Gesprächsforschung - Online-Zeitschrift zur verbalen Interaktion (ISSN 1617-1837)
Ausgabe 27 (2026), Seite 53-68 (www.gespraechsforschung-ozs.de)

Book review of Nora Schönfelder: Multimodale Ressourcen beim Argumentieren im Gespräch: Interaktions- und erwerbsbezogene Analysen. Heidelberg: Verlag für Gesprächsforschung 2025

Oliver Spiess

1. Introduction

Coordinating bodies are the foundation of *social interaction* in face-to-face settings (Kendon 1990:248-249).

On the one hand, this is true from a constitutive-definitional perspective when describing the fundamental organization that *constitutes* what we call social interaction: The space *for* interaction is not simply there – it is continuously produced as an *interactional space* through the relevancies established by interactants (Mondada 2014a:270). The well-known *participation framework,* an account of what constitutes participation in interaction and of the roles that participants can take (Goffmann 1979:11), has been criticized for its logocentrism – and it has been extended to make visible the interactional work that participants accomplish when they are 'just listening' (Goodwin 2006:25-26).

On the other hand, the same claim holds from an ontogenetic-developmental perspective, whose aim is to explain how children *learn* to do talk-in-interaction. *Joint attention* – children's "triadic [...] coordination of their interactions with objects *and* people" (Tomasello 1999:62; emphasis added) – is often cited as a milestone in language acquisition (Kauschke 2012:175). In this realm, approaches such as the *Interactional Acquisition Support System* have been developed to highlight the multimodal nature of how children learn to participate in and accomplish joint activities in interaction (Behrens/Pfänder 2025:418-419).

Schönfelder's book discussed in this review adopts this perspective and focuses on the role of multimodality in (the acquisition of) one specific kind of interactional activity: conversational argumentation. Although early conversation-analytical work on argumentation among children has acknowledged and analyzed the role of embodied conduct in such interaction (Goodwin 2000; Goodwin et al. 2002), Schönfelder is still right in observing that most studies on conversational argumentation have neglected multimodality.

There are thus two striking research gaps that the book aims to address (extensively discussed in Chapter 5): First, no previous study has systematically reconstructed the various functions that gestures serve within the actions recurrently used to process argumentation. Second, findings on the role of multimodality in the discourse acquisition of argumentation are, to date, very limited and fragmented.

Accordingly, the two research questions that the author derives are crucial:

1. Which functions do *multimodal gestalts* (Mondada 2014b:109) – with a focus on gestures – fulfill in and for conversational argumentation?
2. What changes can be observed in the skills of children between the ages of 12 and 15 in using such resources for conversational argumentation?

The next section of this review will present and discuss the chapters that address these questions. A critical appraisal of the work follows in the third section.

Veröffentlicht unter der Creativ Commens Lizenz CC BY 4.0 International
https://creativecommons.org/licenses/by/4.0/legalcode.de

2. Contents of the book

Chapter 2, after Chapter 1 introduced the book, lays out the adopted interactional approach to conversational argumentation. This approach represents a shift away from earlier work that conceptualizes argumentation primarily as a *product* that can be analyzed from a logical, decontextualized perspective toward a more *process*-oriented view of argumentation. This is in line with a recent trend (see Koch/Spiess 2026:189) that traces the interactional and sequential organization of argumentation in conversational data (Deppermann 2006b:20-21). In the 'wild', argumentation often does not follow the ideal principles assumed in product-oriented research (Luginbühl/Müller-Feldmeth 2022:2) and instead consists of the incremental linking and sense-making of actions that may initially appear unrelated or even "anarchistic" (Klein 1980:16). As will be seen throughout the following chapters, the multimodal perspective taken in this book is particularly well-suited for reconstructing how firmly such argumentative order is grounded in the fine-tuned coordination through which participants orient to the argumentative affairs at hand.

Another cornerstone of the book is the conceptualization of argumentation as a specific discursive *type of activity*: Such activity types are typical global practices of action, or *genres*,[1] that have been conventionalized as interactional solutions to recurrent communicative problems (Luckmann 1986:202). While 'problem' is to be understood broadly as a communicative goal participants jointly orient to, the domain of argumentation often serves as an exemplary case in which it might also be conceptualized more literally: It constitutes an orderly participants' method, or *argumentative machinery*, that "provides conventionalized interactional means and strategies to resolve conflicts and reach a joint conclusion" (Spiess 2025a:93; Spiess submitted a). To make the orderliness of this argumentative machinery fruitful for analysis, Schönfelder uses the GLOBE model "as a *tertium comparationis* for describing variance" in how discourse practices are processed by different (age groups of) interactants (Hausendorf/Quasthoff 2005:605-606; translated by O.S.). Its adaption to argumentation by Heller (2012:68) is an empirical reconstruction of the five conversational jobs that interactants usually accomplish when processing an argumentative *quaestio*: *establishing dissent / problematizing*, *establishing the obligation to give reasons*, *justifying*, *concluding*, and *transitioning* (Heller 2012:68; for an extension of this model, see Domenech et al. 2018:19-20; Spiess submitted b). Schönfelder limits her analyses to five argumentative actions that are frequently used in her data to process these five jobs (a comprehensible decision given the richness of multimodal descriptions and the sheer variety of different actions in these contexts; see Heller 2012:69-101; Spiess 2025b:28): *taking a stance*, *justifying*, *agreeing*, *opposing*, and *demanding a response*.

Chapter 3 first discusses different frameworks for describing the orchestration of multimodal resources in interaction. After considering the approaches of Kendon (2004), Enfield (2009), and Goodwin (2000), Schönfelder ultimately adopts the concept of multimodal gestalts developed by Mondada (2014b:109). This choice is well grounded in the holism that Mondada's approach entails: It includes not only the entirety of resources that participants may choose to make relevant in and for

[1] Importantly, as Schönfelder explains, this conceptualization of *practice* differs from the concept of locally organized practices that predominates in conversation analysis (see Schegloff 1997:538-539).

the interaction but also acknowledges that participation falls within the realm of *all* participants, rather than those 'momentarily speaking'. Furthermore, treating multimodal orchestrations as gestalt allows one to "identify the recurring bundles of resources that are involved in the realization of argumentative actions" (p. 203).

The second part of this chapter focuses on the categorization and description of functions of gestures. A critical comparison and integration of four approaches – represented by McNeill (1992; 2005), Bavelas et al. (1992; 1995), Kendon (2004; 2017), and Streeck (2009) – proves fruitful for carving out one commonality: All of them, in one way or another, differentiate between *referential gestures* that serve a (broadly conceived) representational function and *pragmatic gestures* that operate over the interaction and its participants. Since previous studies on gestures in discourse activities such as narrating and explaining have shown that pragmatic gestures become particularly important around age 11 – the age at which the book's subjects are located – Schönfelder narrows the focus to this type of gestures.

The functional typology of pragmatic gestures derived from the above-mentioned approaches consists of four categories: (i) Gestures with *performative function* are used to explicate or emphasize the conducted action. (ii) Gestures with *modalizing function* serve to mark the interactant's stance (affiliative, epistemic, etc.). (iii) Gestures with *utterance-structuring function* support the construction of turns. (iv) Gestures with *interaction-organizing* function are related to turn-taking.

In the remainder of the chapter, Schönfelder presents the few insights available in the existing literature on the specific uses of these types of gestures in argumentation. To avoid redundancy, these will be directly integrated into Schönfelder's insights on the uses of gestures found in her corpus (Chapter 8).

Chapter 4 discusses, in preparation for answering research question 2, different models for distinguishing dimensions of conversational argumentation skills.

Vogt (2002) develops a dual model consisting of a structural (the inner structure of arguments, the links between propositions that lead to complex structures) and a processual component (the sequential-interactive processing of argumentation).

Based on Bühler (1934) and Coseriu (2007), Grundler (2011) presents a complex model of twelve subskills, resulting from a crossing of conversational, cognitive, personal, and social aspects of argumentation with elocutionary knowledge (general principles of thinking), idiomatic knowledge (about the specific language at hand), and expressive knowledge (regarding situational context-of-use).

Krelle (2014) focuses on the different jobs that interactants continuously need to process in argumentation, leading to five areas: action management, identity- and relationship management, language management, multimodality management, and thematic management.

Ultimately, Schönfelder decides to use a fourth model by Quasthoff (2009) that differentiates between three dimensions of conversational argumentation skills: *Contextualization* refers to the skill of orienting to the typical genre-bound structures of discourse, including the capability to detect and initiate the corresponding interactional jobs and to manage the related conditional relevancies. *Textualization* means the cohesive and coherent structuring of turns and longer *discourse units* used to process those jobs. Finally, *marking* consists of the appropriate use of the forms (language, gestures, etc.) to mark and conduct these turns. The choice of this model is justified by the assumption that it "allows for decided observations regarding acquisition phenomena" (p. 6; translation by O.S.). From a practical point of

view – this choice allows comparison with and embedding in a large body of previous studies that use the same model – this choice is reasonable. Nevertheless, a more in-depth discussion of the various strengths and weaknesses of the different models would have strengthened the argument.

On the one hand, the models of Vogt (2002), Grundler (2011), and Krelle (2014) definitely exhibit shortcomings that could have been discussed here: For example, the exclusion of interpersonal aspects, the heavy focus on knowledge that may clash with conversation-analytical principles, or the uncritical separation of 'language' and 'multimodality', respectively. On the other hand, one might ask how well the *chosen* approach ultimately suits a description of *multimodal argumentation* skills (see the extensive discussion in Section 3).

In the next step, Schönfelder presents the developmental trajectories in these three dimensions that have been found in previous studies about children's use of vocal/aural resources in conversational argumentation. For contextualization, the trajectory goes from reactive to autonomous behavior, with increasingly proactive establishment of jobs and conditional relevancies. Regarding textualization, a trend is seen from locally toward globally organized structures, that is, argumentative discourse units spanning multiple sentences and coordinating several thematic and pragmatic aspects. Marking skills develop from implicit toward more explicit displaying of sequential and thematic relations, epistemic modalization, and so on.

Finally, due to the lack of respective insights on argumentation, existing studies on discourse acquisition of visuospatial resources in narrating and explaining are summarized. They show a general trend from the use of referential and iconic gestures to more abstract and metaphorical gestures and ultimately to an increase in pragmatic gestures.

Chapter 5 then derives the desiderata and research questions previewed in the first chapter from the preceding literature review. As they have already been presented in the introduction to this review, they will not be elaborated again here.

Chapters 6-7 discuss the data and the methods used in the study. The longitudinal corpus consists of 99 conversations among 33 parent-child dyads (3 each for age 12, 13, and 15) with the goal of a justified joint decision. Three different tasks were conducted: The participants had to choose one of four options for a price of a competition (age 12), an activity to do with exchange students (age 13), or a venue for the class trip (age 15). The filmed conversations were transcribed following GAT2 conventions (Selting et al. 2009) and Mondada (2001), including video stills.

To answer research question 1 (functions of multimodal gestalts in argumentation), 20 interactions from the third recording time (age 15) that neither contain an overly low nor an overly high frequency of gestures were selected (total duration: 2:17:40h). These were analyzed with multimodal conversation analysis and a stepwise creation of collections of the five recurrent argumentative actions mentioned in Chapter 2. The analyses in Chapter 8 aim to describe the variety of pragmatic functions of the multimodal gestalts used to accomplish these actions.

To answer research question 2 (changes in the skills to use such resources in argumentation), pilot analyses of the corpus were conducted to identify different *interaction patterns* – regularities in the interactional dynamics regarding what children versus parents contribute to the processing of the jobs (Heller/Krah 2015:7). Six dyads (two of the three identified interaction patterns each) were chosen for further analysis of differences in contextualization, textualization, and marking

skills of the children. For the in-depth longitudinal analysis of the conversations, two children with maximally diverging profiles were then selected for contrastive purposes (total duration: 00:30:37h). The corresponding analyses in Chapter 10 focus on the argumentative action *justifying* and aim to reconstruct the multimodal manifestations of the three argumentative skills mentioned above.

Chapter 8 may well be described as the heart of the book, as it represents the pioneering attempt to provide an extensive overview of the variety of forms and functions that multimodal gestalts can take in central argumentative actions. Over 100 pages, Schönfelder gives a detailed account of over 20 different roles that these gestalts play in *taking a stance*, *justifying*, *agreeing*, *opposing*, and *demanding a response*. Since a discussion of the full repertoire of these gestalts would go beyond the scope of this review, I will pick out only some particularly remarkable cases below (see the book's detailed table of contents for an overview of all functions).

Schönfelder shows that *taking a negative stance* toward an option can be gesturally marked by *throwing/holding away gestures* (Bressem/Müller 2014:1581) that metaphorically exclude the discussed option and thus serve a performative function. Interestingly, gesture and vocal/aural parts are not merely redundant. Rather, the gesture can be finely timed to project the negative stance before it is vocally/aurally explicated or to complement syntactically incomplete utterances.

Multimodal gestalts also appear in *agreeing* actions, where the composition varies according to the type and sequential embedding of the affiliation. Schönfelder observes that emphatic agreeing is frequent after preceding opposition. In such cases, a credible agreement that demonstrates epistemic access to the position requires an extended vocal/aural agreement accompanied by iterative or ostensive gestural reference to the option (printed on worksheets). This multimodal gestalt thus serves a modalization function by emphasizing the realized action.

A particularly complex case lies in the marking of concession in *opposing* actions. Schönfelder presents a case of a 'reversed' *cardinal concession* (Couper-Kuhlen/Thompson 2000) in which partial opposition is followed by confirmation. The critical point is the 'hinge' between opposing and agreeing, where the adversative conjunction *aber* ('but') is coordinated with the preparation phase of a *holding away* gesture, marking the opposition as stopped and thus anticipating the change in affiliation. The multimodal gestalt is completed by tilting the head in the opposite direction of the manual gesture's orientation, thereby adding an *utterance-structuring function* to the performative one by representing the transition to a different part of the contribution linked to a change of perspective.

Schönfelder's descriptions of gestalts that go beyond the traditional view of gestures as performed by 'speakers' within 'their' turn are particularly interesting.

One case in point is *palm up gestures* whose function – especially in *justifying* – can go beyond the known performative role of presenting topics (here: arguments) as abstract discursive objects (Müller 2004:254). As the examples show, they can be used both to mark continuity between contributions and to reinforce conditional relevancies when they are held post-stroke at – and beyond – the end of the turn.

Another example is deictic reference to pictures on the worksheet during *other* participants' justifications. These gestures fulfill an *interaction-organizing* function by documenting understanding and displaying engagement to co-participants.

A third case concerns actions for *demanding a response* that also contribute to *structuring* co-participants' turns. The multimodal gestalts in this role consist of the

(vocal/aural) establishment of conditional relevance for a specific follow-up action, supported by *counting gestures* (that is, stretching multiple fingers as if counting). These gestures are not only used to project the expected type of response; they are also held across the following turn, ratifying individual parts of the response through stepwise retraction of the extended fingers.

Finally, Schönfelder presents modalizing gestalts described as independent of the type of action because they can be used with different types of actions. One example is the demonstrable and acoustically observable dropping of hands on the table, combined with epistemically unmitigated utterances. These gestalts mark the utterance as evident, often supported by lexical means such as modal particles.

Chapter 9 discusses the relevance of the findings in Chapter 8 for multimodal discourse studies. It emphasizes the importance of accounting for multimodality to grasp the variety of manifestations of argumentative actions. Gestures that functionally and/or formally go beyond the vocal/aural turn (for example, when structuring others' turns) illustrate this importance particularly well.

It is also plausible to conclude, as Schönfelder does, that multimodal gestalts can serve as indicators of interactional skills, since they are multimodal forms used to mark textualization (for example, through creating cohesion), contextualization (for example, through signaling the type of action conducted), and modalization.

In my view, however, the claim that the analyses help reconstruct the variety of "genre-typical, that is, argumentation-specific, functions of multimodal bundles of resources" (p. 204) is more questionable. Although all the presented gestalts are undeniably crucial for the interactional sequences shown, it remains unclear to what extent they are truly argumentation-*specific* or simply of general importance in interactional activities. A more explicit discussion of this issue, especially in relation to the broader question of how the findings may challenge the GLOBE model, would have helped readers contextualize the results more coherently.

On a related note, an in-depth discussion of how the impressive repertoire of identified multimodal gestalts could contribute to a data-driven re-typologization of argumentative actions and of the functions of multimodal gestalts would have strengthened the book (see also the discussion in Section 3). While the summary of the findings at the end of Chapter 8 is helpful, one might ask how the results call into question the mapping of (vocally/aurally identified) types of argumentative actions and multimodal gestalts. A case in point is agreeing multimodal gestalts that are used in simultaneity with justifications to mark them as evident: As Schönfelder shows, such gestalts already initiate the transition to the job *concluding* before it is vocally indicated, thus blurring the boundaries between both jobs as well as the actions used to process them.

Chapter 10 addresses research question 2, which concerns the changes in conversational argumentation skills across the three time points collected (age 12, 13, and 15). It begins with a discussion of the selection of the two focus children for the following qualitative analyses, conducted through pilot analyses based on their maximally different interaction profiles: Focus child 1 represents the 'stronger' interaction profile – *Explorative stance-taking and justifying* – characterized by the skills to contextualize global conditional relevancies, to follow up on them with globally structured, genre-appropriate discourse units, and to act proactively. Focus child 2 represents the 'weaker' interaction profile – *Confirming stance-taking and*

assessing – characterized by rather reactive behavior oriented toward local relevancies and by shorter, less complex discourse units.

While the differences between the two profiles are illustrated well in a table and become clear in the subsequent analyses, one might still wish for a little more elaboration on the topic. In particular, the text does not clearly state which factors and kinds of observations formed the basis for distinguishing the profiles, nor does it provide a definition of what is exactly meant by *interaction profile*.

The analyses focus on the action *justifying* and are divided into two parts: The first part examines the development of contextualization, textualization, and modalization skills of the two focus children across the three recorded ages from a purely vocal/aural perspective. The second part focuses on the visuospatial resources used in the same data to characterize the multimodal manifestations of the skills mentioned above. For brevity, I will summarize both parts together.

First, some general observations. Regarding multimodal gestalts, the main difference between the children does not lie in the repertoires or frequencies, but rather in *how* they use the gestalts (see discussion below). This is also reflected in the age-based comparison: Both children show a markedly reduced repertoire and frequency of multimodal gestalts at age 15. A plausible explanation supported by other studies (for example, Andrén 2014a; Kinalzik 2024) is that the acquisition of new forms first passes through a phase of excessive trial-and-error before the skills are mastered to the point that they can be selectively used only in contexts where they are most appropriate.

Focusing on contextualization, Schönfelder's analyses show that focus child 2 seems to recognize global conditional relevancies for discourse units already at age 12 but lacks the skill to effectively build such multi-unit turns. Attempts to do so appear at age 13, only to be reduced again by age 15 (in contrast to focus child 1, who exhibits a gradual improvement). Moreover, the ability to proactively justify, to initiate conditional relevancies, and to initiate interactional jobs such as *concluding* are displayed later by focus child 2 than by focus child 1.

These difficulties with producing global units are also reflected in the use of performative gestures that present justifications as discursive objects: Such gestures are less well coordinated with the vocal/aural parts by focus child 2 and seldomly used to contextualize complex arguments beyond the sentence level. Finally, the limited initiative is also evident in the purely reactive gestural rejection of topics (using *away gestures*), in contrast to the proactive realizations by focus child 1.

Textualization shows a similar picture: The arguments of focus child 1 become increasingly globally structured and complex with age. They start as addition-like lists of arguments that are not well connected on the marking level, and they end as multimodally well-structured and explicitly linked multi-unit turns. Focus child 2, on the other hand, uses almost exclusively turns on or below the sentence level at age 12. When trying to produce global discourse units at age 13, these attempts are accompanied by issues of structuring and coherence and are heavily dependent on interactional support from the parent. Eventually, at age 15, focus child 2 returns to local contributions and does not even attempt to form bigger packages any longer.

These findings are especially mirrored by the utterance-structuring gestalts, where focus child 2 struggles to coordinate emphasizing gestures and language (for example, focus accents or transition points between utterance parts) and only uses them with local scope. In contrast, focus child 1 deploys them skillfully to mark

cohesion or coherence across multiple TCUs or even turns. Moreover, focus child 1 later shows the skill to project specific types of actions (rather than establishing unspecific conditional relevance for any uptake) and exhibits the challenging multimodal coordination of contrasting standpoints in concessions.

Regarding modalization, the data are more limited, because focus children 1 and 2 do not use any of the identified gestalts at age 13 and 12, respectively. From a vocal/aural perspective, both use many forms to mark their contributions as evident in the early years. However, while focus child 1 mostly uses gestures to *further* emphasize the force of justifications, focus child 2 uses gestures mainly to make the intended argumentative character explicit in the first place. Focus child 2 also employs more modalizations to signal missing knowledge or uncertainty, especially after conditional relevancies established by the parent. With age, focus child 1 increasingly uses modalization to frame justifications as explorative and open for discussion. In general, the modalizing means of focus child 1 seem more diverse.

I would like to highlight one striking general difference between the two children: While focus child 1 shows a gradual development of argumentation skills in all three dimensions, focus child 2 appears, at age 15, to no longer exhibit some of the behaviors shown at age 13. Focus child 2 stops trying to produce more global justifications and reverts to mainly locally structured, short arguments. Schönfelder interprets this as regression to easier ways of interactional processing, prioritizing the safety of being able to produce turns as planned over further development toward globality. It may have high practical relevance if future studies with larger data sets would confirm that discourse acquisition in these respects and contexts can already stop at age 15.

One thing that might be criticized is the decision to separate the analyses of vocal/aural means and multimodal gestalts. While this reinforces the artificial separation of resources (which the author acknowledges, p. 206), it appears to be mainly driven by practical considerations, since the pilot analyses for selecting the focus children were based solely on vocal/aural resources; presumably due to time constraints. That is understandable, but it does not have to be reflected in the book; where the analyses could have been integrated. Doing so might have reduced redundancy and provided a clearer, more quickly graspable picture of how various resources complement, contrast, and interact with each other.

Chapter 11 discusses the two research questions in light of the insights gained during the study and formulates developmental hypotheses for future research.

Research question 1 (functions of multimodal gestalts in argumentation) was addressed by illustrating the crucial roles these gestalts can play in the accomplishment of argumentative actions: *performative, modalizing, utterance-structuring*, and *interaction-organizing*. The study also showed that these functions are not determined merely by the resources involved; they can be differentiated according to sequential embedding, duration, and temporal unfolding as well.

Research question 2 (changes in conversational argumentation skills with age) was addressed by demonstrating that multimodal gestalts, while used by the children of all observed ages, differ across ages mainly in their configuration: This configuration can be described along three dimensions that also appear especially promising for further research on multimodal argumentation acquisition: (i) *scope*, (ii) *complexity*, and (iii) *coordination*. Schönfelder thus links her findings with developmental hypotheses for future studies.

The scope and the complexity of justifications vary considerably across children and with age. The corresponding hypotheses are that both increase with the development of argumentation skills (except that, as for focus child 2, they can decrease again in favor of the coordination of resources).

While some forms of these developments are evident in the analyses, a clear definition and distinction between the two facets appears to be missing. Scope intuitively seems more related to globality – that is, the length of contributions that are marked as cohesive and coherent – whereas complexity seems more associated with denseness and the variety of resources used simultaneously. However, Schönfelder relates both facets to the establishment of cohesiveness and coherence, and she also ascribes the length of justifications to both facets (p. 286-287).

The Coordination of multiple resources for accomplishing multimodal gestalts is challenging for both children, especially for focus child 2 when trying to produce longer argumentative units, which eventually led this child to return to smaller units in favor of more precise coordination. A derived developmental hypothesis is that multiple resources are easier to coordinate for locally structured arguments. An additional hypothesis, whose confirmation would have even larger implications, is that multimodal gestalts are not acquired holistically but rather stepwise.

Chapter 12 discusses the limitations and perspectives of the study, some of which are presented in the following (see Section 3 for additional points).

Some limitations pertain to the setting of the conversations: The interactants' sitting position may influence the results because it restricts the gesture space (Zima 2013:21). The decision task and available options, the ecology (for example, worksheets), and the social roles (children vs. parents) may also be relevant.

Another perspective lies in extending the approach to other discourse genres and to interactions of younger children to investigate what happens before pragmatic gestures become predominant. This would also involve questions about which resources are acquired at which age and with what associated functions.

Finally, the study demonstrates the importance of visuospatial resources for argumentation skills – yet these are hardly integrated into school curricula. Translating the gained insights into practice would be a desirable step toward teaching argumentation in its full richness.

3. Critical appraisal

The greatest achievement of this book worth reading is that it makes argumentation *visible*. This is reflected in wonderful more-than-just-metaphors such as "the giving-and-taking of arguments" (p. 9; translated by O.S.) or the "dropping of conversational objects" (p. 188; translated by O.S.) – and in an initial attempt to develop an overview of the central recurrent multimodal gestalts and their functions that are used to accomplish various argumentative actions (Chapter 8). Schönfelder's insights into the importance of visuospatial resources for conversational argumentation, especially when vocal/aural means are insufficient to meet the encountered interactional challenges, urge future work to develop a *multimodal model of argumentation skills*.

Pioneering work always carries the burden of having to explore some aspects while excluding others, and of re-examining models and concepts in light of the new perspectives. One area where such reflections could have been deepened is the

discussion of whether the model of argumentations skills that distinguishes between contextualization, textualization, and marking (Chapter 4) is fully suitable for assessing *multimodal* argumentation skills. That model was initially developed on vocal/aural grounds and for examining the acquisition of *narrative* skills, as reflected in the central importance ascribed to discourse units and the corresponding skill of contextualizing them. The limitation to *justifying* in Chapter 10 – effectively excluding all other interactional jobs usually processed in argumentative episodes – shows that such discourse units are primarily relevant only for this particular argumentative job. In addition, because actions in talk-in-interaction are always context-shaped *and* context-renewing (Heritage 1995:398), contextualizing them appears to be a ubiquitous task also involved *within* multi-unit turns.

Moreover, previous research on conversational argumentation among children has shown that argumentative *co-constructions* can be considered a particularly demanding and advanced way of arguing (Kreuz 2021:329; Kreuz/Luginbühl 2020: 109). The question is how such co-constructive stretches of argumentation that lack longer discourse units produced by single participants should be assessed within this framework. This problem is further compounded from a decisively multimodal perspective, where "*sequentially ordered simultaneities*" (Mondada 2018:103) characterize the interactional order. Under this premise, the line between contextualization and textualization becomes blurry, because all participants continuously contribute to writing layers of the shared multimodal text while thereby simultaneously contextualizing other layers.

This, of course, does not call into question the relevance or validity of the findings on the development of multimodal argumentation skills shown in Chapter 10. Instead, it suggests that a modified or different model might be better suited to distinguish the different skills needed to process conversational argumentation and to describe its manifestations. Related to this, and although modalization certainly plays a key role in conversational argumentation (skills) – as Schönfelder convincingly demonstrates in her work (see also Kreuz/Luginbühl 2020:104-105; Morek 2020; Spiess et al. in press) – its conceptual role in the study remains unclear. In Chapter 10, modalization serves as a third dimension of comparison across children and ages alongside textualization and contextualization, while in earlier chapters it is used as one of four pragmatic functions of gestures. Should it be seen as its own dimension of argumentation skills or merely as one possible means to enact one of the others (which one?). Schönfelder argues that modalization's special role, compared to the other pragmatic functions, is justified by its overarching character: It is used in different argumentative actions (p. 102). However, this also applies to other pragmatic functions such as the performative one.

One group of findings that may be particularly transformative for further research on conversational argumentation (skills) is the discovery that multimodal gestalts are not limited to the traditional role of 'the speaker'. For example, Schönfelder nicely shows that counting gestures can be used during the vocal/aural turns of others to participate in structuring them, and that deictic gestures are employed to track grounding processes. Such observations are crucial for how we think about and conduct research on argumentation. While the study does not discuss these implications extensively, the scope of this review also does not allow more than raising a few related questions that need to be addressed in future work:

1. What consequences do the simultaneities of argumentative action (such as an embodied display of dissent during another person's vocal/aural stance-taking) have for the jobs reconstructed in the GLOBE model?
2. Can multimodality be integrated into previous models of conversational argumentation skills (see discussion above), or does it force us to reconceive what counts as competent argumentation (for example, are gestures necessary for skilled argumentation? Are words necessary?).
3. What are the multimodal actions that are *specific to argumentation*, and how should we redefine and categorize them – rather than treating visuospatial compositions merely as "specific forms of argumentative actions [defined on vocal/aural grounds]" (p. 103; translation and addition by O.S.) (for example, by typologizing them according to properties identified by Schönfelder, such as affiliation/disaffiliation, emphasis/mitigation, scope, or established/satisfied conditional relevancies)?

A limitation that is not discussed extensively in the book concerns the limited data set, which such an explorative study, of course, can only partially be blamed for. On the one hand, it varies a lot of variables. It is not clear why Schönfelder did not analyze interactions among younger children, where a richer body of prior work (with a focus on multimodality in narrative and explanatory discourse) could have served as a comparative tool. In addition, the study varies both age (three levels) and interaction profile (two levels) rather than analyzing a larger number of children, which reduces explanatory power and makes it difficult to develop interpretations and hypotheses about which observed phenomena are attributable to which particular dimension. Potential risks of the small data set are also reflected in low frequencies of justifications and multimodal gestalts, or even gaps where the investigated children use none of the analyzed gestalts at all (see p. 280; p. 276).

Future research should definitely test the results and hypotheses presented in this book with a richer data set, following a third trend (Koch/Spiess 2026:189-191) in addition to the longitudinal approach (Deppermann/Pekarek Doehler 2021; Pekarek Doehler et al. 2018) and the multimodal perspective already adopted in this book: the combination of conversation analysis with quantitative methods (Spiess et al. 2026). Such an extension is valuable when studying *recurrent* multimodal gestalts that are assumed to be(come) entrenched and relatively fixed; the identification of empirical regularities lends itself to quantification (Spiess 2025b:7).

Interestingly, and perhaps related to this gap, Schönfelder never makes explicit reference to the current trend toward research on *multimodal constructions* in *construction grammar* (Zima 2025). This may be a missed opportunity, since the construction-grammar framework is widely compatible with interactional approaches (Deppermann 2006a:60) and more advanced in quantifying multimodal bundles (see, e.g., Andrén 2014b; Bressem/Müller 2017; Schoonjans 2018).

To conclude this review, it should be emphasized that most of the critical questions discussed above represent not merely limitations of Schönfelder's book. Rather, they underscore the major contribution of this innovative work by calling previous frameworks into question and prompting such questions in the first place. In this sense, there are several promising perspectives for future research on multimodal argumentation (skills).

One particularly relevant direction is investigating the multimodal support of "the collaborative continuation of talk [that] is potentially at risk" (Schönfelder/Heller 2019:86). This tackles the very core of conversational argumentation: The *argumentative machinery* (Spiess 2025a:93) "provides conventionalized interactional means and strategies to resolve conflicts" (Spiess submitted a) and is thus a specialized tool for such 'risky' interaction. As Schönfelder shows, visuospatial resources contribute significantly to this core function; they can serve to mediate controversial and potentially conflict-escalating points (p. 290) as well as to signal ongoing alignment and engagement in a shared activity, even beyond the point at which vocal/aural means fail (p. 54).

4. References

Andrén, Mats (2014a): Gestures in Northern Europe: Children's gestures in Sweden. In: Müller, Cornelia / Cienki, Alan / Fricke, Ellen / Ladewig, Silva / McNeill, David / Bressem, Jana (eds.), Body – language – communication: An international handbook on multimodality in human interaction (Vol. 2). Berlin, München & Boston: De Gruyter Mouton, 1282-1289. https://doi.org/10.1515/9783110302028.1282.

Andrén, Mats (2014b): Multimodal constructions in children. In: Gesture 14(2), 141-170. https://doi.org/10.1075/gest.14.2.02and.

Bavelas, Janet B. / Chovil, Nicole / Coates, Linda / Roe, Lori (1995): Gestures specialized for dialogue. In: Personality and Social Psychology Bulletin 21(4), 394-405. https://doi.org/10.1177/0146167295214010.

Bavelas, Janet B. / Chovil, Nicole / Lawrie, Douglas A. / Wade, Allan (1992): Interactive gestures. In: Discourse Processes 15(4), 469-489. https://doi.org/10.1080/01638539209544823.

Behrens, Heike / Pfänder, Stefan (2025): Spielend kooperieren lernen: Perspektiven für die Spracherwerbsforschung. In: Proske, Nadine / Weber, Thilo / Dannerer, Monika / Deppermann, Arnulf (eds.), Gesprochenes Deutsch. Berlin & Boston: De Gruyter (Jahrbuch 2024), 401-424. https://doi.org/doi:10.1515/9783111623825-017.

Bressem, Jana / Müller, Cornelia (2014): The family of Away gestures: Negation, refusal, and negative assessment. In: Müller, Cornelia / Cienki, Alan / Fricke, Ellen / Ladewig, Silva / McNeill, David / Bressem, Jana (eds.), Body – language – communication: An international handbook on multimodality in human interaction (Vol. 2). Berlin, München & Boston: De Gruyter Mouton, 1592-1604. https://doi.org/10.1515/9783110302028.1592.

Bressem, Jana / Müller, Cornelia (2017): The "negative-assessment-construction" – A multimodal pattern based on a recurrent gesture? In: Linguistics Vanguard 3(s1), 20160053. https://doi.org/10.1515/lingvan-2016-0053.

Bühler, Karl (1934): Sprachtheorie: Die Darstellungsfunktion der Sprache. Jena: Gustav Fischer.

Coseriu, Eugenio (2007): Sprachkompetenz: Grundzüge der Theorie des Sprechens. Tübingen: Narr.

Couper-Kuhlen, Elizabeth / Thompson, Sandra A. (2000): Concessive patterns in conversation. In: Couper-Kuhlen, Elizabeth / Kortmann, Bernd (eds.), Cause -

condition - concession - contrast. Berlin & New York: De Gruyter Mouton, 381-410. https://doi.org/10.1515/9783110219043.4.381.

Deppermann, Arnulf (2006a): Construction grammar – eine Grammatik für die Interaktion? In: Deppermann, Arnulf / Fiehler, Reinhard / Spranz-Fogasy, Thomas (eds.), Grammatik und Interaktion: Untersuchungen zum Zusammenhang von grammatischen Strukturen und Gesprächsprozessen. Radolfzell: Verlag für Gesprächsforschung, 43-65. http://www.verlag-gespraechsforschung.de/2006/deppermann.htm.

Deppermann, Arnulf (2006b): Desiderata einer gesprächsanalytischen Argumentationsforschung. In: Deppermann, Arnulf / Hartung, Martin (eds.), Argumentieren in Gesprächen: Gesprächsanalytische Studien (2nd ed.). Tübingen: Stauffenburg, 10-26.

Deppermann, Arnulf / Pekarek Doehler, Simona (2021): Longitudinal conversation analysis: Introduction to the special issue. In: Research on Language and Social Interaction 54(2), 127-141. https://doi.org/10.1080/08351813.2021.1899707.

Domenech, Madeleine / Heller, Vivien / Petersen, Inger (2018): Argumentieren mündlich, schriftlich, zweitsprachlich: Verfahren und Anforderungen. In: Massud, Abdel-Hafiez (ed.), Argumentieren im Sprachunterricht. Landau: Verlag Empirische Pädagogik, 15-35.

Enfield, Nick J. (2009): The anatomy of meaning: Speech, gesture, and composite utterances. Cambridge: Cambridge University Press (Language, Culture and Cognition 8).

Goffman, Erving (1979): Footing. In: Semiotica 25(1-2), 1-30. https://doi.org/10.1515/semi.1979.25.1-2.1.

Goodwin, Charles (2000): Action and embodiment within situated human interaction. In: Journal of Pragmatics 32(10), 1489-1522. https://doi.org/10.1016/S0378-2166(99)00096-X.

Goodwin, Charles (2006): Interactive footing. In: Holt, Elizabeth / Clift, Rebecca (eds.), Reporting talk: Reported speech in interaction. Cambridge: Cambridge University Press, 16-46. https://doi.org/10.1017/cbo9780511486654.003.

Goodwin, Marjorie H. / Goodwin, Charles / Yaeger-Dror, Malcah (2002): Multimodality in girls' game disputes. In: Journal of Pragmatics 34(10), 1621-1649. https://doi.org/10.1016/S0378-2166(02)00078-4.

Grundler, Elke (2011): Kompetent argumentieren: Ein gesprächsanalytisch fundiertes Modell. Tübingen: Stauffenburg.

Hausendorf, Heiko / Quasthoff, Uta M. (2005): Konversations-/Diskursanalyse: (Sprach)Entwicklung durch Interaktion. In: Mey, Günter (ed.), Handbuch qualitative Entwicklungspsychologie. Köln: Kölner Studien Verlag, 585-618.

Heller, Vivien (2012): Kommunikative Erfahrungen von Kindern in Familie und Unterricht: Passungen und Divergenzen. Tübingen: Stauffenburg Verlag (Stauffenburg Linguistik 67).

Heller, Vivien / Krah, Antje (2015): Wie Eltern und Kinder argumentieren: Interaktionsmuster und ihr erwerbssupportives Potenzial im längsschnittlichen Vergleich. In: Mitteilungen des Deutschen Germanistenverbandes 62(1), 5-20. https://doi.org/10.14220/mdge.2015.62.1.5.

Heritage, John (1995): Conversation analysis: Methodological aspects. In: Quasthoff, Uta M. (ed.), Aspects of oral communication. Berlin & New York: De Gruyter (Research in Text Theory 21), 391-418. https://doi.org/10.1515/9783110879032.391.

Kauschke, Christina (2012): Kindlicher Spracherwerb im Deutschen: Verläufe, Forschungsmethoden, Erklärungsansätze. Berlin & Boston: De Gruyter. https://doi.org/10.1515/9783110283891.

Kendon, Adam (1990): Conducting interaction: Patterns of behavior in focused encounters. Cambridge: Cambridge University Press.

Kendon, Adam (2004): Gesture: Visible action as utterance. Cambridge: Cambridge University Press. https://doi.org/10.1017/cbo9780511807572.

Kendon, Adam (2017): Pragmatic functions of gestures: Some observations on the history of their study and their nature. In: Gesture 16(2), 157-175. https://doi.org/10.1075/gest.16.2.01ken.

Kinalzik, Noelle (2024): Mündliches Erklären von der Grundschule bis zur Sekundarstufe – Varianzen zwischen Kindern heterogener Ausgangslagen. Göttingen: Verlag für Gesprächsforschung. https://www.verlag-gespraechsforschung.de/2024/pdf/erklaeren.pdf

Klein, Wolfgang (1980): Argumentation und Argument. In: Zeitschrift für Literaturwissenschaft und Linguistik 38/39, 9-57.

Koch, Tamara / Spiess, Oliver (2026): Children's conversational argumentation skills: Methodological advances, changing interactional practices and inferential structures [editorial]. In: Research on Children and Social Interaction 9(2), 189-202. https://doi.org/10.3138/rcsi-2025-0019.

Krelle, Michael (2014): Mündliches Argumentieren in leistungsorientierter Perspektive: Eine empirische Analyse von Unterrichtsdiskussionen in der neunten Jahrgangsstufe. Baltmannsweiler: Schneider Verlag Hohengehren.

Kreuz, Judith (2021): Ko-konstruiertes Begründen unter Kindern: Eine gesprächsanalytische Studie von Kleingruppeninteraktionen in der Primarschule. Tübingen: Stauffenburg.

Kreuz, Judith / Luginbühl, Martin (2020): From flat propositions to deep co-constructed and modalized argumentations: Oral argumentative skills among elementary school children from grades 2 to 6. In: Research on Children and Social Interaction 4(1), 93-114. https://doi.org/10.1558/rcsi.12416.

Luckmann, Thomas (1986): Grundformen der gesellschaftlichen Vermittlung des Wissens: Kommunikative Gattungen. In: Neidhardt, Friedhelm (ed.), Kultur und Gesellschaft: René König, dem Begründer der Sonderhefte, zum 80. Geburtstag gewidmet. Opladen: Westdeutscher Verlag, 191-211.

Luginbühl, Martin / Müller-Feldmeth, Daniel (2022): Oral argumentation skills between process and product. In: Languages 7(2), Article 139. https://doi.org/10.3390/languages7020139.

McNeill, David (1992): Hand and mind: What gestures reveal about thought. Chicago & London: The University of Chicago Press.

McNeill, David (2005): Gesture and thought. Chicago: University of Chicago Press.

Mondada, Lorenza (2001): Conventions for transcribing multimodality. https://www.lorenzamondada.net/multimodal-transcription.

Mondada, Lorenza (2014a): Interactional space and the study of embodied talk-in-interaction. In: Auer, Peter / Hilpert, Martin / Stukenbrock, Anja / Szmrecsanyi,

Benedikt (eds.), Space in language and linguistics: Geographical, interactional, and cognitive perspectives. Berlin & Boston: De Gruyter, 247-275. https://doi.org/10.1515/9783110312027.247.

Mondada, Lorenza (2014b): Pointing, talk, and the bodies: Reference and joint attention as embodied interactional achievements. In: Seyfeddinipur, Mandana / Gullberg, Marianne (eds.), From gesture in conversation to visible action as utterance: Essays in honor of Adam Kendon. Amsterdam & Philadelphia: John Benjamins Publishing Company, 95-124. https://doi.org/10.1075/z.188.06mon.

Mondada, Lorenza (2018): Multiple temporalities of language and body in interaction: Challenges for transcribing multimodality. In: Research on Language and Social Interaction 51(1), 85-106. https://doi.org/10.1080/08351813.2018.1413878.

Morek, Miriam (2020): Learning to modalize is learning to reason: On the role of epistemic modalizations in parent–child-talk and in written argumentation of secondary school students. In: Research on Children and Social Interaction 4(1), 115-141. https://doi.org/10.1558/rcsi.12418.

Müller, Cornelia (2004): Forms and uses of the Palm Up Open Hand: A case of a gesture family? In: Müller, Cornelia / Posner, Roland (eds.), The semantics and pragmatics of everyday gestures: Proceedings of the Berlin conference April 1998. Berlin: Weidler Buchverlag Berlin (Körper, Zeichen, Kultur 9), 233-256.

Pekarek Doehler, Simona / Wagner, Johannes / González-Martínez, Esther (eds.) (2018): Longitudinal studies of the organization of social interaction. London: Palgrave Macmillan. https://doi.org/10.1057/978-1-137-57007-9.

Quasthoff, Uta (2009): Entwicklung der mündlichen Kommunikationskompetenz. In: Becker-Mrotzek, Michael (ed.), Mündliche Kommunikation und Gesprächsdidaktik. Baltmannsweiler: Schneider-Verlag Hohengehren, 88-105.

Schönfelder, Nora / Heller, Vivien (2019): Embodied reciprocity in conversational argumentation: Soliciting and giving reasons with palm up open hand. In: Grimminger, Angela (ed.), Proceedings of the 6th Gesture and Speech in Interaction Conference. Paderborn: Universitätsbibliothek Paderborn, 81-86. https://doi.org/10.17619/UNIPB/1-810.

Schoonjans, Steven (2018): Modalpartikeln als multimodale Konstruktionen: Eine korpusbasierte Kookkurrenzanalyse von Modalpartikeln und Gestik im Deutschen. Berlin & Boston: De Gruyter. https://doi.org/10.1515/9783110566260.

Selting, Margret / Auer, Peter / Barth-Weingarten, Dagmar / Bergmann, Jörg / Bergmann, Pia / Birkner, Karin / Couper-Kuhlen, Elizabeth / Deppermann, Arnulf / Gilles, Peter / Günthner, Susanne / Hartung, Martin / Kern, Friederike / Mertzlufft, Christine / Meyer, Christian / Morek, Miriam / Oberzaucher, Frank / Peters, Jörg / Quasthoff, Uta / Schütte, Wilfried, . . . Uhmann, Susanne (2009): Gesprächsanalytisches Transkriptionssystem (GAT 2). In: Gesprächsforschung – Online-Zeitschrift zur verbalen Interaktion 10, 152-183. http://www.gespraechsforschung-ozs.de/heft2009/px-gat2.pdf.

Spiess, Oliver (2025a): Exploring participatory dynamics in conversational argumentation: Operationalizations, visualizations, and cooccurrences. In: Gesprächsforschung - Online-Zeitschrift für verbale Interaktion 26, 87-131. http://www.gespraechsforschung-online.de/fileadmin/dateien/heft2025/ga-spiess.pdf.

Spiess, Oliver (2025b): Transforming interactional structures into codes: Methodological cycle for a CA-based coding of conversational actions. In: Nota BeNe 2(1), 5-40. https://doi.org/10.1075/nb.00022.spi.

Spiess, Oliver (submitted a): Mismatching frames in argumentative talk: The misunderstanding–disagreement continuum.

Spiess, Oliver (submitted b): Relevance in numbers: Accounting for interactionality when coding and quantifying social interaction.

Spiess, Oliver / Koch, Tamara / Luginbühl, Martin (in press): Being uncertain together: Reconstructing practices of coping with knowledge gaps in children's conversational argumentation. In: Younis, Ramy / Oswald, Steve (eds.), Pragmatic approaches to argumentation. Lausanne: Peter Lang.

Spiess, Oliver / Luginbühl, Martin / Müller-Feldmeth, Daniel (2026): CA and quantitative approaches. In: Burdelski, Matthew / Greer, Tim (eds.), The Routledge handbook of conversation analysis. London: Routledge (Routledge Handbooks in Applied Linguistics), 210-226. https://doi.org/10.4324/9781032720852-15.

Streeck, Jürgen (2009): Gesturecraft: The manu-facture of meaning. Amsterdam & Philadelphia: John Benjamins Publishing Company. https://doi.org/10.1075/gs.2.

Tomasello, Michael (1999): The cultural origins of human cognition. Cambridge & London: Harvard University Press. https://doi.org/10.4159/9780674044371.

Vogt, Rüdiger (2002): Im Deutschunterricht diskutieren: Zur Linguistik und Didaktik einer kommunikativen Praktik. Berlin & Boston: Max Niemeyer Verlag. https://doi.org/10.1515/9783110940794.

Zima, Elisabeth (2013): English multimodal motion constructions: A construction grammar perspective. In: Studies van de BKL: Travaux du CBL: Papers of the LSB, 14-29. https://doi.org/http://uahost.uantwerpen.be/linguist/SBKL/sbkl2013/Zim2013.pdf.

Zima, Elisabeth (2025): Multimodal construction grammar. Cambridge: Cambridge University Press. https://doi.org/10.1017/9781009359856.

Oliver Spiess
Deutsches Seminar der Universität Basel
Nadelberg 4
4051 Basel
Schweiz

oliver.spiess@unibas.ch

Veröffentlicht am 12.11.2026