

Organization of the visual structuring of the voice in a singing lesson

Yasuaki Dan

Osaka University of Economics (Faculty of Human Sciences)

Abstract

Drawing on ethnomethodology and conversation analysis, I examine a lesson for Japanese nursery rhymes as an activity in which the culture of a particular expression is reproduced. More precisely, I investigate the way in which the singing voice becomes the object of instruction in a lesson and the object is visually structured by gesticulating the singing voice, thereby identifying the mechanisms of a lesson that produce better singing. To this end, the analysis focuses on (1) the sequential characteristics of a vocal lesson, (2) the use of the body configuration to demonstrate the singing voice to be taught, (3) and the envisioning of the singing voice to make it accountable. Based on the presented analysis, this article specifies the ways in which the teacher can convey the invisible qualities of singing to their student, with particular attention to the visualization of the voice.

INTRODUCTION

The singing voice is invisible. It is delivered to us with melody, rhythm, and lyrics. In the case of vocal music, the melody, rhythm, lyrics, and direction of expression are written in the score to reproduce the song. Although not every detail for singing is written in the score (e.g., Garfinkel's [2006] accounts for instruction and instructed action), they are taught in lessons for good singing, particularly in specialized singing genres such as vocal music, for all practical purposes. Songs that are seen as music with poetry created based on the language, we usually use everyday are expected to harmonize the linguistic expression of poetry with the expression of melody. On the other hand, melody does not always harmonize with poetry, and sometimes melody has the potential to interfere with the expression of poetry. There is a kind of tension between melody and poetry. Thus, the tension between lyrics (language) and melody (music) is a characteristic issue widely seen in diverse cultures (see Barthes 1977; 1991 for *Le Grain de la voix* [The grain of the voice]; Vaughan Williams 1963) and in the transition from orality to literacy (Havelock 1963; Ong 2002).

This paper explores a lesson for a Japanese nursery rhyme (Doyo song) written by Kitahara Hakushu, a famous poet in the early 1900s who was opposed to the idea of

putting melodies to his poems (Shuto 2008). Kitahara argued that poems for children should be recited in rhyme produced by the child's natural voice, rather than with a melody based on Western music (Shuto 2008). Kitahara was able to leave behind many Doyo songs by working with composer Yamada Kosaku, who was skilled at transforming the characteristics of the Japanese voice into melody (Shuto 2008: 272). This historical background of the song (i.e., the tension between language and music) is relevant to the data analyzed in this article. The teacher in the lesson uses this background as a resource for the instructions and, following the instructions, the singer must perform the song considering the tension. The lesson therefore provides the task for the tension in a way that is accountable—namely, “observable and reportable” (Garfinkel 1967: 1)—not only for the participants but also for the observers. In this way, the culture of the song is taught as something that can be reproduced in the lesson.

In what follows, I will first review the previous praxeological studies on lessons for musical instruments and singing. After discussing the data analyzed in this article, I will report two examples of instruction focusing on (1) the sequential characteristics of a vocal lesson, (2) the use of the body configuration to demonstrate the singing voice to be taught, (3) and the envisioning of the singing voice to make it accountable. Finally, based upon the findings of the current article, I conclude with a discussion confirming that expressive techniques and nuances specific to particular music can be communicated among participants in a lesson as well as the significance of elucidating this practice.

LITERATURE REVIEW

Lessons for various musical instruments and singing, such as masterclasses, rehearsals, and private lessons, are a form of institutional interaction that operates with asymmetrical knowledge, entitlement, and right as well as constraints on turn taking and turn contributions (Heritage & Clayman 2010), with the common feature of containing a mechanism for improving expression. These institutional characteristics in musical lessons have been analyzed in detail in terms of their sequential structure and structuring music practice in previous ethnomethodological and conversation analytic studies.

In terms of the sequential structure of musical lessons, a number of studies have pointed to procedures for the progression of activities when unprepared teaching areas are discovered and remedied. Weeks (1996) analyzed the conductor's instructions for improving the player's performance based on the correction sequence (Schegloff et al. 1977). Reed et al. (2013) demonstrated that instructions in masterclasses are relevant units of interaction, as participants clearly orient to their initiation, implementation, and closure. In addition, studies focus on the various activity units of an instruction.

Reed and Szczepek Reed (2014) pointed out several ways in which learnables require improvement in the interaction of participants in a masterclass and the reciprocal action characteristics of these initiations of teaching. Reed (2015) focused on the transition of activity from teaching by the instructor to performance by the student. Undertaking a multimodal conversation analysis provides a progressive appreciation of the

“relinquishing move” produced by the master. Szczepek Reed et al. (2014) focused on the negotiation between the master’s clusters of directives and performers’ (students and pianists) return to performance. Second, I summarize previous research on structuring music practice (Nishizaka 2006; Tolins 2013). Some previous research has focused on the visual structuring of sounds and performances in music instruction. This is a way to make invisible sound features accountable. Various multimodal resources such as musical quotations, descriptions, gestures, and gazes are used there.

Various studies have pointed out that in teaching situations, the participants achieve the content of the teaching through mutual actions with multimodal resources. Goodwin (2003) found that the senior archeologist and a graduate student in archaeological fieldwork made the ground significant by pointing to it as an object of instruction. On the other hand, the singing voice is not an object placed in front of the participants. There are various ways to make the singing voice as the object, as it is not possible to place the object of instruction in front of a student (see Sakai et al. (2014) for an example of visualizing an object that is not in front of the participant). This is a teaching method characteristic of arts that involve the expression of time, such as music and dance (e.g., Keevallik 2010).

The contrast pair method developed by Weeks (1996) is an important way of structuring the object of instruction. Its pairs illustrative expressions of improving students’ singing and the ideal singing according to the teacher, making the main points that need to be improved accountable. Tolins (2013) drew attention to the vocal reproduction of music in private clarinet lessons, where the teacher uses their own voice to express musical phrases, reproducing the performance and using non-lexical vocalization to evaluate and instruct. In the evaluation, the teacher quotes the student’s performance in vocalization and gesture, while during instruction, the teacher presents the desired music in vocalization.

In addition, Tolins (2013) noted that musical sounds cited in instruction are often coupled with gestures and made meaningful. Nishizaka (2006) examined a lesson on how to use a musical instrument and demonstrated that learning was achieved through concrete details of reciprocal action using the tool of the string, gestures connected to the tool, and the bodies and gazes of the participants to demonstrate them. Based on the methods presented in the previous studies above, this study will focus on the methods used to teach expression related to the tension between words and music in vocal music. There, students’ singing is not only quoted, characterized, or contrasted with the ideal singing, but also structured with multimodal resources such as gestures and gaze movements. In such a combination of methods, I will show two examples of how singing can be visually structured and accountable.

DATA AND TRANSCRIPTION

This study used music lesson video data containing a student (“S” in the transcript) and teacher (“T” in the transcript) recorded in October 2010 whose native tongue is Japanese.

The teacher was a vocalist who wrote choral pieces. The student had been taking lessons from the teacher for 9 years since they were in high school. The student is the same person as the author. The student had won prizes in competitions when they were in high school. At the time of recording, the student was a graduate student. After enrolling in university, the student continued to take specialized lessons from the teacher. These video data were one of these regularly scheduled lessons.

The recorded lesson was approximately 30 minutes long, while normal lessons are usually a little under one hour. The lesson consisted of vocal exercises, including the Japanese songs “Kaya no kiyama no” and “Karatachi no hana” (music by Kōsaku Yamada, lyrics by Hakushū Kitahara). The student used sheet music included in a collection of famous Japanese songs 110 (*Nihon meika 110 kyokushū*), published by Zen-on Music Company Ltd (Zen-on Gakufu Shuppan-sha). The first five minutes were spent on vocal exercises, approximately fifteen minutes were spent singing “Kaya no kiyama no,” and the last ten were spent singing “Karatachi no hana.” The student had sung “Kaya no kiyama no” several times in the past, having read the musical score in high school. In particular, I would like to focus on the instruction for the two points in two bars of “Kon:: yamo ame daro::” (It will rain again tonight). The score is shown in Fig. 1.



Figure 1: Targeted score

Video recording permission was acquired from the teacher. Prior to recording, the author repeatedly participated in the lessons and identified points to be highlighted as well as chose a place to position the camera to capture these points on video. Adhering to the ethical consideration, I anonymized the participants' names and other identifiers and used the drawings from the video recordings. The space for the lesson is a room in the teacher's home. When the lesson begins, the teacher takes a seat at the grand piano. The student stands facing the teacher and sings. The teacher places the sheet music on the piano's music stand and accompanies and guides the student. The student puts the

music on the music stand and sings the song. The extracts have been transcribed as per the Jefferson transcription system. The list of conventions can be found in Appendix 1.

THE VISUAL STRUCTURING OF THE VOICE: TWO CASES

First, this paper presents the interaction characteristics of the “lesson” among “doing music.” Whether the scene was a “lesson” was understood and experienced by the participants in the activity before it was classified by the researcher. For instance, if the scene was not a “lesson” but a “rehearsal,” then there should be coordination between piano accompaniment and singing with little detailed instruction on singing. A “lesson” has a different structure of activities than a “rehearsal.”

The first feature is that the participants in the “lesson” have a relationship of “teacher” and “student.” In masterclasses, the master is at the piano listening to the student sing; however, a piano accompanist is absent. Before recitals and competitions, lessons and rehearsals are held with an accompanist; however, lessons generally focus on practicing a specific piece of music with the teacher accompanying the student. It is common for the teacher to teach only a specific part of a particular piece in one lesson.

In the following sections, I will examine two examples of teaching and clarify the ways in which the music that needs to be improved becomes understandable through visual structuring. I will focus on the structure of the activity, the participation framework in teaching, and the gestures coupled with singing in teaching.

Looking at the structure of the activity, I can see the same characteristics that Reed and Weeks saw in masterclasses and orchestra rehearsals. At the same time, in vocal music lessons, the teacher was shifting to a distinctive framework of participation that was not found in masterclasses. It was done so that a single teacher could both accompany and teach. The teacher listens to the student singing while accompanying, and when they find something to teach, they stop accompanying the student, change their body configuration, and start teaching. In this transition of participation framework, the teacher can show the student their body as an example by keeping the body placement and gaze forward rather than toward the student. Then, as the example ends, the teacher turns to the student and continues the explanation. By this use of their body, various singing and gestures are connected to create a visual structure.

In the two cases, the singing is visually structured by a combination of similar participation framework features and singing and connected gestures. The first case will be analyzed with an emphasis on the structure of the activity and the second case will be analyzed with an emphasis on the participation framework and the visual structuring of singing and connected gestures.

CASE 1: IMPROVE THE WAY SINGING FROM “N” TO DIFFERENT PRONUNCIATION

Case 1 is an example of teaching a student how to sing the “ ん ” in the phrase “konyamo” (see Fig. 1). The student does not make any mistakes in the scales or rhythms written on the sheet music. However, there is room for improvement in the expression of the “n” when pronouncing the Japanese word “konya” (tonight). Specifically, the goal is to improve the relationship between the word “konya” and “n” sound.

There are various ways to pronounce “ ん ” (n) in Japanese. Phonetically speaking, when pronouncing “ ん ” (n) only, it is pronounced with an “N” and a palatal drip. In conversation, the “n” in “konya” is a nasal vowel because it follows an approach sound. Furthermore, in singing, the pronunciation does not always seem to match the pronunciation in conversation. The teacher does not use phonetic knowledge in their instruction, but rather demonstrates the nuances of the sounds using a variety of methods.

In Case 1, I first analyze the structure of the lesson, which can be summarized in three steps: (1) the teacher identifies points that require improvement in the student’s singing, (2) the teacher provides instruction on these points, and (3) the teacher checks to see if the student is able to understand and demonstrate the instruction, and if so, moves onto the next instructional area. This is a similar feature to that shown by Reed et al. (2013) in their examination of vocal masterclasses.

Next, I will examine the development of this sub-activity using a teaching moment as an example, focusing on using the singing voice as an object in the progress of these activities, thereby making it a resource for the activities. Then, in (2), I confirm that the singing is visualized.

(1) Transition from performance to instruction: Identifying areas and reprising singing to be taught

One of the lesson sub-activities that enables the student to sing better is for the teacher to identify areas for improvement. First, a distinction must be made between errors and areas for improvement, which may be referred to as “learnables” (Reed and Szczepek 2014). When singing in conformity with the text of the score, deviations in scale and rhythm are deviations from what is notated in the score, and, thus, are immediately recognized as mistakes.

On the other hand, areas for improvement are beyond what is written in the score and are identified and guided by the instructor during the lesson. Once the student understands the music and can sing without error, the focus of the lesson is on the areas that need improvement. This requires an understanding and mastery of the nuances of expression. Certain songs have nuances that are difficult to express and these parts may be identified through the teacher’s singing experience and analysis of the music. In lessons, teachers tend to teach techniques from their own experience.

The teacher suggests that there are additional areas that require improvement but does not specify these areas. Following this, the student repeats the part. The teacher prompted the student to begin singing by saying “se::no::” (“Ready, go”; Line 7), and the student begins singing from the indicated part (Line 8). The teacher listens to the student singing, states, “sō sono toki ni” (“Right, at that point”; Line 8), and begins instructing the student regarding singing (Line 9). When the student stops singing and listens to the teacher, the student does not make any mistakes regarding scale or rhythm.

However, when the teacher stops the student singing, they do so using an expression indicating that they realized something: sō (“right”). Though the Japanese word “so” has a variety of meanings that an affirmative response, an awareness, or an indication of understanding (Sadanobu 2002), here, “so” is immediately followed by “at that time,” indicating what was sung earlier, and then the instruction follows. This shows that the method is not so much an affirmation and evaluation of the singing as it is an indication of awareness and a way to move onto a specific instruction. As will be discussed in more detail in the next section, here the teacher says, “Yes” and turns their body alignment back toward the student and places their finger in front of their mouth (see Appendix 2). Rather than making an explicitly positive assessment, the teacher is trying to show awareness and gain the student’s attention. The teacher then moves to concrete instruction (2).

(2) Instruction: Bad and good singing and explanations

The teacher stops the student’s singing by saying, “sō” (“right”) in line 9, adding, “at that point,” which suggests that the student’s singing requires instruction. Here, I focus on the method used to teach singing that is not immediately perceptible. The teacher explains the singing of the student by contrasting it with the ideal form of singing.

The teacher then reproduces the student’s singing of the part, embedding that singing in their instructive utterances. In other words, by re-singing ♪kon::♪ while dropping their jaw (part of the ♪kon::ya::mo::♪ that the student sang in Line 8), they characterize the student’s singing (Line 9, see Appendix 2). In addition, the teacher adds an explanation, “♪kon::♪tte ano:nodo de utatteru” (“You’re singing kon with your throat”). This form of teaching was described by Weeks (1996) as an illustrative expression identifying an error followed by a verbal expressions.

Furthermore, in Line 11, the teacher sings the phrase ♪kon::ya::mo::♪. This phrase is the same phrase that the student sung; however, the teacher sings it another way using different bodily movements. The teacher moves their body forward slowly without stopping as they sing. When the teacher is singing, the student sings the same phrase at the same time to ensure it is done in the right way (Arano 2020). In addition, the teacher shows that this way is more appropriate by stating “te yatta hōga yoi” (“It is better to ...”). In other words, in Lines 9 and 11, the teacher characterizes the place that requires instruction and presents an example of how to correct it. The teacher demonstrates the

difference between the two by drawing a contrast, thereby helping the student understand how to improve their singing.

The teacher cannot visually demonstrate and teach the singing voice. Therefore, the teacher quotes the student's singing, characterizes it, identifies the problem, and demonstrates how to improve it. The teacher demonstrates and contrasts several types of singing in a single line, coupling physical movements with singing and visually structuring singing characteristics.

(3) Confirmation of the acquisition of teaching content and re-teaching

In the lesson, after receiving instruction, the student is told to re-sing the relevant parts. This is an important point, characteristic of lesson interactions. A lesson does not simply teach singing as it is in a music practice book. The aim is to improve the student's singing. Therefore, the student must not only understand the teacher's instruction but also reflect it in their singing.

After Line 11, the teacher tells the student that they had taught this part to another student who had entered a competition. This is something that I have learned over a long period of time from my teachers; through their own singing and teaching experiences, they have found the areas that need to be taught for each song. This is not so much that they have prepared their teaching content in advance, but rather that they knew which section they needed to focus on. After that, the teacher returns to the lesson from Line 23, indicating where to begin singing by saying, "kon::kara iku yo" ("Let's start from kon"). This is the section that the student sang in Line 8 and the teacher identified as requiring instruction. In other words, to check if the student is able to do what was instructed, the teacher chooses to begin here. Following this, the student sings in accordance with the timing of the teacher's prompting.

However, after listening to the student singing, the teacher provides additional instruction while the student is singing again, stating, "sō fukuramashite" ("Right, inflate"; Line 25) and singing the indicated part that is coupled with gestures. In fact, the student's second attempt does not adequately reflect the provided instruction; therefore, the teacher adds a different kind of instruction in Line 25.

As lessons aim to improve students' singing, a different kind of instruction may be provided if the student cannot do what the teacher instructed, and the student will be asked to sing again. Listening to the student's singing in Line 27 and singing along, the teacher does not stop at the part they are teaching and starts teaching a new part. The student's singing voice in Line 27 is used to determine whether the content of the previous instruction is embodied (3) and to identify new instruction (1).

Here, the teacher uses three types of singing in their instruction. The first is a quotation of a song where the student had something to learn, the second is an example of a song that should be sung to practice technique, and the third is an example of a song that should be sung in a different way. All three types of singing are accompanied by gestures. Each gesture clarifies the characteristics of the singing, and by contrasting them,

the content that needs to be improved was visualized. In the first case, the teacher lowers their chin when singing the “n” in “konyamo,” indicating only the “n” as foreign and explaining that they are singing it with their throat. In the first example of how singing should be done, there is no verbal explanation, but the phrase “konyamo” is sung smoothly, including the “n.” This contrast indicates that only the “n” sung in the throat should be improved and that the phrase “konyamo” should be sung as one coherent phrase. However, the presentation of this singing contrast is not enough to improve the student’s singing. Hence, the teacher has to explain with words and gestures that the singing should be “inflated.” In the relationship between the three types of singing shown in this sequence, the ideal singing is accountable.

CASE 2: HOW TO BREATHE BEFORE SINGING “AME DARO”

In this section, I examine instruction on the way to breath before singing the phrase “ame,” which improves the “a” sound in “ame.” The technique for better “singing” taught here is often the subject of instruction in vocal and choral music. Choral books state that when a sound begins with a vowel and the next phrase begins without an obstruction, it is better to shape the mouth for the next vowel (Schneider 1972). This method can be described as a way of organizing the previous phrase and the phrase to be sung as one melody. In other words, it is a way to prevent the sound of the two phrases from being changed by breath.

In the previous case, I focus on the sequential structure in my analysis. Therefore, as mentioned above, the teacher uses their own body in various ways to demonstrate the singing and combines them to make the teaching comprehensible and visible. I examine this point in depth below, through two analyses: one concerns the participation framework (Goffman 1981) for structuring singing visually and the other concerns the gestures connected to singing.

Transition of the participation framework from performance to instruction

I first look at the transition of the participation framework from the teacher listening to the student as they sang to teaching. At that point, though the teacher would turn their body toward the student, they do not turn completely and their gaze is not directed at the student. The seemingly strange use of the teacher’s body was rationally organized as a way to teach singing. In Lines 27 and 28, the teacher listens to the student sing and sings along. When the teacher hears the student’s “A:” (Line 27), the teacher says, “Right,” indicating confirmation of the part that required instruction (including Lines 9 and 25) and stopping the student’s singing (Line 28). Following this, the teacher instructs the student on their singing, saying, “At that time.” The teacher holds up an index finger in front of their face and turns their body toward the student (Fig. 2). In other words, the teacher moves beyond the framework of “singing with the student” for confirmation to

find the next “part to be taught.” This transition is clearly recognizable through the teacher’s body language and changes in body configuration.

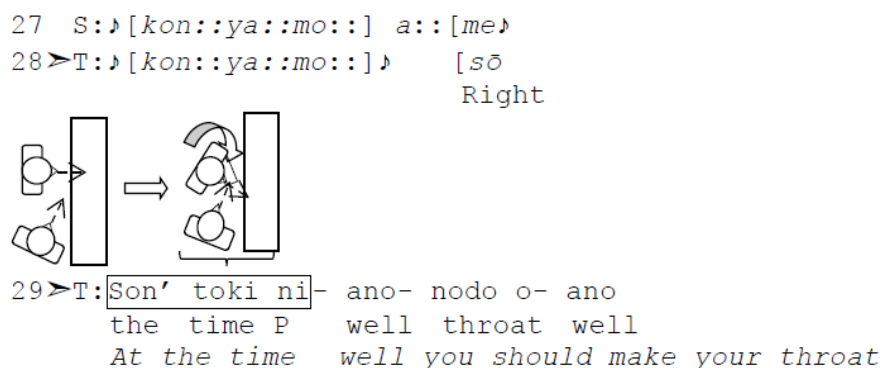


Fig.2 transition of participation framework

Figure 2

After using “sō” in Line 28 to indicate a place requiring instruction, in Line 29, the teacher brings their index finger to the front of their face. At this point, the student’s line of sight is directed toward the teacher’s hand and the area around the teacher’s face where their hand is located. In other words, this utterance and hand placement become the focus of the student’s line of sight. However, the teacher does not turn their line of sight to the student at this point. Until “tte yatta hōga kirei ni naru” (“It’s prettier when done this way”) in Line 30, the teacher’s line of sight is not directed toward the student but the space in front of the student.

This becomes a reasonable way of demonstrating singing technique. Previous studies (Goodwin 2001, Nishizaka 2008) have dealt with cases in which multiple participants manipulate an object (e.g., a tool) that is external to the body and the participants’ gaze and body orientation are directed toward the tool. For instance, when one instructor instructed a student on handling the strings in a violin performance, the focus of the student and instructor overlapped on the tools called strings, and the objects therein were structured in the mutual action (Nishizaka 2008). As such, tools external to the body, such as strings, and the participants’ speech and gestures complemented each other, making the content of the instruction understandable.

However, the body arrangement seen here is different from the method using tools, stemming from the fact that the object of instruction is not a visible object such as a string but a voice. Because instruction is based on the difference in visualizing the voice, which cannot be placed as an object between the student and instructor, the teacher shows the object to be instructed through body movements.

Instructors use linguistic resources such as those presented by Weeks (2010) and multimodal resources such as body placement and gestures (Tolins 2013) to instruct students on how to improve their singing. In singing lessons in particular, exemplary techniques

and body placement are used as the body is the instrument. The singing exemplified by this body arrangement, accompanied by gestures, becomes a technique for conveying complex nuances.

Visually characterizing the singing voice with gestures

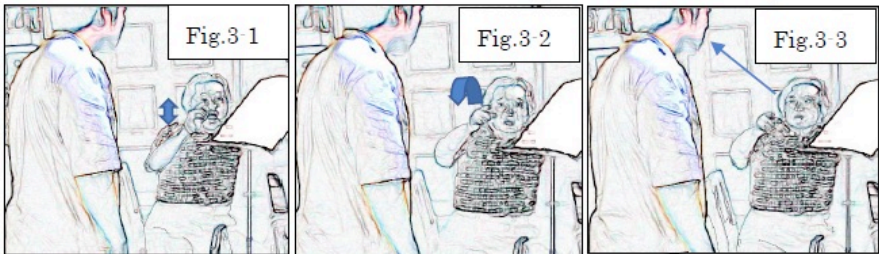
By using a participation framework to exemplify their own bodies, the teacher creates an environment that allows their singing to be visually structured. In this context, the teacher demonstrates this using gestures.

Line 29 stated, “Then, that throat.” This suggests that the content of the instruction is related to the throat. After limiting the content of the instruction, an example of singing is inserted and an explanation is added, stating, “The more you do this, the more beautiful your voice will be.” That means that the way to “become beautiful” is shown through the example. Though mentioned earlier, the teacher shifts their body configuration to illustrate this; however, shifting the body placement does not visually demonstrate the voice.

In Line 29, the teacher positions the throat in a certain way, making the student’s eyes turn to their face overlapped by their own fingers. Through that body placement, the teacher illustrates singing (Line 30), with their hand and fingers placed in front of their face.

The movement of the teacher’s hand and fingers are coupled with the song they are exemplifying. The invisible singing is visually rendered using the bodily environment. First, the teacher sings “kon::ya::mo::” while marking the beat by moving their right hand up and down, which is in the student’s line of sight. The right hand is moved as a baton that visually renders the beat (Fig. 3-1).

29>T:Son' toki ni- ano- nodo o- ano
 the time P well throat well
 At the time well you should make your throat



30>T:kon::ya::mo:: :.hhh a[::] tte yatta hōga kirei ni[naru
 Tonight too a :: P done prettier become
 kon::ya::mo:: :.hhh a[::] it's prettier when done that way
 31 S: [((nod)) [Ah ((nod))
 Yes

Figure 3

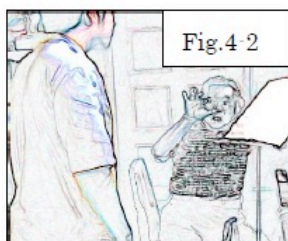
Next, the teacher's instruction focuses on the breath before "a::me" that follows "mo::." The teacher takes their right hand, which had been keeping the song's beat, and thrusts it forward with the phrase "a::me," no longer marking the beat (Fig. 3-2). At this time, the index finger, which is used as a baton, points forward. The teacher sings "a::me" while thrusting their right hand forward.

Normally, when singing, one essentially breathes out after breathing in. Breathing out (singing) and breathing in are thus in a symmetrical relationship. However, here, the teacher breathes in while thrusting their right hand forward, which is coupled with breathing in. Then, with their right hand forward, the teacher begins to sing. The gesture of sticking out the right hand visually renders the inhalation of breath and the beginning of singing as one flow. This is a way of teaching breathing, which prepares the student for the beginning of the song and to be conscious of the beginning of the song.

Assuming the body configuration connected to the singing and bodily movement, the teacher directs their line of sight toward the student and explains, "tte yatta hōga kirei ni naru" ("It's prettier when done this way") (Fig. 3-3). In response, the student utters, "Ah," indicating that they understand and accept the explanation. In other words, while the way of singing rendered here is not expressed as an utterance, the student shows that they understand the content of the teacher's instruction. Following this, the teacher continues and provides more concrete instruction, making the singing structure clearer.

Following the example explained as being prettier, the teacher immediately presents a different example. Using the voice construction shown through this contrast, it is clarified how to construct "pretty" singing. After the example "tte yatta hōga kirei ni naru" ("It's prettier when done this way"; Line 30), the next example with "sore o" (Line 32) is compared with the previous one, that is, in contrast to the "pretty" example. In other words, "not pretty" singing, or a bad example, is shown here. Omitting a diagrammatic representation, the "kon::ya::mo::" here involves different singing and bodily movement than the previous example. The teacher's singing is clearly quieter and their breathing is shallower. From the right hand that is connected to the singing, only the index finger keeps the beat. The right hand is not brought forward for the breath preceding "a::me" but stays in front of their head. In fact, there is no voice rendering in which the hand is brought forward and outlines a peak. In addition, the teacher's facial expression is different from the previous example, with their mouth spread out horizontally in the shape of a rectangle, showing a bad example of weak breathing (Fig. 4-1).

32 T: Sore o, *ko[n::ya::mo::* *a:[:]de yori mo*
 That tonight too *a::* P rather than
 Rather than *kon ya mo a::*
 33 S: *[n::ya::mo::* *[((nod))*



34 T: *Mō., hotondo., a de sutotta hōga yoi[yo*
 Almost a P breathe had better
 it's better to breathe in almost all the way at 'a'
 35 S: *[Hai ((nod))*
 Okay

Figure 4

Furthermore, after showing this bad example, the teacher says “de yori mo” (“Rather than ...”) and shows an example that contrasts with the bad one. This way of comparing demonstrates that the next example is a good one. That, in Line 30, is visually rendered by connecting singing with bodily movement. In Line 32, a bad example is contrasted by visually rendering it with singing and bodily movement. In Line 34, the teacher again explains a good example and provides an additional explanation while looking at the student’s face: “*mō, hotondo, a de sutta hōga yoi yo*” (“It’s better to breathe in almost all the way at *a*”) (Fig. 4-2). The student nods, accepting the teacher’s instruction.

The teacher’s body configuration and movement is important. The teacher turns their line of sight toward the student’s face from “de yori mo” in Line 32, where the teacher is not providing an example but explaining. At this time, the teacher’s right hand, which is used to demonstrate singing in Line 30, remains in front of their face. At the same time as “*mō, hotondo*” in Line 34, the teacher changes the shape of their right hand, bending their index finger and extending their little finger and thumb to open their closed hand (Figs. 3–6). Then, the teacher turns to the student and shows them the shape of their hand, at the same time as uttering, “*a in a de sutotta hō ga yoi*” (“It’s better to breathe in at *a*”). At this time, the teacher’s opened right hand is not a “singing body” but a bodily movement that renders the “*a de sū*” (“breathe in at *a*”). Connecting this explanation, which is difficult to understand just with words, to the bodily movement of opening their right hand, the teacher suggests that this is an issue of a space that needs to be opened up. This explanation is within the framework of the instruction regarding “*sono toki ni, nodo o*” (“At that time, with your throat ...”). Therefore, the space created by the right hand indicates the throat. That is to say, the teacher’s instruction enables the student to understand that the right hand, here connected with “*a de sū*” (“breathe in at *a*”), illustrates turning the throat into the form of “*a*” and breathing in.

First, in the rendition in Line 30, the teacher shows good singing while making motions with their body. By moving their hand forward for the breath and “a::me” sequence, they visually render singing. Next, in Line 32, having shown this example and keeping their body in the same configuration, the teacher presents a bad example that contrasts with the good example. Furthermore, adding the explanation “a de sottotta hōga yoi” (“It’s better to breathe in at a”), which connects the shape of their right hand to their throat and body at the time of singing “a,” exemplified in Line 30, is visually rendered with the body. The latter explanation related to the throat’s construction allows the student to understand more clearly the previous example of the construction of singing as well as the relationship between the breath and “a::me.”

This example uses thrusting the hand forward despite a new breath being taken. This exemplifies the throat being in the shape of “a” at the breath that precedes “a::me,” and singing “a::me” after having taken a breath with the throat in the shape of “a.”

The right hand visually renders the connected nature of the act of vocalizing “a” with the throat in the shape of “a.” While the bodily movement of the teacher in Line 30 at first glance appears contradictory, it is rationally structured. This becomes apparent in the instruction in Lines 28 to 35 (Table 1).

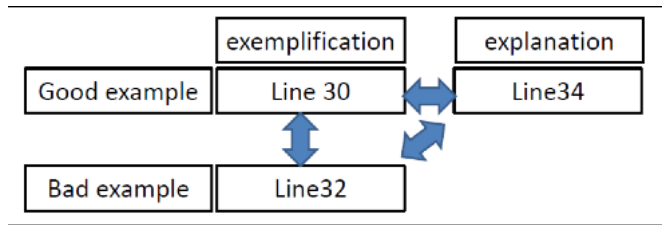


Table 1: Contrasting relationships in the instruction

The teacher proceeds to check that the student followed the instruction. In Line 36, the teacher indicates that the student should re-sing the part for which the instruction was provided, and in Line 37 the student does so. In Line 38, while listening to the student singing, the teacher shares their assessment that the student was able to do what was instructed (sō kirei; “Right, beautiful”), and continues the song. This assessment confirms that the student was able to properly sing the part that required instruction (discovered in Line 28) due to the instruction that continued to Line 35.

The student is instructed to improve their singing of “kon::ya::mo::a::me, da::ro::,” particularly the “a::me” part. As described, various resources are used that are not captured by language and sheet music. These resources are connected in the teacher–student interactions, thereby making the nuances for better singing understandable to the student, including techniques and knowledge on breathing and vocalizing parts of a word (a), to be able to sing the word “ame” better.

CONCLUDING OBSERVATIONS

This study clarifies the ways in which singing is visually structured in teaching vocal music lessons. A number of researchers have focused on teaching practices in music, examining sequential structures and the use of multimodal resources. The present study clarifies how teachers can use a combination of the various methods that these studies have shown to make it possible to understand the issues of singing that arise in the tension between words and music.

Two cases were examined. In the first case, the melody or rhythm was adequate, but the phrase “konyamo” was pronounced as a single letter instead of a single word. This was an area that needed to be improved when singing the lyrics. This instruction was analyzed by focusing on the sequential structure. Previous studies have shown three activities: identifying the problem, conveying the actual instruction, and checking to see if the student not only understood but could demonstrate the content of the instruction. In that instruction, the teacher demonstrated multiple singing techniques through gestures and contrasted them to make accountable that which needed to be improved.

In the second case, the instruction was on how to use their breath between the two phrases. There is a rest between the two phrases, but as lyrics, the two phrases are one sentence; hence, the two phrases must be expressed as one melody. The teacher taught this by focusing on breathing techniques. The second case was analyzed by focusing on the participation framework in teaching and the way the gestures were connected to the singing in the framework. The teacher connected the gestures with the singing examples to highlight the meaning and characteristics of the singing examples. In addition, these examples of singing, juxtaposing good and bad examples, and adding explanations made the main points of instruction understandable.

Both these examples of teaching aimed to sublimate the tension between the linguistic expression technique of lyrics (words) and the musical expression technique of melody and breath to enable better expression. That is to say, it was a device to enable better “vocal fine-tuning.”

The identification of the techniques inherent in the teaching of a particular piece of music, as this paper has shown, also reveals the ways in which the music’s inherent values and culture are transmitted in mutual action. Expressions specific to an artform are often addressed in critical discourse in the arts, but they are an important practical topic for practitioners of expression to master. This paper studied the practices of teachers and students and described the methods used, which is a contribution to the study of ethnomethodology and conversation analysis in the arts. Through the analysis in this paper, I am convinced that practitioners would benefit from knowing the nuances that are not written in the score and the procedures that members use to understand those nuances, as Sudnow’s (1978) study showed. It is also important for understanding the techniques of a specific musical genre or music in general, as it is the transmission of techniques specific to the art of using the body. Moreover, describing the various ways in which the body is used to create musical expression, based on the understanding of its practitioners,

is important for understanding the culture of music in two ways. First, as this paper has shown, it can reveal how the physical techniques of a particular musical expression are handed down in the institution of the lesson. Second, especially with regard to singing, the ways in which various musical expressions based on the uses of the voice are made possible can be clarified. This suggests a policy of understanding the culture of music from the aspect of singing and communicating singing, rather than the work itself.

ACKNOWLEDGEMENT

Early versions of this article were presented at the 2019 Conference of the International Institute for Ethnomethodology and Conversation Analysis in Mannheim, Germany. I am grateful to Yusuke Arano, Yuki Yoshikawa, the members of the seminar on Hongo Conceptual Analysis, and the two reviewers for their detailed, critical, and insightful comments on and the great encouragement for the earlier versions of this paper. I also thank my teacher for teaching me how difficult and wonderful it is to sing.

REFERENCES

- Arano, Yusuke. 2020. 'Doing reflecting: Embodied solitary confirmation of instructed enactment'. *Discourse Studies*, 22(3): 261–290.
- Barthes, Roland. 1977, 'The Grain of the voice', in Roland Barthes, *Image-Music-Text*: 179–189, London: Fontana Press.
- Barthes, Roland. 1991, 'Music, Voice, Language', in Roland Barthes, *The Responsibility of Forms: Critical Essays on Music, Art, and Representation*: 278–285. Berkeley, Los Angeles: University of California Press.
- Garfinkel, Harold. 1967, *Studies in Ethnomethodology*, Englewood Cliffs, NJ, Prentice-Hall.
- Garfinkel, Harold. 2006, *Seeing Sociologically: The Routine Grounds of Social Action*, Boulder, CO, Paradigm Publishers.
- Goffman, Erving. 1981, *Forms of Talk*. Philadelphia: University of Pennsylvania press.
- Goodwin, Charles. 2003 "Pointing as Situated Practice." In *Pointing: Where Language, Culture and Cognition Meet*, edited by Sotaro Kita. Mahwah, NJ: Lawrence Erlbaum: 217–241.
- Havelock, Eric A. 1963. *Preface to Plato*. Cambridge: Harvard University Press.
- Heritage, John, Clayman, Steven E. 2010, *Talk in Action: Interactions, Identities and Institutions*. Oxford, Wiley-Blackwell.
- Keevallik, Leelo. 2010. 'Bodily quoting in dance correction'. *Research on Language and Social Interaction*, 43(4): 410–426.
- Nishizaka, Aug. 2006. 'What to learn: the embodied structure of the environment', *Research on Language and Social Interaction*, 39(2): 119–154.
- Ong, Walter J. 2002. *Orality and Literacy: The Technologizing of the Word* (2nd ed). New York: Routledge.
- Ralph, Vaughan Williams. 1963. "National Music." In *National Music and Other Essays*: 1–82. London: Oxford University Press, Originally published as *National Music*. London: Oxford University Press, 1934.

- Reed, Darren, and Szczepek Reed, Beatrice. 2013. 'Building an interactional project: actions as components of music masterclasses.' In: Szczepek Reed, B., Raymond, G. (Eds.), *Units of Talk: Units of Action*. John Benjamins: 313–342.
- Reed, Darren, and Beatrice Szczepek Reed, 2014. "The emergence of learnables in music masterclasses", *Social Semiotics*, 24(4): 446–467.
- Reed, Darren J. 2015. 'Relinquishing in musical masterclasses: Embodied action in interactional projects'. *Journal of Pragmatics*, 89: 31–49.
- Sadanobu, Toshiyuki. 2002, "'Un" to "So" ni imi ha aruka' [Does the words "un" and "so" have any meaning?]' In *"un" to "so" no gengogaku [Linguistics of the "un" and "so"]*: 75–112, Hituzi-Syobo.
- Schegloff, E., Jefferson, G., Sacks, H. 1977. The preference for self-correction in the organization of repair in conversation. *Language*, 53: 361–382.
- Schneider, W. 1972. *EINSINGE IM CHOR*. Peters. Frankfurt.
- Shinichiro Sakai, Ron Korenaga, Motoko Igarashi, Yoshifumi Mizukawa. 2014. 'Envisioning the plan in interaction: Configuring pipes during a plumbers' meeting', In *Interacting with Objects: Language, Materiality, and Social Activity*, edited by Maurice Nevile, Pentti Haddington, Trine Heinemann, Mirka Rauniomaa: 339–356. Amsterdam / Philadelphia, John Benjamins.
- Shuto Yoshiki, 2008. Douyou no media ron: 1920nenndai ni okeru "koe no bun ka" no saihei (Doyo songs in media studies: The multilayers and reconfiguration of orality in the magazine Akaitori), *Japanese Sociological Review*, 59(2): 262–280.
- Sudnow, David. 1978, *Ways of the Hand: The Organization of Improvised Conduct*, London, Routledge & Kegan Paul.
- Szczepek Reed, Beatrice, Reed, Darren, Haddon, Elizabeth. 2013. 'NOW or NOT NOW: Coordinating restarts in the pursuit of learnables in vocal master classes', *Research on Language and Social Interaction*, 46(1): 22–46.
- Tolins, Jackson. 2013. 'Assessment and direction through nonlexical vocalizations in music instruction', *Research on Language and Social Interaction*, 46(1): 47–64.
- Weeks, Peter. 1996. 'A rehearsal of a Beethoven passage: an analysis of correction talk', *Research on Language and Social Interaction*, 29(3): 247–290.
- Zemel, Alan, Koschmann, Timothy. 2014. "Put your fingers right in here": Learnability and instructed experience'. *Discourse Studies*, 16(2): 163–183.

APPENDIX 1

All the extracts cited in this article are composed of three parts. At each numbered line, there is a romanized original Japanese transcript and below this are phrase-by-phrase glossaries. Finally, a rough English translation is added after each turn. In the original transcript, the most recent version of the transcription system developed by Gail Jefferson is used (see Jefferson, 2004).

The conventions of the Jeffersonian transcription system are listed below:

- [beginning of overlapping
- (.) micro pause (less than 0.1 seconds)

wo::rd	prolonged sound
°word°	soft sound
word.	intonation
word?	rising intonation
Wo-	cut-off
>word<	speedy utterance
(word)	transcriber's best guess
♪word♪	singing is transcribed in italic and eighth note delimited.
(())	transcriber's note

In phrase-by-phrase glosses, the following abbreviations are used:

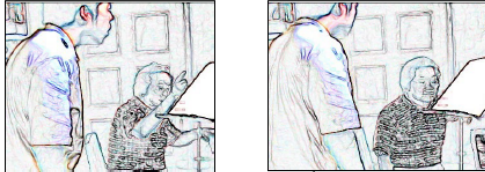
P particle

APPENDIX 2

8 S: [♪kon::ya::mo:::]
Tonight too

9 T: [°Sō son toki ni°.h ♪kon::]tte ano:nodo de >utatteru<
Right, that time P ♪kon:: ♪ P um throat with you sing
Right, at that point, you're singing kon with your throat

10 S: [°A- hai° ((nod))
Ah okay
Ah okay



11 T: ♪ko[n::]ya::mo::]tte yatta hōga[yoi(.) un
Tonight too P to had better right
It is better to do ♪kon ya mo], right

12 S: [♪n::]((nod)) [(hai) ((nod))
n yes
yes

