Presence – Passage

The discussion presented here attempted to show how different forms of articulation and codification between speaking and singing voices – in situations with strongly divergent historical and cultural conditions – are organized in a related manner and performatively transformed into a presence of the voice. Finally, three main threads of this discussion will now be highlighted and linked:

---

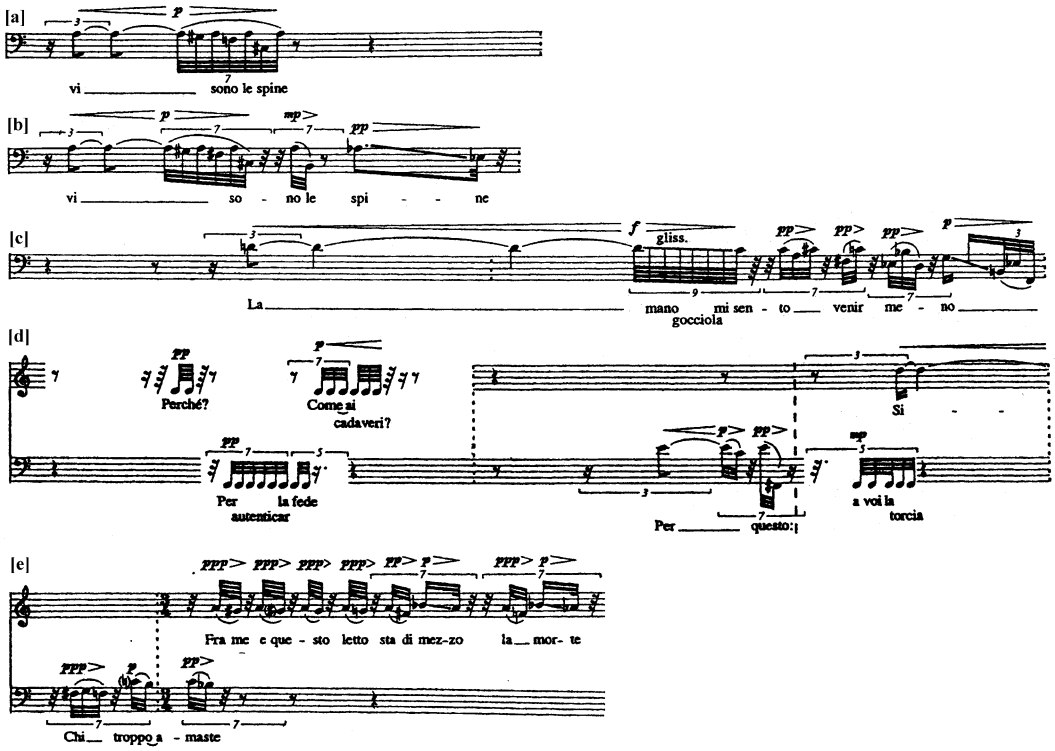
67 Giulio Caccini, *L'Euridice composta in musica in stile rappresentativo* [1600], dedication to Giovanni Bardi, quoted from ibid., 463.

68 Giulio Caccini, *Nuove musiche e nuova maniera di scribere* [1614], preface, quoted from ibid., 464–465.

69 Sciarrino and Petazzi, "Salvatore Sciarrino im Gespräch," 37 ("die Nuancierung nicht temperierter Intervalle, wie dies beim Sprechen üblich ist").

70 See, for more detail, Utz, "Statische Allegorie und 'Sog der Zeit.'"

Example 5.9: Salvatore Sciarrino, ***Luci mie traditrici****, variants of the "sillabazione scivolata": [a]–[c] scene 1, Il Malaspina (baritone), mm. 15, 17, 29–30; [d], [e] scene 8, La Malaspina (soprano)/Il Malaspina (baritone), mm. 67–69, 83*



1. In a wide variety of contexts, types of vocal music can be found that tend to hybridize an abundance of fragmentary individual articulations. It is therefore logical that the role of the performer becomes central, for it is only in the act of vocal execution that it becomes possible to form an overarching musical meaning from this fragmented material. If one compares the field of possibilities of a *tayū* reciter in *gidayū-bushi*, a vocal performer of *Pierrot lunaire*, or a singer of Sciarrino's "sillabazine scivolata," it becomes evident that, despite the clearly different degrees of prescriptive notation and performance traditions involved (that is, the codifications on which these articulations are based), the performers face only gradually different tasks: in all cases they are required to consider a wide range of secondary information in addition to the notation, and to make essential intuitive decisions where the notational or oral codification does not provide sufficiently binding instructions. In this regard, there is an obvious convergence of "traditional," mostly orally transmitted, non-Western practices and vocal music of the twentieth and twenty-first centuries.
2. Such a form of vocal articulation operates in a particularly close relationship with language in which its semantic dimension is often put aside. In this sense, all the examples discussed here can be understood with reference to Roland Barthes as representatives of the "geno song," as "the space where significations germinate 'from within language and its very

materiality'; […] where the melody really works at the language, […] explores how the language works"[71] – to the point, of course, that they obscure or even destroy the language. And by ultimately provoking – for all the inspiration they derive from the spoken language – conflicts with conventional forms of prosody, linguistic intonation and declamation instead of following these structures, they probably represent far more suitable examples of the "grain of the voice" than those Barthes himself tried to grasp (which tend toward a culturalist stereotyping of French and German song traditions that has become increasingly dubious[72]).

3. Carolyn Abbate's narratological expansion of the concept of the voice as part of an "aural vision of music animated by multiple, decentered voices localized in several invisible bodies,"[73] the conception of "voices" of different origins that support musical structures, but can also counteract them and thus destabilize the composer's monological authority, can be assumed for all the examples presented here. Yet the "codification" of the vocal articulation can neither be reduced to the type of written notation nor to the gestural-performative presence of its actual rendering but arises from the interaction of these two "voices" with those of cultural memory and reception history. The passages between speech and song explored here seem to exhibit a particular, obvious tendency toward the divergence of written and performed voices, resulting in an unfolding of composition, terminology, and categorization on the one hand and vocal practice on the other. The nature of this divergence, as is equally evident in *gidayū-bushi*, in Schoenberg's *Pierrot*, Peri's *Euridice*, or Sciarrino's "sillabazione scivolata," allows a specification of the vocal medium, and thus a differentiation from Abbate's metaphor of "multiple voices" applicable to both instrumental and vocal music alike.

Seen in this way, an intercultural history of the voice could appear as a process of condensation in which such contradictions, divergences between individual voices in the conception, written codification, sounding event, and perception are made visible, and physical and metaphorical levels of the phenomenon are envisaged. The outline undertaken here can at best be understood as a first attempt in this direction.

71 Barthes, "The Grain of the Voice," 269.

72 See Meyer-Kalkus, *Stimme und Sprechkünste im 20. Jahrhundert*, 427–444.

73 Abbate, *Unsung Voices*, 13.