

This is a self-archived version of an original article. This version may differ from the original in pagination and typographic details.

Author(s): Lilja, Niina; Piirainen-Marsh, Arja

Title: Recipient Design by Gestures : How Depictive Gestures Embody Actions in Cooking Instructions

Year: 2022

Version: Published version

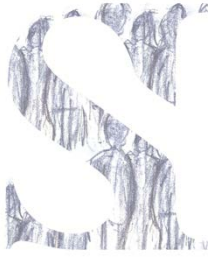
Copyright: © 2022 the Authors and Journal

Rights: CC BY-NC-ND 4.0

Rights url: <https://creativecommons.org/licenses/by-nc-nd/4.0/>

Please cite the original version:

Lilja, N., & Piirainen-Marsh, A. (2022). Recipient Design by Gestures : How Depictive Gestures Embody Actions in Cooking Instructions. *Social Interaction*, 5(1).
<https://doi.org/10.7146/si.v5i2.130874>



Social Interaction. Video-Based Studies of Human Sociality.
2022 Vol. 5, Issue 1
ISBN: 2446-3620
DOI: 10.7146/si.v5i2.130874

Social Interaction

Video-Based Studies of Human Sociality

Recipient Design by Gestures: How Depictive Gestures Embody Actions in Cooking Instructions

Niina Lilja¹ & Arja Piirainen-Marsh²

¹Tampere University

²University of Jyväskylä

Abstract

This paper investigates how depictive gestures, i.e., hand movements that depict actions, scenes or objects, are configured and used for accomplishing instructions. By drawing on video recordings of second language interactions in cooking classes for newcomers in Finland, we focus on instructions that project a certain type of complying bodily action as the relevant next action. We demonstrate that the instructions are designed to be sensitive not only to the contingencies of the material ecology of the kitchen but also to the epistemic and linguistic asymmetries between the participants. The analysis shows how depictive gestures contribute to the forward-feeding function of cooking instructions by visualizing how the instructed action should be appropriately carried out. The findings contribute to the accumulating understanding of how embodied resources further intersubjectivity in second language interactions (Eskildsen & Wagner, 2015; Greer, 2019; Lilja & Piirainen-Marsh, 2019).

Keywords: instruction, depiction, depictive gesture, recipient design, second language

1. Introduction

Depictive gestures, i.e., hand movements that depict actions, scenes or objects, are focal to human meaning making. Like other types of gestures, they are systematically deployed as part of multimodal action packages (gestalts, Mondada, 2016) in achieving, maintaining and restoring intersubjectivity (see e.g., Schegloff, 1984; Goodwin, 2000, 2007; Streeck, 2009; Enfield, 2009; Mondada, 2014b, 2016, 2018; Keevallik, 2018, Lilja & Piirainen-Marsh, 2019). Yet little is known about their role in action formation, especially in situations where talk and embodied activity are shaped by the organization of concrete activities and their local material ecologies.

This paper focuses on analyzing how depictive gestures are used as part of the multimodal design of instructions in cooking classes for participants who are newcomers to Finland, using Finnish as their second language. Instructions are social actions that make a complying second action conditionally relevant (see e.g., Lindwall, Lymer & Greiffenhagen, 2015). The second action, i.e., *instructed action* (Garfinkel, 2002), exhibits how the recipient has understood the instruction and what aspects of the instruction they treat as relevant. The cooking instructions analyzed in this paper project a certain type of complying manual-bodily action as the relevant next action. We set out to scrutinize how the instructions are multimodally designed so that they are recognizable as instructive actions and understandable to their recipients. Our analysis thus deals with the action formation problem as defined by Schegloff (2007: xiv): “*how are the resources of the language, the body, the environment of the interaction, and position in interaction fashioned into conformations designed to be, and to be recognizable by recipients as, particular actions*”. Further, it traces how multimodally assembled utterances get interpreted as instructions to accomplish a next action in a specific way.

From earlier research we know that action formation and ascription; that is, assignment of an action to a turn (or larger stretch of talk) (Levinson, 2013), involve multiple dimensions and interactional work by the participants. In analyzing action formation in its local contexts, *recipient design* (Sacks, 1995) is central: Each action is designed to its recipient(s) in ways that are sensitive to the context of ongoing activities, sequential environment, and local contingencies. In the cooking classes that serve as our data, actions are performed to advance the project of preparing food. The aim is to introduce typical Finnish dishes to the participants and to teach them how to prepare the dishes. As the participants are newcomers in Finnish society, they are neither familiar with the dishes nor with the process of preparing them. In addition, they are new to the language of instruction (Finnish). Accordingly, the classes are characterized by observable asymmetries in the participants’ knowledge about the ingredients, the process of manipulating them, and the language used while cooking. The analysis presented in this paper shows how these asymmetries are reflected in the multimodal design of the instructions and how they are manifested in the interactional work of

making sense of them in order to comply. We will demonstrate that the instructions are designed to be sensitive not only to the contingencies of the material ecology of the kitchen but also to the epistemic and linguistic asymmetries between the participants (see also De Stefani, 2018).

Our study builds both on previous research on the use and functions of depictive gestures in interaction (e.g., Arnold, 2012, Debreslioska & Gullberg, 2020) and on conversation analytic research investigating instructions accomplished in different material ecologies, e.g., crafts education (Lindwall & Ekström, 2012), dance classes (Keevallik, 2010), cooking (Mondada, 2014a, Raevaara, 2017), and driving (De Stefani, 2018). The findings add to earlier work by elucidating how depictive gestures work to enhance the recognizability of cooking instructions that make relevant a complying manual or bodily action. The gestures contribute to the forward-feeding function of cooking instructions by visualizing how the instructed action should be appropriately carried out. The findings contribute to the accumulating understanding of how embodied resources further intersubjectivity in second language interactions (Eskildsen & Wagner, 2015; Greer, 2019; Lilja & Piirainen-Marsh, 2019).

2. Instructions and instructed actions

Previous ethnomethodological and conversation analytic research has analyzed instructions in various interactional environments and situations. In this research, different understandings of the notion “instruction” are identifiable. Lindwall, Lymer & Greiffenhagen (2015) identify three different uses of the term, all of which are relevant for the analysis presented in this paper. “Instruction” can refer to 1) social actions that make a complying second action relevant, 2) the practice of teaching, and 3) directions given in written form.

First, at the core of this paper is the understanding of instructions as social actions that make relevant a complying second action. In this sense, instructions belong to a larger group of actions that are “designed to get someone else to do something” (Goodwin, 1990: 67). Such actions are recurrently referred to as *directives* or *requests* and it is not unambiguously clear how the differently labeled actions are distinct from each other. Building on earlier research, this study approaches instructions as part of a family of directive actions that aim to bring about a future action (Deppermann, 2018; De Stefani, 2018). In Finnish, as well as in many other languages, instructions and other types of directive actions can be linguistically realized in many different ways (see e.g., Etelämäki & Couper-Kuhlen, 2017; Raevaara, 2017; Rauniomaa, 2017; Rouhikoski, 2021; Stevanovic, 2017; VISK § 1645; see also Frick & Palola, 2022/this issue). In our collection, instructions are most often designed with the verb in morphological passive (extracts 1 and 2, VISK § 1655) or as zero person constructions with modal verbs (extract 4). In zero-person constructions the verb is inflected in third person singular, but the expression of a subject is missing. Therefore, the referent

of the construction is formally open and has to be inferred in the situation (see e.g., Laitinen, 1995; Couper-Kuhlen & Etelämäki, 2015), and the same applies for the passive forms. Directives can also be formulated with imperative verbs (Sorjonen, Raevaara & Couper-Kuhlen, 2017), but imperatively formatted instructions are less frequent in our data (see, however, extract 3). The variation in the linguistic design of the instructions in our collections reflects their variation in our larger data base. It is, however, beyond the scope of this paper to provide an overview of motivations for the different linguistic realizations of the instructions. Rather, our aim is to elucidate how the linguistic formulations interact with gestural resources in the multimodal accomplishment of instructions in cooking activities. The focus is on the way instructions are formulated using verbal, gestural, and other bodily resources in ways that are sensitive to the activity context and the contingencies of physical activities of the participants (see also Mondada, 2014a, De Stefani, 2018).

Previous research shows that the linguistic design of directive actions is motivated by factors such as the speaker's entitlement to give instructions or make requests (Curl & Drew, 2008). Their design also reflects other situational contingencies such as the epistemic status, agency, and the willingness of the recipient to comply, all of which provide subtle cues of the participants' social and organizational relationship (De Stefani, 2018; Drew, 2013; Mazeland, 2013). The distribution and design of instructions also display the participants' deontic authority — in other words, the participants' right to determine others' future actions (Stevanovic & Peräkylä, 2012). In our data, the focal instructions are produced by persons who are assigned the institutional role of a cooking instructor. The instructors' task is to give guidance to the participants, who are willing to learn more about preparing Finnish dishes. In this way, the analysis also speaks to the second understanding of "instruction" as teaching; that is, the institutional activities and practices that characterize pedagogical settings. In such settings, the teaching of new knowledge or skills is the main line of activity and it is usually the teacher who is responsible for both providing instructional content and directing or organizing activities.

An important feature of instructions as social actions is that they only become complete in the following action that shows how the instruction is interpreted. The following actions, i.e., the instructed actions (Garfinkel, 2002), show how the recipient has understood the instruction and what aspects of it they treat as relevant. In our data, the instructions make relevant physical and manual actions that advance the process of preparing the dishes. These may be realized either with or without accompanying speech (see also Mondada, 2014a). In settings that focus on manual or physical actions, following instructions and the competencies involved, are realized in specific movements of the body and ways of handling task-relevant objects (see Stuckenbrock, 2014; Kääntä & Piirainen-Marsh, 2013). For example, Mondada (2014a) shows how participants' expertise in cooking shows in their fluency of manipulating the ingredients and cooking utensils. Beginners face more challenges in handling both ingredients and tools

than more experienced cooks as they are unfamiliar with their uses. This is also clearly observable in our data. As the participants in the cooking classes are not familiar with the Finnish dishes they are preparing, they frequently encounter trouble in deciding how to manipulate certain ingredients. This also shows in that they frequently ask for help in deciding how to proceed with the cooking.

The third understanding of instruction, i.e., instruction as written direction, is relevant for the analysis since the preparation of different dishes is also guided by recipes. In our data, recipes are present in the situation as textual objects that are frequently consulted. There are differences, however, in following a written instruction and in being instructed by a more knowledgeable person (see also Lindwall, Lymer and Greiffenhagen, 2015). A written instruction, such as a recipe, is a general guideline that applies to any situation of preparing a certain dish. A physically present instructor, on the other hand, can design the instruction to the recipients in a specific situation. In our data, the participants often rely on the instructor's help in interpreting the written recipes.

A fourth possible understanding of instruction comes from Garfinkel (2002), who observes that specific spatial or material configurations function as instructions to human social actions (see also De Stefani, 2018). For example, the material design of the kitchen can instruct its users in how to use it. This dimension is visible in our data in the ways in which instructions as social actions are fitted to the local circumstances, including spatial arrangements and bodily alignment or movements of the participants relative to the material environment.

The analysis to follow focuses on instructions that make relevant a complying manual-bodily action by the recipient as the next step in the larger project of preparing a dish. It traces how depictive gestures contribute to multimodal design of the instructor's turns as recognizable instructions and how they are ascribed as such. We aim to show how depictive gestures elaborate the meaning of verbal instructions by providing specific information about how exactly the instructed action is to be accomplished. The gestures contribute to action formation and ascription by providing details that are not specified in the linguistic design of turns, but are relevant for (fluent) bodily accomplishment of the expected next action.

3. Depictive gestures in their material ecologies

Gestures are often defined as visible bodily actions (Kendon, 2004), or as symbolic movements related to ongoing talk (Gullberg, 2006), and categorized according to how they accomplish their meaning. In analyzing their meanings and functions, much attention has been paid to the interaction between gestures and the words they are associated with. The term *depiction* has been used in different ways in previous research (see Kendon, 2004; Streeck, 2009; Clark, 2016). Clark (2016), who considers depiction as a basic — but currently understudied — method of communication, sees depiction as a way of *showing others what*

something looks, sounds, or feels like. Depictions are thus different from, for instance, descriptions that often involve language and categorizations (Clark, 2016; see also Hsu, Brone & Feyeraerst, 2021.) In this study we build on Clark's work and define depictive gestures as hand movements that depict actions, scenes, or objects that are referred to in talk.

Conversation analytic research on gesture and other bodily actions has investigated how the material ecology of the interaction matters for what kinds of gestures are produced and how they achieve their meaning. Goodwin (2007) was one of the first to analyze how gestures are not always interpretable "within the skin of the actors" by focusing only on the gesturing hands and connected speech (Goodwin, 2007, p. 195). He showed how the material environment is relevant for the production and interpretation of gestures and launched the term "environmentally coupled gestures" to refer to gestures that cannot be understood without taking into account the structures of the environment to which they are connected and the phenomena that the gestures make relevant (such as those targeted by pointing gestures) (Goodwin 2007).

Like Goodwin, Mondada (2014b, 2016) has emphasized a holistic, multimodal approach to analyzing interaction and the accomplishment of social actions. She uses the term *Multimodal Gestalt* to refer to the way that social actions are accomplished through situated use and intertwinement of gestures, other bodily resources, materials, movement, and verbal resources. The actions in focus of this paper are organized as Multimodal Gestalts: we set out to analyze the specific role of depictive gestures in multimodally constructed instructions. In these actions, gestures are important resources for meaning making but, importantly, they never work alone. They achieve their meaning and interactional function in connection with participants' other bodily conduct and the materials that are relevant in the situation. The sections to follow detail how the cooking instructions in focus emerge as recognizable actions through coordination of verbal, gestural, and other bodily resources in ways that are sensitive to the practical actions that participants are engaged with in the material ecology of a kitchen.

Previous conversation analytic research on bodily practices in second language interaction has shown that gestures are important components of action accomplishment in a range of interactional activities and environments, including turn completion (Olsher, 2004; Mori & Hayashi, 2006), establishing reciprocity (Mortensen, 2009), displaying ongoing understanding (Eskildsen & Wagner, 2015), or willingness to participate (Evnitskaya & Berger, 2017). Recent studies demonstrate how bodily resources feature in actions such as noticing (Greer 2019), initiating and doing repair (Lilja, 2014), instructing (Eskildsen & Wagner, 2018), and explaining (Jakonen & Morton, 2015; Kääntä et al., 2018). This study contributes to this work by demonstrating how depictive gestures work as an integral part of the design of instructions and the interactional process of interpreting them (see also Lilja & Piirainen-Marsh, 2019).

4. Data and methods

The data comes from video recordings of NGO-led project that aimed to improve the possibilities of newcomers in Finland to participate in society. The project activities included cooking classes (21 hours in total). The purpose of the classes was to introduce traditional and contemporary Finnish dishes to the participants and instruct them in preparing the dishes. During the classes, the participants prepared and ate some dishes together. There were 5–10 participants and at least two instructors present in each class. The instructors work for the Martha Organization, a home economics organization that aims to promote well-being and quality of life. The classes did not have any institutional objectives to teach Finnish; rather their purpose was to support the process of integration of newcomers into Finnish society. Nevertheless, as the language of instruction was the second language for the participants, the classes also provide opportunities for language learning (Jokipohja, 2022).

Different kinds of instructions were recurrent social actions in the classes. The participants were often not familiar with the dishes that were introduced, and the instructors had to give advice at all phases of the process of preparing them. The analysis is based on a collection of 34 instruction sequences, where the cooking instructors used depictive gestures as part of multimodally accomplished instructions. After an initial interest in the use of such gestures in general, we observed that depictive gestures were a recurrent feature in instructions. Instructions that are multimodally designed to include depictive gestures were thus included in the collection. The selection of cases was also motivated by the sequential context: Our focus is on instructions that deal with concrete here-and-now matters and project physical or manual action as the relevant next action. Instructions dealing with future actions were thus excluded from the collection. We have transcribed the focal sequences using the transcription conventions developed for multimodal conversation analysis by Mondada (2018, n.d.). In addition, we illustrate the unfolding of the instructions with the help of graphic transcripts (see Laurier 2014). The analysis elucidates how the instructive actions are designed to be recognized as instructions and assigned as such actions by the recipients in the local ecology of cooking activities.

5. Analysis: Recipient design by gestures

In the sections to follow, we present four extracts illustrating the use of depictive gestures as part of the multimodal design of instructions. The focal gestures depict actions that involve objects relevant in the situation and become understandable in relation to these. The actions the gestures depict are also referred to with the verbs used in the instructions. However, as our analysis will illustrate, the gestures do not merely represent the semantic information provided by the verbs; rather, they elaborate the meaning of the turn and contribute to

action formation by providing nuanced information about how exactly the instructed actions should be performed.

The linguistic design of the instructions varies: In extracts 1 and 2, the instructions are designed with the verb in morphological passive, while in extract 3, the imperative verb form is in second person singular. In extract 4, the instruction is designed as a modal verb construction.

5.1 Depictive gestures specify how the instructed action should be performed

We begin with three short extracts that are representative of recurrent ways of designing instructions in our data. In particular, the extracts illustrate how the depictive gesture provides explicit, concrete, and specific information about how to perform the instructed action. In extract 1, the instruction is about how to pour salad dressing on top of a salad. In extract 2, the instruction addresses how to cut a leek, and in 3 it gives information about pressing lettuce leaves so that they fit inside tortilla wraps. In all three cases, the instructed action is a manual action that could be performed in different ways to achieve the same result. However, the gestures in the instructions depict a specific way of performing the actions. The gestures add a layer of meaning that is integral to the design of the turn to be recognized as an instruction and provide important cues about how to go about performing the manual action that the instruction makes relevant. For example, they depict specific hand movements, ways of using tools, and handling ingredients. The recipients' actions show that they attend to these cues in following the instructions.

Before the situation presented in extract 1, one of the cooking class participants, Ahmed, has summoned the instructor and asked what to do next with a sauce that he has been preparing and left to cool on the stove. First, the instructor checks whether the sauce is cool enough and then advises that a vanilla pod has to be removed from the sauce. This step is carried out in collaboration without talk: Ahmed takes hold of the pan and the instructor removes the vanilla pod and throws it in the trash bin (pic 1). After this, the instructor moves a salad bowl placed on the table a bit closer to herself and Ahmed (pic 2). With this action, she projects the next step in the process. As the instructor moves the salad bowl, Ahmed produces a change-of-state token, "ahaa" (pic 2), which displays a new understanding, possibly of the fact that the dressing will go with the salad. The instructor then gives the focal instruction by producing a verbal turn and co-occurring depictive gesture performed above the salad bowl (pic 3). She first turns her gaze to Ahmed and places her right hand above the salad bowl, palm open and downwards. She then says: "and then it is poured" (*ja sitten se kaadetaan*) and simultaneously makes a circular movement above the salad near the sides of the bowl (pics 3, 4, 5). The gestures depict the action of pouring the dressing on top of the salad so that it is spread equally around the bowl.

Extract 1. Salad dressing

+ delimits gestures by the INS

* delimits instructor's gaze

^ delimits gestures by AHM

```
>> Ahmed takes the vanilla stick out of the sauce (pic 1)
>> teacher takes the salad bowl and moves it closer (pic 2)

01 AHM ahaa:

02 INS +ja *sit#ten se # kaade#*taan
      *gaze to Ahmed
      +moves right hand above the salad and starts moving it
      in a circle above the salad->
      *gaze towards hands and the salad

      pic          #pic3
      pic          #pic4
      pic          #pic5

03 AHM °no ^ni:in°+
      ^moves a bit closer to the salad bowl and then pours
      the dressing on top of the salad->
      ins          ->+

04      (.)+# (.)#^
      ins          +starts walking away from the situation, smiling->
      pic          #pic6
      pic          #pic7
      ^Ahmed turns and looks towards
      other participants behind him

05      (.)

06 VID ahaa (.) okei
```



PIC 1: AHM TAKES A VANILLA POD OUT OF SAUCE



PIC 2: INS MOVES THE SALAD BOWL CLOSER



PIC 3 : BEGINNING OF INS' GESTURE



PIC 4: INS GESTURE MOVEMENT CONTINUES



PIC 5: INS' GAZE TO GESTURE



PIC 6: AHM STARTS TO POUR THE DRESSING ON TOP OF THE SALAD



PIC 7

The instruction is linguistically designed with the verb in passive declarative. The verb is preceded with the conjunction “and” (*ja*) and an adverb of time “then” (*sitten*). The extract exemplifies a typical context in which the multimodal instructions in our data are formulated with passive verb forms: the instruction orients to the preparation of the dish as a process that consists of several

successive phases. This orientation is observable especially in the adverb of time (*sitten*) but also in the use of the morphological passive. The linguistic design of the turn does not specify the addressee, which makes the turn hearable as an instruction that is part of the general procedure rather than an action specific to this addressee and this situation only.

The recognizability of the turn as a specific kind of instruction relies on the way that bodily resources, in particular gaze and depictive gesture, are deployed relative to the materials in the field of vision of the recipient. The pouring gesture above the salad bowl is very explicit as the circular movement the instructor performs with her right hand is wide and easy to notice. The gesture depicts how the pouring should be done to ensure that the dressing is spread evenly on top of the salad. The instructor addresses Ahmed with her gaze, which also enables her to observe his reactions. This allows her to monitor whether Ahmed pays attention to her gesturing and thus receives the information that it provides. When the gesture has been performed and the instructor's hand reaches home position (Schegloff, 1984) (pic 5), she turns her gaze to the bowl, away from Ahmed, and withdraws from the close bodily alignment with him. This marks the instruction as complete and opens up the space for Ahmed to comply and perform the instructed action, which he does next (pic 6). The pouring is carried out in the manner that the teacher's action depicted: Ahmed moves the pan around the salad dish carefully while pouring to spread the dressing evenly.

What is interesting in this extract is that the instructed action is accompanied with clear change-of-state tokens (pic 6) that are produced not only by Ahmed but also by other participants who are observing the situation. It seems that they treat the turn as conveying new information and that this information (possibly the fact that the sauce they have prepared is meant to be poured over the salad) is in some way surprising to them. The pedagogical character of the instructor's actions in this extract is also related to the fact that she does not perform the actual pouring of the sauce herself — even though she certainly could have done that. Instead, with her depictive gesture she models the action and lets the cooking class participant perform the actual pouring.

Extract 2 is in many ways similar to extract 1. Here, too, the instruction is linguistically designed as a declarative with the verb in passive voice (lines 2 and 3). It is performed in the context of preparing a dish that includes leek and is designed to instruct the recipient to cut a piece of leek in a specific way. Moments before the extract begins, the participants have been talking about leek and the participant Hussein has told the instructor that he has never eaten leek before. The instructor tells him what parts of the leek are edible and then explains that for the dish that they are preparing they only need “this much” (line 1). She then delivers a multimodal instruction that expresses what to do next (lines 2 & 3). The verbal turn components refer to two different actions to be performed next: the piece of leek needs to be cut and rinsed. Our focus is on the depictive gesture that makes explicit how the leek should be cut.

Extract 2. Cutting the leek in half

+ delimits gestures by the INS

^ delimits instructor's gaze

* delimits gesture by HUS

```
01 INS  +nyt (.)^ nyt +me kä- tar+vitaan vain #tämän verran (.)
        need-PASS
        now      now we us- need      only this much
        +raises  the leek with right hand to shoulder level
        and turns to Hussein->
        ^gaze to HUS->
        +touches the leek with left hand
pic     #pic1

02      ^(.)+(.)+#leikataan# hal*ki. + (.)
        cut-PASS
        is cut   in two
-> ^gaze to chopping board->
   ->+places the piece of leek on the chopping board
        +places the RH index on top of the leek, moves it down to
        the other end of the leek+
hus     *leans to take the knife
pic     #pic2      #pic3

03      ja (.)*(.)^ huuh+^#dellaan      +
        rinse-PASS
and     rinse
        -> ^gaze to water tap
        +points to water tap with left hand, open palm+
        ^gaze to HUS
hus     *takes the piece of leek in hand->
pic     #pic4

04      ^ (0.4)
ins     ^gaze down (presumably to the leek in Hussein's hands)->

05 INS  juu +kokonaan hal+ki (.) ja
        yes completely in two and
        +brings right hand forward
        +points to water tap->

06      #huuhdel+laan (.) ^hanan alla (.) vedellä (.)
        rinse-PASS
        is rinsed      under the tap with water
        -->+
        ^gaze to HUS
pic     #pic5

((INS leaves and HUS starts cutting the leek as instructed, pic6))
```



Just before initiating the verbal turn (line 2) the instructor turns her gaze from Hussein to the chopping board. This already signals the relevance of forthcoming actions that need to be attended to. She then puts the piece of leek on the chopping board and places her right-hand index finger on top of the piece of leek (pic 3). She produces the passive verb form with an adverb (is cut in two) and

simultaneously moves her index finger from the top of the leek to the other end (pic 3). Already during the gesture Hussein demonstrates understanding by preparing for the next action: He leans to a knife that is placed next to the chopping board and takes it in his hands.

After finishing the depictive cutting gesture, the instructor raises her gaze towards the faucet in front of her and points to it with her left arm while producing the first syllable of the second passive verb form of her instruction (*huuh-dellaan*, is rinsed) (pic 4). After the first syllable she turns her gaze back to Hussein who has taken the piece of leek in his hands. Similarly to extract 1, gaze direction is an important resource in designing and delivering the instruction. Here it indicates the focus of attention to the different features of the environment that are relevant for the two different actions that are referred to.

Next, the instructor gazes down, apparently towards the leek in Hussein's hands (which are not observable in the video because Hussein is standing behind a pillar). After this, she produces a second instruction that is formulated as a response to some action by Hussein (that is not observable here). The responsive character of the turn is evident in the turn initial acknowledgement token *juu*, which is followed by a statement confirming that the leek is to be cut completely and then rinsed (pic 5). It thus seems that Hussein's prior action sought for confirmation before performing the cutting. Mondada (2014) has shown that such checking sequences after cooking instructions are commonplace, especially before instructed actions that will be irreversible, such as cutting something in pieces. After the instructor has provided this confirmation, Hussein proceeds to perform the action and cuts the leek exactly as instructed.

In extract 2, the depictive gesture models the action of cutting. It is performed above the leek in the exact position where the knife is to be placed and traces the action of cutting the leek in half lengthwise. The instruction thus depicts a specific way of cutting the leek and displays the instructor's specialist knowledge that the recipient does not have. While Hussein's action of picking up the knife when the instruction is still in progress demonstrates his understanding of what he is expected to do, a full understanding of the specifics of the instruction only becomes visible in the way that he performs the cutting. The recognizability of the instruction emerges as an embodied, interactional process that involves orientations to epistemic state and authority as well as to local circumstances and features of the environment.

Similarly to extracts 1 and 2, extract 3 shows how the depictive gesture in the instruction is performed in a very explicit way. In this extract, however, the linguistic formulation and sequential environment of the instruction is different. The verbal turn is formulated with an imperative verb form and it responds to Hussein's question. Hussein is preparing tortilla wraps and asks the instructor what to do next (l. 1). The question is formulated with a modal verb and first-person reference (*mitä voin tehdä nyt*, what can I do now). The instructor gazes towards the tortillas and then instructs Hussein to press down the lettuce leaves

that are placed on the tortillas (line 3, pics 2 & 3). The use of the imperative verb is grammatically fitted to Hussein's question in that it specifies him as the agent of the instructed action. Simultaneously with the verbal turn, the instructor performs a depictive gesture, which is repeated three times. The multimodal design of the instruction highlights the local character of the instruction and directs the recipient to address a visible problem: the lettuce leaves need to be pressed down here and now because they do not seem to be flat enough to fit into the tortilla. This is evidenced also by the instructor's gaze which is directed to the lettuce prior to the instruction. She thus observes the tortilla before giving her instruction and only formulates it after having seen what needs attention.

Extract 3. *Press the lettuce*

+ delimits gestures by the INS

^ delimits instructor's gaze

* delimits gesture by HUS

```

01 HUS  Aija* (.) mitä (voin ^#tehä nyt)
        NAME-F      what (can I do now)
        *points towards tortillas on the table
        ins          ^turns gaze towards tortillas->
        pic          #pic1

02      (.)*(.)
        -->*

03 INS  +ää paina ^(.) #sa*laatteja ala[s*#päin
        press-IMP
        press      the lettuce      down
        +raises and lowers hands three times, palms open downwards->
        ^gaze towards HUS->
        hus          *moves hands in
        a similar manner twice*
        pic          #pic2          #pic3

04 HUS          [olen painanut kyllä
                I have pressed yes

05 INS  +=^joo
        yes
        ->+
        ->^turns gaze to tortillas

06      (.)

07 HUS  +mut[ta: *painan taas: (0.4) uudelleen
        but      I press again      once more
        *starts walking towards tortillas->
        ins      +starts walking towards tortillas->

08 INS  [voit pa-
        you can pr-

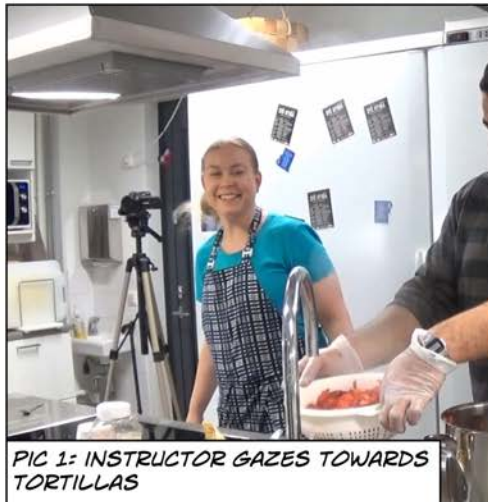
09      (.)+ #(.)
        -->+ stops and presses lettuce on tortillas with both hands->

10 HUS  uu+delleen
        again

11      (.)*#
        hus    -> *stops and presses the tortillas with both hands
        pic    #pic4

12 INS  joo (.) hyvä (.) ää sitte seuraavana
        yes      good          then next

```

The instructor performs the co-occurring gestures with both her arms. Preparation for the gesture begins before the verbal turn: The instructor raises her arms to chest level, palms facing down. She moves her arms down and up again, modeling the movement of pressing the lettuce leaves to flatten them (pics 2 & 3). Again, the gesture provides information that is not conveyed by the verb

alone: It embodies how the instructed action should be done in order to proceed to the next step. Hussein's actions show that he is closely monitoring the instructor's actions. Already during the instructor's turn, he performs what looks like a return gesture by moving his hands in a way that resembles the instructor's gesturing. He then produces a verbal claim of understanding by stating that he has already done what the instructor asks (line 4). However, following the shift of the instructor's gaze towards the tortillas, he agrees to do it again and prepares for complying by walking towards the tortillas and by then pressing the lettuce leaves exactly as instructed (l. 7–11, pic 4).

In extracts 1, 2 and 3, the depictive gestures in the instructions provide details about how to perform the instructed action that are not specified in talk. In this way they specify how the instructed action is to be carried out. In all extracts, the gaze of the instructor was important in guiding the attention to the focal gesture and then in handing the turn to the recipient. The gestures make the instruction very explicit and model the next action in such a way that the instruction is accessible to the recipient. The depiction makes the action so transparent that it might be possible to grasp how to perform the instructed action even without understanding or hearing the verbal turn. In all extracts, the recipient demonstrates understanding of the instruction by performing the action exactly as instructed. The recipients thus pay close attention to the gestures and orient to them in their actions.

5.2 Instruction becomes complete in the interpretation

The instruction in extracts 1–3 concerned single actions of pouring, cutting, and pressing. Sometimes, however, instructions deal with more complex matters and require more interactional work to deliver, make sense of and comply with. Extract 4 exemplifies one such situation. It illustrates how the instruction becomes complete in the interactional process of negotiating how it is understood and interpreted. Here the recipients' different orientations to the instructor's bodily conduct reveal what parts of the instruction they treat as relevant for the accomplishment of the next action.

Two participants (Musa and Hassan), and the instructor are figuring out the next steps in preparing a filling for a sandwich cake. The filling will contain some feta cheese and the focal instruction is about cutting the cheese up into smaller pieces so that it can be mixed with the other ingredients. This instruction is reproduced altogether 5 times in extract 4 and each realization is adjusted to the local circumstances. Our focus is on the gestures that depict the action of crumbling the cheese.

The sequence begins as the instructor walks towards Hassan and Musa holding the feta cheese, puts it on the table, points to it and simultaneously proposes cutting it up with a fork ("these could be made smaller with a fork") (l. 1, pic 1). The turn is designed as a zero-person construction and a modal verb (*vois*, could)

prior to referring to the main action. These linguistic features of the turn indicate that the instructor introduces the talked-about action as a possibility. Nevertheless, it draws the recipients' attention to the materials to be handled next, as is also visible in the direction of their gaze (pics 1 and 2).

Extract 4. Feta cheese

01 INS >>*(näitä) vois vä*hän *#<pienentää> vaikka haarukan kanssa (.)
 these could be made a bit smaller for example with a fork
**walks towards Musa and Hassan*
**puts the cheese on the table*
**point to the cheese with RH*
 pic #pic1
 >>all participants gaze towards the material on the table->

02 *tää #o fetajuustoo (.) *ja sitte se #menee tonne (.)
 this is feta cheese and then it goes there
**taps the sides of the cheese container with fingers of both hands*
 pic #pic2 **point to cheese and then to bowl with RH*
 pic #pic3

03 mut sitä *vois vähä ^pienen#tää ^[haarukan kanssa
 but it could be made a bit smaller with a fork
**moves the hands above the cheese pot, fingers*
assembled as if holding a fork or spoon
 pic #pic4

04 HAS [(- -) ((says something and
 turns to Musa))
 ^gaze to INS ^gaze toward table

05 +#(.)
 mus +points first to the cheese pot and then to bowl
 pic #pic5

06 INS joo *oikeestaan voisit pistää vaikka sen kaikki
 yes actually you could put all of it
**points to the cheese pot*

07 ^(.) pistetään kaikki (.)
 let's put all of it

08 *mut sitä vois ^haarukan kanssa #pienen [tää (.)
 but it could with a fork be made smaller this
**gesture depicting the same up and down movement as gesture in*
line 3; hand in different configuration
 mus ^takes the cheese pot in hands
 pic #pic 6

09 [(- -)

10 INS *haarukoita# on ^tossa (.) päässä (.)
 forks are there at the end
**points with left arm to left*
 has ^gaze to the direction of INS's point
 pic #pic7

11 **(--)** **(.)** **tossa missää tää nainen seisoo**
 there where that woman stands

12 ***niin siel laatikossa on ^*<haa#rukoita>**
 so there are forks in the drawer
 **retracts the pointing gesture, turn gaze to AHas*
 **depicts the cation of crumbling,*
 hands up and down ->
 pic #pic8
 ^all gaze toward the materials
 on the table

13 **(.)+**
 +Musa starts to walk towards the direction where INS pointed

14 HAS **(+^ - -)#**
 +depicts the action of crumbling with the spoon in right hand
 ^gazes toward instructor
 pic #pic9

15 INS ***+saat vähän hienon*nettua ^sitä**
 you can get it a bit crumbled
 **continues the up*
 *and down movement with RH**
 HAS *+leans towards the cheese pot, holds it with LH and*
 starts to crumble the cheese with the spoon in RH

16 **(.)**

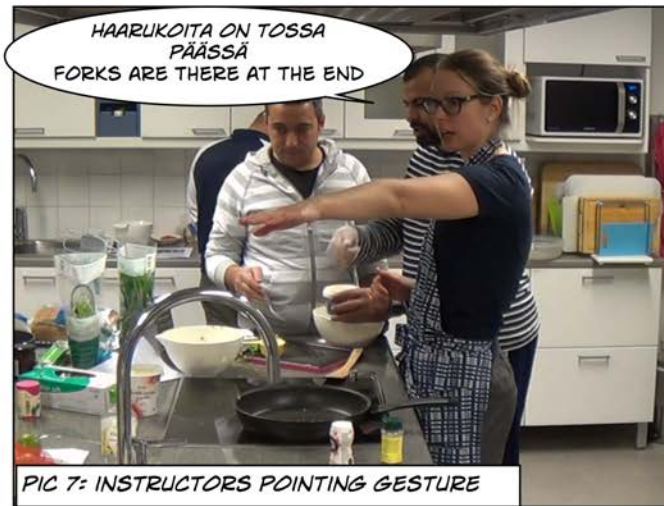
17 INS **niin (.) #jos sä saat sen kans niin onnistuu mutta**
 yeah if you get it [to crumble] with that then it works but
 pic #pic10

18 **vähän pienemmäksi sitä ***
 a bit smaller
 **starts walking away*

19 **(.)**

Hassan is the participant in the left, Musa in the middle and the instructor on the right in the pictures below.





Both Hassan and Musa direct their gaze to the materials on the table (pic 1), but do not take action to respond in any other way. In line 2, the instructor continues the turn by providing information that contextualizes the instruction. First, she taps on the sides of the pack of cheese and introduces the type of cheese (“this is feta cheese”, see pic 2). The multimodal design of this turn-constructural unit

(TCU) seems to orient to the recipient's lack of both linguistic and activity-specific knowledge. It implies that the recipients (Musa and Hassan) have not recognized the cheese and are not necessarily familiar with it. The instructor then continues her turn verbally and indicates what happens to the cheese in the next step of the cooking procedure: "it goes there" (pic 3). The deictic pronoun "it" in the beginning of the TCU is anaphoric and refers to the cheese. The deictic adverbial *tonne* (there) gets its meaning in relation to the instructor's pointing gesture: As she produces this utterance, she first points to the feta cheese and then moves her stretched index finger to point to the bowl in which the participants are mixing all the ingredients. Explaining what will happen with the cheese implies that the recipients do not know or cannot infer this in the situation. In this way, this contextualization orients to the participants as novice cooks not necessarily familiar with all the procedures.

After this, the instructor repeats the focal instruction almost verbatim (line 3, pic 4). As she verbalizes the instruction, she simultaneously moves her right hand up and down above the cheese pot on the table holding her fingers together as if she was holding a fork or some similar tool (pic 4). The gesture can be interpreted as depicting the action of crumbling the cheese. The gesture is closely intertwined with the material ecology of the situation and becomes understandable in relation to the cheese pot, in particular.

At this point, Hassan turns to Musa and says something inaudible and Musa performs a deictic gesture that the instructor attends to (pic 5). He points to the cheese pot and then to the bowl. This moving point could be interpreted as showing understanding of the instructor's earlier turn unit (l. 2) indicating that the cheese is to be mixed with the other ingredients in the bowl. The instructor, however, reacts to this more as an inquiry about the quantity of the cheese that is going to be used with her response: She suggests that Musa puts all of the cheese into the bowl (l. 6).

Next, Musa prepares to perform the instructed action by taking the cheese pot in his hands, while the instructor repeats the instruction (l. 8, pic 6) once again. This time the verbal turn is formed using a different word order: She starts with the pronoun referring to the cheese (*sitä*) and continues with the modal verb in conditional. After the modal verb, the turn continues with the adverbial "with the fork", followed by the main verb "make smaller" (*pienentää*). Simultaneously with the verbal instruction, she performs a gesture in the position where the cheese pot was previously (as it is now in Musa's hands). The arrangement of the instructor's hands is also different this time: Now, her right hand is flat, palm facing downwards (pic 6). She moves her hand up and down, and her left hand accompanies this movement on the side. The change in the form of the gesture compared to the first realization of the action (l. 3) is interesting: This time, the gesture is performed without depicting the use of a tool and it draws attention to the up-and-down movement. The second realization of the gesture is bigger and possibly also more noticeable. This suggests that it attends to the possibility that

the instruction has still not been fully understood. Again, the gesture becomes meaningful in relation to its material ecology. It is observably related to the previous realization of the gesture as it is performed exactly in the same position. It can be argued that the changes both in the word order and the gestural action serve to foreground the main action, i.e., cutting up the cheese, while treating other elements of the instruction (e.g., the tool) as less focal.

After this, the instructor redirects her focus to the tool: She informs Musa and Hassan where forks can be found. She points to her left with her left arm and states that forks are kept in a drawer at the other end of the table at which they are working. As she produces the word fork (*haarukka*) after the pointing gesture, she repeats the gesture depicting the action of crumbling (l. 11, pic. 8), this time with fingers again assembled as if holding a fork. The gesture clearly connects the informing with the instruction and orients to epistemic asymmetry. By performing the gesture while telling the recipients where forks can be found, the instructor highlights the relevance of this utensil to the manual action that she makes relevant.

At this point, Musa starts walking to the direction of the instructor's pointing gesture (l. 12). Just before he starts walking, he passes a spoon he is holding to Hassan. He thus orients to the relevance of the fork as the specific tool needed for the task and treats the instructor's turn as a request to fetch one. Hassan, for his part, visibly orients to the projected action: He says something (that is again inaudible to the analyst), raises his gaze towards the instructor and performs a similar depicting action with his right hand (pic 9). Importantly, he does not perform the actual action yet, even though the cheese pot is now again on the table. With this gesture, he thus checks his understanding of what should be done (see also Mondada 2014). The instructor joins the gesture and moves her right hand up and down, too. Hassan seems to interpret the instructor's gesture as a confirmation and next, he takes the cheese pot and starts to mash the cheese with the spoon that he is holding in his right hand (l. 15). At the same time, the instructor verbally restates the instruction, this time referring to the outcome ("get it crumbled", l. 15). Observing Hassan's actions, the teacher then expresses tentative acceptance of Hassan's use of the spoon to do the task (lines 16–17). Musa, on the other hand, is still searching for the fork (l. 20).

In this extract, the focal gesture co-occurs with verbal turns that refer to cutting feta cheese into pieces and depicts the action of mashing or crumbling the cheese. The instructor's turns get their sense as instructions from not only the way that the linguistic resources combine with the gestures to accomplish a first action, but the way that the action is embodied within the material ecology and relative to the displayed orientations of the other participants. At first, the instructor's turn that refers to the task is hearable as a suggestion that informs the recipients about the relevant next action, but also bodily instructs how it can be accomplished. This does not generate a response from the participants, which occasions further actions that both inform and make relevant the instructed

action. Each time the focal instruction is performed, it is elaborated with a depictive gesture that is performed on top of the actual cheese pot or in the position where the cheese pot previously was. The gesture thus indexes the cheese, making the instruction understandable in relation to it. The way that the different resources are assembled to communicate the instruction is sensitive to the recipients' conduct and attends to the possibility that the instruction has not yet been understood. The verbal turns repeatedly make reference to using a fork, and the first and the last versions of the gesture depict the use of a utensil, as is observable in the way that the instructor assembles her fingers. On the other hand, an observably bigger gesture performed with a flat hand palm facing down gesture puts emphasis on the projected action rather than the tool. Interestingly, the two participants orient to different parts of the instruction and gestures. While Hassan focuses on the action and finally starts to crumble the cheese with a spoon, Musa orients to the need to fetch the fork first. Both thus orient to carrying out the relevant action as instructed, but go about it in different ways. In sum, the extract illustrates the embodied interactional process through which cooking instructions get formulated, recipient designed and ascribed as such by the recipients.

6. Concluding summary and discussion

The analysis presented in this paper has illustrated how depictive gestures contribute to action formation and ascription by providing detailed guidance on how the instructed (manual) action is to be accomplished. The gestures we have analyzed depict actions that involve objects relevant in the situation, such as basic ingredients, knives, forks, and pots, and become understandable in relation to these. The actions the gestures depict are also referred to with the verbs used in the instructions, but the verbs used as semantically rather generic. The gestures affiliated with them elaborate their situated meaning and provide information related to the manner in which the instructed action is to be carried out that is not specified in the linguistic design of the instructions.

We showed how the multimodal design of instructions involves guiding the attention of the recipients with their gaze: While performing the focal gestures, the instructors also gaze toward their own gestures, thus orienting to the relevance of paying attention to the gestures. The gestures were performed in a very explicit way and coordinated closely with their verbal affiliates. By depicting the manual actions that the instruction makes relevant, the gestures visualize the meaning of the verbs used in verbal turns and tie the meaning to the local contingencies of the situation. In other words, the gestures provide specific information about how the leek needs to be cut or how the lettuce on the tortilla should be pressed for it to become flatter. By elaborating the verbal information in visual form, the gestures work towards making the instructions easily accessible also for recipients who do not hear or understand the verbal contents.

In this way they support the forward-feeding function of instructions and contribute to their recognizability as actions that move the interaction forward.

The analysis has shown that the multimodal instructions are recipient designed to orient to both epistemic and linguistic asymmetries between the participants. Sometimes the asymmetries, and especially epistemic asymmetries connected to the participants' familiarity with the ingredients and their function in the cooking process, are verbalized and readily observable in the interactions. This happens in extract 2, in which the cooking course participant, Hussein, declares he has no previous experience of eating or cooking leek. In other situations, the epistemic asymmetries are not necessarily verbalized but are still oriented to and observable in the actions of the instructor. For example, in extract 4, the instructor presented the feta cheese as an ingredient that the participants have not recognized and explained what the cheese is going to be used for (i.e., the sandwich cake filling). Presenting the cheese and explaining its use implies an orientation to the recipients as not having this understanding and knowledge.

While the epistemic asymmetries are observable in participants actions, their orientation to linguistic asymmetries is more subtle. Importantly, the participants do not explicitly orient to any clear non- or misunderstanding arising from the language of interaction. Instead, the multimodal instructions formulated in this way in our data are recurrently understood and complied with without any trouble. This shows in that the recipients usually proceed to performing the instructed actions in the way that was depicted by the gestures. It can thus be argued that the gestures work towards securing action ascription and towards securing the progressivity of the interaction (see also Lilja & Piirainen-Marsh, 2019). However, as extract 4 illustrates, actions can be layered and complex, and recipients may interpret them in different ways. The analysis shows that action ascription involves tracing the linguistic unfolding and embodied enactment of turns as well as their content, the underlying project, features of the surround and epistemic asymmetries (Levinson, 2013).

The depictive gestures are tightly connected to the material ecology of the situation as well as the bodily configurations of the participants and the sequential and temporal unfolding of the verbal instructions. The analyzed gestures are thus clearly environmentally coupled (Goodwin, 2007) and not conventional in the sense that they could be understood without the materials in the environment. Goodwin (2007) has suggested that environmentally coupled gestures may be more pervasive in certain interactional environments than in others. Kitchens are materially rich environments with a wide array of ingredients, tools and equipment needed in the process of preparing different dishes. Because of this, they may be considered as environments that motivate the participants to use gestures — environmentally coupled gestures, in particular. However, we hope to have shown that it is not only the material environment of the interaction that motivates the use of gestures, but their use is shaped by local and interactional factors such as epistemic asymmetries and displayed understandings of the participants.

References

- Arnold, L. (2012). Dialogic embodied action: using gesture to organize sequence and participation in instructional interaction. *Research on Language and Social Interaction*, 45(3), 269–296. <https://doi.org/10.1080/08351813.2012.699256>
- Clark, H. (2016). Depicting as a Method of Communication. *Psychological Review*, 123(3), 324-347. <https://doi.org/10.1037/rev0000026>
- Couper-Kuhlen, E. & M. Etelämäki (2015). Nominated actions and their targetted agents in Finnish conversational directives. *Journal of Pragmatics*, 78, 7–24. <https://doi.org/10.1016/j.pragma.2014.12.010>
- Curl, T. S., & Drew, P. (2008). Contingency and action: A comparison of two forms of requesting. *Research on Language and Social Interaction*, 41(2), 129–153. <https://doi.org/10.1080/08351810802028613>
- Debreslioska, S. & Gullberg, M. (2020). The semantic content of gestures varies with information status, definiteness and clause structure. *Journal of Pragmatics*, 168, 36 - 52. <https://doi.org/10.1016/j.pragma.2020.06.005>
- Deppermann, A. (2018). Editorial: Instructions in driving lessons. *International Journal of Applied Linguistics* 28, (2), 221 - 225. <https://doi.org/10.1111/ijal.12206>
- De Stefani, E. (2018). Formulating direction: Navigational instructions in driving lessons. *International Journal of Applied Linguistics*, 28(2), 283–303. <https://doi.org/10.1111/ijal.12197>
- Drew, P. (2013). Turn Design. In T. Stivers & J. Sidnell (Eds.), *The handbook of conversation analysis* (pp. 131-149). Hoboken, NJ: Wiley-Blackwell.
- Enfield, N. J. (2009). *The Anatomy of Meaning: Speech, Gesture, and Composite Utterances*. Cambridge: Cambridge University Press.
- Eskildsen, S. W. & Wagner, J. (2015). Embodied L2 construction learning. *Language Learning*, 65, 419–48. <https://doi.org/10.1111/lang.12106>
- Eskildsen, S. & Wagner, J. (2018). From Trouble in the Talk to New Resources: The Interplay of Bodily and Linguistic Resources in the Talk of a Speaker of English as a Second Language. In S. Pekarek Doehler, J. Wagner & E. González-Martínez (eds.), *Longitudinal Studies on the Organization of Social Interaction* (pp. 143–171). London, Palgrave Macmillan.
- Etelämäki, M. & Couper-Kuhlen, E. (2017). In the face of resistance: A Finnish practice for insisting on imperatively formatted directives. In M-L. Sorjonen,

- L. Raevaara and E. Couper-Kuhlen (eds.), *Imperative Turns at Talk: The design of directives in action*, 215–240. Amsterdam: Benjamins.
- Evnitskaya, N. & Berger, E. (2017). Learners' multimodal displays of willingness to participate in classroom interaction in the L2 and CLIL contexts. *Classroom Discourse* 8 (1), 71–94. <https://doi.org/10.1080/19463014.2016.1272062>
- Frick, M., & Palola, E. (2022/this issue). Deontic Autonomy in Family Interaction: Directive Actions and the Multimodal Organization of Going to the Bathroom. *Social Interaction. Video-Based Studies of Human Sociality*, 5(1). <https://doi.org/10.7146/si.v5i2.130870>
- Garfinkel, H. (2002). Instructions and instructed actions. In H. Garfinkel (Ed.), *Ethnomethodology's program* (pp. 197–218). Oxford: Rowman & Littlefield.
- Goodwin, M. H. (1990). *He-said-she-said. Talk as social organization among black children*. Bloomington, IN: Indiana University Press.
- Goodwin, C. (2000). Action and embodiment within situated human interaction. *Journal of Pragmatics*, 32 (10), 1489–1522. [https://doi.org/10.1016/S0378-2166\(99\)00096-X](https://doi.org/10.1016/S0378-2166(99)00096-X)
- Goodwin, C. (2007). Environmentally coupled gestures. In S. D. Duncan, J. Cassel & E. T. Levy (eds.), *Gesture and the Dynamic Dimension of Language: Essays in Honor of David McNeill* (pp. 195–212). Amsterdam / Philadelphia, John Benjamins.
- Gullberg, M (2006). Some reasons for studying gesture and second language acquisition (Hommage a Adam Kendon). *International Review of Applied Linguistics*, 44, 103-124. <https://doi.org/10.1515/IRAL.2006.004>
- Greer, T. (2019) Noticing words in the wild. In J. Hellermann, S. Eskildsen, & S. Pekarek Doehler & A. Piirainen Marsh (Eds.). *Conversation analytic research on learning-in-action: The complex ecology of L2 interaction in the wild* (pp. 131-158). Dordrecht: Springer.
- Hsu, H., Brône, G. & K. Feyaerts (2021). When gesture “takes over”: Speech-embedded nonverbal depictions in multimodal interaction. *Frontiers in Psychology*. <https://doi.org/10.3389/fpsyg.2020.552533>
- Jakonen, T. & Morton, T. (2015). Epistemic Search Sequences in Peer Interaction in a Content-based Language Classroom. *Applied Linguistics* 36(1), 73–94. <https://doi.org/10.1093/applin/amt031>
- Jokipohja, A-K. (2022). Kieltä kokkauksen lomassa – Aloittelevien suomen kielen käyttäjien sanastokysymykset. In N. Lilja, L. Eilola, A. Jokipohja & T.

Tapaninen (eds). Aikuisten maahanmuuttajien kielellinen arki – suomen kielen oppimisen mahdollisuudet aja mahdottomuudet. Tampere: Vastapaino.

Keevallik, L. (2010). Bodily quoting in dance correction. *Research on Language and Social Interaction*, 43(4), 401–426.

<https://doi.org/10.1080/08351813.2010.518065>

Keevallik, L. (2018). What does embodied interaction tell us about grammar? *Research on Language and Social Interaction*, 51, 1–21.

<https://doi.org/10.1080/08351813.2018.1413887>

Kendon, A. (2004). *Gesture: Visible action as utterance*. Cambridge University Press.

Kääntä, L., & Piirainen-Marsh, A. (2013). Manual Guiding in Peer Group Interaction: A Resource for Organizing a Practical Classroom Task. *Research on Language and Social Interaction*, 46 (4), 322-

343. <https://doi.org/10.1080/08351813.2013.839094>

Kääntä, L., Kasper, G. & Piirainen-Marsh, A. (2018). Explaining Hook's law: definitional practices in a CLIL Physics classroom. *Applied Linguistics*, 39 (5), 694-717. <https://doi.org/10.1093/applin/amw025>

Levinson, S. C. (2013). Action formation and ascription. In T. Stivers & J. Sidnell (Eds.), *The handbook of conversation analysis* (pp. 103–130). Hoboken, NJ: Wiley-Blackwell.

Laitinen 1995. Nollapersoonaa. *Virittäjä* 99, 337–358.

Laurier, E. (2014). The graphic transcript. Poaching comic book grammar for inscribing the visual, spatial and temporal aspects of action. *Geography Compass*, 8(4), 235–248. <https://doi.org/10.1111/gec3.12123>

Lilja, N. (2014). Partial repetitions as other-initiations of repair in second language talk: re-establishing understanding and doing learning. *Journal of Pragmatics*, 71, 98–116. <https://doi.org/10.1016/j.pragma.2014.07.011>

Lilja, N. & A. Piirainen-Marsh (2019). How hand gestures contribute to action ascription. *Research on Language and Social Interaction*. <https://doi.org/10.1080/08351813.2019.1657275>

Lindwall, O. & A. Ekström (2012). Instruction-in-Interaction: The Teaching and Learning of a Manual Skill. *Human Studies*, 35(1), 27-49. <https://doi.org/10.1007/S10746-012-9213-5>

- Lindwall, O., G. Lymer and C. Greiffenhagen (2015). The Sequential Analysis of Instruction. In Markee, Numa (ed.), *The Handbook of Classroom Discourse and Interaction*, 142-157. Oxford: Wiley.
- Mazeland, H. (2013). Grammar in Conversation. In T. Stivers & J. Sidnell (Eds.), *The handbook of conversation analysis* (pp. 475–491). Hoboken, NJ: Wiley-Blackwell.
- Mondada, L. (2014a). Cooking instructions and the shaping of things in the kitchen. In M. Nevile, P. Haddington, T. Heinemann and M. Rauniomaa (eds.), *Interacting with Objects*, 199–226. Amsterdam: Benjamins.
- Mondada, L. (2014b). The local constitution of multimodal resources for social interaction. *Journal of Pragmatics*, 65, 137–156.
<https://doi.org/10.1016/j.pragma.2014.04.004>
- Mondada, L. (2016). Challenges of multimodality. Language and the body in social interaction. *Journal of Sociolinguistics*, 20(3), 336–366.
https://doi.org/10.1111/josl.1_12177
- Mondada, L. (2018). Multiple temporalities of language and body in interaction. Challenges for transcribing multimodality. *Research on Language and Social Interaction*, 51(1), 85–106.
<https://doi.org/10.1080/08351813.2018.1413878>
- Mondada, L. (n.d.) Conventions for multimodal transcription.
<https://www.lorenzamondada.net/resources>
- Mori, J. & Hayashi, M. (2006). The achievement of intersubjectivity through embodied completions: A study of interactions between first and second language speakers. *Applied Linguistics* 27, 195–219.
<https://doi.org/10.1093/applin/aml014>
- Mortensen, K. (2016). The body as a resource for other-initiation of repair: Cupping the hand behind the ear. *Research on Language and Social Interaction*, 49(1), 34–57. <https://doi.org/10.1080/08351813.2016.1126450>
- Olsher, D. (2004). Talk and gesture: the embodied completion of sequential actions in spoken interaction. In Gardner, R. & J. Wagner (eds.), *Second Language Conversations* (pp. 221–245). London: Continuum.
- Raevaara, L. (2017). Adjusting the design of directives to the activity environment: Imperatives in Finnish cooking club interaction. In M-L. Sorjonen, L. Raevaara and E. Couper-Kuhlen (eds.), *Imperative Turns at Talk: The design of directives in action*, 381–410. Amsterdam: Benjamins.

- Rauniomaa, M. (2017). Assigning roles and responsibilities: Finnish imperatively formatted directive actions in a mobile instructional setting. In M-L. Sorjonen, L. Raevaara and E. Couper-Kuhlen (eds.), *Imperative Turns at Talk. The design of directives in action*, 325-355. Amsterdam: Benjamins.
- Rouhikoski, A. (2021). *Direktiivien variaatio. Pyynnöt, neuvot ja ohjeet asiakaspalvelutilanteessa*. Helsingin yliopisto.
- Sacks, Harvey (1995) *Lectures on Conversation*. Oxford: Wiley-Blackwell.
- Schegloff, E. A. (1984) On some gestures' relation to talk. In M. Atkinson & J. Heritage (eds.) *Structures of Social Action. Studies in Conversation Analysis*. Cambridge: Cambridge University Press, 266–295.
- Schegloff, E. A. (2007) *Sequence Organization in Interaction. A Primer in Conversation Analysis*. Cambridge: Cambridge University Press.
- Sorjonen, M-L, Raevaara, L. & E. Couper-Kuhlen (2017) (eds.), *Imperative Turns at Talk: The design of directives in action*. Amsterdam: Benjamins.
- Stevanovic, M. (2017). Managing compliance in violin instruction: The case of the Finnish clitic particles -pA and -pAs in imperatives and hortatives. In M-L. Sorjonen, L. Raevaara and E. Couper-Kuhlen (eds.), *Imperative Turns at Talk: The design of directives in action*, 357–380. Amsterdam: Benjamins.
- Stevanovic, M. & Peräkylä, A. (2012). Deontic authority in interaction: The right to announce, propose, and decide. *Research on Language and Social Interaction* 44, 843–862. <https://doi.org/10.1080/08351813.2012.699260>
- Streeck, J. (2009). *Gesturecraft: The manufacture of meaning*. John Benjamin.
- Stukenbrock, A. (2014). Take the words out of my mouth: Verbal instructions as embodied practices. *Journal of Pragmatics*, 65, 80–ExExE102. <http://dx.doi.org/10.1016/j.pragma.2013.08.017>
- VISK= Hakulinen, Auli & Vilkuna, Maria & Korhonen, Riitta & Koivisto, Vesa & Heinonen, Tarja Riitta & Alho, Irja (2008) *Iso suomen kielioppi*. Helsinki: Suomalaisen Kirjallisuuden Seura.