frontiers | Frontiers in Communication

TYPE Review
PUBLISHED 02 June 2023
DOI 10.3389/fcomm.2023.1187977

OPEN ACCESS

EDITED BY
Claudia Lehmann,
University of Potsdam, Germany

REVIEWED BY
Elisabeth Zima,
University of Freiburg, Germany
Lucien Brown,
Monash University, Australia

*CORRESPONDENCE
Geert Brône
✉ geert.brone@kuleuven.be

†These authors have contributed equally to this work and share first authorship

RECEIVED 16 March 2023
ACCEPTED 03 May 2023
PUBLISHED 02 June 2023

CITATION
Andries F, Meissl K, de Vries C, Feyaerts K, Oben B, Sambre P, Vermeerbergen M and Brône G (2023) Multimodal stance-taking in interaction—A systematic literature review. *Front. Commun.* 8:1187977. doi: 10.3389/fcomm.2023.1187977

Multimodal stance-taking in interaction—A systematic literature review

Fien Andries[1†], Katharina Meissl[2†], Clarissa de Vries[2†], Kurt Feyaerts[2], Bert Oben[2], Paul Sambre[1], Myriam Vermeerbergen[1,3] and Geert Brône[2*]

[1]Department of Linguistics, Faculty of Arts, KU Leuven, Antwerp, Belgium, [2]Department of Linguistics, Faculty of Arts, KU Leuven, Leuven, Belgium, [3]Department of Afrikaans and Dutch, Faculty of Arts and Social Sciences, Stellenbosch University, Stellenbosch, South Africa

Stance-taking, the public act of positioning oneself toward objects, people or states of affairs, has been studied in many fields of research. Recently, its multimodal realization in interaction has received increasing attention. The current contribution aims to take stock of research on multimodal stance-taking so far, and to present possible avenues for future research. We systematically gathered and appraised 76 articles that investigate the involvement of bodily-visual resources in stance-taking in interaction. The critical appraisal focused on two dimensions of the stance act: form-function relations constituting it, and its dynamic organization in interaction. Regarding form-function relations, we found systematic involvement of specific bodily-visual resources in different stance acts, as well as patterns of multimodal intensification and mitigation of stances. As for its dynamic organization, the review discusses how stance-taking is organized temporally throughout an interaction, with all participants involved carefully negotiating and adapting their stances to one another. Finally, attention is paid to the broader context of stance-taking, including its role in different social and societal contexts. Based on this review, we were able to identify several gaps in the literature, and avenues for future research. We argue that much potential for broadening the scope of research lies in increasing the methodological diversity in approaching multimodal stance-taking, as well as in cross-linguistic studies and varying settings and participant constellations. In conclusion, research into multimodal stance-taking is vibrant, with ample opportunities for future work. This review can be considered as a call to action to move beyond the premise that stance-taking is multimodal, and further investigate this intriguing and fundamental human capacity.

KEYWORDS

systematic literature review, stance, stance-taking, multimodality, interaction

1. Introduction

One of the key properties of language is its ability to simultaneously represent subjects, objects or events, and express a stance or viewpoint toward these representations. As a matter of fact, it is hard to conceive of human communication, in spoken, written and signed language alike, as not expressing a form of attitude, viewpoint, belief or evaluation, even if this is largely a neutral positioning (Jaffe, 2009). Not surprisingly, then, *stance-taking* as a socially contextualized and recognized interpersonal phenomenon has received substantial attention in various subbranches of linguistics and neighboring fields (Iwasaki and Yap, 2015; Nir and Zima, 2017; Iwasaki, 2022).

Over the past years, several authors have presented overviews of the existing research on this topic, zooming in on the so-called "types" of stance (epistemic, affective, deontic)[1], its realization by means of lexico-grammatical resources, its embeddedness in context (sequentiality and simultaneity, alignment and disalignment of stance, etc.), but also on the various methodological and terminological approaches that have been developed across different fields (see e.g., Englebretson, 2007; Takanashi, 2018). However, a systematic overview on multimodal stance-taking is missing to date.

An important and inevitable consequence of the broad interest in stance is that different terminology is used to refer to the same (set of) phenomena as well as different definitions of the same terms. Among the labels used to refer to what we described above as stance-taking are: *evaluation, assessment, attitude, appraisal, viewpoint, perspective,* and *subjectivity*. Each of these concepts is couched in a particular theoretical framework, which entails a specific epistemological and methodological positioning. Takanashi (2018) presents an overview of the major strands in the research of stance, independent of these labels and theoretical premises, including the forms and functions of stance, and its relation to the inherent dialogicity of language.

One of the key frameworks that has been developed for the study of stance, the so-called "Stance Triangle", defines stance-taking as "a public act by a social actor, achieved dialogically through overt communicative means, of simultaneously evaluating objects, positioning subjects (self and others), and aligning with other subjects, with respect to any salient dimension of the sociocultural field" (Du Bois, 2007, p. 163). By explicitly adopting a modality-neutral approach ("overt communicative means"), Du Bois paved the way for multimodal analyses of stance-taking.

Whereas the large majority of studies have focused on lexico-grammatical means (see e.g., Biber and Finegan, 1989; Du Bois, 2007; Wang et al., 2022) and to a lesser extent prosodic resources (e.g., Couper-Kuhlen, 2012) for expressing stance, the stance-taking potential of embodied resources has only recently attracted systematic attention. This can be explained, at least in part, by a more general multimodal turn in Conversation Analysis (Goodwin, 2000; Deppermann, 2013; Mondada, 2019), usage-based linguistics (Feyaerts et al., 2017; Perniss, 2018) and other fields. In these fields, language is increasingly being studied as an inherently complex system in which different semiotic resources[2], situated in different modalities (visual and acoustic), interact in construing and negotiating meaning. In this article, we build on the notion of stance as "a multimodal and polysemiotic phenomenon, [meaning] that it involves various semiotic resources in different (perceptual) modalities. We do not limit ourselves to the mere sum of linguistic and gestural, but include any element from any channel or modality that (can be) combine(d) to create meaning" (Andries et al., in press). By opting for this approach, we subscribe to a non-logocentric view of language, treating on a par various embodied resources that contribute to meaning-making in social interaction, including gesture, gaze, facial expressions, head movements, etc. (for overviews, see Cienki, forthcoming; Müller et al., 2013; Bateman et al., 2017; Jewitt, 2017). Including all research that develops a multimodal take on any of the aforementioned concepts related to stance would generate a sample that extends the scope of a systematic review. For this reason, we opted for an approach that is semasiological, focusing only on research that explicitly uses the term "stance".

The present systematic review has two aims: First of all, we provide a systematic overview of existing research on multimodal stance, including a full-fledged analysis of form-function relations as well as the dynamic organization of multimodal stance in interaction. Secondly, we identify gaps in the research and pave avenues for future work.

This article is structured as follows: Section 2 contains an overview of the approach adopted for the collection and screening of texts, eligibility criteria, the coding process and critical appraisal. In Section 3, we provide an overview of the 76 selected articles as well as the synthesis of two major dimensions that surfaced in the critical appraisal. The findings and relevant gaps in the research are summarized and contextualized in a broader scope in Sections 4–6.

2. Method

In this section, we describe the databases and search strategies chosen, the eligibility criteria, our selection process, as well as the synthesis methods we used to arrive at our critical appraisal. Throughout this systematic review, the PRISMA statement was followed (Page et al., 2021). The PRISMA statement contains a set of detailed guidelines to facilitate transparent and complete reporting of systematic reviews and includes guidelines concerning, for instance, the reporting of the Methods, e.g., "Specify the methods used to decide whether a study met the inclusion criteria of the review, including how many reviewers screened each record and each report retrieved, whether they worked independently, and if applicable, details of automation tools used in the process" (Page et al., 2021, p. 4).

2.1. Databases and search strategies

Based on consultation with experts on systematic literature reviews (E. Vancoppenolle and T. Vandendriessche, personal communication, February 2022), as well as a guidance article on conducting systematic literature reviews in the field of Linguistics (Macaro, 2019), we included the following databases in our search: Web of Science (henceforth WoS), JSTOR, Linguistics and Language Behavior Abstracts (henceforth LLBA), and four databases from EBSCO, namely the eBook collection; MLA International Bibliography; Communication and Mass Media; and CINAHL. These databases reflect a wide range of information sources from different fields.

1 Epistemic stance is concerned with the distribution of knowledge, whereas affective stance has to do with attitudes, feelings and emotions that are part of the public record of interaction. Deontic stance, finally, concerns the positioning with regard to the desirability and necessity of an action (Couper-Kuhlen and Selting, 2017).

2 In this article, we distinguish (semiotic) resources from articulators. With articulators, we refer to parts of the body such as eyes, shoulders, hands, etc. Resources are the activities of these articulators, i.e. how these articulators are being used specifically to convey meaning, e.g. gaze, shrug, gesture.

In order to extract all relevant articles from the selected databases, we developed the following search strategy: The two primary concepts that needed to be reflected in our search string were *stance*, and *interaction*. As mentioned in the introduction, we adopted a semasiological approach and only included articles in which the authors use the word 'stance' in the title, abstract or keywords. This means that we did not specifically include articles that investigate 'appraisal', 'attitude' or other terms that can be argued to refer to the same concept (see Takanashi, 2018, p. 174 for a list of alternative terminology connected to stance). The rationale behind this was that broadening the scope of the review would lead to an increasingly fuzzy notion of the concept at stake. To capture the concept of *interaction*, we proposed the following quasi-synonyms as search terms: 'conversation,' 'interaction,' 'talk,' 'discourse,' 'communication,' and 'sign language.' By adding 'sign language' to our search string, we aimed at explicitly including research on signed interaction.

To be included in the search, articles should contain both 'stance' and one of the synonyms for 'interaction' This resulted in the following search string:

> *(stanc* AND (conversat* OR interact* OR talk* OR discourse* OR communicat* OR "sign* language*" OR spoken))*

We applied further filters, depending on the options for each database. For a complete overview of the search strings per database, including filters, see Supplementary Table 1. After an initial search, we inspected the search results in order to ensure that our search string had an accurate scope (i.e., whether other terms should be added or removed), and found that an adaptation was not necessary.

2.2. Eligibility criteria

We defined and continuously adapted the eligibility criteria in an iterative process throughout the screening. In our final selection, we included peer-reviewed book chapters and articles that presented original research (excluding literature reviews, secondary literature, introductions to special issues, as well as conference proceedings). Based on the language proficiencies within our author team, we decided to include articles in English, Dutch, French, German, and Italian[3]. Given our primary focus on stance-taking in the bodily-visual modality, we included articles if they concerned spoken or signed interaction between humans, and excluded articles investigating stance in written or non-human (e.g., human-AI-interaction) forms of communication.

Another set of criteria related to the meaning and interpretation of *stance*. We included articles that investigate the form and function of stance in interaction, and excluded articles in which stance was used in relation to (a) the stance of the authors of an article (e.g., "in this article, we take the stance that…"), (b) intentional stance, which pertains to the philosophical concept that ascribes intentions or agency to people, AIs or other systems (Dennett, 1987), (c) physical stance, when not connected to the notion of stance as expression of evaluations, and (d) lexico-grammatical markers in which the multimodal aspects of stance-taking were not considered. Lastly, we excluded articles concerning (e) the content of stance, without focusing on its form as such. For example, an article was excluded which investigates "levels of paternal involvement during the decision-making process for childbirth method" (Andreou et al., 2022), and the stances fathers have about certain types of childbirth methods, without discussing how exactly these stances are expressed. Articles on identity construction or political discourse often fell into this content-related category.

2.3. Screening process

Figure 1 shows a flowchart of the screening process. The final search in all databases yielded a total of 8,423 references. After duplicate removal, 6,776 articles were left. Abstracts were screened based on the eligibility criteria described above. For the initial screening, we used the free online software Rayyan (Ouzzani et al., 2016). Rayyan allows the creation of exclusion criteria and labels as well as highlighting key words in texts to aid the screening process. An inter-coder agreement test for 6.8% of the articles ($n = 459$), was performed by two junior researchers and one senior researcher (KM, CV and GB). The test showed almost perfect agreement (Fleiss' Kappa = 0.819, %-agreement = 95.0). The complete author team revised conflict cases together and the criteria were refined. The remaining abstracts were divided and screened individually by the three coders above, who also resolved any remaining cases of doubt together.

After the abstract screening, 244 articles were included for a full-text screening. Using the eligibility criteria described above, we limited the selection of articles to those that concerned stance in the visual modality, be it in combination with other resources or not. As a consequence, we excluded texts that concerned (a) only prosodic markers of stance, and (b) only liminal signs such as laughter, clicks or other vocalizations (Dingemanse, 2020). Another inter-coder agreement test was performed for 8.6% of the articles ($n = 21$), by the three junior researchers (FA, CV and KM), which showed substantial agreement (Fleiss' Kappa = 0.795, %-agreement = 85.7). The remaining texts were divided and screened individually, and cases of doubt were resolved between the three coders. In this phase, a total of 168 articles was excluded, resulting in the final selection of 76 articles for data collection and critical appraisal.

2.4. Coding process

76 articles were imported into Nvivo (QSR International Pty Ltd, 2022). All co-authors participated in this phase and coded the articles adhering to the codebook, which can be found in Supplementary Table 2. The codes and definitions were negotiated in several steps, and sub-codes were added to specify or standardize codes such as 'analytical framework', or 'multimodal package'. We defined the latter as passages that describe examples and results containing information regarding the different semiotic resources at stake. Importantly, the codes were all based on what is explicitly stated in the texts.

3 Note that we did not translate our search string into different languages, leading to search results with articles that had English titles and/or abstracts.

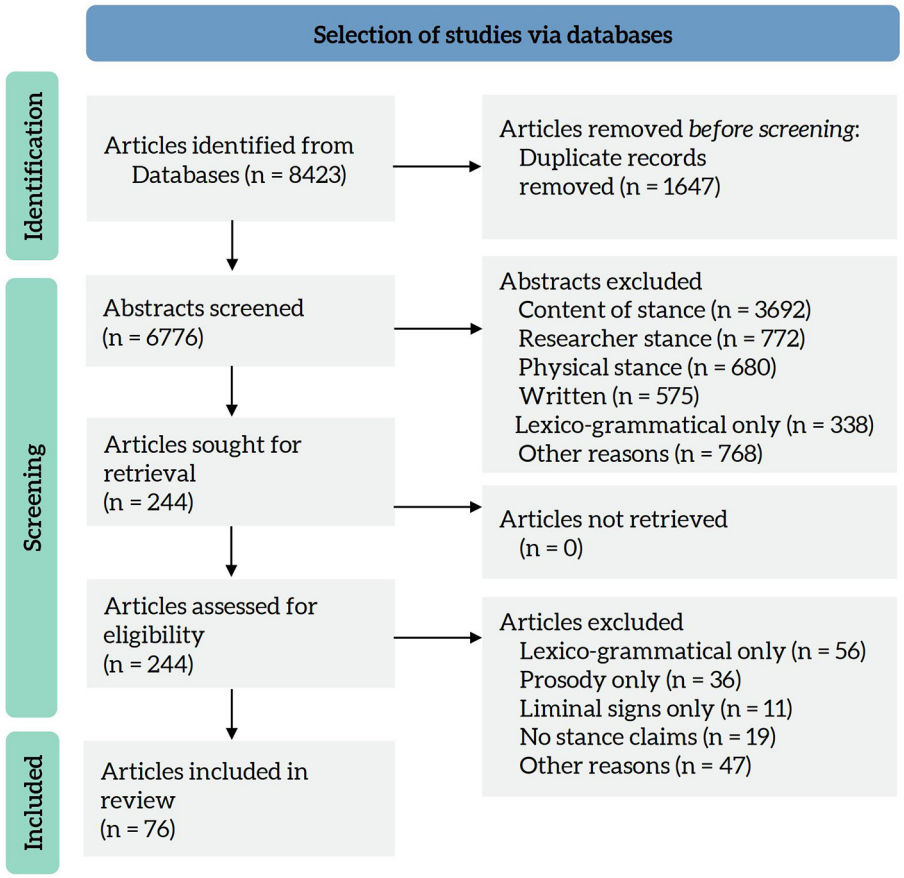

FIGURE 1
A flow chart showing four steps for the selection of relevant studies. First, relevant articles were identified in databases ($n = 8423$) and duplicates were removed ($n = 1647$). Next abstracts were screened. Here, abstracts were excluded for various reasons: "content of stance" ($n = 3692$), "researcher stance" ($n = 772$), "physical stance" ($n = 680$), "written" ($n = 575$), "lexico-grammatical only" (338), and other ($n = 768$). In the third step, the full text was retrieved for 244 articles. Fourth, the full texts were assessed for eligibility. Articles were excluded for various reasons: "lexico-grammatical only" ($n = 56$), "prosody only" ($n = 36$), "liminal signs only" ($n = 11$), "no stance claims" ($n = 19$), and other ($n = 47$). In total, 76 articles were included in the review.

The goal of this process was to gain an overview of general information about the selected articles, which can be found in Supplementary Table 3. In the next phase of coding and annotating, we went beyond explicit information given in the texts and strived for more synthesis and interpretation to identify common themes in research aims or claims pertaining to multimodal stance.

2.5. Critical appraisal

All coders identified research aims or questions as well as claims made in the selected publications pertaining to stance in the bodily-visual modality. To identify claims, coders decided which text passages contained not only specific analytical observations but more general assertions about multimodal stance. Three authors (FA, CV, and KM) detected common themes connecting the articles, taking a bottom-up approach. The rest of this contribution is structured along the following dimensions:

(1) Form-function relations, including specific form-function relations and the intensification and mitigation of stance. (2) Stance as a dynamics process, more specifically, the temporal and collaborative organization of stance as well as stance in a broader context.

For each of these dimensions, specific claims per article were revisited and related to the analyses and multimodal packages. Importantly, we acknowledge that the dimensions listed above are in no way mutually exclusive. However, they mirror higher-level theoretical questions connected to stance-taking and multimodal interaction in general, and are therefore discussed separately in this review.

3. Results

3.1. Overview of the included articles

Let us start by summarizing the selection of articles that feature in the current review. As can be seen in Supplementary Table 3, our selection includes 76 articles, spanning a range of analytical frameworks and languages of interaction. Although we did not limit our search according to date of publication, the earliest

publication in our review dates back to 2005, highlighting the recency with which the term stance has gained prominence in relation to multimodality. Almost two thirds of the selected articles place themselves in the analytical framework[4] of Conversation Analysis ($n = 44$), an overwhelming majority of the articles took a qualitative approach ($n = 60$), with few articles using mixed methods ($n = 9$) or quantitative methods ($n = 7$). With respect to the languages under scrutiny, English is the most frequently investigated language ($n = 32$) followed by Swedish ($n = 9$) and Finnish ($n = 8$). As regards the stance types being studied in the articles, most articles focus on affective stance ($n = 39$), while less attention goes to epistemic stance ($n = 28$) and very few articles focus on deontic stance ($n = 5$). Finally, almost half of the selected articles ($n = 36$) do not use the labels above.

Not surprisingly given our selection criteria, almost all articles highlight that lexico-grammatical resources do not constitute a stance act by themselves, but instead cooperate with bodily-visual resources, also taking into account the embodied participation framework and local material context of the interaction. Stance acts can be accomplished in the bodily-visual modality only (e.g., Berger and Rae, 2012; Van De Mieroop, 2020) as well as in combination with verbal resources (e.g., Goodwin C., 2007; Stoenica and Fiedler, 2021). In most cases, stance acts arise not by isolated resources, bodily-visual or other, but as a "gestalt" (Mondada, 2014) in which the stance is expressed through a combined effect of all resources working in concert (Jensen, 2014; Pinto and Vigil, 2019). See as an example the case of the shrug: The associated stance of indexing obviousness emerges as a compounded effect of the opening of the hands, raising of the shoulders, tilting and shaking of the head and raising of the eyebrows, or any combination of these components (Jehoul et al., 2017; Marrese et al., 2021). Sometimes the verbal and bodily-visual modality are described to have a contrastive relation, and the stance taken in the verbal modality is different from the one in the bodily-visual modality (e.g., Deppermann and Gubina, 2021). In Section 3.2, we will dig deeper into the possible relations between different articulators and resources at stake.

Another recurring claim in the selection of articles is that bodily-visual resources present affordances of stance-taking unavailable in the verbal modality. The possibility to use multiple semiotic resources (bodily-visual or other) simultaneously, gives rise to a wide range of options for participants to time their stance display, and continuously adapt their stance to that of their interlocutor, without interrupting talk. This will be discussed in more detail in Sections 3.3.1 and 3.3.2.

Furthermore, the bodily-visual modality affords a variety of possibilities to stay "off the record" (Ford et al., 2012, p. 209), as in avoiding explicitly stating potentially inappropriate assessments verbally. As Ford and colleagues put it: "by using a bodily-visual modality, [the participant] displays an orientation to the action as one that should be obvious and, therefore, inferable" (Ford et al., 2012, p. 209). Due to this inferable nature of the bodily-visual modality, it is particularly suited for the expression of stance on various occasions such as to playfully withhold information (Ford et al., 2012), to avoid face-threatening actions (Berger and Rae, 2012), to motivate other-completion of delicate assessments (Li, 2021) or to display a stance in institutional contexts like courtrooms (Matoesian, 2018). This aspect will be revisited in Section 3.3.3. We now present the critical appraisal of the literature, with a focus on two main dimensions: form-function relations, and dynamicity of stance-taking.

3.2. Form-function relations

3.2.1. Specific form-function relations

One of the main questions many scholars who study multimodal stance are driven by is which semiotic resources speakers and signers use to convey stance and how they are related to each other. Many articles in this review have, amongst other things, analyzed different resources and provide partial responses to this overarching question. The following paragraphs are based on Table 1 which has a two-fold purpose: first, it provides an overview of resources that have already been linked to specific function types, and second, it serves as reference list for scholars interested in specific stance expressions[5].

As far as relations between multimodal packages and specific functions are concerned, we were able to identify three established patterns which are mentioned in multiple articles. Firstly, we observed that a shrug, which can contain head tilting, raising of the shoulders and eyebrows, a stretched down mouth and/or a palm up gesture, can be used to express either low certainty (Roseano et al., 2016) or obviousness (Jehoul et al., 2017). Although contradictory at first glance, in both cases, the shrug indicates a nuance of distancing, either from a stance object or co-participant(s).

Secondly, we found a relation between several bodily-visual resources and the expressed proximity or distancing toward stance subjects or objects. For instance, proximity may be displayed through gazing at (Mondada, 2009; Kääntä, 2014; Churchill, 2022), moving toward (Cekaite, 2007; Mondada, 2009; Kääntä, 2014; Pauletto et al., 2017), pointing at (Mondada, 2009; Butler and Edwards, 2018) or touching (Pauletto et al., 2017) a stance object or co-participant. These activities can be tied down both to epistemic access (in reference to the stance object) as well as affiliation (regarding stance subjects). Also note that the same form—for example gazing at the stance object—can have differing functions such as expressing disengagement (Heller, 2021) or directing attention toward it (Mondada, 2009; Kääntä, 2014; Churchill, 2022), depending on its sequential positioning (see Section 3.3.1).

A third form-function relation that emerges from our analysis is the expression of a change-of-state by the recipient through

4 The label "not specified" was assigned whenever authors did not explicitly place their work in any specific analytical framework. It should be noted, however, that our review includes several articles by pivotal figures in their respective areas of research, which may result in analytical frameworks not being specifically mentioned in articles by, for example, M.H. or C. Goodwin as well as Couper-Kuhlen or Mondada.

5 Note that this table does not include any information on the specific sequential position of these resources in the course of interactions, nor does it take into account the full multimodal packages. For example, raised eyebrows, which can constitute part of a compound 'shrug' are listed separately in the table. The temporality of stance is discussed separately in Section 3.3.1.

10.3389/fcomm.2023.1187977

TABLE 1 Specific form-function relations.

Semiotic resource		Form/design feature	Interactional function	Relevant studies
Head		Head tilt	Negative stance	Clift, 2014
		Head tilt	Request	Hübscher et al., 2019
		Head tilt	Obviousness	Jehoul et al., 2017
		Head tilt	Questioning stance	Kääntä, 2014
		Head shake	Resistant compliance	Clift, 2021
		Head shake	Refusal	Evaldsson and Melander, 2017
		Head shake	Obviousness	Jehoul et al., 2017
		Head shake	Distancing	Kärkkäinen, 2012
		Head shake	Intensification of semantic content of stance	Kärkkäinen, 2012
		Head shake	Positive stance	Mondada, 2009
		Neck roll	Confrontational, argumentative stance	Goodwin and Alim, 2010
		Head nod	Doing thinking	Heller, 2021
		Head nod	Acknowledgment	Mondada, 2009
		Head nod	(Re)alignment	Muntigl et al., 2012
		Head nod	Shared affective stance	Pfänder and Couper-Kuhlen, 2019
		Head nod	High certainty	Roseano et al., 2016
		Head nod	(Dis)affiliation	Stivers, 2008
		Head down	Request	Hübscher et al., 2019
		Head orientation	Epistemic access	Mondada, 2009
		Head movement	Accompanying pragmatic or discourse marker	Bolly and Boutet, 2018
		Head movement	Intensification of stance	Bolly and Boutet, 2018
Face	Mouth	Smiling	Backchanneling	Feyaerts et al., 2022
		Smiling	Amused or humorous stance	Ruusuvuori and Peräkylä, 2009; Kaukomaa et al., 2013; Feyaerts et al., 2022
		Smiling	Request	Hübscher et al., 2019
		Smiling	Marking as delicate	Li, 2021
		Smiling	Affiliation	Ruusuvuori and Peräkylä, 2009
		Smiling	Expectant stance	Soulaimani, 2018
		Pressed lips	Negative stance	Li, 2021; Hou, 2022
		Pouted lower lip	Disapproval	Li, 2021
		Mouth stretched down	Low certainty	Roseano et al., 2016
		Mouth movement	Accompanying pragmatic or discourse marker	Bolly and Boutet, 2018
	Eyes	Eye roll	Resistant compliance	Clift, 2021
		Eye roll	Confrontational, argumentative stance	Goodwin and Alim, 2010
		Squinting	Amusement	Feyaerts et al., 2022
		Squinting	Doing thinking	Heller, 2021
		Squinting	Puzzlement	Hou, 2022
		Squinting	Distancing	Kärkkäinen, 2012
		Squinting	Intensification of semantic content of stance	Kärkkäinen, 2012
		Widened eyes	Surprise	Kaukomaa et al., 2015
		Widened eyes	Shock	Schröder, 2020
		Stirring gaze	Heightened emotional intensity	Evaldsson and Melander, 2017

(Continued)

TABLE 1 (Continued)

Semiotic resource		Form/design feature	Interactional function	Relevant studies
	Eyebrows	Raised brows	Reaction to obviousness	Feyaerts et al., 2022
		Raised brows	Request	Hübscher et al., 2019
		Raised brows	Obviousness	Jehoul et al., 2017
		Raised brows	Surprise	Kaukomaa et al., 2015
		Raised brows	Negative stance	Li, 2021
		Raised brows	Shock	Schröder, 2020
		Eyebrow flash	Change of knowledge state	Mondada, 2009
		Frown	Marking as problematic	Ruusuvuori and Peräkylä, 2009; Kääntä, 2014; Heller, 2021; Li, 2021
		Frown	Puzzlement	Hou, 2022
		Eyebrow movement	Accompanying pragmatic or discourse marker	Bolly and Boutet, 2018
Gaze		Gaze aversion	Resistance	Cekaite, 2007; Evaldsson and Melander, 2017
		Gaze aversion	Divergent stance	Haddington, 2006; Ekström, 2012; Toomaneejinda and Harding, 2018; Pekarek Doehler et al., 2021
		Gaze aversion	Doing thinking	Heller, 2021
		Gaze at stance object	Directing attention to stance object	Mondada, 2009; Kääntä, 2014; Churchill, 2022
		Gaze at stance object	Doing thinking	Heller, 2021
		Shared gaze at stance object	Shared attention to stance object	Haddington, 2006; Mondada, 2009; Churchill, 2022
		Mutual gaze	Shared stance	Haddington, 2006; Jensen, 2014; Toomaneejinda and Harding, 2018
		Mutual gaze	Equal epistemic rights	Pfänder and Couper-Kuhlen, 2019
		Gaze at co-participant	Affiliation	Vranjes et al., 2018; Stoenica and Fiedler, 2021
		Gaze at co-participant	Change of knowledge state	Vranjes et al., 2018
		Gaze at co-participant	Distancing	Vranjes and Brône, 2021
		(Accelerated) gaze alternation	Invitation for affiliation	Rühlemann et al., 2019
Arms or shoulders		Shoulder shrug	Request	Hübscher et al., 2019
		Shoulder shrug	Obviousness	Jehoul et al., 2017
		Shoulder shrug	Low certainty	Roseano et al., 2016
		Arm-flapping	Attracting visual attention	Butler and Edwards, 2018
		Arm-flapping	Negative stance	Butler and Edwards, 2018
Trunk or body		Orientation toward referent or co-participant	Directing attention to stance object	Cekaite, 2007
		Orientation toward referent	Directing attention to stance object	Mondada, 2009; Kääntä, 2014
		Orientation toward co-participant	Preparation for turn-sharing	Pfänder and Couper-Kuhlen, 2019
		Change of posture	Change of knowledge state	Mondada, 2009
		Movement toward co-participant	Display of closeness	Pauletto et al., 2017

(Continued)

TABLE 1 (Continued)

Semiotic resource		Form/design feature	Interactional function	Relevant studies
		Movement toward co-participant	Stance or sentiment questions	Trujillo and Holler, 2021
		Movement away from co-participant	Disalignment, disagreement	Honda, 2010
		Lateral or forward leaning	Request	Hübscher et al., 2019
		Forward and backward swaying	Shared affective stance	Jensen, 2014
		Sychronized posture	Shared affective stance	Pfänder and Couper-Kuhlen, 2019
		Frozen body movement	Change of state	Schröder, 2020; Heller, 2021
		Deflation	Negative stance	Clift, 2014
		Forceful body movement	Refusal	Evaldsson and Melander, 2017
		Enlarged torso movement	Stance or sentiment questions	Trujillo and Holler, 2021
Hands		Open hand(s)	Obviousness	Jehoul et al., 2017; Marrese et al., 2021
		Open hand(s)	Distancing	Kärkkäinen, 2012
		Open hand(s)	Intensification of semantic content of stance	Kärkkäinen, 2012
		Palm-up gesture	Attracting visual attention	Butler and Edwards, 2018
		Palm-up gesture	Negative stance	Butler and Edwards, 2018
		Hand position toward object	Epistemic rights	Norén et al., 2021; Churchill, 2022
		Hand position toward co-participant	Affiliation	Trujillo and Holler, 2021
		Pointing gestures	Attracting visual attention	Butler and Edwards, 2018
		Pointing gestures	Negative stance	Butler and Edwards, 2018
		Pointing gestures	Evidentiality	Roseano et al., 2016
		Pointing gestures	Epistemic stance	Stirling et al., 2022
		Pointing gestures to stance object	Epistemic access	Mondada, 2009
		Acoustic pointing gesture	Authorative stance	Gilbert, 2018
		Hand covering face	Shame	Schröder, 2020
		Emphatic hand gestures	Negative stance	Butler and Edwards, 2018; Schröder, 2020
		Pragmatic hand gestures	Stance	Bolly and Boutet, 2018
Physical action		Forceful physical action	Intensification of negative stance	Evaldsson and Melander, 2017
		Manipulating object	Epistemic access	Mondada, 2009
		Moving toward/touching co-participant	Display of closeness	Pauletto et al., 2017
		Pushing away co-participants' hands	Epistemic rights	Norén et al., 2021

eye and eyebrow movements (Mondada, 2009; Kaukomaa et al., 2015; Schröder, 2020), which may be accompanied by frozen body posture (Schröder, 2020). Here, the widened eyes and raised eyebrows can index shock (Schröder, 2020) or surprise (Kaukomaa et al., 2015), and eyebrow flashes may indicate a change in knowledge state (Mondada, 2009). All of these displays can index that a verbal elaboration on the stance may follow (see also Section 3.3.1), and are epistemically as well as affectively loaded, on a continuum ranging from more negative to more positive. At this point, it may be interesting to note that only few articles in our selection explicitly mention the combination of affective and epistemic stance ($n = 3$). In contrast, affective stance on its own appears in almost half of the claims ($n = 33$), followed by epistemic stance in about a quarter of the claims ($n = 21$). For a more elaborate description of the distribution of stance types over the selection of articles, see Supplementary Table 4.

On the level of individual forms, regardless of the multimodal package they appear in, several patterns emerge. We found multiple claims in our selected articles stating that head nods are used for acknowledgments (Mondada, 2009) or alignment (Muntigl et al., 2012; Pfänder and Couper-Kuhlen, 2019) and that head shakes can be used to express distancing or disbelief (Kärkkäinen, 2012; Evaldsson and Melander, 2017). Moreover, a frown is described to problematize a topic or stance object (Ruusuvuori and Peräkylä, 2009; Kääntä, 2014; Heller, 2021; Li, 2021) while interactants smile

to express amusement (Ruusuvuori and Peräkylä, 2009; Kaukomaa et al., 2013; Feyaerts et al., 2022). Note again that these stance acts contain both affective and epistemic dimensions of multiple polarities. Zooming out, about two thirds of the articles link a cluster of multiple articulators to one function, whereas one third link only one resource to a function, highlighting the 'gestalt' nature of stance acts. For a more elaborate description of the distribution of the articulators involved in stance expression over the selection of articles, see Supplementary Table 4.

3.2.2. Intensification and mitigation of stance

A subset of articles in our selection describes form-function relations with reference to the intensity of a stance. Multimodal construction plays an important role in the intensification and mitigation of stance, surfacing in three aspects: (1) the relation between the verbal and bodily-visual modality; (2) their temporal organization; and (3) effort and goal-orientation.

Similar to what was mentioned in the overview, the relation between the verbal and bodily-visual modality can take different forms. It can be one of convergence, with multimodal resources "boosting" (Van De Mieroop, 2020, p. 605), or intensifying a stance (Evaldsson and Melander, 2017; Marrese et al., 2021). On the other hand, it can be one of contrast, with multimodal resources either distancing the speaker from the stance object and another stance subject (Kärkkäinen, 2012) or implicitly resisting a line of action by another speaker (Cekaite, 2020).

With regard to the temporal organization, several studies have documented sequences of intensification within one participant ('intra-participant') (Cekaite, 2007; Keel, 2015; Evaldsson and Melander, 2017), in which marked bodily conduct is contrasted with previous unmarked bodily conduct (Matoesian, 2005). In these cases, an initially primarily verbal stance act may be upgraded in a 'second, multimodal attempt', which then serves as an intensification of the 'first attempt'. In other cases, stances are intensified locally, when they occur in moments of "heightened embodiment" (Kärkkäinen, 2012, p. 497). Here, intensification can be achieved either by drawing on more resources (Kärkkäinen, 2012; Muntigl and Horvath, 2014; Evaldsson and Melander, 2017; Hübscher et al., 2019; Van De Mieroop, 2020), by a larger or more animated use of resources (Goodwin and Alim, 2010; Keel, 2015; Cekaite, 2020), or by temporally aligning the use of manual gesture with speech (Matoesian, 2005), see also Section 3.3.1. In light of these observations, multimodal intensification presents itself as a gradual phenomenon (Pinto and Vigil, 2019) offering a multitude of ways of expression.

Multimodal intensification and mitigation seem to be a gradable, flexible effort, challenging participants to continuously adapt the amount and nature of their multimodal marking in the most efficient way to the ever-evolving communicative situation. We found this both in cases of sequential intra-participant intensification of stance, as well as in cases of interpersonal (or inter-contextual) intensification. For instance, children tend to use more multimodal resources during requests in more face-threatening situations (requesting from an unknown experimenter) compared to less face-threatening situations (requesting from a classmate) (Hübscher et al., 2019). By continuously adapting to the situation at hand, the phenomenon of multimodal intensification of stance qualifies as a 'goal-oriented process'. This is manifested in various ways, such as by obtaining a personal goal or effect, e.g., when requesting an object (Hübscher et al., 2019) or attention (Cekaite, 2007); resisting someone else's goal, e.g., in non-compliance (Evaldsson and Melander, 2017; Cekaite, 2020); and increasing or decreasing distance from other participants, i.e., other stance subjects (Goodwin and Alim, 2010; Ekström, 2012; Kärkkäinen, 2012; Keel, 2015), see also Section 3.3.3.

3.3. Stance as a dynamic process

3.3.1. Temporal organization of stance

A large number of articles in the selection contains claims pertaining to the temporal organization of stance-taking. The following synthesis emerges from the combination of four parameters: phenomena that are primarily realized as actions of one participant ('intra-participant') or across participants ('interpersonal'), at the same time ('simultaneously') or temporally unfolding ('sequentially'), resulting in the four subsets discussed in the following paragraphs. Articles containing claims about several phenomena, also in overlap, will reappear in the respective paragraphs.

Firstly, a subset of articles discusses specific combinations of simultaneously occurring use of different semiotic resources expressed by one participant (intra-participant), which, in their combination, fulfill a certain stance function. Such multimodal packages occur in different combinations of semiotic resources. Some are linked to specific verbal utterances (Pekarek Doehler et al., 2021; Stoenica and Fiedler, 2021), e.g., the turn-final French phrase *tu vois* ("you see") with a rising intonation occurs simultaneously with the speaker's gaze directed at the recipient (Stoenica and Fiedler, 2021). Other packages consist of liminal signs, such as clicks combined with head movements (Pinto and Vigil, 2019). These gestalts may also be purely visual, as in the example of thinking displays (Heller, 2021). In all these cases, one participant draws on multiple resources simultaneously to express stance(s). Some articles mention the use of polyphonic strategies in stance expressions, i.e., the way speakers or signers present multiple stances from different viewpoints at the same time, which may or may not be contradicting (Kärkkäinen, 2012; Soulaimani, 2018; Groenewold and Armstrong, 2019; Donzelli, 2020). When speakers draw on enactment to represent someone else's stance, they can simultaneously use visual resources, such as manual gestures (Kärkkäinen, 2012; Soulaimani, 2018; Groenewold and Armstrong, 2019; Donzelli, 2020), squinting of the eyes, and head shakes (Kärkkäinen, 2012) to disaffiliate with the reported stance (see also Section 3.3.3).

Secondly, intra-participant stance displays are studied in relation to their sequential unfolding. These can be split up into stance displays that occur in 'turn-initial,' 'mid turn,' or 'turn-final' position. For example, turn-initially, cut-off gaze can precede an utterance with a divergent stance (Haddington, 2006; Pekarek Doehler et al., 2021). Manual gestures and facial expressions are described as having the potential to announce or negotiate

stance visually before it is done so verbally (Berger and Rae, 2012; Kaukomaa et al., 2013, 2015; Norén et al., 2021), similar to specific visual resources co-occurring with clicks before stanced utterances (Ogden, 2013). Skogmyr Marian (2021) observes that in complaining sequences, visual displays, such as covering one's eyes with the hands and shaking the head, can signal stance early on. In a mid-turn position, Marrese et al. (2021) observe palm-up gestures at the same time as verbal utterances to express or intensify stances of obviousness (Marrese et al., 2021). Iwasaki (2015) describes that speakers of Japanese pause their turn and establish mutual gaze in order to be able to include their co-participant's reaction in finishing their own turn. For the turn-final position, the potential for withholding information and potentially invoking other-completion after incomplete syntax is discussed (Ford et al., 2012; Li, 2021). There are also stance displays which can potentially occur at any of the three given points with reference to a participant's turn (Jehoul et al., 2017; Deppermann and Gubina, 2021). For instance, depending on the sequential order of permission-seeking and action completion, different deontic stances are expressed (Deppermann and Gubina, 2021). Relating to longer stretches of interaction, an increase in stance expressions and recruitment for alignment are observed around climaxes and endings of storytelling (Ruusuvuori and Peräkylä, 2009; Niemelä, 2010; Rühlemann et al., 2019).

Thirdly, some articles deal with stance phenomena across participants (inter-participant) occurring simultaneously, thus closely linked to the collaborative nature of stance-taking (see Section 3.3.2). For instance, recipients are observed to perform thinking displays during a speaker's turn (Heller, 2021), or participants perform synchronized actions at the same time (Katila and Philipsen, 2019; Pfänder and Couper-Kuhlen, 2019).

Fourthly, stance displays are studied on an inter-participant (interpersonal) level in their sequential unfolding. Some articles discuss specific stance displays with respect to their sequential position in an interaction (Stivers, 2008; Clift, 2014; Kääntä, 2014; Pillet-Shore, 2018, 2020). Others observe certain reactions following stance displays: facial expressions responding to expressions of obviousness (Feyaerts et al., 2022), the timing of head nods (Stivers, 2008; Muntigl et al., 2012) and turn-opening smiles (Kaukomaa et al., 2013) depending on alignment or disalignment with the previous participant's stance. Other effects of stance displays on the sequential unfolding of an interaction are observed with respect to storytelling and promoting alignment (Ruusuvuori and Peräkylä, 2009; Bateman, 2020), student-teacher interactions (Sharma, 2013; Evaldsson and Melander, 2017), as well as collaborative reasoning (Heller, 2021). Finally, Churchill (2022) studies the diachronic development of multimodal stance-taking in an instructional setting, showing how changes in knowledge of the student are reflected in the recipient design of instructions by the teacher.

3.3.2. Collaborative organization of stance

The fact that stance-taking is a collaboratively achieved phenomenon has been demonstrated in several previous studies (cf. Takanashi, 2018). In this review, the notions of collaboration, cooperation and co-construction surface in about half of the articles.

First of all, several general observations are made with regard to the visual modality: a shared (embodied) participation framework and epistemic access are prerequisites for any collaborative stance-taking. Epistemic access and rights are shaped by, for example, body positioning (Goodwin M. H., 2007; Goodwin and Goodwin, 2012; Melander, 2012; Melander Bowden, 2019; Norén et al., 2021), gaze behavior (Haddington, 2006), and deictic gestures (Mondada, 2009). Deontics and agency can also be managed through embodied actions, for instance, in presuming the intersubjectivity of the permissibility of actions (Deppermann and Gubina, 2021). Some articles explicitly mention a division of labor between speaker and addressee with regard to who initiates an expression of stance which can then be taken up by another participant (Kaukomaa et al., 2015; Heller, 2021).

Secondly, linked to the notion of initiation and uptake, several articles describe how one stance elicits another stance with the goal of alignment. Some do so without mentioning specific form-function relations (Matoesian, 2005; Cekaite, 2007; Muntigl and Horvath, 2014; Tainio and Laine, 2015). In the setting of storytelling, narrators' visual stance displays are observed to assure fit responses by recipients and eventual reciprocation of the stance expressed (Ruusuvuori and Peräkylä, 2009; Niemelä, 2010; Soulaimani, 2017), rendering affective stance a "systematic collaborative achievement" (Bateman, 2020, p. 646). From the recipient's perspective, Pillet-Shore (2020) describes that stance displays along with context information aid to infer which response is adequate or expected. Specific forms of stance displays and reactions to them are observed, for example, by Iwasaki (2015) concerning the use of gaze and turn suspension, see Section 3.3.1. Other examples include the reactions to expressions of obviousness (Jehoul et al., 2017; Feyaerts et al., 2022) or whining (Butler and Edwards, 2018). So called "thinking displays" are found to emphasize the joint and emergent character of decision-making (Heller, 2021) and often result in reciprocation of the same form. Another action described as an invitation for collaboration, and more specifically collusion, is the eye roll (Clift, 2021).

Thirdly, an uptake of the same or similar forms as shared stance-taking plays an important role in alignment, for instance, in the case of reciprocation of smiles (Kaukomaa et al., 2013) and recycling of manual gestural material (Goodwin M. H., 2007; Soulaimani, 2017). Uptake of verbal elements of a previous participant can also be accompanied by mutual gaze (Haddington, 2006) in the case of alignment. Some studies mention processes of synchronization, where participants perform the same stance expression simultaneously, such as manual gesture (Katila and Philipsen, 2019) or laughter as a bodily gestalt (Jensen, 2014). Two special cases of turn-sharing which serve affective stance-taking are discussed by Pfänder and Couper-Kuhlen (2019): "choral performance" and "chiming in", accompanied by visual markers, such as gaze and head movement.

Fourthly, while most articles discuss the collaborative process of alignment, some do explicitly mention how disalignment can take form. For instance, Honda (2010) describes that movement away from co-participants signals disalignment. An important aspect in

this regard is the timing of stance displays: in environments of disalignment, a delay of reciprocated smiling (Kaukomaa et al., 2013) or nodding (Muntigl et al., 2012) is observed. Evaldsson and Melander (2017) show that disalignment through the expression of negative emotions is still mutually organized and involves constant change and adaptation.

A final aspect with respect to the collaborative nature of stance-taking is the distinction between (dis)alignment and (dis)affiliation. Alignment, on the one hand, is described as a "structural endeavor" (Kunitz and Jansson, 2021, p. 45), influencing the sequential progressivity of an (inter)action. Affiliation, on the other hand, is linked to stance-sharing, identity and relationship work and can be reached both by aligning with another participant or disaligning with a third party (Melander, 2012; Kunitz and Jansson, 2021). Similarly, Stivers (2008) distinguishes different recipient reactions for alignment and affiliation. More large-scale implications of stance on, for example, emotion work, as well as contextual factors influencing stance will be discussed in the next section.

3.3.3. Stance in a broader context

Given that stance is expressed by social actors in an intersubjective manner, there is an interplay between the form of stance expressions and (the consequences of) its broader function in a social and societal context. Several articles in our selection discuss this interplay, in relation to interpersonal relations, identity construction, as well as the normative organization of interaction.

The aspect of interpersonal dynamics is discussed in several studies both with regard to immediate and long-term effects (Sharma, 2013; Jensen, 2014; Keel, 2015; Tainio and Laine, 2015; Katila and Philipsen, 2019). Regarding long term effects, Tainio and Laine (2015) describe how repeated negative stance displays by teachers responding to incorrect student answers can affect the general attitude toward a subject. In a similar way, aligning or re-enacting together can help building longer-term relations based on the local dynamics in interaction (Jensen, 2014; Katila and Philipsen, 2019).

Furthermore, by examining multimodal stance-taking, other patterns, related to inter- or cross-cultural encounters as well as identity construction can be uncovered (Goodwin M. H., 2007; Mondada, 2009; Goodwin and Alim, 2010; Sharma, 2013; Gilbert, 2018; Schröder, 2020; Van De Mieroop, 2020; Kunitz and Jansson, 2021; Liang, 2021). For instance, Gilbert (2018) describes how police officers use acoustic gestures (finger snaps) and pointing gestures as a way of enacting their authoritative identity during stance-taking.

Lastly, there are various norms and expectations with regard to how stance ought to be expressed. These are either more general, societally embedded (Goodwin M. H., 2007; Berger and Rae, 2012; Rekittke, 2017; Butler and Edwards, 2018; Hübscher et al., 2019; Pinto and Vigil, 2019; Bateman, 2020) or more local, linked to a specific institutional setting such as classrooms (Cekaite, 2007, 2020; Kääntä, 2014; Evaldsson and Melander, 2017) or courtroom interaction (Matoesian, 2018). In a classroom setting, students can time stance displays within the familiar temporal organization of Initiation-Response-Evaluation sequences without disrupting the activity (Cekaite, 2007). Similarly, in courtrooms, participants may draw on visual stance displays given existing restrictions on when to speak and what to say (Matoesian, 2018). Connected to more general societal norms, some articles mention the notion of responsibility for stance expressions. Certain visual resources may be employed to take less responsibility or none at all for a certain stance (Kärkkäinen, 2012; Janzen and Shaffer, 2013; Jones et al., 2015; Vranjes et al., 2018; Li, 2021). For example, interactants may draw on visual stance expressions after incomplete syntax to avoid verbalizing assessments and thereby stay "off the record" (Ford et al., 2012, p. 209; Li, 2021).

4. Summary of the findings

With this systematic literature review, we aimed firstly, to provide an overview of existing studies on multimodal stance-taking, including a full-fledged analysis of relevant analytical dimensions in these texts, and secondly, to identify gaps in the research and pave avenues for future work. In what follows, we provide a summary of our findings. Through the critical appraisal, we were able to map out and analyze two major dimensions, *Form-function relations* and *Stance as a dynamic process*, within our set of 76 articles.

Various articles in this review make claims on form-function relations connected to multimodal stance-taking. The review provides an overview of the ways in which all articulators that were coded (face, gaze, head, arms or shoulders, hands, and trunk or body) can be combined to express stance.

Interactants may also intensify, foreground or mitigate a stance multimodally. This intensification can be regarded as a flexible investment, that is adapted to the interactional needs. Bodily-visual resources can be used to upgrade an initially primarily verbal stance-act, resulting in a 'second, multimodal attempt'; or they can be used simultaneously, in moments of "heightened embodiment" (Kärkkäinen, 2012).

With regard to the temporal organization of stance, the review points to two relevant parameters: First, in intra-participant stance expressions, semiotic resources can either be stacked simultaneously or over the course of a turn. Second, inter-participant stance expressions are built collaboratively, again, either simultaneously or across longer stretches of an interaction. Zooming out from individual analyses, this foregrounds the potential for stance-taking at any point, reiterating that, indeed, "[…] there is never a time out from the social action of taking stances and adopting positions" (Du Bois and Kärkkäinen, 2012, p. 438).

The review shows that collaborative stance-taking benefits greatly from the dynamic affordances of the bodily-visual modality: during turns-at-talk, bodily-visual resources allow for ensuring fit responses, opening the floor for negotiation, and synchronization of stance expressions. Additionally, a shared (embodied) participation framework and shared epistemic access are a prerequisite for any form of shared stance-taking, making the visual modality a key factor in stance co-construction.

Lastly, there is an interplay between the form of stance and its broader function in a social and/or societal context. Stance

expressions may impact attitudes and identity construction or affect interpersonal relations. Conversely, (societal) norms in different settings impact the multimodal organization of stance-taking as well as the extent to which participants shape their responsibility for these stances.

5. Discussion

The discussion of our findings is structured in three sections. We start with a reflection on the limitations of our selection and review process. This is followed by a general reflection on the fundamental nature of (multimodal) stance. We close with some suggestions on broadening the scope of research, touching on methodological as well as linguistic and cultural diversity.

5.1. Limitations

In this review, we aimed at synthesizing research within a 'stance paradigm' (see Section 1). Unavoidably, this choice has resulted in a somewhat biased selection of articles with regard to the represented methodological and linguistic diversity. Potential for future research including systematic reviews will be addressed further in Section 5.3.

More than half of the selected articles are situated in the Conversation Analysis framework, which leads to a predominance of qualitative research. A quantitative perspective is less represented. Moreover, the predominance of conversation analytic work that stems from an established tradition and prominent research groups in specific geographical areas is mirrored in the languages of interaction studied: 32 articles are based on English data[6], 8 articles on Finnish and 9 on Swedish interactions, with a number of authors or labs represented in our selection more than once.

Another implication of the eligibility criteria is an imbalance in the representation of types of stance. While affective and epistemic stance are scrutinized in a large part of our selection, deontic stance remains less prominent. This might be due to the differences in traditions, linking deontics to the term *modality*, rather than *stance*. Keeping these limitations in mind, we will proceed to discussing our findings and potential implications for future research in the following paragraphs.

5.2. The fundamental nature of (multimodal) stance-taking

The fundamental nature of (multimodal) stance-taking surfaces in two aspects. First of all, while it is often argued that multimodal stance-taking is still a largely-understudied phenomenon, there is in fact a considerable body of work approaching the topic from different angles. This leaves us with the impression that the inherent multimodality of stance-taking has become widely accepted, which enables researchers to move beyond this premise.

Secondly, the set of 76 reviewed articles in its entirety shows the important role of stance-taking in human interaction. The study of multimodal stance-taking touches upon a wide range of interactional phenomena and dimensions that might not appear as evidently when consulting the articles in isolation. While multimodality is often regarded as a mere stacking of different resources within one participant, the review shows that the temporal organization of these resources is an essential aspect of the study of multimodal interaction. Next to intra-speaker aspects such as polyphonic strategies, the sequential organization of interaction seems indivisibly connected to the processes of (dis-)alignment and (dis-)affiliation with other participants, depending on the context. This highlights the collaborative character of interaction as a whole as well as the social and cultural embeddedness of multimodal stance expressions.

5.3. Broadening the scope of research

Based on this review, we are able to identify several relevant avenues to broaden the scope of research, on both a methodological and thematic level.

Regarding methodology, this review shows—in part because of its focus on the specific term *stance*—a predominance of qualitative approaches to stance-taking, rooted mainly in Conversation Analysis. While revealing a wide range of interactional practices for the expression of multimrodal stance, the principally qualitative approach leaves us with little statistical evidence for recurrent multimodal patterning in relation to stance-taking. Examples of such patterns include the shrug, the expression of distancing or proximity, as well as change-of-state markers in the face (see Section 3.2.1). Notable exceptions are the more quantitatively-oriented studies of Roseano et al. (2016), Hübscher et al. (2019), and Feyaerts et al. (2022). To the best of our knowledge, there is to date very little empirical work that has uncovered such entrenched or conventionalized form-function pairings, neither in other subbranches of linguistics zooming in on the notion of *stance,* nor in studies using different conceptual frameworks, such as *appraisal*, *subjectivity*, *viewpoint*, etc[7]. This raises the important methodological question to what extent a form of methodological triangulation is needed to provide a full-fledged account of multimodal stance-taking. In what way can fine-grained micro-analyses pave the way for large-scale corpus-based studies on particular co-occurrence patterns or for experimental studies

6 As was pointed out by the editor, the search criteria arguably skewed the results toward English data, considering that there is a rich tradition of publishing in languages other than English.

7 Given the broad range of terms and paradigms that appear in the literature connected to stance phenomena, as already mentioned in the introduction of this article, we have to acknowledge that relevant (mixed method or quantitative) work may not be included in this review. Therefore, future systematic reviews tackling terms associated with stance would be beneficial in mapping out the landscape of existing studies and connecting them across paradigms.

that tap into the underlying cognitive processes involved in the production and processing of multimodal stance expressions?[8] And how can corpus-based or experimental work raise questions for which a micro-analytical approach is warranted? In sum, the current review shows a steady increase in the number and a broadening of the scope of studies on multimodal stance-taking. However, the development of a research program on this phenomenon requires careful methodological considerations, including the consistent use of mixed-methods approaches as well as a reflection on the status of entrenched multimodal patterns a part of language users' repertoire (e.g., can they be considered as form-function pairings on a par with grammatical constructions, lexical units, etc.?). We hope that the current review may serve as a starting point for this type of methodological reflection.

A second methodological gap relates to the represented linguistic diversity. Apart from the lack of studies on signed languages, non-European languages or non-Western varieties of, e.g., English or French, we found that there lies potential in cross-linguistic research on multimodal stance. A notable example in this regard is a cross-linguistic study by Pekarek Doehler et al. (2021). Given the fundamental role of stance in interaction highlighted by this review, approaching the phenomenon across different languages, settings and demographics[9] seems very promising.

On a more thematic level, we are able to identify the following suggestions for future research: First of all, apart from the local dimensions pertaining to stance-taking (i.e., form-function relations and temporal relations), several studies in our selection point to the interpersonal effects of stance-taking, the relation to emotion work, the development of interpersonal relations, and identity construction (see Section 3.3.3). A relevant avenue is, thus, to systematically address this interplay between local and global aspects of stance-taking.

Regarding the higher-level social and societal context of stance-taking, some studies explicitly investigate stance expressions in institutional settings such as therapy sessions (Muntigl et al., 2012; Muntigl and Horvath, 2014) or courtroom interactions (Matoesian, 2018), others scrutinize 'mundane' settings such as dinner talk. Investigating preferences and permissibility of different stance displays in relation to their interactional context could reveal more about the multimodal organization of stance. Some first steps in this regard have already been taken (e.g., Hübscher et al., 2019). Furthermore, the settings described in the selection of articles go way beyond the 'standard' constellation of dyadic interaction, including among others the institutional settings mentioned above, as well as classroom discourse (Cekaite, 2007, 2020; Kääntä, 2014; Evaldsson and Melander, 2017), and interpreted interactions (Janzen and Shaffer, 2013; Vranjes et al., 2018). The inherent complexity of settings in which stance-taking occurs perhaps asks for a revision of the "Stance Triangle" (Du Bois, 2007), which is based on interactions involving two subjects, to be able to take into account more participants as well as stances involved in an interaction. Specifically, the way in which semiotic resources are organized to coordinate stance-taking might depend considerably on the constellation, showing another interesting avenue for future research.

Finally, based on the following observations, we argue that potential lies in broadening the often-used classification of stance into three types (affective, epistemic, and deontic). In this review, the majority of the articles that explicitly mentioned the type of stance under investigation featured affective stance, followed by epistemic stance. In contrast, only a handful of articles investigated deontic stance. Regarding the expression of stance in relation to this categorization, our review did not show differences in the use of semiotic resources according to types of stance. Instead, it seems that almost all resources can be involved in the expression of all sorts of stance. Furthermore, the categorization of stance types has previously been criticized, as they so often overlap (e.g., Chindamo et al., 2012, p. 620), which is supported by the findings in this review: The expression of 'surprise', for instance, is both affective and epistemic (the same applies to 'shock', but with a different polarity). A shrug can have both epistemic and deontic functions. Instead of trying to assign clear types to stance, it might therefore be more accurate to classify stance expressions based to several features, allowing for overlap.

6. Conclusion

The current contribution presents a systematic review of 76 articles on multimodal stance-taking and a critical appraisal of the literature. The size of this review highlights the considerable growth of work on the topic over the past years.

The critical appraisal was conducted along two major dimensions: form-function relations, and aspects of dynamicity. With regard to the first dimension, the review shows that the whole body can be part of stance expressions, and that different semiotic resources are used flexibly, depending on the interactional needs. The overview based on Table 1 and Supplementary Tables 3, 4 is meant to serve as a starting point for researchers interested in specific form-function relations related to stance-taking. Findings with regard to stance as a dynamic process highlight the fundamental nature of stance-taking in interaction, touching on a wide range of interactional phenomena occurring in a wide range of forms and contexts.

For future research, the following potentials are identified: Choosing mixed method approaches could contribute to the study of multimodal stance by enriching and underpinning largely qualitative results. Much potential lies in cross-linguistic studies and varying settings and participant constellations. Finally,

8 An interesting example of research on stance uptake is the study by Peräkylä et al. (2015) who investigate the link between physiological arousal and affiliation during storytelling.

9 A small number of studies in this review deals with developmental aspects of stance expressions in children and the elderly (Bolly and Boutet, 2018; Hübscher et al., 2019; Cekaite, 2020). Older people would tend to use pragmatic markers slightly differently from a younger demographic as they have a larger pragmatic competence. This is observed in an imbalanced use of pragmatic markers across modalities or a reduction in gesture amplitude (Bolly and Boutet, 2018). In the context of politeness strategies, preschoolers are found to have an increasing repertoire of mitigation strategies to signal politeness as they get older (Hübscher et al., 2019).

future work could benefit from attempts in (at least punctually) reconciling different conceptual frameworks and terminologies surrounding the phenomenon of multimodal stance-taking.

Author contributions

CV, KM, and GB contributed to conception and design of the study. FA, CV, KM, and GB conducted the article search and screening and wrote the first draft of the article. All authors contributed to the article coding, analysis, and critical appraisal. All authors participated in the manuscript revision, read, and approved the submitted version.

Funding

This study reported in the present article was conducted as part of the MUST: Multimodal Stance Taking in Interaction project (C14/20/040) funded by KU Leuven.

Acknowledgments

We would like to thank the members of KU Leuven Libraries 2Bergen, Eline Vancoppenolle and Thomas Vandendriessche, as well as Artes, Leah Budke and Nele Noppe, for their advice in the initial steps of this project. We also greatly appreciate the feedback from the reviewers and the editor, who have helped us improve this publication significantly.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fcomm.2023.1187977/full#supplementary-material

SUPPLEMENTARY TABLE 1
Search strings.

SUPPLEMENTARY TABLE 2
Codebook.

SUPPLEMENTARY TABLE 3
Overview of included articles.

SUPPLEMENTARY TABLE 4
Overview of form-function relations.

References

Andreou, N., Ioannou, Y., Petronda, S., and Ferreira, N. (2022). Exploring cypriot fathers' attitudes, beliefs and level of involvement around the decision-making process for childbirth method. *Mediterr. J. Clin. Psychol.* 10, 1–21. doi: 10.13129/2282-1619/mjcp-3317

Andries, F., Brône, G., and Vermeerbergen, M. (in press). Stance in Flemish sign language: a multimodal and polysemiotic phenomenon. *Belg. J. Linguist.* 36, 16–45.

*Bateman, A. (2020). Young children's affective stance through embodied displays of emotion during tellings. *Text Talk* 40, 643–668. doi: 10.1515/text-2020-2077

Bateman, J., Wildfeuer, J., and Hiippala, T. (2017). *Multimodality: Foundations, Research and Analysis – A Problem-Oriented Introduction*. Berlin, Boston: De Gruyter.

*Berger, I., and Rae, J. (2012). Some uses of gestural responsive actions. *J. Pragmat.* 44, 1821–1835. doi: 10.1016/j.pragma.2012.04.012

Biber, D., and Finegan, E. (1989). Styles of stance in English: lexical and grammatical marking of evidentiality and affect. *Text Interdiscip. J. Study Discourse* 9, 93–124. doi: 10.1515/text.1.1989.9.1.93

*Bolly, C. T., and Boutet, D. (2018). The multimodal CorpAGEst corpus: keeping an eye on pragmatic competence in later life. *Corpora* 13, 279–317. doi: 10.3366/cor.2018.0151

*Butler, C. W., and Edwards, D. (2018). Children's whining in family interaction. *Res. Lang. Soc. Interact.* 51, 52–66. doi: 10.1080/08351813.2018.1413893

*Cekaite, A. (2007). Soliciting teacher attention in an L2 classroom: affect displays, classroom artefacts, and embodied action. *Appl. Linguist.* 30, 26–48. doi: 10.1093/applin/amm057

*Cekaite, A. (2020). Subversive compliance and embodiment in remedial interchanges. *Text Talk* 40, 669–693. doi: 10.1515/text-2020-2078

Chindamo, M., Allwood, J., and Ahlsén, E. (2012). "Some Suggestions for the Study of Stance in Communication," in *2012 International Conference on Privacy, Security, Risk and Trust and 2012 International Conference on Social Computing* (Amsterdam), 617–622.

*Churchill, E. (2022). The education of attention in enskillment. *J. Pragmat.* 193, 155–172. doi: 10.1016/j.pragma.2022.03.006

Cienki, A. (forthcoming). *Cambridge Handbook of Gesture Studies*. Cambridge: Cambridge University Press.

*Clift, R. (2014). Visible deflation: embodiment and emotion in interaction. *Res. Lang. Soc. Interact.* 47, 380–403. doi: 10.1080/08351813.2014.958279

*Clift, R. (2021). Embodiment in dissent: the eye roll as an interactional practice. *Res. Lang. Soc. Interact.* 54, 261–276. doi: 10.1080/08351813.2021.1936858

Couper-Kuhlen, E. (2012). "Exploring affiliation in the reception of conversational complaint stories," in *Emotion in Interaction*, editors A. Peräkylä and M.-L. Sorjonen (Oxford: Oxford University Press), 113–146.

Couper-Kuhlen, E., and Selting, M. (2017). "Stance and footing," in *Interactional Linguistics: Study Language in Social Interaction* (Cambridge, New York: Cambridge University Press). Available online at: cambridge.org/interactional (accessed February 13, 2023).

Dennett, D. C. (1987). *The Intentional Stance*. Cambridge, Massachusetts: MIT Press.

Deppermann, A. (2013). Multimodal interaction from a conversation analytic perspective. *J. Pragmat.* 46, 1–7. doi: 10.1016/j.pragma.2012.11.014

*Deppermann, A., and Gubina, A. (2021). When the body belies the words: embodied agency with darf/kann ich? ("may/can I?") in German. *Front. Commun.* 6, 661800. doi: 10.3389/fcomm.2021.661800

Dingemanse, M. (2020). Between sound and speech: liminal signs in interaction. *Res. Lang. Soc. Interact.* 53, 188–196. doi: 10.1080/08351813.2020.1712967

*Donzelli, A. (2020). The act of reading aloud: animating the neoliberal speaking subject in post-Suharto Indonesia. *Discourse Soc.* 31, 459–477. doi: 10.1177/0957926520914688

Du Bois, J. W. (2007). "The Stance Triangle," in *Stancetaking in Discourse: Subjectivity, Evaluation, Interaction*, editor R. Englebretson (Amsterdam, Philadelphia: John Benjamins Publishing Company), 139–182.

Du Bois, J. W., and Kärkkäinen, E. (2012). Taking a stance on emotion: affect, sequence, and intersubjectivity in dialogic interaction. *Text Talk* 32, 433–451. doi: 10.1515/text-2012-0021

*Ekström, M. (2012). Gaze work in political media interviews. *Discourse Commun.* 6, 249–271. doi: 10.1177/1750481312452200

Englebretson, R. (2007). *Stancetaking in Discourse: Subjectivity, Evaluation, Interaction*. Amsterdam, Philadelphia: John Benjamins Publishing Company.

*Evaldsson, A.-C., and Melander, H. (2017). Managing disruptive student conduct: negative emotions and accountability in reproach-response sequences. *Linguist. Educ.* 37, 73–86. doi: 10.1016/j.linged.2016.05.001

Feyaerts, K., Brône, G., and Oben, B. (2017). "Multimodality in Interaction," in *The Cambridge Handbook of Cognitive Linguistics*, editor B. Dancygier (Cambridge: Cambridge University Press), 135–156.

*Feyaerts, K., Rominger, C., Lackner, H. K., Brône, G., Jehoul, A., Oben, B., et al. (2022). In your face? Exploring multimodal response patterns involving facial responses to verbal and gestural stance-taking expressions. *J. Pragmat.* 190, 6–17. doi: 10.1016/j.pragma.2022.01.002

*Ford, C. E., Thompson, S. A., and Drake, V. (2012). Bodily-visual practices and turn continuation. *Discourse Process* 49, 192–212. doi: 10.1080/0163853X.2012.654761

*Gilbert, K. E. (2018). Evaluating evaluation: interactional positioning and embodied narration in police interviews. *NI* 28, 1–29. doi: 10.1075/ni.17005.gil

Goodwin, C. (2000). Action and embodiment within situated human interaction. *J. Pragmat.* 32, 1489–1522. doi: 10.1016/S0378-2166(99)00096-X

*Goodwin, C. (2007). Participation, stance and affect in the organization of activities. *Discourse Soc.* 18, 53–73. doi: 10.1177/0957926507069457

*Goodwin, M. H. (2007). Participation and embodied action in preadolescent girls' assessment activity. *Res. Lang. Soc. Interact.* 40, 353–375. doi: 10.1080/08351810701471344

*Goodwin, M. H., and Alim, H. S. (2010). "Whatever (neck roll, eye roll, teeth suck)": the situated coproduction of social categories and identities through stancetaking and transmodal stylization. *J. Linguist. Anthropol.* 20, 179–194. doi: 10.1111/j.1548-1395.2010.01056.x

*Goodwin, M. H., and Goodwin, C. (2012). Car talk: integrating texts, bodies, and changing landscapes. *Semiotica* 2012, 257–286. doi: 10.1515/sem-2012-0063

*Groenewold, R., and Armstrong, E. (2019). A multimodal analysis of enactment in everyday interaction in people with aphasia. *Aphasiology* 33, 1441–1461. doi: 10.1080/02687038.2019.1644814

*Haddington, P. (2006). The organization of gaze and assessments as resources for stance taking. *Text Talk* 26, 281–328. doi: 10.1515/TEXT.2006.012

*Heller, V. (2021). Embodied displays of "doing thinking." epistemic and interactive functions of thinking displays in children's argumentative activities. *Front. Psychol.* 12, 636671. doi: 10.3389/fpsyg.2021.636671

*Honda, A. (2010). "Storytelling in a Japanese television talk show," in *Storytelling Across Japanese Conversational Genre Studies in Narrative*, editor P. E. Szatrowski (Amsterdam, Philadelphia: John Benjamins Publishing Company), 183–209.

*Hou, L. (2022). LOOKing for multi-word expressions in American Sign Language. *Cogn. Linguist.* 33, 291–337. doi: 10.1515/cog-2020-0086

*Hübscher, I., Garufi, M., and Prieto, P. (2019). The development of polite stance in preschoolers: how prosody, gesture, and body cues pave the way. *J. Child Lang.* 46, 825–862. doi: 10.1017/S0305000919000126

*Iwasaki, S. (2015). Collaboratively organized stancetaking in Japanese: sharing and negotiating stance within the turn constructional unit. *J. Pragmat.* 83, 104–119. doi: 10.1016/j.pragma.2015.04.007

Iwasaki, S. (2022). Stancetaking in motion: stance triangle and double dialogicality. *Text Talk*. 1–24. doi: 10.1515/text-2020-0222

Iwasaki, S., and Yap, F. H. (2015). Stance-marking and stance-taking in Asian languages. *J. Pragmat.* 83, 1–9. doi: 10.1016/j.pragma.2015.04.008

Jaffe, A. M. (2009). *Stance: Sociolinguistic Perspectives*. Oxford, NY: Oxford University Press.

*Janzen, T., and Shaffer, B. (2013). "The interpreter's stance in intersubjective discourse," in *Sign Language Research, Uses and Practices Crossing Views on Theoretical and Applied Sign Language Linguistics*, editors L. Meurant, A. Sinte, M. Van Herreweghe, and M. Vermeerbergen (Berlin; Boston, MA: De Gruyter) 63–84.

*Jehoul, A., Brône, G., and Feyaerts, K. (2017). The shrug as marker of obviousness: corpus evidence from Dutch face-to-face conversations. *Linguist. Vang.* 3, 1–9. doi: 10.1515/lingvan-2016-0082

*Jensen, T. W. (2014). Emotion in languaging: languaging as affective, adaptive, and flexible behavior in social interaction. *Front. Psychol.* 5, 720. doi: 10.3389/fpsyg.2014.00720

Jewitt, C. (2017). *The Routledge Handbook of Multimodal Analysis*. 2nd Edn. London, New York: Routledge.

*Jones, G. M., Semel, B., and Le, A. (2015). "There's no rules. It's hackathon.": negotiating commitment in a context of volatile sociality: negotiating commitment in a context of volatile sociality. *J. Linguist. Anthropol.* 25, 322–345. doi: 10.1111/jola.12104

*Kääntä, L. (2014). From noticing to initiating correction: students' epistemic displays in instructional interaction. *J. Pragmat.* 66, 86–105. doi: 10.1016/j.pragma.2014.02.010

*Kärkkäinen, E. (2012). On digressing with a stance and not seeking a recipient response. *Text Talk* 32. doi: 10.1515/text-2012-0023

*Katila, J., and Philipsen, J. S. (2019). The intercorporeality of closing a curtain: Sharing similar past experiences in interaction. *P C* 167–196. doi: 10.1075/pc.19030.kat

*Kaukomaa, T., Peräkyl,ä, A., and Ruusuvuori, J. (2013). Turn-opening smiles: facial expression constructing emotional transition in conversation. *J. Pragmat.* 55, 21–42. doi: 10.1016/j.pragma.2013.05.006

*Kaukomaa, T., Peräkyl,ä, A., and Ruusuvuori, J. (2015). How listeners use facial expression to shift the emotional stance of the speaker's utterance. *Res. Lang. Soc. Interact.* 48, 319–341. doi: 10.1080/08351813.2015.1058607

*Keel, S. (2015). Young children's embodied pursuits of a response to their initial assessments. *J. Pragmat.* 75, 1–24. doi: 10.1016/j.pragma.2014.10.005

*Kunitz, S., and Jansson, G. (2021). Story recipiency in a language café: integration work at the micro-level of interaction. *J. Pragmat.* 173, 28–47. doi: 10.1016/j.pragma.2020.11.017

*Li, X. (2021). Multimodal practices for negative assessments as delicate matters: incomplete syntax, facial expressions, and head movements. *Open Linguist.* 7, 549–568. doi: 10.1515/opli-2020-0164

*Liang, M.-Y. (2021). Multilingual and multimodal mediation in online intercultural conversations: a translingual perspective. *Lang. Awareness* 30, 276–296. doi: 10.1080/09658416.2021.1941069

Macaro, E. (2019). "Systematic reviews in applied linguistics," in *The Routledge Handbook of Research Methods in Applied Linguistics*, editors H. Rose and J. McKinley (London, New York: Routledge), 230–239.

*Marrese, O. H., Raymond, C. W., Fox, B. A., Ford, C. E., and Pielke, M. (2021). The grammar of obviousness: the palm-up gesture in argument sequences. *Front. Commun.* 6, 200. doi: 10.3389/fcomm.2021.663067

*Matoesian, G. (2005). Struck by speech revisited: embodied stance in jurisdictional discourse. *J. Sociolinguistics* 9, 167–193. doi: 10.1111/j.1360-6441.2005.00289.x

*Matoesian, G. (2018). This is not a course in trial practice: multimodal participation in objections. *J. Pragmat.* 129, 199–219. doi: 10.1016/j.pragma.2018.03.022

*Melander Bowden, H. (2019). Problem-solving in collaborative game design practices: epistemic stance, affect, and engagement. *Learn. Media Technol.* 44, 124–143. doi: 10.1080/17439884.2018.1563106

*Melander, H. (2012). Knowing how to play the game of jump rope: participation and stancetaking in a material environment. *J. Pragmat.* 44, 1434–1456. doi: 10.1016/j.pragma.2012.06.018

*Mondada, L. (2009). The embodied and negotiated production of assessments in instructed actions. *Res. Lang. Soc. Interact.* 42, 329–361. doi: 10.1080/08351810903296473

Mondada, L. (2014). The local constitution of multimodal resources for social interaction. *Body Resour. CA Stud. Soc. Conduct* 65, 137–156. doi: 10.1016/j.pragma.2014.04.004

Mondada, L. (2019). Contemporary issues in conversation analysis: embodiment and materiality, multimodality and multisensoriality in social interaction. *J. Pragmat.* 145, 47–62. doi: 10.1016/j.pragma.2019.01.016

Müller, C., Cienki, A., Fricke, E., Ladewig, S., McNeill, D., and Tessendorf, S. (2013). *Body – Language – Communication: An International Handbook on Multimodality in Human Interaction*. Berlin, Boston: De Gruyter.

*Muntigl, P., and Horvath, A. O. (2014). "I can see some sadness in your eyes": when experiential therapists notice a client's affectual display. *Res. Lang. Soc. Interact.* 47, 89–108. doi: 10.1080/08351813.2014.900212

*Muntigl, P., Knight, N., and Watkins, A. (2012). Working to keep aligned in psychotherapy: using nods as a dialogic resource to display affiliation. *LD* 2, 9–27. doi: 10.1075/ld.2.1.01mun

*Niemelä, M. (2010). The reporting space in conversational storytelling: orchestrating all semiotic channels for taking a stance. *J. Pragmat.* 42, 3258–3270. doi: 10.1016/j.pragma.2010.06.015

Nir, B., and Zima, E. (2017). The power of engagement: stance-taking, dialogic resonance and the construction of intersubjectivity. *Funct. Lang.* 24, 3–15. doi: 10.1075/fol.24.1.01nir

*Norén, N., Melander Bowden, H., and Evaldsson, A.-C. (2021). Young students' treatment of synthetic voicing as an interactional resource in digital writing. *Classroom Discourse* 13, 241–263. doi: 10.1080/19463014.2020.1814367

*Ogden, R. (2013). Clicks and percussives in English conversation. *J. Int. Phon. Assoc.* 43, 299–320. doi: 10.1017/S0025100313000224

Ouzzani, M., Hammady, H., Fedorowicz, Z., and Elmagarmid, A. (2016). Rayyan—a web and mobile app for systematic reviews. *Syst. Rev.* 5, 210. doi: 10.1186/s13643-016-0384-4

Page, M. J., Moher, D., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., et al. (2021). PRISMA 2020 explanation and elaboration: updated guidance and exemplars for reporting systematic reviews. *BMJ* 372, n160. doi: 10.1136/bmj.n160

*Pauletto, F., Aronsson, K., and Galeano, G. (2017). Endearment and address terms in family life: Children's and parents' requests in Italian and Swedish dinnertime interaction. *J. Pragmat.* 109, 82–94. doi: 10.1016/j.pragma.2016.12.014

*Pekarek Doehler, S., Polak-Yitzhaki, H., Li, X., Stoenica, I. M., Havlík, M., and Keevallik, L. (2021). Multimodal assemblies for prefacing a dispreferred response: a cross-linguistic analysis. *Front. Psychol.* 12, 689275. doi: 10.3389/fpsyg.2021.689275

*Peräkylä, A., Henttonen, P., Voutilainen, L., Kahri, M., Stevanovic, M., Sams, M., et al. (2015). Sharing the emotional load: recipient affiliation calms down the storyteller. *Soc. Psychol. Q.* 78, 301–323. doi: 10.1177/0190272515611054

Perniss, P. (2018). Why we should study multimodal language. *Front. Psychol.* 9, 1109. doi: 10.3389/fpsyg.2018.01109

*Pfänder, S., and Couper-Kuhlen, E. (2019). Turn-sharing revisited: an exploration of simultaneous speech in interactions between couples. *J. Pragmat.* 147, 22–48. doi: 10.1016/j.pragma.2019.05.010

*Pillet-Shore, D. (2018). Arriving: expanding the personal state sequence. *Res. Lang. Soc. Interact.* 51, 232–247. doi: 10.1080/08351813.2018.1485225

*Pillet-Shore, D. (2020). When to make the sensory social: registering in face-to-face openings. *Symbolic Interact.* 44, 10–39. doi: 10.1002/symb.481

*Pinto, D., and Vigil, D. (2019). Searches and clicks in Peninsular Spanish. *PRAG* 83–106. doi: 10.1075/prag.18020.pin

QSR International Pty Ltd (2022). *NVivo*. Available online at: https://www.qsrinternational.com/nvivo-qualitative-data-analysis-software/home (accessed May 10, 2023).

*Rekittke, L.-M. (2017). Viewpoint and stance in gesture: how a potential taboo topic may influence gestural viewpoint in recounting films. *J. Pragmat.* 122, 50–64. doi: 10.1016/j.pragma.2017.08.013

*Roseano, P., González, M., Borràs-Comes, J., and Prieto, P. (2016). Communicating epistemic stance: how speech and gesture patterns reflect epistemicity and evidentiality. *Discourse Process* 53, 135–174. doi: 10.1080/0163853X.2014.969137

*Rühlemann, C., Gee, M., and Ptak, A. (2019). Alternating gaze in multi-party storytelling. *J. Pragmat.* 149, 91–113. doi: 10.1016/j.pragma.2019.06.001

*Ruusuvuori, J., and Peräkylä, A. (2009). Facial and verbal expressions in assessing stories and topics. *Res. Lang. Soc. Interact.* 42, 377–394. doi: 10.1080/08351810903296499

*Schröder, U. (2020). Zwischen den Welten: zur kognitiven und kommunikativen Ko-Konstruktion von Alteritätserfahrung. *LO* 104, 137–165. doi: 10.13092/lo.104.7321

*Sharma, B. K. (2013). Enactment of teacher identity in resolving student disagreements in small group peer interactions. *Ling. Educ.* 24, 247–259. doi: 10.1016/j.linged.2012.09.002

*Skogmyr Marian, K. (2021). Initiating a complaint: change over time in French L2 speakers' practices. *Res. Lang. Soc. Interact.* 54, 163–182. doi: 10.1080/08351813.2021.1899709

*Soulaimani, D. (2017). Embodiment in Moroccan Arabic storytelling: language, stance and discourse analysis. *Text Talk* 37, 335–357. doi: 10.1515/text-2017-0008

*Soulaimani, D. (2018). Talk, voice and gestures in reported speech: toward an integrated approach. *Discourse Stud.* 20, 361–376. doi: 10.1177/1461445618754419

*Stirling, L., Gardner, R., Mushin, I., Blythe, J., and Possemato, F. (2022). On the road again: displaying knowledge of place in multiparty conversations in the remote Australian outback. *J. Pragmat.* 187, 90–114. doi: 10.1016/j.pragma.2021.10.026

*Stivers, T. (2008). Stance, alignment, and affiliation during storytelling: when nodding is a token of affiliation. *Res. Lang. Soc. Interact.* 41, 31–57. doi: 10.1080/08351810701691123

*Stoenica, I.-M., and Fiedler, S. (2021). Multimodal practice for mobilizing response: the case of turn-final Tu Vois 'you see' in French talk-in-interaction. *Front. Psychol.* 12, 659340. doi: 10.3389/fpsyg.2021.659340

*Tainio, L., and Laine, A. (2015). Emotion work and affective stance in the mathematics classroom: the case of IRE sequences in Finnish classroom interaction. *Educ. Stud. Math.* 89, 67–87. doi: 10.1007/s10649-015-9591-5

Takanashi, H. (2018). "Stance," in *Handbook of Pragmatics*, editors J.-O. Östman and J. Verschueren (Amsterdam, Philadelphia: John Benjamins Publishing Company), 173–200.

*Toomaneejinda, A., and Harding, L. (2018). Disagreement practices in ELF academic group discussion: verbal, nonverbal and interactional strategies. *J. English Lingua Franca* 7, 307–332. doi: 10.1515/jelf-2018-0016

*Trujillo, J. P., and Holler, J. (2021). The kinematics of social action: visual signals provide cues for what interlocutors do in conversation. *Brain Sci.* 11, 996. doi: 10.3390/brainsci11080996

*Van De Mieroop, D. (2020). A deontic perspective on the collaborative, multimodal accomplishment of leadership. *Leadership* 16, 592–619. doi: 10.1177/1742715019893824

*Vranjes, J., and Brône, G. (2021). Interpreters as laminated speakers: gaze and gesture as interpersonal deixis in consecutive dialogue interpreting. *J. Pragmat.* 181, 83–99. doi: 10.1016/j.pragma.2021.05.008

*Vranjes, J., Brône, G., and Feyaerts, K. (2018). Dual feedback in interpreter-mediated interactions: on the role of gaze in the production of listener responses. *J. Pragmat.* 134, 15–30. doi: 10.1016/j.pragma.2018.06.002

Wang, Z., Fan, W., and Fang, A. C. (2022). Lexical input in the grammatical expression of stance: a collexeme analysis of the introductory it pattern. *Front. Psychol.* 12, 762000. doi: 10.3389/fpsyg.2021.762000

∗ Articles included in the review.