Leelo Keevallik*

Vocalizations in dance classes teach body knowledge

<https://doi.org/10.1515/lingvan-2020-0098>

Received November 11, 2020; accepted January 11, 2021

Abstract: Language is believed to be a central device for communicating meaning and knowledge between humans. It is superb in its capacity to code abstract ideas and displaced information, which can be conveyed from person to person, sometimes across centuries. When it comes to instructing a bodily skill in co-present situations, language is used along with other multimodal resources. This paper focuses on the role of vocalizations in dance teaching, syllables that express simultaneous body movement rather than abstract lexical content. While being essentially a vocal resource, the meaning of vocalizations arises in the simultaneously moving bodies. By carrying indexical and only partially conventionalized meaning, vocalizations constitute a puzzle for linguistic theory that preferably targets the arbitrary, symbolic and conventionalized aspects of human vocal production. The meanings conveyed from one body to another through a vocalization are experiential rather than intellectual. Vocalizations provide a solution to the problem of transferring body knowledge from one autonomous organism to another, and can even be embedded in syntax. The analysis is based on an occasion of teaching a jazz routine to a larger group of students.

Keywords: body knowledge; conventionalization; dance instruction; indexicality; non-lexical vocalization

Learning to dance involves trying out the moves with your own body, copying and mirroring experienced practitioners. Therefore, dance classes predominantly feature practice. There may, however, also be surprising amounts of language involved (Keevallik 2010, 2013). Instructors explain, dissect, clarify, historicize, correct, and segment relevant aspects of the movements through at least partially verbal means. Other aspects of the dance, such as the temporalities (rhythms and sequences), phrasing, and muscle tension can hardly be conveyed merely through a verbal code. These skills, sedimented in expert bodies, are necessarily acquired through repeated emulation and still, the individual sensations of your moving body constitute one of the biggest challenges when trying to teach other bodies to “feel the same”. Verbal devices often fall short, even though descriptive mental concepts, such as “try to think X”, are frequently used. In present-day Western societies learning a skill is frequently institutionalized as a course or a class where expert practitioners supervise and facilitate its acquisition by novices (interactional studies include Ivaldi 2016; Råman 2019; Reed and Szczepek Reed 2014; Stevanovic and Kuusisto 2019). This is also the case for swing dancing, which originated in informal communal occasions. This study targets swing classes in order to dissect the role of vocal behavior at the far edge of language, focusing on vocalizations that are used to convey bodily meanings from experts to less experienced dancers.

The vocalizations involve single sounds, separate syllables, or syllables in a series produced with prosodically coherent contours, as required by the logic of the dance moves. In contrast, words have a long history of being defined as the smallest meaningful units of speech where specific phoneme sequences are conventionally tied to semantics (Bloomfield 1926). Lexical meanings are mostly understood to be symbolic, with only a subset of words featuring iconic and depictive content, even though some languages feature a larger share of those than others (Dingemanse 2011). Similar to some ideophones analyzed in Siwu (Dingemanse 2014), the vocalizations analyzed here are not fixed in their phonetic or phonological composition but can be invented on

*Corresponding author: Leelo Keevallik, Department of Culture and Society, Linköping University, 581 83 Linköping, Sweden, E-mail: leelo.keevallik@liu.se. <https://orcid.org/0000-0003-2175-8710>

the spot. They do not feature conventionalized meaning. Recently, a strong argument has been made to include depiction in standard models of language processing (Clark 2016). Depictions are “physical scenes that people stage for others to use in imagining the scenes they are depicting” (Clark 2016: 324). In contrast to these stagings that can make sense across contexts and over time, vocalizations used during dance teaching are rendered meaningful through the moving bodies here and now. Rather than targeting a mental image they reflect the immediate bodily action, “what my own body is currently doing”, and thereby facilitate a similar action in others’ bodies. They thus feature an intimate connection with the body, hitherto undescribed in linguistic literature. As will be shown below, vocalizations are nevertheless involved in meaning-making, similar to the role of words in a language. Even though they do not feature conventionalized meanings that can be detached to a variety of contexts, the participants in dance classes orient to them as instructive aids to practice as well as having referential meaning here and now, through the unification of body with sound. This points at a fuzzy boundary between lexical and non-lexical, language and non-language in what could be seen as “liminal signs” in communication (Dingemanse 2020).

Vocalizations have been mentioned in interactional studies of musical instruction, such as master classes (Haviland 2007), or orchestra practice (Weeks 1996, 2002), where they can have both assessing and directive functions (Tolins 2013). The current paper will look further into the meaning potential of vocalizations, based on video recordings from a different setting where the resulting sound is not in focus, as is the case when a musical instrument is played. A vocalization in a dance class is not uttered to improve the sound but to teach body movements. Furthermore, differently from earlier interactional research, the paper will also target vocalizations during actual practice, in addition to those produced before or after actual performances. It will discuss the functioning of vocalizations in relation to the moving body on the one hand, and words in a language on the other. The method is multimodal interaction analysis (e.g., Goodwin 2017) that enables a close scrutiny of sense-making processes as evidenced in the co-present participants verbal and bodily actions.

The dataset comes from lindy hop¹ classes in Swedish, Estonian, and English (32.5 h in total), all of which feature vocalizations. In order to provide a qualitative and functional understanding of their situated meaning, the argument here is built on a sample instance of teaching a jazz routine in a class in English, which enables maintaining an emic participant perspective on the phenomenon throughout the analysis. The people in the recording have agreed to the materials to be used for research.

1 Vocalizing for rhythm and coordination

We will start by looking at the beginning of the class in order to exemplify the simplest form of vocalizations first and discuss the basic fit of these vocalizations to the activity. The teacher couples present themselves and immediately launch the first step (as shown in Excerpt 1). As this is an advanced class, the teachers can rely on a certain amount of prior knowledge, which is also evident in line 3 when one of them states the students have done this before. The two dancer roles within a lindy hop couple are called lead and follow, and in the transcripts the teachers are called TeaL and TeaF, respectively. The transcripts are organized around the conventional eight-beat patterns in swing music and swing dances, which provide the temporal organization for the activity. Measured pause lengths otherwise common in interactional studies would not be meaningful here, also because the patterns are danced to different musical pieces with different tempos.

In line 1 the lead teacher launches the first demonstration as well as the first segment of practice – the students are expected to join in as soon as they get the step pattern. In line 2 the teachers begin to dance and already during the first word “rock” and the first vocalization “qa” in line 2 the students are joining in. By beat 6 everybody in the picture is dancing (Figure 1).

¹ Lindy hop is a partner dance from the Swing era that originated in Harlem and incorporates elements of charleston, tap, African dances, and authentic jazz.

Figure 1: Teachers and students dancing.

Excerpt 1 Rhythm (Vinter2D02 1:23)

1 TeaL: here we go guys. we're just gonna start to do this.

beats	1	2	3	4	5	6	7	8
2 TeaL:	rock	step	kick		rock	step	kick	
TeaF:	qa		qa		qaqa	qu	qa:	

beats	1	2	3	4	5	6	7	8
3 TeaL:	rock		you've done	this	before.			
TeaF:	qa	qu	qa		qa#	qu	qa:	
					#fig.1			

beats	1	2	3	4	5	6	7	8
4 TeaF:	qa	qum	qa		qa	quw	qa:	

beats	1	2	3	4	5	6	7	8
5 TeaF:	step	step	Kick	step	step	step	Kick	step

Video 1:

What we see in this excerpt is that the follow teacher TeaF starts producing syllables as soon as the couple begins to dance. The syllables consist of a glottal stop (marked as *q* in the transcripts to ease readability) and a vowel, either an open low one (*a*) or a narrow high vowel (*u*). Most of them have the format *qa*, while *qu* is used intermittently, also occasionally as a closed syllable variant (*qum*). The syllables form patterns of three, such as *qa qu qa* (line 3), covering exactly one bar in the imaginary music (there is no music playing), with either the last syllable lengthened to cover the fourth beat (as in *qa qu qa:*), or a pause after the third beat. In one case (line 2) the pre-beat is also accompanied by a syllable (*qaqa*). In short, the vocalizations mark the rhythm of the dancing bodies, making it audible beyond the sound of the steps on the floor. The syllables fall arbitrarily on different kinds of steps but they alternate and the stronger beat is always marked with the open vowel *a*.²

In closely studying the video, it becomes clear that each syllable falls on a step on the floor or a kick in the air. They are uttered in synchrony with the follow teacher's demonstrating body. The lead teacher TeaL, in the meanwhile, utters step names that likewise fall exactly on the steps he is currently performing (*rock, step, kick*). These monosyllabic words can be uttered in chorus with the vocalizations (line 2) or on their own (line 5). Since the two teachers dance a pre-learned choreography in perfect synchrony, their syllable nuclei also fall on the very same moments, apart from the very first step and syllable by TeaF

² The author is indebted to Oliver Ehmer for this latter insight.

(in line 2, slightly delayed *qa*), when she follows TeaLs lead in launching the exercise. The vocalizations and the step names are equally functional when it comes to marking the rhythm in the dancing bodies and they can be uttered either on their own, in chorus, or during non-rhythmic talk by the other (line 3). The words, in contrast to syllables, also convey basic features of the steps, i.e., whether it is just a step, a kick, or a rock (step), which provides slightly more conceptual information to the practicing students, but still does not fully specify the nature of every move (which leg, how far, which direction). The specification of the step can only be discovered by observing the teachers' bodies and in that way both the words and vocalizations are highly indexical.

In Excerpt 1, the two teachers are repeating a four-beat step sequence, over and over, to get all the students on board. In principle, it would also be possible to do this in silence, with only the feet sounding on the floor. As can be seen in the transcript, the teachers may also talk in a normal pace during the demonstration, without marking the rhythm that their body is performing at the same time (line 3, TeaL). Speech is also used to launch the demonstration and to frame what the students need to do next (line 1). We can thus see a variety of vocal resources used in this excerpt:

- talk is used to frame the teachers' demonstration and to guide the students to participate rather than watch;
- step names are used in parallel or interchangeably with vocalizations to mark the rhythm (other data show that this also holds for beat numbers);
- step names and vocalizations are uttered in synchrony with the dancing body;
- step names and vocalizations coordinate the movements of the whole class;
- vocalizations are qualitatively arbitrary in regard to particular steps.

Reinforcing rhythm in demonstrations and coordinating class activities constitute the basic functions of vocalizations in a lindy hop class. They have, however, also an intricate role in conveying more precise knowledge about what the body should be feeling and doing at the current moment in the dance.

2 Vocalizing for pedagogical contrast

Let us look at what happens 40 s later, when TeaL adds some styling advice to the students. Being advanced dancers, they can “get fancy”, as the teacher says in Excerpt 2. In contrast to Excerpt 1 the students are now standing in a circle and observing the teacher in the middle. His demonstration is fitted into instructive talk and he is doing the steps alone. Beat numbers are marked (above the transcription line) only in places where he actually does a dance step.

Excerpt 2 Contrast (Vinter2D02 2:11)

```

1 TeaL:      for YOU guys,
2            if you wanna make it L:I:ttle bit fancier,

3            .hhh when you take your first (.) one T1Wo, (0.2) 2
4            tak- you can lift your right foot above the ground,
5            so you're not going

6            1ro::ck 2ste::p 3ki::ck 4down. 5qa:: 6qa:: 7qa::--

7            (it's) gonna look fancy by going

8            1DZ2higidi 3DZ4ha 5ki 6DZ7ha 8ki Dah

9            (you) can REally get fancy,
10           IF you want.
```

Video 2:

The teacher launches his demonstration with a verbal explanation: “if you wanna make it fancier” (line 1). He then describes the moment in the dance where this becomes relevant “when you take your first one two” and

what needs to be done “tak-you can lift your right foot above the ground” (line 4). Only then does he perform what this will actually look like. In line 6 he is showing the simple way of dancing the step, introduced by “so you’re not going”. In line 8 he shows the fancy version, introduced by “it’s gonna look fancy by going”. The regular version is accompanied by step names and baseline vocalizations that we already heard at the beginning of the class. They are all lengthened to reflect the similarly slow and clumsy step performance, where the pattern even remains unfinished. In contrast, the fancy version is done at a considerably quicker pace, elegantly, and accompanied by intricate syllables (line 8). Their production is not monotonous but has clear phrasing, stressed syllables interchanging with less prominent ones and the pattern completed. The syllable combinations are likewise arbitrary in relation to what exactly the body is doing but the prosodical prominences index the synchronous accents in the teacher’s body. In addition, they clearly mark the pedagogical contrast between the two demonstrations for the benefit of the students: one plain and dull, the other one lively and exciting. This is conveyed through the vocalizations, the precise choice of sounds and the complexity of syllable onset (*k* vs. *dzh*), as well as the prosody of the production. Crucially, the vocalizations are not only depicting the movement, which implies detachment from the here-and-now (Clark 2016; Dingemanse 2018), but are also essentially tied to the performing body in this very moment (Kunst 2009), and thus function as natural “response cries” (Goffman 1981: 100–115) to what the body is currently feeling.

In summary, we can add the following observations:

- talk is used to explain and project a contrast in the instructed dance;
- vocalizations (and similarly prosodically organized step names or beat numbers) reflect the synchronous body moves, depicting their nature and the pedagogical contrast.

3 Vocalizing for complex choreography

As the class moves on and the choreography gets more complex, so do the syllables. Excerpt 3 shows an instance when the teachers demonstrate a 32-beat pattern. The follow teacher introduces it with “we(’re) gonna go” and the two teachers perform the sequence, with students occasionally trying to join in.

Excerpt 3 Choreography (Vinter3A01 8:00)

1 TeaF:	we gonna go?							
2 TeaF:	1	2	3	4	5	6	7	8
TeaI:	wah	pu	qa:		qa:		↑ZA	pu
	(back)							
3 TeaF:	1	2	3	4	5	6	7	8
teaF	qa::		qa::		qa::		HEY	qa
		x		x		x		((snaps))
4 TeaF:	1	2	3	4	5	6	7	8
	Bum		haq		haq		WAP	pum
5 TeaF:	1	2	3	4	5	6	7	8
	↑Za	pum	pum	pum	pa	pu	qa	qu qa

Video 3:

The variation of syllables in this excerpt is much larger than in Excerpt 1, as is the variation of steps. The sequence is now complex and the vocalizations form as if a long prosodic phrase with the pitch not falling until the end of line 5, which is also the end of the step sequence. In addition to the vocalizations, snaps are used to mark the rhythm in line 3. Together, the feet landing on the floor, the vocalizations, the snaps (or sometimes claps) still create the audible rhythm of the dance. While the sound of the feet is a necessary byproduct of the dance itself, the follow teacher’s expressive vocalizations and snaps reinforce the rhythm, also for her dance partner.

As to the syllables, in addition to the basic *qa* and *qu* (as perhaps favored by this particular teacher),³ there is a whole variety of sounds involved. Their relationship to the steps is not completely arbitrary. For

³ Within the communities of practice, vocalizations may of course be copied from dancer to dancer and teacher to teacher.

example, the lengthened vowels in line 3 are uttered during longer holds in the steps, thus indexically marking length. The syllables closed with glottal stops (*haq*, line 4) occur during kicks, i.e., distinct sharp moves, likewise indexically marking sharpness. The velar fricative *h* co-occurs only with kicks in the air, thus possibly reflecting airiness. Finally, there may be some onomatopoeia involved in the two instances of a sibilant *z* (lines 2,5), as they are produced during steps where the feet are sliding on the floor, actually producing a shish sound. Rises in pitch mark the highlights in the dance: a step to the side (line 2) and a slide (line 5).

Importantly, the syllables are not used to represent a melody of a specific song (which it might be at other times; the teachers can sing during demonstration or practice). Rather, the syllables accompany the abstracted step pattern, a choreography that can and will be danced to different melodies and tempos. The vocalizations are thus not merely a substitute for music but expressions of the sensing body in the dance. In fact, they are often produced to the music, in order to reinforce what the body is or should be doing. Apparently, it was also quite common for the original dancers of lindy hop in Harlem in the 1920s and 1930s to sound while exploring movement and treat their voice as inseparable from the rest of the body (Lennart Westerlund, pers. comm.). This supports the argument that vocalizations are intimately tied to the moving body.

Indeed, the vocal tract is also an organ and this unity of voice and movement can be deployed for a variety of pedagogical purposes in dance classes. Among other things, vocalizations constitute a mnemonic tool for dancers. Teachers may explicitly ask students to first master the vocalized version only, as if a poem, before trying out the steps and only then proceed to vocally enforced bodily attempts. This practice is used with particularly complex dance rhythms, as they are believed to be easier to master in the vocal tract than with the whole body.

Thus, vocalizations are:

- indexically tied to the character of a synchronous move;
- also produced to music (thus not a music substitute);
- used as pedagogical tools and mnemonics.

4 Vocalizing for students' practice

Excerpt 4 is taken from the continuation of the class. It is time for the students to practice what the teachers showed in Excerpt 3, with partners holding each other. TeaF is now standing in the middle of the circle and watching the students around her, not dancing herself (Figure 2). Still, she utters the same type of syllables, albeit not articulatorily identical, to the step sequence.

Figure 2: Teacher vocalizing in the middle of the circle, students dancing.

Excerpt 4 Students' practice (Vinter3A01 37:15)

```

1 TeaF: hey (0.2)
2 teaF five (0.2) six (0.2) and five six se:ven and,
      x ((snaps))

3 TeaF: 1 2 3 4 5 6 7 8
      qa qu qa: qa: Wap pum

4 teaF 1 2 3 4 5 6 7 8
      qa:: qa:: qa:: HE^Y# qou^
      ^shoulder lift^
      #fig.2

5 1 2 3 4 5 6 7 8
   qa qu qa: qa: hej qu

6 1 2 3 4 5 6 7 8
   ^↑WU^ pa za pu qa qu qa qu qa::
   ^shoulder lift^

```

Video 4:

The teacher first counts in the students with a partially conventional format (line 2; Broth and Keevallik 2014: 117–119), with the word *and* characteristically placed on the very last beat before the dance should begin. After that the students begin to dance, accompanied by the vocalizing teacher. The syllables are similar to those used in other excerpts, with major stress and quality change on beat 7 in the first three eight-beat patterns (lines 3–5), and beat 1 in the fourth pattern (line 6). The basic rhythm is provided by similar simple open syllables across excerpts. A kick is accompanied by a *hey* in both the teacher demonstration (Excerpt 3 line 3) and the student performance (Excerpt 4 line 4) but many steps have different accompaniments in the two sets. For example, ↑*Za pu* and *Wap pum* are uttered during identical steps in Excerpts 3 and 4, respectively, as are *WAP pum* and *hej qu*, and ↑*ZA* and *WU*. At the same time, the syllable combination *wap pum* does not fall on a similar step in the two sets. This is a general characteristic of vocalizations in this activity – they are not conventionally tied to particular steps. There is no stable or symbolic connection between form and function, as is the case with lexical items in a language. Rather, the meaning of these vocalizations is in their indexical relationship with the dancing body: A switch to a different syllable marks a change in the body dynamics, a prosodic prominence reflects a highlight in the dance, and a sound stretch falls on longer holds or slides as opposed to short staccato steps.

By vocalizing, the teacher makes audible her body knowledge to the students in real time. Her vocalizations provide information about the character of the current step, necessary changes in the bodily trajectories, more relaxed vs. energetic moments, and sharper vs. smoother moves. This knowledge is at the same time firmly anchored in the teacher's own experience of these steps: how the body changes its trajectory, what the energy level is at every moment, how the move should feel from within the dancing body. The teacher's brief shoulder lifts in lines 4 and 6 reflect her bodily understanding of the necessary energy at this very moment in the step sequence, even though these two steps do not involve lifting shoulders. By conveying an intimate bodily experience without simultaneous movements by the speaker, the vocalizations can furthermore be understood to depict a "distal scene" (Clark 2016). The teacher's sensations cannot in this case be perceived in her own body in the "here and now", even though they have been repeatedly experienced earlier in a distal scene. Furthermore, the vocalizations depict what it should feel like for the students if and when they perform the steps correctly ("then and there", in a future scene), as the teacher is modelling what sensations ought to be salient at that moment in the dance.

The intimate tie between a move and the vocalization it occasions in one's own body provides them with the intercorporeal power to instruct other bodies and to connect them into a joint practice of very similar, if not identical, moves. Vocalizations furthermore have the capacity to connect bodies before they are actually assembled into couples, since energy levels, move qualities and trajectories will later on be transferred through non-vocal modalities within the dancing couple. They establish shared knowledge about the movements during practice, while in the actual dance this is achieved through the interconnected bodies.

5 Vocalization as reference

In addition to expressing current bodily experience, the vocal tract is also used for producing language and these two modes are intertwined. Vocalizations in a dance class regularly take the form of syllables or multisyllabic prosodic structures, comparable to words or phrases. This, among other things, enables their usage as part of utterances. Vocalizations, together with the embodied demonstration that they are a part of, can be temporally incorporated into the evolving syntactic structures (Keevallik 2013, 2014), thereby constituting a boundary phenomenon between language and non-language.

Excerpt 5 shows a student question during the same class, concerning a moment in the choreography when the dance partners have to move towards each other. The question in line 4 is about whether the partners could crash into each other.

Excerpt 5 Reference (Vinter3A01 30:37)

```

1 TeaF:      =it's just to do it.=
2 TeaL:      =yeah,=
3 TeaF:      =where's the music.
4 Stud: -> you don't crash when you do the zup ↑p(r)um,
5 TeaF:      o::h no. because we're stepping-
6            +okay. +let's do it.
            teaF:      +turns to teaL
                        +hand to teaL hand, gaze to him
7 Stud:      heh
8 Stud:      heehe
9 TeaF:      let's do it. (.)
10           so ^we have^,
            teaF,L      ^gaze down, preparation to dance^

```

Video 5:

In lines 1–3 the two teachers are negotiating whether to start with or without music. At this moment a student launches his question “you don’t crash when you do the”. The last element in this grammatical structure, a definite article, projects a noun, which is not forthcoming. Instead, there is a vocalization *zup ↑p(r)um*, produced with a marked prosody, exactly like it would have been uttered while dancing, and most probably the student also does the step (unfortunately he happens to be out of the picture). The student deploys the vocalization to refer back to a particular move that he believes will lead to a collision. The follow teacher has no problem identifying what is being referred to. She immediately launches the response in line 5 but abandons it midway to suggest to her partner that they demonstrate the whole choreography, which they do after line 10. A vocalization has thus taken the position of a noun in a noun phrase in the syntactic structure but also accomplished a reference to a dance move in this local setting. Over time, some of these references can of course sediment into form-meaning packages and become words, step names. There is no evidence for this particular vocalization being used across contexts, but here and now it has nevertheless constituted a fully functional device for reference within a grammatically well-formulated sentence.

6 Conclusion

One of the biggest challenges in instructing other bodies is conveying how performing a skill should be felt from “the inside” of the body. This paper looked at how vocalizations – sounds and syllables that do not constitute conventional words – are used to convey meanings across the dancing bodies. Vocalizations during a dance class function to structure the activities and maintain the rhythm but are also essential in embodied instruction, i.e., making salient the individual experience of one’s own moving body to others. In contrast to step names and numbers they provide distinctly indexical information. Like words, vocalizations can be extracted from their original environment and uttered without the speaker’s body doing the moves herself,

such as when coaching other bodies, when asking questions about past moves, and in practicing complex rhythms ahead of trying to perform them through the body. In all of those uses, the indexical relationship between the vocalization and the move is maintained. Vocalizations may also be embedded in emerging syntactic structures. While dance teachers use grammatical and lexical means to explain, direct, argue, and structure, vocalizations are used for depicting moves that are, or can be, performed. In contrast to conventional words, the sounds and prosodies of vocalizations can be molded for each moment, and thus constitute a more direct source of body knowledge, immediately tied to the body and its local version of the dance move. As compared to lexical items, they feature essentially *ad hoc* structure (with some degree of conventionalization in the community of practice) and receive their meaning only in relation to the moving bodies.

Vocalizing is a regular part of teacher demonstrations of dance steps, indexically making hearable the experience from within the teacher's performing body in real time. The students thus not only see what they should be learning but also hear it as a hint at what the bodily experience should be. Voice is an efficient pedagogical tool for direct access to the students' practicing bodies as it makes audible some information about what the body is or should be doing at this very moment. This is especially useful when the teacher cannot be constantly observed, such as when the students are changing directions in space and when their gaze needs to be organized within the dance. Furthermore, a dance instruction emerges as a collaborative achievement where one party may be vocalizing the bodily experience that the others are currently performing. Vocalizations at the boundary between lexical and non-lexical phenomena constitute a central real-time strategy of the intercorporeal acquisition of a bodily skill and should therefore be analyzed on a par with other sense-making devices in achieving intersubjectivity in human interaction. Furthermore, vocalizations could constitute a sample case for studying how meanings originating from embodied behavior make their way to becoming referential entities in a language.

Acknowledgements: The author is grateful for the supportive suggestions by Emily Hofstetter and the two anonymous reviewers who will recognize several of their ideas incorporated into the final version of the paper.

Research funding: This study was carried out in the project "Vocal coordination of human action", funded by the Swedish Research Council, VR 2016-00827.

Transcription conventions: <https://www.lorenzamondada.net/multimodal-transcription>, CAPS – stress; the number of caps mark the degree of prominence.

References

- Bloomfield, Leonard. 1926. A set of postulates for the science of language. *Language* 2(3). 153–164.
- Broth, Mathias & Leelo Keevallik. 2014. Getting ready to move as a couple: Accomplishing mobile formations in a dance class. *Space and Culture* 17(2). 107–121.
- Clark, Herbert H. 2016. Depicting as a method of communication. *Psychological Review* 123(3). 324–347.
- Dingemanse, Mark. 2011. Ideophones and the aesthetics of everyday language in a West-African society. *The Senses & Society* 6(1). 77–85.
- Dingemanse, Mark. 2014. Making new ideophones in Siwu: Creative depiction in conversation. *Pragmatics and Society* 5(3). 384–405.
- Dingemanse, Mark. 2018. Redrawing the margins of language: Lessons from research on ideophones. *Glossa: A Journal of General Linguistics* 3(1). 4.
- Dingemanse, Mark. 2020. Between sound and speech: Liminal signs in interaction. *Research on Language and Social Interaction* 53(1). 188–196.
- Goffman, Erving. 1981. *Forms of talk*. Philadelphia: University of Pennsylvania Press.
- Goodwin, Charles. 2017. *Co-operative action*. Cambridge: Cambridge University Press.
- Haviland, John B. 2007. Master speakers, master gesturers: A string quartet master class. In Susan D. Duncan, Elena T. Levy & Justine Cassell (eds.), *Gesture and the dynamic dimension of language: Essays in honor of David McNeill*, 147–172. Amsterdam: John Benjamins.
- Ivaldi, Antonia. 2016. Students' and teachers' orientation to learning and performing in music conservatoire lesson interactions. *Psychology of Music* 44(2). 202–218.

- Keevallik, Leelo. 2010. Bodily quoting in dance correction. *Research on Language and Social Interaction* 43(4). 1–26.
- Keevallik, Leelo. 2013. The interdependence of bodily demonstrations and clausal syntax. *Research on Language and Social Interaction* 46(1). 1–21.
- Keevallik, Leelo. 2014. Turn organization and bodily-vocal demonstrations. *Journal of Pragmatics* 65. 103–120.
- Kunst, Bojana. 2009. *The voice of the dancing body*. Zagreb: Frakcija.
- Râman, Joonas. 2019. Budo demonstrations as shared accomplishments: The modalities of guiding in the joint teaching of physical skills. *Journal of Pragmatics* 150. 17–38.
- Reed, Darren & Beatrice Szczepek Reed. 2014. The emergence of learnables in music masterclasses. *Social Semiotics* 24(4). 446–467.
- Stevanovic, Melisa & Arniika Kuusisto. 2019. Teacher directives in children's musical instrument instruction: Activity context, student cooperation, and institutional priority. *Scandinavian Journal of Educational Research* 63(7). 1022–1040.
- 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.
- Weeks, Peter. 2002. Performative error-correction in music: A problem for ethnomethodological description. *Human Studies* 25(3). 359–385.

Supplementary Material: The online version of this article offers supplementary material (<https://doi.org/10.1515/lingvan-2020-0098>).