

**To be cited:**

Stevanovic, M., & Frick, M. (2014). Singing in interaction. *Social Semiotics*, 24(4), 495–513.  
<https://doi.org/10.1080/10350330.2014.929394>

# Singing in Interaction

Melisa Stevanovic & Maria Frick, University of Helsinki

## Abstract

Drawing on a database of 26 hours of video-recorded Finnish conversations from three different settings - everyday conversations among family and friends, instrumental lessons and church workplace meetings - we consider the ways in which singing can be used as an interactional resource to enact the three basic communicative motives of humans: requesting, informing, and sharing (see Tomasello 2008). Singing has the potential to initiate joint activities, which allows the participants to *share* their emotional stances. The usage of singing is, however, more limited in *requesting* or *informing* – a disadvantage which people, especially in musical settings, need to deal with. There are, nevertheless, situations where the possibility to choose to perform such actions through singing can help maintain mutual solidarity between participants. We argue that people's distinct ways of interpreting spoken and sung utterances can be best understood from the perspective of participants' orientations to agency and accountability.

Keywords: singing, agency, conversation analysis, instrumental lessons, communicative motives, heteroglossia, multimodality

## 1. Introduction

Human communicative actions are essentially multimodal; they are built by combining different kinds of semiotic resources (e.g. linguistic structures, prosody, gaze, body postures, gestures, facial expressions, material artifacts, etc.), whose abundance has become increasingly clear by recent empirical interaction studies (see e.g., Kendon 2004; Goodwin 2007; Mondada 2007; Enfield 2009; Heath & Luff 2012; Peräkylä & Ruusuvuori 2012; Barth-Weingarten et al. 2010; Rossano 2012). From the point of view of people's vocal conduct, there are two central modalities, speech and singing, which differ from each other in terms of prosodic features such as pitch movement, duration and rhythm. While people's spoken conduct has been subjected to empirical scrutiny already for several decades, the interactional functions of singing in naturally occurring face-to-face interactions have been dealt with only in a few empirical studies. Weeks (1996) has explored correction sequences in orchestra rehearsals and observed how the limitations of talk were displayed in the conductor's frequent use of sung, chanted, or hummed instructions, which were used in conjunction with verbal expressions. Rampton (2006) has described the social identity work that teenagers in the classroom do through singing, and Frick (2013) shows how singing can be used as a means to initiate sequence closure in everyday interaction. These studies indicate that singing can be used as a vehicle for performing communicative actions, while their exact nature

remains understandable only within the interactional context in which the singing is embedded. This study contributes to this emerging line of research by describing some of the ways in which singing can be used in interaction and on this basis discussing the strengths and weaknesses of this particular interactional resource in different action environments.

In what follows, we will first outline our conceptual framework. Then, after describing our data and method, we will turn to the empirical analysis of data extracts in which one or more participants sing. In our analysis of singing in everyday conversations, we will describe the potentials of singing in enabling joint actions. Thereafter, we will analyze participants' orientations to singing in musical settings and discuss how these differ from those in everyday interactions. Finally, we will consider singing in a setting in which the interactional import of singing is ambiguous and subject to negotiation. We will draw on Tomasello's (2008) model of cooperative communication, according to which all human communicative actions are generated by three basic motives, which emerge early in ontogeny and have plausible evolutionary roots in human social interaction. The first such motive is *requesting* - getting others to do what one wants them to do; the second motive is *informing* - helping others by telling them about things they will potentially find useful or interesting; and the third motive is *sharing* - wanting others to take part in their emotional stance towards something (cf. Svennevig 2004). In the following, we will discuss the advantages and disadvantages of singing in performing communicative actions associated with these three motives.

## 2. Agency and accountability in singing in social interaction

We will start by introducing the reader to the conceptual framework within which the examples of singing<sup>1</sup> in interaction will be subsequently analyzed.

One of Goffman's (1981) most important contributions to the study of interaction was the idea that speakerhood consists of several distinct components: the person who is in charge of the physical act of speaking (*animator*), the one who has composed the spoken message (*author*), and the one who commits to what is being said, taking responsibility for its causes and effects (*principal*). These components of speaker agency can be distributed among different actors, as is the case, for example, when the police spokesman reads a prewritten message. Importantly, however, when people speak, their utterances are normally interpreted by applying an *agent unity heuristic* (Enfield 2011) - that is, by assuming that the animator, author and principal are one and the same. Because hearers normally have access only to the animatorship, they tend to assume that if someone is the animator of the message, then s/he is also the author and the principal of the message. Spoken interaction therefore usually involves each speaker claiming "a sole entitlement to voice that speech" (Lerner 2002, 250). If people want a hearer to suspend his/her standard application of the agent unity heuristic in his/her interpretation of a spoken utterance, formal means are required to signal that something is to be heard, for example, as reported speech (cf. Goffman 1981; Couper-Kuhlen and Klewitz 1999; Günthner 1999, Holt and Clift 2007). Otherwise, people tend to assume that the different components of agency are embodied in the same person.

People's orientations to agency have implications regarding the extent to which they treat each other as accountable for their actions or utterances. In general, the greater a person's degree of agency, the more s/he can be held accountable for his/her action or utterance. Accordingly, the less freedom and control a person has over the various aspects of his/her action or utterance, the less

---

<sup>1</sup> In this paper, singing is defined as the act of vocalizing parts of songs and musical compositions (with or without lyrics). This means that not all series of sounds and words that have been produced in musical tones (e.g. sing-song prosody) "count" as instances of singing.

s/he can be praised or blamed for it, the less s/he can feel pride or shame for it, and the less s/he can be rewarded or punished for it. In sum, agency and flexibility tend to scale with accountability (Kockelman 2007). When it comes to spoken utterances, speakers are usually held as having control over their utterances. Therefore, they are regularly also held accountable for each increment they produce in a communicative sequence. However, if the agency in an utterance is distributed among different actors, some of whom may be temporarily and/or spatially far away from the contingencies of the current interaction, the producer of that utterance is commonly held as less accountable for what comes out of his/her mouth.

Given everything that has been said above, we may ask: What happens when a person, in the middle of a spoken conversation, ceases to talk and starts to draw on *singing* as a resource? What impact does such a shift have on people's orientations to agency? What assumptions about accountability underlie people's interpretations of sung utterances? Let us consider the following instance, in which one participant starts to sing a song, the lyrics of which suggest that the action performed is not quite what the co-participant might have hoped for. The recipient, however, does not respond to the proposition of the sung utterance; instead, he responds by singing the very same lyrics.

Before Extract 1, Timo has told his boyfriend, Tapio, that he is going to go buy some refreshments for himself, after which Tapio has made a rather straightforward request for Timo to bring him something too: some beer. (See Frick, 2013 for a full transcript of the extract.)

**Extract 1**

01 Tapio: Tuot mulleki,  
**Bring me too**

((insert sequence omitted))

11 Timo: mä tuon sulle yhen vähemmän,  
**I'll bring you one less,**

12 (.)

13 Timo: sä oot ottanu yhen enemmän.  
**you've had one more.**

14 (3.0)

Musical notation for Timo's first line of lyrics. The melody is written on a single staff in 2/4 time with a key signature of one flat (B-flat). The notes are: quarter note G4, quarter note A4, quarter note Bb4, quarter note C5, quarter note Bb4, quarter note A4, quarter note G4, quarter note F4, quarter note E4, quarter note D4. The lyrics are: Sil - lä se on oi - ke - us ja koh - tuus.

15 Timo: *♩* sillä se on oikeus ja kohtuus *♩*  
**because it's right and just** *♩*

16 (3.0)

Musical notation for Timo's second line of lyrics. The melody is written on a single staff in 4/4 time with a key signature of two sharps (F# and C#). The notes are: quarter note G4, quarter note A4, quarter note B4, quarter note C5, quarter note B4, quarter note A4, quarter note G4, quarter note F#4, quarter note E4, quarter note D4. The lyrics are: Se on oi - ke - us ja koh - tuus.

- 17 Tapio: ♪ see-e [on oikeus ja koh[tuuuus. ♪  
 ♪ **it is right and jus[t'** ♪ ((waving his finger))
- 18 Timo:                    [(hihi)                    [(hihihi)]
- 19 Timo:    (hi[hi])
- 20 Tapio:            [eikä sitä enää ees lauleta tollee  
 [and it's not even sung like that anymore

As a delayed response to Tapio's request, Timo states a decision to bring him one beer less, and justifies this with the fact that Tapio has already had one beer (lines 11-13). So doing, Timo calls for Tapio to accept his plan. Tapio, however, remains silent (line 14), which makes the situation appear somewhat tense (for a more detailed analysis of this extract, see Frick 2013). This is the point where Timo starts to sing: he sings the last phrase of the Eucharistic prayer (*sillä se on oikeus ja kohtuus* "for it is right and just" line 15). From the verbal perspective, Timo's turn clearly fits the participants' ongoing activity: it is a justification for Timo's plan. However, the fact that Timo actually *sings* his utterance with a known melody from a church setting, changes the dynamics of the situation: he invokes the context of church, which is quite far away from the couple's dispute in the here-and-now. Thereby Timo reduces his agency with respect to his verbal message,<sup>2</sup> which mitigates the directiveness of his previous line of action. And indeed, when Tapio subsequently takes a turn, he does not address the issue with beer at all but, instead, starts to sing the same prayer that has been sung by Timo (although with a different melody, line 17). Tapio's response is received by Timo with some amusement (lines 18-19), which ultimately breaks the tension between the participants.

In this case, a participant sings a turn that, in principle, could also have been spoken. Even if such a practice might play a relatively marginal role in everyday life, it, however, serves as a particularly clear demonstration of how people may interpret sung utterances. When people speak, they usually author their own lines, which is reflected in the interpretation of spoken utterances according to the agent unity heuristic (Enfield 2011). However, when people sing, they usually sing "something," a composition, which has been created by someone else and which is recognised as belonging to a particular genre, and perhaps associated with a particular social institution (e.g., church). What we argue, therefore, is that, whenever people sing, their musical utterances are interpreted according to a principle that we will call an *agent discontinuity heuristic*. Similarly to idiomatic expressions (Drew and Holt 1988), singing represents one form of heteroglossia which distances the speaker from what is being said (cf. Bakhtin 1982; Gumperz 1982, 34; Günthner 1999; Frick 2013).<sup>3</sup> Since this is something that can happen even in a situation where a participant fits the lyrics of his song to the verbal content of the preceding speech (Extract 1), we may assume that there is even more such distancing in sings that are not lexically fitted to the context.

In spoken interaction, participants are usually positioned quite asymmetrically with regard to each other: it is always the *speaker*, not the recipient, who has full control of what is being said and is thus accountable for that (Lerner 2002). Such asymmetries might not be a problem when people

<sup>2</sup> Something similar might have also been achieved with other heteroglossic devices, such as by voicing the utterance as a re-enactment of another person or as a quote from some text by prosodic means. In this case, the switch to singing a prayer marks a shift away from the serious mode of the interaction.

<sup>3</sup> There is another aspect to using songs amidst spoken interaction, which is that the change of mode acts as a contextualisation cue much like the usage of extreme prosody or codeswitching would (cf. Gumperz 1982). This is especially evident when people improvise songs ad hoc, meaning that there is no prior text that is voiced.

enact the basic communicative motives of requesting or informing. Indeed, these motives even presuppose asymmetries; a requester and a requestee or an informer and an informee should certainly be positioned quite differently with regard to what is being said (cf. Heritage and Raymond 2005; Stevanovic and Peräkylä 2012). However, such asymmetries can become problematic when people enact the communicative motive of sharing (Tomasello 2008: 86-87). As demonstrated by Heritage (2011), sharing an emotional stance can present quite a dilemma when people's orientations to each other's epistemic rights impose constraints on their expressions of empathy and affiliation. As we will demonstrate in the following pages, it is here that singing can become advantageous: when agency is spread between the current singer, the creator of the song and prior users of it (agent discontinuity heuristic), the participants are positioned much more symmetrically relative to each other.

In sum, singing in interaction is a special case of separation of the three components of speakerhood put forth by Goffman. In the following, we will examine the ramifications of the reduced/distributed agency that results from this separation. Through the analysis of singing in different settings, we will demonstrate (1) how singing in interaction can relax people's sole accountability for their utterances, and (2) how this results in singing having advantages in the enactment of the communicative motive of sharing and disadvantages in the enactment of the communicative motives of requesting and informing.

### 3. Data and methods

Our data consist of video-recorded Finnish interactions in three different contexts: everyday conversations within families and among friends (16 h), instrumental lessons (4h) and church workplace meetings (10h). These three data sets differ in terms of how singing relates to the surrounding talk. In everyday conversations, singing is often embedded in story telling and in music-related topics of conversation, or it is motivated by something that has just happened in the spoken interaction (see Extract 1). In instrumental lessons, singing is an essential part of the process of teaching and learning the music to be played. In church workplace meetings, the interactional functions of singing are more ambiguous: since music is often a topic of conversation and a matter of decision-making, singing can be used to endorse proposals and clarify ideas, but it can also be used to initiate the joint practicing of musical items. Altogether, our data contain 90 sequences in which one or more participants sing.

The data were transcribed according to the conventions developed by Gail Jefferson (Schegloff, 2007, 265-269) and analysed within the methodological framework of conversation analysis - an inductive, data-driven approach to the study of how people perform mutually intelligible actions in the sequential unfolding of interaction (Heritage 1984; Schegloff 2007).

### 4. Sharing: Singing in everyday conversation

In Section 2 we discussed how singing can help participants distance themselves from the contingencies of the spoken interaction. In the following, it is shown that it also puts forward something that the participants can potentially share.<sup>4</sup> The following instance from an everyday interaction exemplifies this capacity of singing: singing enables joint action and the sharing of

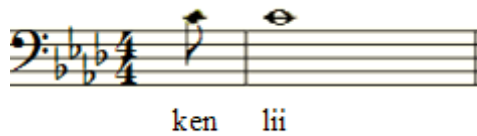
---

<sup>4</sup> Something like this happened also in Extract 1, where singing evoked the context of a church service as a "common ground" (Clark 1996) for the participants to construct their subsequent actions (singing the same song).

“emotional stance” (for the term, see, e.g., Svennevig 2004). Extract 2 is drawn from a gathering of members of a students’ association. Previously, the participants have reminisced about an unlucky attempt of a candidate in a Music Idol TV-show to perform the well-known Badfinger song *Without you* and have laughed especially at the candidate’s incomprehensible pronunciation of its refrain lyrics (*I can’t live if living is without you*).<sup>5</sup> Thereafter, the participants have turned to a discussion of other topics related to the game of dice they are playing - something that is also apparent at the beginning of the fragment (lines 1-6). Subsequently, however, one of the participants, Tuija, initiates a return to the previous singing-related topic: she starts to sing the song in question, imitating the singer’s pronunciation (line 7).

**Extract 2**

- 01 Teppo: kaks parii tai kolmiluku.  
**two pairs or a triple.**
- 02 Sari: joo-o. Pa kaks parii.  
**uhhuh. Put two pairs.**
- 03 Tuija: kaks parii.  
**two pairs.**
- 04 ?: ju-huu?  
**yay ((Teppo marks down the score))**
- 05 Teppo: **whiuu? whiuu? whiuu. ((whistling))**<sup>6</sup>
- 06 (1.0) ((Raili throws the dice.))
- 07 Tuija: **ken [lii:**
- 08 Raili: **[lii:**
- 09 Sari: [( )]



<sup>5</sup> The original lyrics of the refrain that the participants are singing are: *I can’t live, if living is without you. I can’t live, I can’t give (anymore).*

<sup>6</sup> Teppo’s whistling consists of three impulses of gliding pitch movement, which do not share any characteristics of the song soon to be sung.

10 Alex: mm-m.

11 Tuija: [♪ libutibu daut juu-u-u♪]

12 Raili: [♪ libutibu daut juu-u-u♪]

13 Teppo: [♪ libutibu daut juu-u-u♪]



li - bu - ti - bu daut juu - u - u

14 Tuija: [♪ ken liiiiii♪]

15 Raili: [♪ ken liiiiii♪]



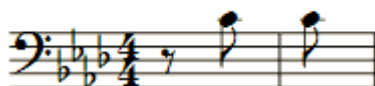
16 Alex: mitä,  
what,

ken li - ii

17 Sari: minkä maa[lainen se oli,  
from what [country did he come from,

18 Tuija: [♪ ken gii♪]

19 Teppo: se oli (.) oliko se=  
it was (.) was it=



ken gii

20 Raili: =Kroatia [(tai joku)  
=Croatia [(or some)

21 Teppo: [Kroatiasta tai jostai  
[Croatia or from some

In this case, singing is initiated at a point at which there is not much verbal interaction going on. Tuija's singing (line 7) is almost immediately joined in by Raili (line 8). Thereafter, also Teppo joins in in the singing and the phrase *libudibu daut juu* is sung together by the three participants (lines 11-13). The next phrase is sung by the two women – Tuija and Raili - only (lines 14-15). The last one to stop singing is Tuija who sings line 18 alone.

While Tuija's first sung utterance involves no signs of having been constrained by the participants' previous utterances,<sup>7</sup> the situation is quite different with respect to the participants' *next* sung utterances. From the perspective of language use, singing is extremely "formalized:" it involves constraints on the choice of words and on the prosodic parameters of pitch movement and rhythm - matters that have been determined already by the composer of the song (Bloch 1974). But this is not all. After someone has initiated singing, the rest of the interactional sequence becomes predetermined also in terms of absolute pitch and tempo (see Extract 2, lines 11-13). This means

<sup>7</sup> Speakers have been shown to deploy different kinds of markers, such as *oh*, and *by the way*, to display a distant relationship between the current talk and prior talk and thus to account for the potential inappropriateness of their talk in its current location (on "disjunct markers" and "misplacement markers," see Jefferson 1978; Schegloff and Sacks 1973). Even in those instances where people continue a topic after a temporary shift away from it, this retopicalization is frequently done by deploying certain lexical markers, such as *but anyway*, to indicate that the turn is of little relevance to the just preceding one (Adato 1979). In this case, however, there are no such markers. This makes one ask whether singing, in general, might be justifiable on bases different from those that are required to justify speaking (e.g., "earworms"; see Rampton 2006, 102-105). As an example, we may think about singing and speaking in solitude: while someone might get embarrassed by being caught singing to him/herself, it is certainly way much more fatal for a person's reputation to be caught having an animated and energetic self-conversation (Goffman 1981, 78-123).

that the unfolding of a singing sequence is highly predictable. This makes it particularly easy for several participants to construct the activity of singing as a joint one.<sup>8</sup>

The singing participants are in a situation in which none of them has full responsibility for the utterances produced – none of them is its *author* in the Goffmanian sense (cf. discussion of agent discontinuity heuristic in Section 2). They are thus positioned relatively symmetrically in relation to each other. Their mocking emotional stances towards the the Music Idol’s non-native pronunciation of English<sup>9</sup> involve no differentiation of epistemic priority inherent to conversational sequences; in these latter, the first speakers may claim epistemic rights in the matter at hand simply on the basis of the “firstness” of their turns (Heritage and Raymond 2005). In the case of singing, each participant’s sole entitlement to voice his/her utterance is relaxed, and all the participants are given the possibility to cooperate in the construction of the ongoing activity as a joint one (Lerner 2002) - something that can also happen through laughter (Jefferson et al. 1987) or concurrent agreeing assessments (Goodwin and Goodwin 1987).<sup>10</sup>

In sum, our analysis in this section has pointed to two features of the uses of singing in interaction:

1. Singing sequences are highly predetermined, which makes it easy for the participants to sing together and thus to share the responsibility for their utterances.
2. This shared responsibility enables the participants to share their emotional stances without having to deal with the asymmetries in agency inherent to spoken interaction.

## 5. Informing and requesting: Singing in musical settings

If singing is a special case of separation of Goffman’s three components of speakerhood, we may ask what that means from the perspective of those interactions in which singing is deeply intertwined with the overall activity framework of the interaction. This is the case in musical settings such as in instrumental lessons, in which singing is an essential part of the process of teaching and learning the music to be played. Even if singing in musical settings can be (and often is) embedded in some joint activity of the participants, it is not about the kind of spontaneous sharing of emotional stance that was seen in the extract from an everyday setting. Instead, singing is used by participants in musical settings to perform actions generated by the two inherently asymmetric communicative motives: informing and requesting (Tomasello 2008; see our discussion in Section 2).

Extract 3 represents an instance in which singing is associated with the communicative motive of *informing*; it is used as a way to demonstrate to the recipient the characteristics of the musical tones she is supposed to play. This instance is drawn from a violin lesson of five-year old Pinja (P). The fragment starts in the middle of a violin piece that Pinja has played *pizzicato* (a playing technique in which the player plucks the strings of his/her string instrument; indicated by \* in the transcript), while the teacher (T) has assisted her by telling her the letter names of the notes to be played next.

---

<sup>8</sup> There can, however, be limits to different people’s ability and willingness to keep with the initially set musical coordinates.

<sup>9</sup> The original lyrics of the refrain that the participants are singing are: *I can’t live, if living is without you. I can’t live, I can’t give (anymore)*.

<sup>10</sup> From this perspective, the situation is essentially different for those who do not participate in the joint activity of singing; even if Sari’s spoken contribution (line 17) indicates an appreciation of her co-participants’ joint display of stance, still, by refraining from singing, she refrains from taking that stance.



This can be seen to happen also at the beginning of the fragment, where the teacher’s “aa” (line 1) is followed by the pizzicato played by Pinja (line 2). Thereafter, the teacher’s speaking and Pinja’s playing occur simultaneously (lines 3-4).

**Extract 3**

01 T: aa:?

02 P: \*

03 T: [aa aa] °tauko°  
[aa aa] °pause°

04 P: [\* \* ]

05 T: sitten on,  
then there is,

06 (.)

07 T: <pi:tkä.>  
<lo:ng.>

08 (.)

09 T: ♪ dee:[::::: dee:::] °tauko°  
♪ dee:[::::: dee:::] °pause°<sup>11</sup>



10 B: [\* \* ]

11 T: †SIInä.  
†THEre you go.

After line 4, there is only one bar left for Pinja to play. Pinja, however, seems to have problems in producing the next tone of the piece. Pinja is silent while the teacher starts her description of the forthcoming note – something that she, after a micropause (line 6), characterizes as “long” (line 7). The word *pitkä* “long” is articulated slowly and with a lengthening of the vowel *i*, which can be heard as an iconic representation of a musical tone being long. But thereafter, most probably due to the lack of Pinja’s reaction to her verbal description of the tone, the teacher ends up singing her description of the tone in question (line 9).

Immediately after the teacher has started to sing, Pinja starts to play (line 10). Singing seems thus to have helped solve Pinja’s problem. This may have to do with the “lyrics” of the teacher’s singing, which contained information about the letter name of the note she was supposed to play next. Also, the teacher’s singing offered an audible demonstration of how the tone should sound in terms of its pitch. Besides, it seems that the teacher’s singing conveyed information about the length of the tone, too; the next, and final, tone of the piece is produced by both of the participants at exactly the same time – although the previous note was “exceptional” in its length (it is the only half note in the piece that otherwise consists of quarter notes and quarter pauses). It is as if the teacher’s singing has demonstrated to Pinja exactly how long “long” is.

<sup>11</sup> In the transcription of the sung letter notes we have applied the Finnish language-based convention according to which double vowels (e.g. *aa*, *dee*) mark “ordinary” long vowel sounds, as opposed to single vowels which denote short vowel sounds. In the transcripts, but not in the music notation frames, we have additionally used colons to indicate the singing-related “extra” lengthening of these vowel sounds.

Even if the teacher’s singing was able to convey information about the exceptional length of a musical tone in a piece, however, this was done only after the same information had already been conveyed verbally. Most probably, mere singing would not have been enough to convey the same message – at least, not equally effectively. This is consistent with what has been previously suggested by Weeks (1996). Weeks pointed out that the “illustrative expressions” (singing, humming, and chanting) through which conductors in orchestra rehearsals frequently demonstrate their views on the music to be played are to be seen as indexical in that their specific sense always has to be articulated verbally. Verbal expressions, which usually precede the illustrative expressions, instruct orchestra musicians prospectively how to hear the illustrative expressions. This dependence of illustrative expressions on the verbal expressions is also manifested in the pattern that only verbal expressions can be used to change the topic (Weeks 1996).

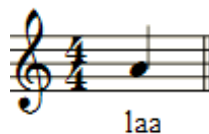
In speech and conversation, one of the most central functions of prosody is to structure the information provided by spoken utterances in terms of its relevance: the most relevant parts of the utterances are spoken slowly and with a relatively high number of accents, while parenthetical comments are spoken fast and with a relatively low number of accents (Uhmann 1992). Thereby, speakers can truly underscore their agency with regard to their utterances. In singing, however, where the prosodic parameters of pitch movement and rhythm are predetermined, it is more difficult to indicate what part of the singing should be paid special attention to. This is why, inasmuch as singing is used to convey information, it is usually only used to *complement* information that has already been given verbally.

It can be argued that, in Extract 3, the interactional function of singing was not only to convey information, but also to serve as an encouragement – or even as a request - for the co-participant to engage in music-making. Next, we will discuss an instance in which singing is even more clearly associated with the communicative motive of *requesting*. This case (Extract 4) is drawn from the same violin lesson as the previously case. Immediately before Extract 4, the teacher (T) has told Pinja (P) her plans about what should happen next: they should sing together the violin piece that Pinja is later supposed to learn to play. However, Pinja seems to have found something more interesting to do on one of the sofas in the room, and the teacher needs to order her to come back and sit down in her chair. The fragment begins right after this, at a point at which the teacher tries to guide Pinja’s attention to the notes on the music stand (line 1).

**Extract 4**

01 T: kato. (.) seurataan tarkasti tosta nuotista, h  
see. (.) we’ll follow carefully that sheet music, h

02 T: .hhh ♪ 1:aa:♪ *Figure 1*



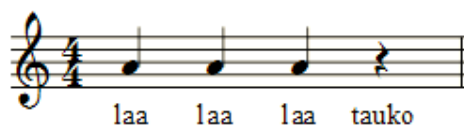
03 (0.2) ((T turns to look at P.)) *Figure 2*

04 P: ♪ la[a 1]aa♪



05 T: [voidaan laulaa.]  
[we can sing. ]

06 T: [ *laa:* laa laa ] tauko  
 [ *laa:* laa laa ] **pause**



07 P: [ *laa:* laa laa ]

08 T: [ *laa:* laa laa ] tauko  
 [ *laa:* laa laa ] **pause**



09 P: [ *laa:* laa laa ]

In line 2, the teacher starts to sing the violin piece in question while pointing to the notes on the music stand (see Figure 1). By singing *laa*, she indicates that the piece is supposed to be sung by using the “standard lyrics” (*la la la*) applicable to any piece of music. First and foremost, however, the teacher’s singing indicates that the exact point in time at which Pinja should start singing is *now*. Given the fact that the activity of singing has already been announced by the teacher (not shown in the transcript), the teacher’s singing implements an actual request for action – at the same time as it also demonstrates how the requested action should be.



Figure 1: The teacher starts singing (*laa*, line 2).



Figure 2: The teacher turns to look at Pinja (line 3).

Pinja, however, does not start singing immediately. It is only after the teacher has turned her head and gaze to look at Pinja (line 3; see Frames 2) that she starts to sing (line 4) - something that is “accompanied” by the teacher’s overlapping talk in which she makes her request for Pinja to join her in singing quite explicit (“we can sing.”<sup>12</sup> line 5). Thereafter, Pinja and the teacher sing together in unison (lines 6-9).

Recall our discussion in Section 2 that, when people try to share their emotional stances through speech, they encounter difficulties because of the inherent asymmetries of spoken interaction (Enfield 2011; Heritage 2011). The other side of the coin is, however, that when people try to request or inform through singing, they encounter difficulties because of the *lack* of such asymmetries in the sung interaction: none of the participants is the sole agent of the sung utterance but, instead, singing is something in the production of which each singer merely participates. It is from this particular perspective that singing as an interactional resource is indeed limited. While it

<sup>12</sup> The teacher’s request is made in the passive voice, which is commonly used in different kinds of instructions.

might be risky to rely on mere singing as a resource when one tries to inform one's co-participants about something (see Weeks 1996), it is even more uncertain when one wants to perform a request for the co-participant to act. While singing as a form of sharing emotional stances is essentially based on the idea of people joining in in the singing voluntarily and spontaneously, singing itself lacks a way to indicate the obligatoriness of such joint activity.

The problems of using mere singing as a way to inform or request is reflected in the fact that singing is indeed often complemented with other resources.<sup>13</sup> For example, in Extract 4, the teacher complemented her singing with three different other resources. First, she deployed spoken utterances both before she started to sing and after she had realized that her singing did not have the desired effect on her little student (cf. Weeks 1996). Second, she used a pointing gesture to establish and maintain the participants' joint attention toward the notes on the music stand, as well as to make visible her own engagement with them (cf. Mondada 2007, 199). And third, she made use of her gaze to increase the recipients' pressure to respond to her sung utterance (cf. Stivers and Rossano 2010). Even if we might have assumed that the communicative function of singing would have been understandable already on the basis of the participants' overall activity, it was still through all these means that the teacher increased her own agency with regard to her sung utterances to make them more effective in accomplishing the tasks at hand.

In sum, our analysis of singing in a context where it is used to enact the "high-agency" communicative motives of informing and requesting has shed light on certain limits of singing as an interactional resource:

1. While sung utterances can be used to convey information (e.g., demonstrate diverse aspects of musical sounds), the specific meaning of the sung utterances usually needs to be conveyed also verbally.
2. While sung utterances make it possible for the co-participants to join in in the production of the utterances, singings are not enough to request the co-participants to do this.

## 6. Negotiating the interaction frame

In the previous section, we concluded that in musical settings, in which singing is often an essential part of the "business at hand", participants may use singing when trying to convey information or to implement requests for actions, but that singing alone is often not enough to accomplish this, and the participants end up resorting to talk and other modalities as well. However, there are also situations where the possibility of choosing to perform such actions through singing can be used to maintain mutual solidarity of the participants. This is the case especially when the ultimate frame of the interaction is still to be negotiated. As a final instance to be analyzed in this paper, we will discuss a sequence from a church workplace meeting between a priest and a cantor. This setting can be characterized as "semi-musical" in that the topics of conversation are often related to the music, hymns and choir songs that are to be performed in the upcoming church events; in this setting, singing is frequently used to endorse proposals and clarify ideas. As in musical settings, however, singing can also serve as an invitation to practice musical items; in the end, it is often the participants themselves who are also supposed to perform the music that they talk about.

---

<sup>13</sup> In this respect, singing appears similar to the "bodily quotes" described by Keevallik (2010). By re-enacting others' body movements, people regularly combine bodily and vocal devices, which are used to accomplish partly complementary tasks.

In the episode from which Extract 5 is drawn, a pastor (P) and a cantor (C) are discussing the Pentecost mass. Previously, the cantor has told the pastor about his plans for the music of the mass. The extract starts in the middle of the cantor's explanation of how the transition from the opening hymn to the opening greeting should be done (lines 1-3). Subsequently, however, the pastor displays doubt whether she will manage to get the right tone for the opening greeting from the end of the opening hymn (line 4). As a reassurance for the pastor, the cantor points out that the opening hymn and the opening greeting are in the same key (lines 5-7) – something that the pastor subsequently acknowledges (line 8).

#### Extract 5

01 C: =lähtisit ninku suoraan tosta ei oom mitää a-,  
**=you would start like directly from here there are no p-**

02 (0.2) alkusoittoja eikä mitääv vaan (.) kun (.)  
**(0.2) intonations or anything just (.) when (.)**

03 ↑virsi loppuu nii, (0.3) .hhh  
**the ↑hymn ends then, (0.3) .hhh**

04 P: saanks mä siitä viimesestä, (0.3) sävelestä kiinnis s[en.]  
**will I get it from that last, (0.3) tone.**

05 C: [juu.]  
**[yea.]**

06 se on ku se on menee sa- siis  
**it is because it is goes in the sa- I mean**

07 tässä sävellaj[issa] missä,  
**in the same ke[y ] as,**

08 P: [joo.]  
**[yea.]**

09 (0.3)

10 P: .hh elikä tosta, .hhh @uu↑dista@,  
**.hh so from that, .hhh @re↑new@,**

11 C: nii,=  
**yea,=**

12 P: =tosta ↑viimesestä, (0.5) tavusta saan  
**=from that ↑last, (0.5) syllable I'll get**

13 k[iinnis sitte,]  
**i[t then, ]**

14 C: [ninku se lop]puu niin sit lähdet? omia  
**[yes when it en]ds then you'll start it in**

15 aikoj[as s]it siitä,  
**your [own ] time then,**

16 P: [joo,]  
**[yea,]**

17 (0.4)

18 P: elikä, (1.0) siinä (.) juu. ilman äänen, (0.8)  
so, (1.0) there (.) yea. without intonation, (0.8)

19 C: [nii. ↑joo.]  
[yea. ↑yea.]

20 P: [a<sub>nto</sub>a t]osta viimesest tavustah,  
[ f]rom that last syllable,

21 C: joo.  
yea.

22 (2.0)

23 P: °mä kirjotan,°  
°I'll write,°

24 (3.4) ((P is writing.))

25 C: °#kyl kyllä#.°  
°#yes yes#.° ((P is writing.))

26 (1.3) ((P is writing.))

27 P: °lähtee (--).°  
°starts (--).° ((P is writing.))



ta - di - da - ti - ta - da - dii - taa - ti - taa - daa - dai - daim

28 C: ♪ tadidatiitadadiitaatiitaadaadaida[i::m::]:♪ ((P is writing.))

29 P: [°joo,°]  
[°yea,°] ((P is writing.))

30 (0.8) ((P is writing.))

31 P: hyvä.  
okay.



tai - ta - da - dai - tai - TAA

32 C: ♪ tai:tadadaitaiTAA:::::♪

33 (1.0)

34 C: ♪ isä::n ja poja:n♪  
♪in the na:me of the fa:ther♪



I - sän ja Po - jan

36 P: joo.  
yea.

37 (.)

38 P: .hh annaksää sittev vaikka mullekki tommosen,  
.hh say will you then give me too that kind of,

39       (.) .hhh ## tommosen tota sitte,  
          (.) .hhh ## that kind of (paper) then,

Next, the pastor formulates the cantor's plan in her own words: she highlights the importance of the final word of the opening hymn *uudista* "renew" as central for her attempts to get the right tone for the opening greeting (lines 10 and 12-13). Thereafter, the cantor, instead of responding with the Finnish particle *joo* "yea," which would have confirmed the pastor's formulation (see Sorjonen 2001, 45-48), responds with the particle *nii* "yea" (line 11), which only reasserts the cantor's previously expressed plan (see Sorjonen 2001: 195-198). Remarkably, the centrality of the word *uudista* is also ignored in the cantor's subsequent utterance. Instead, the cantor points to the lack of any intonation for the opening greeting: the pastor should start it in her "own time" (lines 14-15). Again, the pastor acknowledges the cantor's explanation (line 16).

After that, the pastor makes yet another formulation of the cantor's plan (lines 18 and 20). This time, the pastor mentions both the idea that the opening greeting should start without intonation – something that has been stated by the cantor – and her own view according to which she would get the right tone from the last syllable of the final word of the opening hymn. Now, the cantor's response is more confirmatory (lines 19 and 21), which encourages the pastor to write down the plan (lines 23-30).

However, as will become apparent in the subsequent unfolding of interaction, the pastor has still not really understood how she is supposed to get the right tone for the opening greeting: the opening hymn ends with *c*, which is the tonic in C major, but, even if the opening greeting is also in C major, the pastor should start singing from the note *e*, which is one major third above the final *c* of the opening hymn. Still, to start from the right tone is something that people in these kinds of situations are likely to do right – instinctively. To explicate how they do that is something that might generate confusion, unless the person is well educated in musical theory. In this case, however, where the pastor has repeatedly stated the "wrong" strategy to find the right tone (lines 10, 12, and 20) and has possibly even written it down (line 27), one might ask whether the pastor's musical instincts are strong enough to overcome her mistaken knowledge about the matter. This is a problem that can be most easily solved by trying to sing the transition in practice.

Indeed, while the pastor finishes her writing, the cantor starts to sing the melody of the opening hymn (line 28). Then, after the pastor has finished writing, and indicated the end of the writing activity also verbally (line 31), the cantor makes a leap to the final phrase of the opening hymn and sings it with strong accents and a loud voice (line 32). Especially loud is the last note of the musical phrase, which invites the pastor to try if she can find the first tone of the opening greeting on that basis. The cantor has, however, not explicitly suggested that the pastor should practice the transition. After a silence (line 33), during which the pastor does not indicate in any way that she would want to try to find the tone, the cantor himself, with a very soft voice, sings the beginning of the pastor's part, as it were, for himself (line 34).

As pointed out in the conversation analytic research literature, human social interaction involves different kinds of threats to solidarity but, at the same time, it is ordered in ways that allow the participants to minimize these threats (Heritage 1984, 265-280; Clayman 2002). Extract 5 highlights the usefulness of *singing* from this point of view. It is used in a situation which involves subtle balancing between (1) proposing a plan to the co-participant, (2) giving her advice on how to perform her part in the realization of the plan, and (3) suggesting that she should practice that part. While all these issues can sometimes be delicate matters for the participants (on proposals, see

Houtkoop 1990; Stevanovic, submitted; on advice-giving, see e.g., Vehviläinen 2009), in this context, the last suggestion is most problematic and gives the cantor motivation to decrease his own agency and accountability in the matter. It is here, therefore, that the “limitation” of singing (decreased agency) becomes really advantageous: while the cantor’s singing offers the pastor the possibility to practice her musical performance, no one threatens her “face” (Goffman 1955) by explicitly suggesting that she should do this - and no one could hold the cantor accountable for making such a bold suggestion. When singing, the authorship of the utterance is not entirely his, the cantor’s agency is decreased and he can thus promote his motive of requesting in a much less directive manner than would have been possible through talk.

## 7. Conclusions

While human action is essentially multimodal, in this paper, we have considered the features of one particular semiotic resource that people may deploy to perform social actions: singing. Thereby, we have described how the singing-related changes in people’s orientations to agency and accountability influence the deployment of singing as a resource in enacting the three basic communicative motives: requesting, informing, and sharing (Tomasello 2008). Our analysis of singing in everyday conversations showed the ease with which singing can be constructed as a joint activity. Distributed agency and accountability during sung utterances (with the exception of people improvising songs ad hoc) makes it possible for participants to share their emotional stances without getting involved in the difficulties caused by the inherent asymmetries of spoken interaction. Then again, our analysis of singing in instrumental lessons shed light on its limitations in connection with the communicative motives of informing and requesting: to be able to enact such motives, singing usually needs to be complemented with other interactional resources, such as spoken utterances, pointing, and gaze – means through which one can underscore one’s agency with regard to the sung utterances. However, as demonstrated in our analysis of singing in a church workplace meeting, such “limitations” can also be advantageous from the perspective of the participants’ mutual solidarity.

The paper points to the need for further study in empirical interaction research: Firstly, there is much work to be done in mapping out the multiple ways in which musical performances (singing, humming, whistling, and playing an instrument) may be embedded in social interaction. We may assume that the interactional functions of such performances can be dissimilar in different contexts, everyday conversations, workplace meetings, musical settings, and so on. Secondly, much research needs to be carried out to outline the big picture of the usability of different kinds of interactional resources in connection with the three basic communicative motives described by Tomasello. For example, it is possible that embodied communication – for example, iconic representations of body movement (Keevallik 2010) - can have similar communicative functions to those that we have found singing to have. Thirdly, we should continue to develop our understanding of the distinct mechanisms through which the abundance of different semiotic resources can be used in complementary ways to maintain the mutual solidarity between the participants in interactional encounters.

## Acknowledgments

We want to thank Elizabeth Couper-Kuhlen and Mikko Kahri for their discussions on the topic and their insightful comments on earlier versions of this paper. We are also thankful for the reviewers and editors of this special issue for their invaluable help.



## Appendix: Transcription conventions

.	pitch fall
?	pitch rise
,	level pitch
↑↓	marked pitch movement
<u>          </u>	emphasis
-	truncation
[ ]	overlap
=	latching of turns
(0.5)	pause (length in tenths of a second)
(.)	micropause
:	lengthening of a sound
hhh	audible out-breath
.hhh	audible in-breath
(h)	within-speech aspiration, usually indicating laughter
£	smiley voice quality
#	creaky voice quality
◦	whisper
♪	humming or singing
*	violin pizzicato
@	other change in voice quality
tch, krhm	vocal noises
<word>	slow speech rate
>word<	fast speech rate
( )	unclear talk

## References

- Adato, Albert. 1979. Unanticipated topic continuations. *Human Studies* 2: 171-186.
- Bakhtin, Mikhail Mikhailovich. 1982 [1934]. Discourse in the novel. In *The dialogic imagination: four essays*, 259-422. Austin: University of Texas Press.
- Barth-Weingarten, Dagmar, Elisabeth Reber, and Margret Selting, eds. 2010. *Prosody in Interaction*. Amsterdam: Benjamins.
- Bloch, Maurice. 1974. Symbols, song, dance and features of articulation: Is religion an extreme form of traditional authority? *European Journal of Sociology* 15: 55-81.
- Clark, Herbert H. 1996. *Using language*. Cambridge: Cambridge University Press.
- Clayman, Steven E. 200). Sequence and solidarity. In *Advances in group processes: Group cohesion, trust, and solidarity*, ed. Edward J. Lawler and Shane R. Thye, 229-253. Oxford: Elsevier Science.
- Couper-Kuhlen, Elizabeth, and Gabriele Klewitz. 1999. Quote – unquote? The role of prosody in the contextualization of reported speech sequences. *Pragmatics* 9: 459-85
- Drew, Paul, and Elizabeth Holt. 1988. Complainable matters: The use of idiomatic expressions in making complaints. *Social Problems* 35: 398–417.

- Enfield, Nicholas. 2009. *The anatomy of meaning: Speech, gesture, and composite utterances*. Cambridge: Cambridge University Press.
- Enfield, Nicholas. 2011. Sources of asymmetry in human interaction: Enchrony, status, knowledge and agency. In *The morality of knowledge in conversation*, ed. Tanya Stivers, Lorenza Mondada, and Jacob Steensig, 285-312. Cambridge: Cambridge University Press.
- Frick, Maria. 2013. Singing and codeswitching in sequence closings. *Pragmatics* 23(2).
- Goffman, Erving. 1955. On face-work: An analysis of ritual elements in social interaction. *Psychiatry: Journal for the Study of Interpersonal Processes* 18: 213-31.
- Goffman, Erving. 1981. *Forms of talk*. Oxford: Basil Blackwell.
- Goodwin, Charles and Marjorie Harness Goodwin. 1987. Concurrent operations on talk: Notes on the interactive organisation of assessments. *IPRA Papers in Pragmatics* 1: 1-54.
- Goodwin, Charles. 2007. Environmentally coupled gestures. In *Gesture and the dynamic dimensions of language*, ed. Susan Duncan, Justine Cassell, and Elena Levy, 195-212. Amsterdam: Benjamins.
- Gumperz, John J. 1982. *Discourse strategies*. Cambridge: Cambridge University Press.
- Günthner, Susanne. 1999. Polyphony and the “layering of voices” in reported dialogues: An analysis of the use of prosodic devices in everyday reported speech. *Journal of Pragmatics* 31: 685-708.
- Heath, Christian, and Paul Luff. 2012. Embodied action and organizational activity. In *The handbook of conversation analysis*, ed. Jack Sidnell and Tanya Stivers, 283-307. Boston: Wiley-Blackwell.
- Heritage, John, and Geoffrey Raymond. 2005. The terms of agreement: Indexing epistemic authority and subordination in talk-in-interaction. *Social Psychology Quarterly* 68: 15-38.
- Heritage, John. 1984. *Garfinkel and ethnomethodology*. Cambridge: Polity Press.
- Holt, Elizabeth, and Rebecca Clift, eds. 2007. *Reporting talk: Reported speech in interaction*. Cambridge: Cambridge University Press.
- Houtkoop, Hanneke. 1990. Accounting for proposals. *Journal of Pragmatics* 14: 111-124.
- Jefferson, Gail, Harvey Sacks, and Emanuel A. Schegloff. 1987. Notes on laughter in the pursuit of intimacy. In *Talk and social organisation*, ed. Graham Button and John R.E. Lee, 152-205. Clevedon: Multilingual Matters.
- Jefferson, Gail. 1978. Sequential aspects of storytelling in conversation. In *Studies in the organization of conversational interaction*, ed. Jim Schenkein, 219-48. New York: Academic Press.
- Keevalik, Leelo. 2010. Bodily quoting in dance correction. *Research on Language and Social Interaction* 43: 401-26.
- Kendon, Adam. 2004. *Gesture: Visible action as utterance*. Cambridge: Cambridge University Press.
- Kockelman, Paul. 2007. Agency: The relationship between meaning, power, and knowledge. *Current Anthropology* 48: 375-401.
- Lerner, Gene. 2002. Turn-sharing: The choral co-production of talk-in-interaction. In *The language of turn and sequence*, ed. Cecilia Ford, Barbara Fox, and Sandra Thompson, 225-256. Oxford: Oxford University Press.

- Mondada, Lorenza. 2007. Multimodal resources for turn-taking: Pointing and the emergence of possible next speakers. *Discourse Studies* 9: 194-225.
- Peräkylä, Anssi, and Johanna Ruusuvuori. 2012. Facial expression and interactional regulation of emotion. In *Emotion in interaction*, ed. Anssi Peräkylä and Marja-Leena Sorjonen 64-91. Oxford: Oxford University Press.
- Rampton, Ben. 2006. *Language in late modernity: Interaction in an urban school*. Cambridge: Cambridge University Press.
- Rossano, Federico. 2012. Gaze in conversation. In *The Handbook of Conversation Analysis*, ed. Jack Sidnell and Tanya Stivers, 308-29. Boston: Wiley-Blackwell.
- Schegloff, Emanuel A. 2007. *Sequence organization in interaction: A primer in conversation analysis I*. Cambridge: Cambridge University Press.
- Schegloff, Emanuel A., and Harvey Sacks. 1973. Opening up closings. *Semiotica* 8: 289-327.
- Sorjonen, Marja-Leena. 2001. *Responding in conversation: A study of response particles in Finnish*. Amsterdam: Benjamins.
- Stevanovic, Melisa (submitted). Constructing a proposal as a thought: A way to manage problems in the initiation of joint decision-making in Finnish workplace interaction.
- Stevanovic, Melisa, and Anssi Peräkylä. 2012. Deontic authority in interaction: The right to announce, propose, and decide. *Research on Language and Social Interaction* 45: 297-321.
- Stivers, Tanya, and Federico Rossano. 2010. Mobilising response. *Research on Language and Social Interaction* 43: 3-31.
- Svennevig, Jan 2004. Other-repetition as display of hearing, understanding and emotional stance. *Discourse Studies* 6: 489-516.
- Tomasello, Michael. 2008. *Origins of human communication*. Cambridge, MA: MIT Press.
- Uhmann, Susanne. 1992. Contextualizing relevance: On some forms and functions of speech rate changes in everyday conversation. In *The contextualization of language*, ed. Peter Auer and Aldo di Luzio, 297-336. Amsterdam: Benjamins.
- Vehviläinen, Sanna. 2003. Avoiding providing solutions: Orienting to the ideal of students' self-directedness in counselling interaction. *Discourse Studies* 5: 131-56.
- Weeks, Peter. 1996. A rehearsal of Beethoven passage: An analysis of correction talk. *Research on Language and Social Interaction* 29: 247-90.