



Multimodal Assemblies for Prefacing a Dispreferred Response: A Cross-Linguistic Analysis

Simona Pekarek Doehler^{1*}, Hilla Polak-Yitzhaki², Xiaoting Li³, Ioana Maria Stoenica¹, Martin Havlík⁴ and Leelo Keevallik⁵

¹ Center for Applied Linguistics, Institute of Language Sciences, University of Neuchâtel, Neuchâtel, Switzerland,

² Department of Hebrew Language, University of Haifa, Haifa, Israel, ³ Department of East Asian Studies, University of Alberta, Edmonton, AB, Canada, ⁴ Czech Language Institute, Czech Academy of Sciences, Prague, Czechia, ⁵ Division of Language, Culture and Interaction, Department of Culture and Society, Linköping University, Linköping, Sweden

In this paper we examine how participants' multimodal conduct maps onto one of the basic organizational principles of social interaction: preference organization – and how it does so in a similar manner across five different languages (Czech, French, Hebrew, Mandarin, and Romanian). Based on interactional data from these languages, we identify a recurrent multimodal practice that respondents deploy in turn-initial position in dispreferred responses to various first actions, such as information requests, assessments, proposals, and informing. The practice involves the verbal delivery of a turn-initial expression corresponding to English 'I don't know' and its variants ('dunno') coupled with gaze aversion from the prior speaker. We show that through this 'multimodal assembly' respondents preface a dispreferred response within various sequence types, and we demonstrate the cross-linguistic robustness of this practice: Through the focal multimodal assembly, respondents retrospectively mark the prior action as problematic and prospectively alert co-participants to incipient resistance to the constraints set out or to the stance conveyed by that action. By evidencing how grammar and body interface in related ways across a diverse set of languages, the findings open a window onto cross-linguistic, cross-modal, and cross-cultural consistencies in human interactional conduct.

OPEN ACCESS

Edited by:

Marianne Gullberg,
Lund University, Sweden

Reviewed by:

Kobin Kendrick,
University of York, United Kingdom
Jan Svennevig,
University of Oslo, Norway

*Correspondence:

Simona Pekarek Doehler
simona.pekarek@unine.ch

Specialty section:

This article was submitted to
Language Sciences,
a section of the journal
Frontiers in Psychology

Received: 31 March 2021

Accepted: 28 July 2021

Published: 27 September 2021

Citation:

Pekarek Doehler S,
Polak-Yitzhaki H, Li X, Stoenica IM,
Havlik 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

Keywords: preference organization, gaze, epistemic markers, conversation analysis, turn-prefacing, multimodality

INTRODUCTION

Participants in face-to-face social interaction deploy multiple resources for building actions in mutually accountable ways. A pioneer in the field, Charles Goodwin (e.g., Goodwin, 1981, 2017), has demonstrated that human social interaction is a complex ecology of vocal and bodily visual resources that people draw on to build joint actions. While some multimodal resources are used in concert in *ad hoc* manners to accomplish local interactional tasks, there are also more or less routinized configurations of language and the body ("multimodal packages," Hayashi, 2005; Kärkkäinen and Thompson, 2018, or "gestalts," Mondada, 2014) – that is, recurrent constellations of vocal and bodily conduct that are routinely deployed to accomplish particular interactional tasks. In line with Goodwin's (2013, p. 12) observations of how participants to social interaction

dynamically “assemble” such multiple resources to create actions, we refer to such configurations as “multimodal assemblies.” We are interested in exploring the *methodic* character of such assemblies, i.e., recurrent body-language constellations that ‘go together’ and are put to use in systematic ways as part of interactants’ accountable practices for conducting social interaction (see also Keevallik, 2013; Li, 2014; Pekarek Doehler, 2019; Skogmyr Marian, 2021, *forthc*), and we seek to uncover how these compare cross-linguistically. It is our belief that a better documentation of the commonalities of multimodal behavior of humans around the globe may bring us closer to understanding the possibly shared organizational principles of social interaction in particular, and of human social conduct in general.

In order to develop such an understanding, research has started to explore the ways speakers shape their conversational actions cross-linguistically and cross-culturally (cf. Stivers et al., 2009; Dingemans and Floyd, 2014; Floyd et al., 2020; see below). In this paper, we focus on speakers of five geographically and culturally distinct languages, from five different language (sub)families: Czech (Slavic branch of the Indo-European language family), French and Romanian (Romance branch of the Indo-European language family), Hebrew (Semitic branch of the Afro-Asiatic language family), Mandarin (Sino-Tibetan language family).¹ We analyze how speakers of these languages employ equivalents of the construction ‘I don’t know’ in combination with precise bodily conduct, namely gaze averted from the recipient, in a precise sequential environment, namely turn-initial position in responsive actions. We document that this multimodal assembly serves as a preface to dispreferred responses, such as refusing a proposal or not answering an information request. And we shed light on the sequential location of respondents’ averting their gaze – relative to the prior speaker’s turn and to the response itself – and then possibly returning it to the prior speaker in the course of the response (which we refer to as ‘gaze trajectory’).

The Estonian audio **Excerpt (1)** provides an initial sense of the *linguistic* phenomenon at hand, i.e., how ‘I don’t know’ type of expressions can work to launch a dispreferred response. In this

excerpt, Ene responds to Anni’s question-word question as to the whereabouts of a crib.

The response comes in late – a typical trait of dispreference (cf. Pomerantz, 1984; Sacks, 1987) – and is prefaced by a morphophonologically reduced *mai tea* ‘I dunno’ (the full form would be *ma ei tea*), which is furthermore prosodically latched to the following phrase; it hence stands as a preface to what follows rather than a turn-constructional unit (TCU) in itself. In what follows, Ene offers a strongly non-committal response of the type that Stivers (2010) has treated as a non-answer response, and hence a dispreferred response.² While the *mai tea* may here also work as a pre-positioned hedge, downgrading the speaker’s commitment to her response (see Weatherall, 2011, on English), in this sequential environment it prominently projects that the incipient response departs from the relevancies issued by the sequence-initiating question: It alerts the recipient to the upcoming of a dispreferred next action.

A related working is shown in the French **Excerpt (2)**, where *ch’pas* ‘dunno’ projects a dispreferred response (note that here it cannot be heard as an epistemic hedge), conjointly with the preceding *be:n* ‘well.’ In this excerpt, Marie asks Julie why she would rather choose to learn languages by means of immersion than standard instruction (l.01):

Julie’s response (l.03) bears typical traits of a dispreferred response (Sacks, 1987), i.e., one that does not align with the terms of the sequence-initial action: It starts with a delay, and is *be:n* ‘well’-prefaced (cf. Davidson, 1984; for French ‘ben’ see Bruxelles and Traverso, 2001; Persson, 2020). This is followed by a morphophonologically reduced and prosodically backgrounded (speedup of tempo) variant of *je sais pas* ‘I don’t know,’ namely *ch’pas* ‘dunno,’ which is prosodically packaged together with what follows into a single prosodic unit, rather than standing as an independent unit. These features concur to display the *ch’pas* as a preface to the subsequent response *je trouve c’est mieux* ‘I think it’s better’ (the turn-final *he* is a non-lexical vocalization, produced through strong exhalation). Importantly, this response does not conform to the terms of the question, as Julie simply reaffirms that she prefers immersion (l.03): She offers a non-answer response, and hence a dispreferred response (Clayman, 2002;

¹Some of the studied languages have so far remained largely unexplored in an interactional linguistics approach (Romanina; Czech; but see Oloff and Havlik, 2018).

²Stivers (2010, p. 2778) writes: “non-answer responses included not only ‘I don’t know’ and ‘I can’t remember,’ but also less committal ‘Maybe’ or ‘Probably’ responses and less direct answers”.

Ex. (1) 34_K2B2 (From Keevallik 2003)

01 ANN **aga kus on sõim.**
 but where is crib
 ‘but where is the crib’

02 (1.2)

03 ENE **mai tea peaks olema Varjus kuskil.=**
 I.NEG know must.COND be.SPN NAME.INS somewhere
 ‘I don’t know (it) ought to be somewhere at Varju’

Ex. (2) (Coll_JSP_90)

01 MAR et- pourquoi.
and why

02 (0.8)

03 JUL be:n >ch'pas< je trouve c'est mi↑teux^he.
well dunno I think it's better

Stivers and Robinson, 2006). In this sequential environment, the > *ch'pas* <, here conjointly with the *be:n* 'well,' alerts the recipient to the dispreferred nature of the upcoming next action.

While the above excerpts illustrate the verbal side of the interactional practice that we investigate in this contribution, we show in the remainder of this paper that speakers deploy such verbal conduct recurrently as part of a multimodal assembly. The assembly comprises a morphosyntactic construction corresponding to English 'I don't know' and its variants (henceforth: IDK) that is prosodically produced as part of the emerging TCU and is coupled with the speaker's gaze averted from recipient. The respondent's gaze itself follows an on-line trajectory, comprising aversion most often either within the transition space between turns or coinciding with response onset (and the production of IDK), and then return to the prior speaker in the course of the response. Sequentially, the grammar-body-assembly is recurrently found in turn-initial position in dispreferred responsive actions. Based on multimodal conversation analysis, we provide cross-linguistic evidence showing that [IDK + gaze aversion] serves as a routinized multimodal resource for prefacing a dispreferred response: It indexes incipient resistance to the constraints set out or to the stance conveyed by the coparticipant's prior action. While our findings converge with earlier studies that documented an association of gaze aversion with dispreference in the specific context of responses to polar questions in English (Kendrick and Holler, 2017; Robinson, 2020; see section "Background" below), they also amplify these by showing that the focal practice (a) pertains to a wide range of sequence types, i.e., is not limited to responses to polar questions, and (b) holds across a diverse set of languages.

In what follows, we first present the background (see section "Background") and the data used for this study (see section "Data and methods"), and then turn to multimodal analysis of selected data excerpts (see section "Analysis") documenting the occurrence of the focal multimodal assembly as a preface to dispreferred responses to questions (see section "Prefacing dispreferred responses to questions") and to a range of other types of sequence-initial actions (see section "Prefacing dispreferred responses to proposals, assessments, and informings"). We end with discussing our findings and drawing conclusions regarding the interactional logic of the practice as well as how grammar and body interface across languages (see section "Discussion and conclusion").

BACKGROUND

Preference organization is a basic organizational feature of social interaction – a "formal apparatus" (Sacks, 1987) pertaining to the sequential concatenation of actions. Basically, there are sequence initial actions that call for responses of two opposed types among which one tends to be favored over the other. So-called 'dispreferred' next actions are those that are "uncooperative, disaffiliative or otherwise discordant" (Clayman, 2002, p. 230), such as refusing an offer or not answering a question. So-called 'preferred' next actions are those that are most cooperative, allow the interaction to move forward and align with the action which they respond to Sidnell (2011, p. 81), such as granting a request or answering a question.³ The preference status of actions is reflected in their design, most centrally in their timing and more generally in what Sacks (1987) referred to as their 'contiguity': Preferred next actions tend to be delivered right away, while dispreferred actions tend to be delayed and pushed further back in the turn by means of pre-beginnings (such as *uhm* or inbreaths), hedging, and turn-prefaces (Sacks, 1987), including turn-initial particles such as *well* in English (Davidson, 1984; Heritage, 2015), *ben* in French (Bruxelles and Traverso, 2001; Persson, 2020) or *nu/nå* across a range of European languages (Auer and Maschler, 2016). That is, participants deploy a set of conventionalized practices through which they accomplish alternative types of responsive actions. Turn-initial position is particularly relevant to the preference status of responsive actions, as it is in this position that the dispreferred nature of these may be projected (e.g., Davidson, 1984; Pomerantz, 1984; Deppermann, 2013; Heritage, 2015; Heritage and Sorjonen, 2018).

In addition to turn-initial elements, the timing between the end of an initiating action and the beginning of its responsive action has been documented to project the preference status of responsive actions. Measuring the silence between the initiating and the responsive turns at talk in invitation, offer, request, suggestion, and proposal sequences in a corpus of English telephone conversations, Kendrick and Torreira (2015) show that dispreferred responses are delayed significantly longer than preferred responses (median = 561 ms vs. 269 ms; Kendrick

³Other factors may play into preference organization (Pomerantz and Heritage, 2013), such as the sub-type of sequence-initial action (e.g., for certain offers, accepting might be a dispreferred action; Schegloff, 2007, p. 60) or the speaker's entitlement to perform that action (Curl and Drew, 2008).

and Holler's, 2017 findings suggest that these times are shorter for responses to polar interrogatives). Bögels et al. (2015, using experimental methods) and Bögels et al. (2020, studying telephone conversations) further demonstrate that recipients change their expectations regarding the valence of a response based on the duration of silence between initiating actions and responsive actions: Recipients expect preferred responses after a short gap and dispreferred responses after a longer gap (roughly after 300 ms). Furthermore, there is initial evidence for cross-linguistic convergences in this patterning of preferred and dispreferred response-timing and the related expectations. Roberts et al. (2011) used experimental methods to show that in American English (see also Roberts et al., 2006), Italian, and Japanese, listeners perceive a response that comes in late as an indicator of a disagreement with an assessment or of an unwillingness to comply with a request.

In a nutshell, then, as Pomerantz and Heritage (2013, p. 228) put it: "The turn shapes through which initiating and responding actions are enacted represent interpretive resources" that index the preference status of an action. However, how multimodal conduct is relevant to preference organization has remained largely unexplored.

There is some evidence that disaligning or disaffiliative⁴ responsive actions in general are associated with respondent's gaze aversion. Kidwell (2006) reports that recipient's gaze-withdrawal after a directive may be seen as an act of resistance, and Haddington (2006) finds evidence that respondent's gaze withdrawal from a prior speaker is associated with divergent stance. This is in line with Kendon's (1967, p. 48) earlier observation that mutual gaze between speaker and recipient appears to decrease in non-affiliative and non-cooperative interactions. There are, to our knowledge, only two investigations directly addressing gaze in dispreferred responsive actions, both focusing on responses to polar questions in English. Based on quantitative and qualitative analysis of conversations recorded in a soundproof room, and using eye-tracking methodology, Kendrick and Holler (2017) found that preferred responses (defined as those responses that conform to the preference set up by the grammatical format of the yes-no question), tend to be produced with gaze toward the recipient (61%), while dispreferred responses are most frequently produced with gaze averted from the recipient (82%); in the latter case recipients' gaze aversion tends to start after the first possible completion point of the question. Robinson (2020) presents a quantitative and a qualitative study of responses to information-seeking polar questions that confirms the association of gaze aversion with dispreference: Respondents producing dispreferred responses shift gaze away from questioners in 73% of the cases at some point between the beginning of the question and the

beginning of the response (while they do so only in 21% of the preferred responses).

In light of the above findings on gaze, three issues deserve to be explored in more detail. One key open question is (i) whether the documented association of gaze aversion and dispreference can be found with dispreferred responses in contexts other than polar questions. Also (ii) the on-line trajectory of gaze conduct in dispreferred responses deserves attention, namely in terms of the sequential start of aversion (during the sequence-initial turn, in the transition space between turns, or coinciding with the response onset) and of possible return of gaze to recipient. And (iii) it remains to be investigated how far findings based on (American) English data may be valid for other languages and cultures.

While existing work has evidenced that gaze plays a crucial role in organizing and regulating interaction (Goodwin, 1981), for instance in regard to turn-taking (Kendon, 1967) or sequence organization (Rossano, 2012), cross-cultural consistency or variation in gaze conduct has so far gained only limited attention in research on social interaction. And this is so despite important intensification of comparative conversation analytic research across situations (e.g., Heritage and Clayman, 2010), cultures (Sidnell, 2009; Floyd et al., 2020), and time (Pekarek Doehler and Deppermann, 2021). The existing cross-linguistic/cultural studies of gaze in interaction have mainly focused on question-answer adjacency pairs. In their comparative analysis of question-answer sequences in conversational data from 10 languages, Stivers et al. (2009) found that questioner's gaze affects the timing of responses: Responses were delivered earlier if the questioner was looking at the recipient while the question was asked. Rossano et al. (2009) investigated gaze in question-answer sequences in conversational data in Italian (from northern Italy), Yéli Dyne (Papua New Guinea), and Tenejapan Tzeltal (Mexico). They found that questioners look at recipients more often than vice-versa, but that recipients' gaze conduct is less consistent cross-culturally. One of their results directly relates to our focal issue, namely that « lack of [. . .] recipient gaze [on questioner] is a good predictor of lack of response after questions » (p. 220). This again begs the question as to whether recipient's gaze averted from prior speaker is a more general feature of dispreference, i.e., if it can be found across a variety of sequence types.

As to the linguistic construction investigated in this paper, IDK has been shown to be associated with disaffiliative or dispreferred actions in two ways. For one thing, in its literal uses as a claim of no knowledge, IDK has been considered as a non-answer response, i.e., a dispreferred response in itself, that offers an account for not providing an answer (Stivers and Robinson, 2006). As such, it can be a means for indicating a problem with the question (Keevallik, 2011), resisting a line of questioning (Hutchby, 2002; Keevallik, 2011), resisting offers, proposals and invitations (Keevallik, 2011), avoiding an overt disagreement (Tsui, 1991, pp. 612–617), or deferring a dispreferred response (Helmer et al., 2016). For another thing, and different from the fully epistemic uses, there are also particle- (or: marker-)like uses of IDK that have been suggested to be associated with dispreference. These typically (see ex. 1 and 2 above) show formal reduction (e.g., English *dunno*) as well as semantic

⁴Following Stivers et al. (2011) we use alignment to refer to the "structural level of cooperation" (2011, p. 20) and see aligning responses as ones that "cooperate by facilitating the proposed activity or sequence; accepting the presuppositions and terms of the proposed action or activity; and matching the formal design preference of the turn" (2011, p. 21). Affiliation in turn relates to the "the affective level of cooperation," in that "affiliative responses are maximally pro-social when they match the prior speaker's evaluative stance, display empathy and/or cooperate with the preference of the prior action." (Stivers et al., 2011, p. 21).

bleaching (concretely: loss of epistemicity) as typical features of routinization (for English: Bybee and Scheibmann, 1999; Scheibman, 2000). Schegloff (1988, p. 445; see also Schegloff, 1996) observed that *I dunno* can be one of a range of delay components in a dispreferred response. In a similar vein, Keevallik (2003, 2011) observed that Estonian *mai tea* 'I don't know' (and its variant *ei tea*) may be used in disaligning actions giving dispreferred answers in an indirect manner, as well as signaling the speaker's uncertainty toward the produced content. Maschler (2012, 2017) found that Hebrew ('*ani*) *lo yode'a/yoda'at* '(I) don't know.M/F' (mainly its reduced form *loyde'a/loydat*) may function as discourse markers changing the course of talk. Lindström and Karlsson (2016) point out that Swedish *jag vet inte* 'I don't know' in doctor-patient interactions is a pragmatic marker that frames responsive turns as resisting the interlocutor's question. Pekarek Doehler (2016) found that in French turn-initial *chais pas* 'dunno' projects a dispreferred response, whereas in mid-turn position it tends to function as an epistemic hedge and in turn-final position it serves as a turn- or sequence-closing device.

The above work provides evidence for turn-initial uses of IDK in responsive turns in Estonian, Hebrew, Swedish and French, for projecting a disaligning or dispreferred response. A systematic analysis of the uses of IDK in other languages, however, is still lacking. Also, we know little about the bodily visual behaviors possibly concurrent with IDK. Pekarek Doehler (2019) documents how participants' embodied conduct, and specifically gaze, systematically differs between two distinct particle-like uses of French *chais pas* 'dunno' in turn-final position in sequence-initial actions: IDK plus gaze aversion serve as a practice to withdraw one's own prior action, while IDK plus gaze on recipient, in the very same sequential position, is a practice to invite recipient's response. To our knowledge this is the only existing study scrutinizing how IDK interfaces with bodily conduct in social interaction. In addition to the three issues raised under (i) through (iii) above, the question remains as to (iv) how the linguistic construction IDK combines with gaze as part of a practice for prefacing dispreferred responses. The present study documents the recurrence and robustness of this practice, and shows that it is generalizable across a variety of sequential contexts and languages.

DATA AND METHODS

The data for this study consist of video-recorded interactions between two to five participants. Most stem from conversations between friends/colleagues and family. The Czech data comes from TV talk shows but converges with the data in the other languages in that it consists of unscripted interactions that unfold spontaneously. As the latter data comes from the public domain, it neither needs anonymization nor consent for publication.⁵

⁵The Czech data come from Korpus DIALOG 2.0 (2020). Online: <<http://ujc.dialogy.cz>>. The corpus, consisting of 200 recordings (and related transcriptions), contains a total of 1,188,513 words. There is a total of 856 speakers, 305 of which uttered more than 1000 words.

Names in all other data have been anonymized, and informed consent has been obtained from all participants for use and publication of the materials, including the video frames. Cross-linguistic comparability across the datasets is based on our focus on a precise sequential environment: IDKs in turn-initial position in responsive actions.

We do not have space here for a structural description of the languages studied (but see Dryer and Haspelmath, 2013), and therefore limit ourselves to a succinct presentation of the IDK construction in these languages. **Table 1** shows the conventional full and the reduced forms of IDKs across the languages (in alphabetic order)⁶.

Table 2 sums up the extent of the data (hours of recording per language) and the IDK tokens found therein. For Romanian, due to unavailability of existing data transcribed according to CA standards, we have so far only a small corpus that has been specifically recorded and transcribed for the purpose of this study; the results for this language hence call for further confirmation based on a larger dataset.

⁶Czech, Mandarin, and Romanian are pro-drop languages with possible null subjects (e.g., Huang, 1989, for Mandarin). That is, the subject pronoun 'I' may be dropped, creating reduced form such as *bu zhidao* 'don't know.' The French *je ne sais pas* has a discontinuous NEG morpheme (*ne...pas*). In spontaneous spoken language use, the *ne* is commonly dropped, and the expression can further be reduce based on amalgamation of *je* and *sais*. In Hebrew the present tense does not mark person, but only gender and number.

TABLE 1 | Forms of IDKs in the data in five languages.

| Language | 'Canonical' form | Reduced form |
|----------|---|--|
| Czech | (já) nevím I NEG-know.1SG.PRS canonical form [nei:m] with a long vowel i:, which is a phonological distinction in Czech | [nevim] - > change in vowel pronunciation and reduced length, often change in consonant pronunciation: [v] instead of [v] |
| French | je (ne) sais pas I NEG know.1SG.PRS NEG | j'sais pas chais pas ch'pais - > 1 person pronoun and the start of 'know' are amalgamated to various degrees |
| Hebrew | 'ani lo yode'a/yoda'at I NEG know.SG.PRS.M/F | 'ani lo yde'a 'anloyde'a loyde'a lo yoda'at loyda'at loydat - > various degrees of reduction, including elision of first person pronoun and amalgamation of the negator and 'know' |
| Mandarin | (wo) bu zhidao I NEG know.1SG.PRS | bu zhidao |
| Romanian | (eu) nu știu I NEG know.1SG.PRS | nu știu |

TABLE 2 | Overview of IDK tokens (without complement) occurring in the data from the five languages.

| Language (and hours of recording) | (I) Total IDK (without a complement) | (II) Turn-initial <i>particle-like</i> IDK in responsive actions | (IIIa) Turn-initial <i>particle-like</i> IDK in <i>dispreferred</i> responsive actions | (IIIb) Thereof: tokens with respondent's gaze averted from prior speaker |
|-----------------------------------|--------------------------------------|--|--|--|
| Czech (130 h) | 776 | 47 | 43 | 39 |
| French (20 h) | 184 | 32 | 21 | 19 |
| Hebrew (8 h) | 108 | 29 | 22 | 20 |
| Mandarin (17 h) | 83 | 36 | 31 | 28 |
| Romanian (1 h35) | 31 | 4 | 2 | 2 |

The focal tokens for this study appear in bold.

We established an initial inventory, for each language, of all instances of IDK occurring without a complement (column I, **Table 2**). We then identified amongst these all instances containing a turn-initial *particle-* (marker-)like use of IDK in *responsive actions*, independently of a precise sequence type. That is, we discarded all IDKs that were full epistemic disclaimers (e.g., IDKs accomplishing a non-answer response through a claim of no-knowledge), and this was motivated by the fact that full epistemic disclaims can accomplish dispreferred actions in themselves and are hence functionally different from the particle-like uses analyzed here, which preface such actions but do not accomplish them (see section “Background” above). We included as ‘turn-initial’ tokens those occurrences that were either not preceded by any vocal tract sound or were preceded by pre-starts (e.g., inbreaths, clicks, *guh*: or *hum* and the like) or particles (e.g., *ben* ‘well’). We then conducted multimodal sequential analysis of the collection for each language and discarded all instances that were either not or not clearly categorizable as dispreferred responses (see Robinson, 2020 for continua of preference) or in which the respondent’s gaze conduct was not identifiable from the video captures. This left us with the tokens indicated in column IIIa, **Table 2**. Among these, we identified those instances that were coupled with respondent’s gaze averted from the prior speaker (column IIIb), which amounted to roughly 90% of the cases for each language (except: 100% of the two tokens for Romanian). The association of particle-like IDK plus gaze aversion occurring in turn-initial position in dispreferred responses was hence found to be a strong tendency across the languages, with only a few exceptions for each of these (4 for Czech, 2 for French, 2 for Hebrew, 3 for Mandarin, 0 for Romanian). The quantitative results regarding the recurrence of respondent’s gaze aversion in dispreferred responses converge with the tendencies observed by Kendrick and Holler (2017) and Robinson (2020) for the specific case of responses to polar questions in English (see above). We then organized, for each language, the collection in terms of sequence types; the association of gaze-aversion with dispreference was found across a range of sequence types that we discuss below (except for Romanian for which we have only 2 target occurrences due to a limited dataset; see above).

In section “Analysis,” we present selective excerpts from our data to illustrate the convergent ways in which IDK and gaze tend to be coupled in projecting a dispreferred response. Data excerpts in Czech, French, and Romanian use

CA conventions for transcription of verbal conduct (Jefferson, 1984); excerpts in Mandarin are transcribed using the GAT-2 transcription system (Selting et al., 2011); excerpts in Hebrew are transcribed according to the Santa Barbara transcription method (Du Bois, 2012). The multimodal transcription conventions follow Mondada (2018).

ANALYSIS

In this section we document speakers’ use of IDK in conjunction with gaze averted from the recipient in prefacing a dispreferred response. The following sequential pattern is observed, in which gaze aversion may coincide with the IDK or precede it:

A: sequence-initiating action

(gap)

B: *IDK and gaze aversion* + other turn elements that form a responsive action

The findings across the languages studied show that IDK in this use is semantically bleached (it does not work as a claim of no knowledge and is not treated as such) and tends to be morpho-phonologically reduced or otherwise prosodically downgraded (e.g., by speed-up of tempo or lower volume). These features suggest grammaticalization into a particle-like element (Bybee and Scheibmann, 1999 for English; Maschler, 2012, 2017 for Hebrew; Keevallik, 2003 for Estonian, Pekarek Doehler, 2016 for French). The IDK-prefaced turn is often delayed and/or sometimes prefaced by particles such as *no/ben* ‘well,’ sound objects such as *pff* or clicks as further typical traits of dispreferred responses. This provides cross-linguistic evidence for [IDK + gaze aversion] working as a composite multimodal resource prefacing a dispreferred response. We first discuss question-answer sequences (see section “Prefacing dispreferred responses to questions”) as a case of canonical adjacency pair, and then extend the analysis to other action sequences, namely those including proposals, assessments and informing (see section “Prefacing dispreferred responses to proposals, assessments, and informings”).

Prefacing Dispreferred Responses to Questions

Question–answer sequences have attracted much attention in research. Questions can implement a range of actions (request

for information or confirmation, repair, etc.; e.g., Stivers, 2010). Answers are preferred over non-answers (Clayman, 2002; Stivers and Robinson, 2006); with polar questions, confirming answers are preferred over disconfirming answers (Sacks, 1992), and type-conforming answers (i.e., structurally 'fitted' answers: yes/no) are preferred over non-conforming ones (Raymond, 2003). Overall, preferred responses are delivered faster than dispreferred responses (e.g., Raymond, 2003; Stivers et al.,

2009; Stivers, 2010; Heritage, 2012). In the data, IDK-prefaced dispreferred responses are found in responses to both question-word and polar questions that work as requests for information.

Responses to Question-Word Questions Seeking Information

The Romanian Excerpt (3) provides a first illustration. Greta and Ana are sharing a break at the end of their working day.

Ex. (3) [Corpus Stoenica (G&G)_02m14-02m37 warm water]

01 GRE f#no și?
 PRT and
 fGRE and ANA gaze at each other----->1.06
 fig. #fig.1



fig.1

02 (0.4)

03 ANA și:: cred că (0.9) ceva angajat
 and (I) think that some employee

04 era în concediu Δ#că: .h[h]&
 was in vacation because
 gre Δretracts left hand from her face--->1.06
 fig. #fig.2

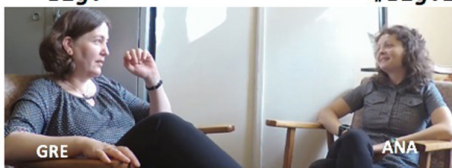


fig.2

05 GRE [unde?]
 where

06 ANA &ă::fΔ *nu ↑știu*# undeva unde era apa cal*dă^HA.h.h.h::[:
 u::h don't know somewhere where the water was warm
 --->f
 gazes up, raising both hands from elbow
 keeps gaze up and moves raised hands up&down
 *lowers hands and gazes at GRE--->>

gre ----->Δ
 fig

#fig.3

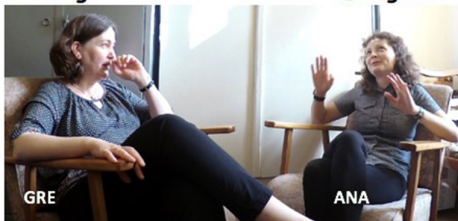


fig.3

07 GRE: [ha: ha

Reporting on an event that happened during the day, Ana mentions the fact that an employee was on vacation (1.03–04), upon which Greta asks ‘where’ (1.05). Up to line 6, participants maintain mutual gaze.

At line 06, Ana’s responsive turn starts with an initial hesitation marker and a subsequent *nu știu* ‘don’t know,’ both concurring to delay the response proper. This is followed by the type of strongly non-committal response discussed above that stands as a dispreferred response (cf. Stivers, 2010, p. 2778): ‘somewhere where the water was warm.’ While the IDK possibly also functions here as a pre-positioned hedge (Keevallik, 2011; Weatherall, 2011), downgrading the speaker’s commitment to her incipient response, it is centrally implicated in the sequential concatenation of turns and actions: it projects that the incipient response departs from the relevancies established by the sequence-initiating question. Noteworthy is the fact that the respondent shifts her gaze away from the co-participants, up into the air; the onset of the gaze shift coincides with the *nu știu* ‘don’t know,’ and gaze is maintained averted throughout the responsive action, turning back at Greta only at its very end (1.06). Interestingly, in the course of the response, the semiotic quality of the gaze appears to change: The gaze, coupled with Ana’s hands spread out raising up (Figure 2) and then down (1.06), heightens the non-committal nature of her response as it is combined with a shrug (see also her raised eyebrows, cf. Streeck, 2009, p. 189) – a pragmatic gesture (Kendon, 2004) conferring disengagement (Streeck, 2009, p. 189): It is as if Ana was enacting ‘don’t ask me’.⁷ The search ends on Ana laughing (1.06), which is met by Greta’s laughter (1.07).

A similar IDK-gaze pattern can be observed in the Hebrew data shown in **Excerpt (4)**. Alex and Dotan, two friends at the beginning of their master’s degree in biology, are talking about Alex’s new position as a practice lecturer. Bracha, a friend of Dotan, sits behind the camera. After Dotan had inquired about how Alex’s office hour with the students went, Alex responds ‘it’s not such fun’ (1.01) upon which Dotan inquires ‘why’ (1.03):

Just like in the preceding excerpt, Dotan’s ‘why’ (1.03) is a question-word question that accomplishes a request for information. Alex responds with a reduced (see **Table 1** above) *anloyde’a* ‘I dunno’-prefaced response. Even though the *anloyde’a* token occurs in a separate intonation unit (and is marked as such according to the Santa Barbara transcription conventions, Du Bois, 2012), there is no pause following it and it is produced with a continuing intonation contour signaling that there is more to come, thus the two intonation units ‘...I dunno, it’s no fun.’ (1.04–05) are delivered as a single TCU. Alex’s response clearly does not conform to the terms of the question: Instead of providing an account for why it’s not fun, Alex just repeats that it is no fun (1.05), which is an upgraded version of his prior negative assessment ‘it’s not such fun’ (1.01). He then elaborates in a non-serious way on the conditions under which it would have been ‘a bit more fun,’ namely if he had known

what he was talking about (1.07). So, again we have a dispreferred response that is prefaced by a morphophonologically reduced IDK, which is part of the same TCU as the subsequent response. This is coupled with the respondent averting his gaze from the prior speaker, right in the transition space after the ‘why’ question (1.03–04), and maintaining it through the *anloyde’a* and the rest of the turn, only at the end of which he re-directs his gaze to the co-participant (1.07).

Similar features are observable in **Excerpt (5)**, in French. Marie just mentioned that the condition for her to move abroad for a job would be that her boyfriend comes along. Pat then delivers a question-word interrogative, asking what if her boyfriend (*il* ‘he’) did not come along (1.01). After Marie starts a delayed and *bon* ‘well’-prefaced response (1.03–04), Pat adds that the job in question would represent Marie’s chance of a lifetime (1.05), thereby re-launching his inquiry in more dramatic terms: His *c’est la chance de ta vie* ‘it’s the chance of your life’ works to renew the relevance of his initial question, thereby increasing the pressure on the recipient to provide an answer.⁸ While Marie’s and Pat’s gazes meet during this re-launch (1.05, Figure 1), Marie averts her gaze already during the ensuing gap, slightly rising it to gaze over Pat’s right shoulder, then keeps it averted while delivering her IDK-prefaced response (1.07, Figure 2), and returns it to Pat only in the further course of her response (1.08):

Marie’s response (1.07–08) comes in late and is *ben* ‘well’-prefaced (cf. Davidson, 1984; Heritage, 2015), both of which concur to indicating an upcoming dispreferred response. The IDK (here: *chais p→as*) itself shows strong morphophonological reduction (see **Table 1** above), and slight rise of pitch projecting continuation. These features concur to displaying it as a preface to some incipient action, rather than as accomplishing an action in itself, and specifically as a preface to a dispreferred next action: Marie provides a non-answer response that resists the terms of the question by claiming inability to answer (1.07; note that in this context French *répondre* corresponds to English ‘answer’); thereby she explicitly claims her response to be a response, and treats the immediately preceding action, preferred in a declarative format (1.05), as pursuing an answer on her part to Pat’s prior question. This is followed by an account for her inability to offer an answer (1.08), which is ultimately accepted by Pat (1.09).

The way the turn is designed is quite conclusive. The turn-initial *ben*: ‘well’ and the subsequent *chais p→as* operate a division of labor. French ‘ben’ is a multifunctional particle, which, in turn initial position in second pair parts, has been found to accomplish such various things as introducing dispreferred responses, indexing contestation of the relevancy of a prior question, prefacing an incipient topical shift or opening a conclusive remark (Bruxelles and Traverso, 2001; Persson, 2020). It is hence an “elusive” (Heritage, 2015) particle similarly to English ‘well’ that, in responses to questions, can signal various types of departures, ranging from dispreference or non-straightforwardness, through resistance regarding the relevance of a question, to steering away from what precedes. In the present case, we have a response that clearly does

⁷Note, that contrary to the typical shrug described by Streeck (2009, p. 189sq), Ana here deploys what Kendon (2004, p. 251) refers to as ‘vertical palm’ « used in contexts where the speaker indicates an intention to halt the current line of action ». This gesture may here be seen as conferring Ana’s disengagement from the ongoing line of action and topic, and not only from her response.

⁸On multiple questions see Clayman and Heritage, 2002; Linell et al., 2003; Pekarek Doehler, 2021a, forthc.

Ex. (4) Hebrew 'Office hour'

01 ALE %ze lo ka-ze kef#. #fig. 1
 this.M.SG NEG like-this.M.SG fun
 it's not such fun
 ale %gazes at Dotan--->1.04
 fig



Fig. 1

02 ...t=o^de'a?
 you=know.PRS.M.SG
 ...y'know?
 dot ^ shifts gaze down--->1.05

03 DOT lama?
 why?

04 ALE (0.11)%(0.11) 'an=lo=y#de'a,
 I=NEG=know.PRS.M.SG
 (0.11) (0.11) I dunno,
 ale ----->%gazes down in front of him-->1.07
 fig #fig.2}



Fig. 2

05 ze lo kef.^
 this.M.SG NEG fun
 it's no fun.
 dot ----->^

06 DOT ^.../tov *ma/
 PRT what
 .../well what/
 dot ^turns gaze to Alex--->>

07 ALE 'i% hayiti yode'a 'al ma 'ani m..daber#,
 if be.PST.1SG know.PRS.M.SG on what I talk.PRS.M.SG
 if I had known what I was talking about,
 %.....gazes at Baracha---->>

fig

#fig. 3



Fig. 3

08 ...ze haya ktsat yoter kef.
 this.M.SG be.PST.3M.SG a bit more fun
 ...it would have been a bit more fun.

Ex. (5) (Pauscaf_13_cam2_1. 590, 0:11:01)

01 PAT **et si: il ne devait pas venir?**
and what if he didn't come

02 (2.5)

03 MAR **bon moi j'crois pas du tout aux relations à distance,**
well I don't believe at all in distant relations

04 **donc euh: c'est==**
so it's

05 PAT **=∞*c'est#la chance de ta vie.***
it's the chance of your life

mar ∞gazes at MAR--->>
fig *gazes at PAT----->*

06 ***(1.1)**
mar *gazes up over PAT's shoulder--->

07 MAR **ben: #chais#p?as j'peux pas te répondre comme ça,***
well dunno I can't answer you like that

fig ----->*blinking, gaze still averted----->*

08 ***c'est un truc qui (s'règle) quand on y [réfléchit?]**
it's a thing that gets figured out when you think about it
mar *gazes at PAT ---->>

09 PAT [(yeah.)]



Fig 1

Fig.2

not conform to the terms of the question. In this context, the IDK more specifically than the *ben:* alerts the recipient to the dispreferred nature of the upcoming response and not only to its non-straightforwardness or to some other moving away from the expected next. The delayed turn-start, the *ben:*, the *chais p?as* prefacing, and the speaker's gaze aversion hence work in concert to project the incipient response as departing from the agenda set up by the preceding request for information.

Excerpts 3 through 5 illustrate a recurrent [IDK + gaze aversion] pattern projecting a dispreferred response in a precise sequential location and action context, namely in responses to seeking information: 3 and 4 in response to a question-word question, in 5 in response to a declarative format that re-does a prior question-word interrogative. The assembly of IDK plus gaze conduct shows a distinct on-line trajectory: gaze aversion either occurs prior to the delivery of IDK (in the transition space) or coincides with its start, but not later than IDK; the respondent's gaze is then maintained averted from the prior speaker during IDK and into the responsive turn, and typically returns to the prior speaker toward the end of the response. As illustrated in the next sub-section, this temporal assembly of multiple resources and its interactional working is also found in responses to polar questions seeking information (and in other action environments, see section "Prefacing dispreferred responses to proposals, assessments, and informings").

Responses to Polar Questions Seeking Information

Excerpt (6), from a Czech TV talk-show, illustrates the case of a dispreferred response to an information-seeking polar question, formatted as a declarative ending in a tag (l.01). Marek Eben interviews the Czech writer Ivan Klíma, who, when he was a child, had been imprisoned in a concentration camp for 4 years, together with his siblings and parents. When Klíma asks Eben whether he had also experienced torment in his life (l.01), Eben responds that his suffering was in no way comparable to Klíma's (l.02).

Through its declarative format, the turn-initial *já myslím* 'I believe' and the final tag *ne* 'no,' Klíma's question (l.01) projects a confirming response as a preferred next (see Sacks, 1992; Heritage, 2010)⁹. Furthermore, the lexical element *trápení* 'torment' confers a sense of strong misery that the speaker suggests the recipient has suffered. Yet, the respondent's response is neither type-conforming, nor does it do confirmation. Rather, the respondent downplays the importance of the misery he himself experienced, thereby discarding the presupposition encoded in the question, while, at the same time conferring sympathy with the profound torment Klíma must have gone through.

⁹Furthermore, it formally displays a preference for a type-conforming response, i.e., a response of the yes-no type. Yet, given its situational context, the question works as an invitation to elaborate rather than ask for a simple 'yes' or 'no' answer.

Ex. (6) Eben, Plovár 24

01 KLI **já** myslim že **ste** měl taky v životě trápení ***ne#**,
 I think that you had PRT too in life torment TAG
 'I believe you had also torment in your life, didn't you'
 †gazes at EBE --> *gazes at KLI -->
 ebe # fig. 1+2
 fig



Fig.1: Eben



Fig.2: Klíma

02 EBE **#no: +tak já nevím*#** **†mně** to tak př- **#přijde (.)**
 well PRT I dunno to.me it so comes
 squints his eyes
 †gaze up-->>
 † moves his head sideways -->>
 fig #fig.3 #fig.4 #fig.5



Fig.3



Fig.4



Fig.5

03 **přijde** mi to v **dost** **značném** **nepoměru**.
 comes to.me it in rather important disproportion
 'well I don't know i'm getting it in quite a disproportionate way
 ((compared to you))'

Just like in the excerpts cited above, the response bears typical traits of a dispreferred action: While it comes in without a gap, the generic turn-initial *no*: 'well' prefacing¹⁰ is here lengthened through sound-stretch, and is followed by the particle *tak*, both contributing to the delay of the actual response. This is enhanced by the subsequent *já nevím* 'I dunno,' which is reduced in form, as Eben pronounces it as [nɛɪm] instead of [nevɪ:m]. The *nevím* is here preceded by the first person pronoun *já* 'I' – albeit Czech is a pro-drop language – , which works as a further delay component, in addition to the preceding *no tak* 'well so,' all of which concur with the

formally reduced [nɛɪm] to project the incipient response as dispreferred. Just as in the preceding excerpts, the IDK is uttered with the respondent's gaze averted from the questioner: During Klíma's question, participants had established mutual gaze (1.01; Figures 1, 2).¹¹ However, exactly at the onset of his reply, Eben stops looking at Klíma, squints his eyes (Figure 3), raises his head slightly and looks away – first up (Figure 4), then down (Figure 5). At the same time, he performs a pragmatic gesture by raising his left hand while further turning his gaze up (Figure 4) and, when doing so, proclaims *no tak já nevím* 'well so I dunno.' (The camera shifts from Eben to Klíma

¹⁰According to Müllerová, 1996, p. 226: "In response to a question or to a preceding turn in general the speaker signals [i.e., projects] by the word *no* (...) his or her doubt, reservation, resignation; negative statement, rejection." (our translation).

¹¹Figures 1 and 2, in the transcript, show shows each participant in a different camera capture but from the preceding and subsequent interaction it becomes clear that these captures present the participants' gazing at each other.

after *poijde* in line 03, so that Eben's further gaze conduct remains undocumented).

The Hebrew **Excerpt** (7) shows Eden and Lital, two friends, sitting together at a café, drinking cold coffee. Suddenly Eden suggests that they will have something to eat (1.01):

Eden's initial question may be seen as a suggestion to order food, yet Lital responds straightforwardly declaring that she is not hungry (1.02), thereby conveying that she is not interested in ordering food. She then immediately adds a polar question asking Eden 'are you hungry' (1.03). Eden's response to this

Ex. (7) Hebrew_ 'Hungry'

```
01 EDE      ...%rotsa      le'exol%  ^mashehu?^
            want.PRS.F.SG eat.INF  something
            ...want to eat something?
            %turns gaze to the right%
ede
ede                                       ^turns gaze to Lital^
```

```
02 LIT      %...'an=lo re'eva,
            I=NEG hungry.F.SG
            ...I'm not hungry,
ede         %gazes at Lital---> 1.04
```

```
03          'at      re'ev#a?
            you.F.SG hungry.F.SG
            are you hungry?
            #fig.1
fig
```



Fig. 1

```
04 EDE      (0.4)%^(1.0) lo=ydat,
            NEG=know.PRS.F.SG
            (0.4) (1.0) dunno,
ede         ---->%
            ^lowers her eyes, gazes at table--->1.07
            #fig.2
fig
```



Fig. 2

```
05          lo 'axalti      klum,
            NEG eat.PST.LSG nothing
            I haven't eaten anything,
```

```
06          'ulay noxal      mashehu?
            maybe eat.FUT.1PL something
            maybe we'll eat something?
```

```
07          ...ba-ktan^~%tana?%
            in.the-little~ATT.F
            ...just a little little bit?
ede         ----->^
ede         %shifts gaze to Lital%
```

```
08          ...m[a yesh      po      bi-xlal?
            what there_is here in-general
            ...what do they have here anyway?
```

```
09 LIT      ['ani be-diyeta.
            I      in-diet
            I'm on diet.
```

question comes in with a strong delay during which she lowers her gaze, averting it from Lital, and subsequently produces the morphophonologically reduced token of *loydat* ‘dunno’(l.04) further delaying her response (cf. Schegloff, 1988, p. 445), while keeping her gaze averted and returning it to the questioner only way into her turn (l.07). Note that even though *loydat* is in a separate intonation unit (Du Bois, 2012 transcript), it is still delivered as a single TCU with the subsequent intonation unit. Here again the IDK plus gaze aversion occur as a preface to a dispreferred response. For one thing, Eden’s response is not type-conforming with regard to the yes-no question (Raymond, 2003). Furthermore, some kind of opposite stance between the participants is also in play, as Eden’s admitting that she has not eaten anything (l.05) and suggesting that they could order ‘just a little little bit’ contrast with Lital’s prior affirmation that she herself is not hungry, and by implication, does not ‘want to eat’ (l.01), i.e., order food. This is further foregrounded by the fact that Eden repeats her (implicit) suggestion from line 01, this time using first person plural denoting both her interlocutor and herself (l.06): ‘maybe we’ll eat something?’.

Intermediate Summary

The excerpts discussed in this section showed a recurrent verbal-embodied practice for projecting dispreferred responses to information- and confirmation-seeking questions: speakers use IDK in turn-initial position, typically in a morphophonologically reduced form, combined with gaze aversion, thereby foreshadowing the non-conformity of the upcoming response to the sequence-initiating question. In the data, these responsive turns show consistently dispreferred action-turn-shapes: delayed turn-starts, particles such as *ben*, *no* ‘well,’ lengthening, hesitation markers such as *u:h* and/or vocalizations such as *phhh* or clicks preceding the IDK; all these elements push the actual response further back into the turn. The precise sequential location of IDK in turn-beginnings as “sequence-structurally important places in conversation” (Schegloff, 1987, p. 71) is decisive for its working as a preface through which respondents alert co-participants to the dispreferred nature of their incipient responsive action. The morphophonologic reduction and semantic bleaching of IDK indicate that it is being used as a routinized (or: grammaticized, cf. Hopper and Traugott, 2003) particle-like element rather than a subject-verb-negation combination. We will return to this in the discussion (see section “Discussion and conclusion”).

Respondents’ verbal and gaze conduct are assembled in time in a way that gaze aversion either starts in the transition space, i.e., prior to the delivery of IDK, or simultaneously with the onset of IDK, but not later than that. More precisely: When there is no gap between the sequence initiating action and the response, gaze aversion coincides with the IDK preface of the response (ex. 3), i.e., with the verbal start of the responsive turn; by contrast, when gaps or other elements such as ‘well’-prefacing further delay the production of IDK and the response, then gaze aversion tends to start prior to IDK (ex. 4, 5, 6, 7), that is: It tends to start shortly after the end of the question turn (for an exception see ex. 10 below). This observation, though in need of fine-grained analysis based on a larger amount of cases, is roughly in line with

Kendrick and Holler’s (2017) finding that gaze aversion begins most frequently 100 ms after the first possible completion point of the question (even in the case where questioning turns reach multiple possible completion points). In all of the examined cases, gaze then remains averted during IDK and into the responsive turn, and typically returns to the prior speaker toward the end of the response. As we will see in what follows, the observed gaze trajectory and its temporal relation to verbal conduct is recurrent across the languages and action contexts studied. Though there are exceptions to this, it is a strong tendency observed in the data.

Prefacing Dispreferred Responses to Proposals, Assessments, and Informings

While we have so far focused on the question-answer adjacency pair, particle-like uses of IDK combined with gaze aversion are also found in other contexts of incipient dispreferred actions. In this section, we show that the practice of prefacing a dispreferred response with [IDK + gaze aversion] is generalizable across a range of sequence-types, being recurrently found in our data in responses to proposals, assessments, and informing. For reasons of space, we here limit ourselves to illustrating each of these action sequences by one or two examples taken from the languages studied.

Proposals

Let us start with two illustrations of IDK in responses to proposals. Couper-Kuhlen (2014) suggests that proposals can be distinguished from similar types of actions (such as suggestions, requests, and offers) in that they are used when the activity is framed as benefiting both speaker and hearer. **Excerpt (8)** shows an example from French. Daniela and Penny are talking about a joint assignment they have to do for one of their university professors. Daniela proposes that they should go and make an appointment with the professor to discuss some details of the assignment (l.01–02). Daniela gazes at Penny throughout the excerpt.

In the face of Penny’s lack of response to her proposal (l.03), Daniela increments it (l.04–06) in pursuit of a response, providing further details about when, in the course of their work-process, they should go to see the professor. It is only here that Penny reacts, yet she does so with a disaligning action: instead of responding to the proposal, she asks if the professor has reception hours (l.07), which ensues in a complex side-sequence (l.07 and following) extending beyond the cited excerpt. Her *chais pas* here introduces an action that is structurally disaligned with the preceding proposal; therefore it can be seen as signaling a sequential disjuncture. Note that Penny’s response is not preceded by any hesitation markers or silence or turn-initial particle other than the IDK, but still comes in a much delayed way in regard to the recognition point of the Penny’s sequence initiating action, the first formulation of which was offered in lines 01–02 and then re-cast in line 04. While Penny had turned her gaze on Daniela in the course of the latter’s proposal (Figure 1) and had maintained it on her throughout part of the re-launch (Figure 2), she then first turns her gaze up over Daniela’s right shoulder right before Daniela’s turn extension reaches a transition relevance place (l.05), and subsequently turns

Ex. (8) Pauscaf 9, 06:55

01 DAN *mais >j'pense< que <U#ne fois> il faudrait juste prendre
 but I think that one time we would just need to take
 pen *->gazes down at her coffee cup-->

02 un moment av*ec #la prof?
 a moment with the professor
 pen *gaze at DAN, establishing mutual gaze --->1.05
 fig #fig.1

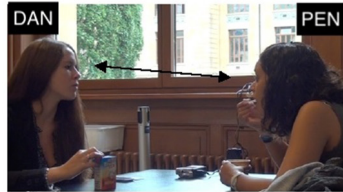


Fig.1

03 (0.2)

04 DAN tu penses PAS?
 don't you think

05 >ge*nre<# avant de commencer les ana*ly:ses,
 like before starting the analyses
 pen *gaze up over DAN's right shoulder-->
 fig #fig.2



Fig.2

06 ou #avant de [commencer un truc concret?]
 or before starting a concrete thing
 fig #fig.3

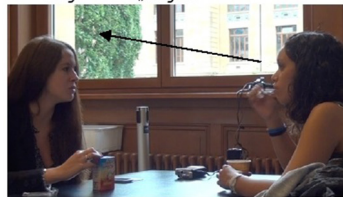


Fig.3

07 PEN [*chais# pas c'est quand ses]
 dunno it's when her
 --->*turns gaze down to her right-->
 fig #fig.4



Fig.4

08 heures de rend*ez vous à elle?
 reception hours
 --->*turns gaze back at DAN-->

09 (0.5)

10 DAN °elle a des heures de rendez-vous tu penses°?
 she has reception hours you think

11 (0.5)

12 PEN °j'pense°.
 I think

it down toward her right simultaneously with her production of *chais pas* (Figure 4, 1.07). Her gaze aversion starts here quite in advance of her response (1.05), and the relatively late delivery of that response in relation of this gaze conduct may be due to Penny's momentary inability to respond verbally, as she is

licking off her coffee spoon (see Figure 2) and takes it out of her mouth only immediately before the delivery of *chais pas* (Figure 3). In other words, gaze aversion is here deployed in a much premonitory manner to the later verbal projection of a dispreferred response through IDK.

So far, we have discussed excerpts in which the recipient actively averts her gaze away from the prior speaker before or concurrently with the IDK token. There are also instances in our data where the respondent's gaze is already averted from the prior speaker during the speaker's sequence-initial action, especially when participants are engaged in other activities; in these cases, the respondent's gaze simply remains averted with the delivery of IDK. **Excerpt (9)** shows such a case from the Mandarin data where Qun and Str are playing a puzzle (see Figure 1) and are making plans about how to go to a recording appointment together afterward (lines 01–08). The sequence starts with Qun

checking whether she understood their plan correctly, namely that they are 'going there together' (l.01).

After Qun requests confirmation of her understanding that they are going to the recording studio together (l.01), Str first offers such confirmation (l.02) but then proposes an alternative arrangement which is that they could go separately and directly meet there (l.03). Overlapping with Str's turn extension (l.04), Qun produces *wo bu zhidao* 'I don't know' (l.05) prefacing an account of her rejection to Str's proposal (line 06; the 'wo' [uo] is here reduced from a diphthong to a schwa). Here, the account in itself does the rejection: When mentioning that her

Ex. (9) BJ_R04_FF_20180621_V_ZO01_25'14"

01 QUN *women shi yiqi zou shima.
 1PL be together go be PRT
 We are going there together right
 *gazes down--->l.05

02 STR +keyiya-
 OK PRT
 OK
 +gazes down--->

03 huozhe ni bu fang+bian ye keyi zai nar jian.
 or you NEG convenient also may at there meet
 or we can also meet there if it's not convenient for you
 ----->+gazes at Qun--->

04 dao[shihou?+
 th[en
 th[en
 ----->+

05 QUN [#wo buzhidao yinwei wo zhuzai shang*+di #neibian;*+
 [I NEG know because I live in (NAME) that side
 [I don't know because I live near Shangdi
 gazes at STR
 +gazes at QUN+
 str #fig.1. #fig.2



Fig.1.



Fig.2

place (the neighborhood named Shangdi, 1.06) is far away from the recording studio, Qun implies that going there separately would incur more transportation costs than sharing a taxi. It is noteworthy that, when Qun seeks Str's confirmation and when Str makes the alternative proposal, Qun looks down at the puzzle (1.01–04). She keeps her gaze down away from Str when she produces *wo bu zhidao* 'I don't know' in line 5 (Figure 1) but then briefly looks at Str toward the end of her account (Figure 2). Here, then, the respondent's gaze is not actively averted, but merely *remains* averted from the recipient, throughout most of the dispreferred response, yet just as in the prior excerpts returns to the recipient toward the end of that response. So, again, the excerpt shows the same response-initial configuration involving IDK plus gaze averted from recipient as a preface to a dispreferred response. Here, it prefaces the rejection of a proposal.

Assessments

Turning now to disagreeing responses with assessments, consider **Excerpt (10)** from the Czech data. This is taken from a TV talk show in which the host (E. Kočíčková, in the middle of **Figure 1**) leads a talk on homosexuality. The excerpt comes from the beginning of the show, after the host had mentioned that she would have liked to have both genders represented among her guests, but that her 'female adventures' did not have the courage to show up (1.01–03). This is produced as an informing containing a negative assessment, to which the male guest Špaček reacts (1.05–06) with a *já nevím* prefaced response.

Clearly, the host's display of disappointment (1.01–02) about not having found a female to sit in the show projects recipients' affiliation (a shared stance) as a relevant next. Furthermore, it contains a negative assessment. Her *žádná z my:ch bývalý:ch avantýr ženských nebyla té odvahy, h aby... (1.01–02)* – literally 'none of my former female adventures was of the courage to... ' – involves both an informing and an assessment: While she informs about her former adventures not having come to the show, she also qualifies their not coming as a lack of courage: The assessing element in her turn can be heard as synonymous with the canonical assessment 'they were too coward to...'. This is what Špaček orients to in his response, in line 05. Rather than responding to the informing, for instance by

means of a display of change of knowledge (Heritage, 1984) or by 'acknowledging' the informing (Thompson et al., 2015), Špaček displays strong disaffiliation with the host's stance as well as disalignment with her negative assessment: He counters that assessment by displaying understanding of the women that purportedly refused to participate in the show: Had he been in the same situation as these women, i.e., having had a liaison with the host, he would not have participated either. He thereby disqualifies the idea that these women lacked courage, i.e., were cowards. The *já nevím* 'I don't know' hence occurs as a preface to both a disaffiliating (in terms of stance) and a disaligning (in terms of the structural organization of actions) response, of which the dispreferred nature is further highlighted by the turn-initial *no*, roughly corresponding to English 'well' (Müllerová, 1996; see footnote 11 above; Auer and Maschler, 2016), itself preceded by a lengthy and heavy inbreath on the part of the recipient. Just like in **Excerpt (6)** above, the speaker's use of the first person pronoun *já* 'I' – usually dropped in spoken Czech – works here as a further delay component to the dispreferred response. Furthermore, the whole stretch *no já nevím* is produced with notable speed-up of tempo, being prosodically downgraded, which further adds to its being heard as a preface¹². The response is then treated as disaligning, and even disaffiliative by the host (1.08), who overtly reproaches to Špaček to have offended her.

The *já nevím* is part of a particularly prominent verbal-bodily assembly here, involving not only the respondent's gaze but also his hands. During her informing, the host's gaze wanders between her two guests and at the table in front of her, but toward the end of her turn, she turns her gaze to Špaček, apparently recruiting him as the next speaker (cf. Lerner, 2003; see already Kendon, 1967), which ensues in the establishment of mutual gaze (Figure 2, 1.10). In an anticipatory manner, Špaček starts turning his gaze away before the end of Kočíčková's turn (Figures 3, 4, 1.03/4), but after the gist of her turn (the negative assessment) has become recognizable (see Broth and Keevallik, 2014; Pekarek Doehler, 2021b, *forthc*, for how the recognition point of a turn/action in progress may affect the timing of responsive actions, both verbal and embodied). Toward the end of the host's turn, he shortly closes his eyes (Figure 3, 1.03), pulls his hands toward his face, opens his palms and starts gazing at them (Figures 4, 5, 1.04/5). His gaze aversion hence precedes the delivery of the *já nevím*, is maintained during that delivery and further into the turn and returns to Kočíčková only in the further course of that turn (Figure 6, 1.05). This gaze trajectory converges with the evidence provided in the prior excerpts (see in particular section "Intermediate Summary"). In ways similar to what we have observed in dispreferred responses to questions (see section "Prefacing Dispreferred Responses to Questions" above),

MP M. Putna; KOC E. Kočíčková ; SPA R. Špaček



FIGURE 1 | Spatial arrangement of participants in the TV show.

¹²It is important to note that the subsequent if-clause, in line 5, cannot grammatically stand as the complement clause to IDK: in Czech, *jestli* (if or whether) is a typical conjunction for complement clause following (*já*) *nevím*; the conditional conjunction *kdyby* used here, by contrast, does not work to introduce a complement clause after (*já*) *nevím*; here, it marks the start of a stretch of talk that is syntactically independent of the preceding *já nevím*.

Ex. (10) Špaček, Dvojhlas 5

01 KOC =.hh jenomže žádná (.) eh z mý:ch bývalý:ch avantýr ženských
 but none of my former adventure female

spa >>igazes at KOC---->1.03

02 eh nebyla té odvahy .h aby sem přišla *se mn#ou#>,
 NEG.was of.the courage to here came.F with me
 had the courage to join me

*gazes at SPA -->
 #fig.1

fig



Fig. 1

03 (.) nahlas poho#vořit. o tě#ch [vě#cech.
 loud speak of these things
 to speak out about these things

spa --->igazes down, closing eyes
 spa igazes at his hands---->1.05

fig

#fig.2

#fig.3

#fig. 4

04 SPA

[.hhhh



Fig. 2



Fig. 3



Fig. 4

05 SPA >no já nevím<# ikdybysme# *my dva spolu něco měli*
 well I NEG.know if we two together something had
 well I don't know had we had something going on together

--->i.....gaze at KOC-->>

waves his right hand between himself and KOC

#fig.5

#fig.6



Fig.5



Fig. 6

06 *tak nevím jestli sem přijdu.*
 then NEG.know if here I.come
 I don't know if I come here
 points to himself with both hands

07 *(0.2)# *(1.4)

koc *gaze at SPA*turns gaze away -->>

Fig

#fig.7



Fig. 7

08 KOC aha, no to si mě urazil.
 oh PRT PRT you.AUX me offended
 oh there you offended me

distinctly strong inbreath, gaze aversion, gesture, *no* 'well' and *já nevím* 'I don't know' work in a minutely, step by step assembled way, to incrementally build up the incipient dispreferred response.

Informings

Further contexts where we find the focal multimodal assembly is in responses to informing, especially those that confer the speaker's stance toward the reported state of affairs. Following

Thompson et al. (2015, p. 51), we “use the term ‘informing’ to designate the action done when a speaker’s turn is constructed to provide information to a non-knowing recipient such that they become (more) knowing.” ‘Informing’ is hence a cover-term that may include actions such as announcements or news deliveries. As the authors show, respondents generally react to informing by indexing a shift in their epistemic stance, from not or less knowing to more knowing, for instance by means of the delivery of a newsmark (Heritage, 1984, pp. 339–349).

The Romanian **Excerpt (11)** provides a first illustration. Ana, who is world amateur champion of paragliding and skydiving, just informed Greta that a skydiving contest would take place in August in the city of Dara, where both are working. Ana’s informing then continues with her naming the competitions that she would take part in later, in September (l.01 ff.):

From line 01 to 07, Ana informs Greta about two competitions that she is planning to do in September. Most of the informing is produced in a rather non-committal way, Ana gazing either down (Figure 1) or at her left-hand fingers, while enumerating the respective contests. Her informing though becomes livelier when she announces that the skydiving national championship should take place also in Dara (l.07), just as the other contest planned for August. Here, Ana switches her gaze to Greta, smiling at her, displaying satisfaction with the latter piece of news (Figure 2). Vocally, also, the delivery of this piece has been made in a dramatic fashion, as Ana changes her voice quality and produces the relative pronoun *care* ‘which’ (l.05) with a prosodic emphasis, marked high rise in pitch and sound stretches, projecting thereby a noteworthy information to come up next. In short, Ana displays a clearly positive, if not enthusiastic stance with regard to the news she is delivering (on news delivery sequences, see Maynard, 1997).

Greta, however, does not affiliate with Ana’s stance, and the pause at line 08 projects a potential dispreferred reaction on her part (note the contrast to her response in line 06). Through her phrasal repeat of *tot la Dara* ‘also in Dara’ (l.09) with rising intonation, she first initiates repair on Ana’s turn by asking for a confirmation of the place where the championship should be held, and, after receiving such confirmation (l.10), keeps silent again (l.11), displaying no verbal, prosodic or embodied affiliation with her interlocutor’s stance.

Greta’s dispreferred reaction becomes clear after this short side sequence as she takes up the floor again and produces an extended turn [l.13–18 – an unrelated clausal response (see Thompson et al., 2015, p. 61)], asking Ana if there are not any other places in the country, apart from Dara, where this championship could be held. Her response is dispreferred not in the sense that it does not treat the information provided as news, but in that it denies the positive valence of the news presented. The dispreferred nature of her turn is foreshadowed by the particle *da* ‘but’, morphophonologically reduced (from *dar* to *da*), itself preceded by a marked in-breath (l.13). Together with IDK (*nu știu*), which is delivered with speed-up tempo and prosodically latched to what follows, these push the response proper further back into the turn. Also, Greta averts her gaze from Ana exactly with the strong inbreath and keeps it averted until line 14, while she was before constantly looking at Ana (l.01–12).

Thompson et al. (2015, p. 14) note: “There exists [...] a strong norm for at least acknowledging, or ‘receipting,’ an informing; in this sense, a ‘dispreferred’ response to an informing would be to not acknowledge the information as an informing at all.” This is exactly what Greta does: Rather than acknowledging the informing, Greta goes on questioning the relevance of holding the championship at Dara. Finally, note that Greta herself orients to the dispreferred nature of her own turn as she provides two parenthetical comments on the fact that she appreciates nevertheless that the championship is organized in Dara (l.15–16), thereby somewhat mitigating her rather strong disaffiliative reaction to her coparticipant’s stance.

Excerpt (12) illustrates a disagreement with what can be qualified as an informing that is strongly asserted, i.e., offered as a claim about a state of affairs; as such, it resembles what Vatanen et al. (2020, p. 6) qualify as an ‘assertion’. It occurs in a context of prolonged disagreement between the coparticipants, where it conveys the speaker’s stance toward the talked-about issue. Luo and Yan are commenting on the performance of a well-known Chinese actress Zhang Yizi (referred to as ‘she’ in line 01) in a movie they have recently watched together.

In lines 01–06, Luo and Yan affiliate with each other in producing positive assessments of the boldness and versatility of the actress Zhang Ziyi. But in lines 07–12, Luo pivots his evaluative stance. He produces a qualified assessment that the actress plays particular types of characters with large internal conflicts well (l.07–08), thereby somehow disaligning with Yan’s prior assertion that ‘(she) can do any type of acting’ (l.03). In line 9, Luo makes a syntactically incomplete ‘off-record’ critical assessment of the performance of Zhang Ziyi in the specific movie that Luo and Yan recently watched. Syntactically incomplete assessments in Mandarin conversation have been documented to perform ‘off-record’ negative assessments (Li, 2016; see also Park and Kline, 2020 for a similar use in English), which can here possibly also be observed from Luo’s gaze aversion from Yan. Then, in lines 10–11, Yan asserts that the actress plays an ‘entirely’ minor role in the movie, possibly in an attempt to discard the idea that one could solidly judge the actress’ talents based on that one movie. The extreme-case formulation *zhengge* ‘entirely’ (Pomerantz, 1986; Whitehead, 2015) expresses Yan’s “unstated disagreement” with Luo (Pomerantz, 1984, p. 76). Immediately after Yan’s turn, Luo produces *bu zhidao* ‘(I) don’t know’ (l.12) followed by another negative assessment of the actress (ll. 12, 14, and 16). It is interesting that Luo averts her gaze from Yan already from line 03 on, possibly in a premonitory way to her subsequent verbally expressed disagreeing stance, which she builds up incrementally: During Yan’s turn in line 03, Luo shifts his gaze from Yan down at the mug on the table and keeps his gaze down during almost the entire sequence (Figure 1), lowering it further immediately after *bu zhidao* ‘(I) don’t know’ (Figure 2, l.12). So, here, Luo also seems to deploy *bu zhidao* ‘(I) don’t know’ (l.12) and a lack of mutual visual engagement with Yan (Figure 4) to project his disagreement with Yan’s immediately preceding assertion and the evaluative stance expressed therein (ll. 10–11).

In this section, we expanded our prior observations documenting that the focal assembly of gaze aversion plus IDK in turn-initial position represents a practice that can be

Ex. (11) ‘Da’ nu știu’ [Corpus Stoenica_10m36-11m19]

01 ANA \$*și în septem[brie::] ∞(0.9)∞ deci avem #ă:: campionatul&
 and in September so we have uh the world
\$gazes down, in front of her----->1.03
∞swallows∞

gre *gazes at ana--->1.13

fig.

#fig.1



Fig.1

02 GRE [°că mergi°]
 that you go

03 ANA &mondial la Vranje cu- (0.4) \$de precizia aterizării cu
 championship at Vranje with- of landing precision with
----->\$gazes at her left-hand fingers->1.05

04 parapanta și după aceea .h ä:: o sa fie și campionatul
 the paraglider and after that uh there will also be the skydiving

05 național de parașutism \$ca:↑re:=-
 national championship which
----->\$gazes up in front of her-->1.07

06 GRE =a[oleu.]
 oh boy

07 ANA [ar trebui] să se \$țină ∞#tot la Dara.
 should take place also in Dara
--->\$gazes at gre--->
∞smiling, opens both palms--->1.08
 #fig.2

fig

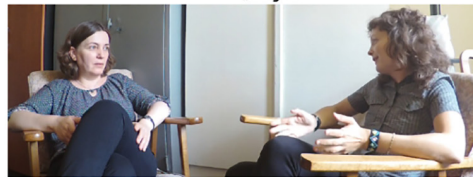


Fig.2

08 (0.5)∞
 ana ----->∞

09 GRE tot la Dara?
 also in Dara

10 ANA da.
 yes.

11 (0.7)

12 ANA [și: m][:-
 and



Fig.3

13 GRE *[[.h]][da' #>nu știu:<da' în țară nu mai există așa un club
 but dunno but in the country there is no other club
**gazes away in front of her--->1.14*

fig.

#fig.3

14 *ă: știi să zici bă uită-te,
 uh you know to say man look
**gazes back at ana--->>*

15 bun ä ii foarte bine că-i la noi Δștii?Δ
 ok uh it's very good that it's at our place you know
 ana Δnodes x2Δ

16 no chiar mă: bucur că că-i la noi,
 PRT I'm even glad that it's at our place

17 da' așa zic nu mai există să mai mergi în țară nu știu,
 but I just say couldn't you still go in our country dunno

found not only with the prototypical adjacency pair of the type question-answer, but also in a range of other action contexts, such as responses to proposals, assessments and informing –

specifically informing that convey the speaker's stance. The practice is hence deployed in locally functional ways across a range of sequence types. The excerpts cited have also shed

Ex. (12) The Chinese actress Zhang Ziyi_CS_VL_01:20:55 audio 01:33:29

01 LUO ta danr tebie da.=
3sg hardy very much
She (Zhang Ziyi) is very bold

02 YAN = .h dui (.) ta?
= .h yes (.) 3sg
Yes she

03 zenme *yan +dou keyi*
how perform all OK
(she) can do any type of acting
+gazes at Luo--->1.06
luo *gazes at Yan--->*gazes down--->1.12

04 ni zenme (.) jiushi[rang ta zhege;
you how (.) just be [ask 3sg this-CL
No matter how you ask her (to play) this

05 LUO [(dui ta.)
[(right, she.)

06 YAN rang ta neige juese +(douxing:.
ask 3sg that-CL character[all fine:.
(or) ask her (to play) that character she's all fine
---->+gazes down-->

07 LUO [yiding you neizhong renwu jiushi;
[must have that character just
There must be the type of character...

08 neizhong: chongtu tingda de renwu ta yanqilai ye ting hao de.
that type conflict very big NOM character 3sg perform also very good PRT
that type of character with very big (internal) conflict. She can act that
very well.

09 xiang xiang zheige zheige nude ta yande jiu (click)
like like this this women 3SG act-COM just
Like like this this female character, her performance was just

10 YAN zheilimian ta ye zhengge jiushi yige yige chuanr ba.
this inside 3sg also entirely just be one-CL one-CL minor character PRT.
In the movie she entirely played a minor role

11 yige peijue.
one-CL supporting role.
a supporting role

12 LUO #<<creaky>bu zhidao>#*[fanzheng wo juede yande jiu (ye);
NEG know [anyway I feel perform just
(I) don't know anyway I think (her) acting is just
----->*gazes at the table--->1.17
#fig.1 #fig.2

Fig.1

Fig.2

13 YAN [ye mei banfa;
[also NEG solution;
[There is nothing (she could do).

14 LUO [<<p>bushi te te.>
[NEG be very very
[not very, very,

15 YAN [bu xian tade.
[NEG show her
[(It) doesn't (allow) her (talent) to be shown

16 LUO <<pp>tebie(.)tebie nasha.>
very (.)very that
very, very na sha

17 YAN <<pp>en.>*
<<pp>umm.>
Umm
----->*

18 LUO *xiang #<NAME:liming> *neiyang;
like <NAME:liming> that type
Like Li Ming
*glances at Yan-----*down----->

19 <NAME:liming> yizhi yan neizhong juese (wo juede);
<NAME:liming> always perform that-CL character (I think);
(I think) Li Ming always plays that type of role

further light on the trajectories of recipients' gaze respective to the prior speaker and their own verbal conduct. They confirmed the consistent (re)turning of recipients gaze to the prior speaker

toward the end of the responsive turn – even when participants are involved in multiactivity, such as playing a puzzle (ex. 10). The excerpts also showed that recipients' gaze aversion

from prior speakers typically occurs in the transition space or simultaneously with the response onset; only rarely does it occur during the preceding speaker's turn, but in any case after the recognition point of the prior action, and hence of the conditionally relevant next action. The onset of such gaze conduct has practical interactional import: For instance, when it precedes the end of the sequence initial turn (ex. 10), it may be a way for respondents to project a dispreferred response in a premonitory way while circumventing overlap; when it occurs in the transition space while the recipient is unable to speak (e.g., while eating, ex. 8), it may be a way to warrant early projection of aspects of the incipient responsive action. In this sense, respondents can be seen to minutely assemble, on-line, their gaze conduct and their vocal conduct in locally functional ways for all practical purposes.

DISCUSSION AND CONCLUSION

In this paper we examined how an assembly of verbal and embodied conduct is related to one of the basic organizational principles of social interaction: preference organization. We documented a recurrent bimodal practice in which speakers deploy IDK in combination with gaze aversion in turn-initial position to project an incipient dispreferred response: Through this practice, speakers retroactively display resistance to the constraints set by and/or disagreement with the stance conveyed by the immediately preceding action, and prospectively project a dispreferred response. We showed that this practice occurs across a diverse set of languages and a variety of sequence types. The findings call for further detailing based on a more extensive collections for some of the languages (namely: Romanian), and comparison across a larger set of languages, and specifically of culturally more diverse participant groups: Despite the diverse language (sub-)families represented in this study, all of our data stem from post-industrial societies. With these limitations in mind, we spell out, in the following paragraphs, some implications of our findings.

We started out by choosing to work on a specific negative epistemic expression involving 1st person and the negated verb of knowing. It turned out that in the focal sequential position – turn-initial position in responses to various initiating actions –, the structure could be variably prefaced with additional materials, such as clicks, hesitation tokens, and particles such as *ben* in French or *no* in Czech. Importantly, it was often morphophonologically reduced and sometimes prosodically downgraded (by lower volume or speed up of tempo) in all of our languages. These features suggest a particle-like working of the IDK construction, which appears to have routinized in all of the studied languages into an interaction-organizational device.

This finding adds to existing research on IDK, evidencing how prefacing uses, far from being limited to functioning as epistemic hedges (e.g., Weatherall, 2011), are implicated in the management of the multimodal infrastructure of social interaction, and specifically in the prefacing of dispreferred responses. Heritage and Sorjonen (2018) show that turn-initial

objects in first position manage the connection of the current turn to its immediately preceding one, and those in responsive position may be used for resisting the constraints set by the first turn on the second position speaker. This study develops this line of research in two ways. First, we showed that in addition to particles such as *well*, *ben*, *nu/no* documented in previous research, phrases such as 'I don't know', concurrent with gaze aversion, are a common occupant in turn-initial position across five distinct languages. This adds to our knowledge of the *type* of turn-initial objects, and particularly suggests that these may include not only linguistic but also bodily visual aspects as part of methodic turn-construction practices. Second, by showing that and how the multimodal practice consisting of IDK and gaze aversion is deployed to preface an incipient dispreferred response in a range of sequence types, this study demonstrates that a practice like this can be applicable across a variety of responsive actions, possibly irrespective of any precise type of the initiating action.

The findings further add to existing knowledge on gaze in interaction, expanding specifically on the results offered by Kendrick and Holler (2017), and more recently Robinson (2020), who demonstrate that dispreferred answers to polar questions tend to correlate with respondent's gaze averted from the questioner. Clearly, gaze aversion also in animals is related to submission and avoiding confrontation, so this gaze aversion in dispreferred sequences is valid across species and has its natural origin in non-confrontational behavior in general (see Kendrick and Holler, 2017). The findings presented here amplify prior observations that gaze-aversion is found with dispreferred responses by showing that this association is valid across a range of sequence and action types, extending to responses beyond those provided to polar questions, and that this is the case across genetically and typologically different languages. Based on these findings, we suggest that by paying close analytic attention to the multimodal make-up of turn formats we might arrive at a more fine-grained understanding of the methodic multimodal practices involved in turn construction and action formation, such as the temporal unfolding of gaze behavior *in relation* to verbal and vocal conduct, and the uttering of specific turn-initial verbal phrases *in concert* with precise embodied conduct.

Finally, the type of multimodal sequential analysis we conducted here allowed us to evidence gaze *trajectories* associated to dispreference. The data show that gaze aversion most typically starts after the end of the sequence initial action – a result that converges with Holler and Kendrick's (2017) earlier finding for responses to polar questions. Additionally, the present study evidences that such aversion starts either in the transition space if there is a gap between turns, or else concurrent with response onset; it also shows that respondents' gaze tends to revert to the prior speaker toward the end of the responsive turn, and that these gaze trajectories hold across different sequence types and languages. The fact that gaze aversion is rarely found to overlap with the end of the sequence initial turn is intriguing in light of prior research showing that response planning begins as early as possible, and sometimes even during the turn-in-progress

(e.g., Levinson and Torreira, 2015) – which might ensue in the production of responsive turns in overlap with sequence initiating turns (Pekarek Doehler, 2021b, forthc) or responsive embodied action (such as affirmative nods, De Stefani, 2021) before the end of initiating turns and actions. Ultimately, by evidencing how grammar and body interface in related ways across languages, the findings open a window onto cross-linguistic, cross-modal, and cross-cultural regularities in human interactional conduct.

TRANSCRIPTION CONVENTIONS FOR EMBODIED CONDUCT

- **/± Symbols such as these indicate start and end of embodied conduct
- *—> l.12 Continuation of the described embodied conduct until line 12 of transcript.
- >* End of the described embodied conduct
- *—>> Continuation of the described embodied conduct until end of excerpt
- # Indicates the location of a figure in the verbal transcript

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/supplementary material, further inquiries can be directed to the corresponding author/s.

ETHICS STATEMENT

Ethical review and approval was not required for the study on human participants in accordance with the local legislation and

REFERENCES

- Auer, P., and Maschler, Y. (eds) (2016). *NU/NÁ: A Family of Discourse Markers Across the Languages of Europe and Beyond*. Berlin: De Gruyter.
- Bögels, S., Kendrick, K. H., and Levinson, S. C. (2015). Never say no... How the brain interprets the pregnant pause in conversation. *PLoS One* 10:e0145474. doi: 10.1371/journal.pone.0145474
- Bögels, S., Kendrick, K. H., and Levinson, S. C. (2020). Conversational expectations get revised as response latencies unfold. *Lang. Cogn. Neurosci.* 35, 766–779. doi: 10.1080/23273798.2019.1590609
- Broth, M., and Keevallik, L. (2014). Getting ready to move as a couple: accomplishing mobile formations in a dance class. *Space Cult.* 17, 107–121. doi: 10.1177/1206331213508483
- Bruxelles, S., and Traverso, V. (2001). “La particule ‘ben’ dans deux type d’interactions,” in *Special Issue of Cuadernos de Filología*, ed. S. Pons Borderia (Valencia: Université de Valence).
- Bybee, J., and Scheibmann, J. (1999). The effect of usage on degrees of constituency: the reduction of *don’t* in English. *Linguistics* 37, 575–596.
- Clayman, S. (2002). “Sequence and solidarity,” in *Advances in Group Processes: Group Cohesion, Trust and Solidarity*, eds E. J. Lawler and S. R. Thye (Oxford: Elsevier Science), 229–253. doi: 10.1016/s0882-6145(02)19009-6
- Clayman, S., and Heritage, J. (2002). Questioning presidents: journalistic deference and adversarialness in the press conferences of U.S. Presidents Eisenhower

institutional requirements. Czech data comes from TV talk shows but converges with the data in the other languages in that it consists of unscripted interactions that unfold spontaneously. As the latter data comes from the public domain, it neither needs anonymization nor consent for publication. Names in all other data have been anonymized, and written informed consent has been obtained from all participants for use and publication of the materials, including the video frames.

AUTHOR CONTRIBUTIONS

SPD has done the main research work on this manuscript, in collaboration with HP-Y. All co-authors contributed with data analysis and detailed discussions of the various versions of the manuscript. All authors contributed to the article and approved the submitted version.

FUNDING

The French and the Romanian parts of this study were conducted with the generous support of the Swiss National Science Foundation, as part of the project *The Emergent Grammar of Clause Combining in Social Interaction*, grant no. 100012_178819. The Hebrew part of the study was supported by the Israel Science Foundation, grants nos. 1233/16 and #941/20 to Yael Maschler.

ACKNOWLEDGMENTS

We thank the two reviewers for their extraordinarily constructive and stimulating comments on a prior version of this manuscript.

- and Reagan. *J. Commun.* 52, 749–775. doi: 10.1111/j.1460-2466.2002.tb02572.x
- Couper-Kuhlen, E. (2014). What does grammar tell us about action? *Pragmatics* 24, 623–647.
- Curl, T. S., and Drew, P. (2008). Contingency and action: a comparison of two forms of requesting. *Res. Lang. Soc. Interact.* 41, 129–153. doi: 10.1080/08351810802028613
- Davidson, J. (1984). “Subsequent versions of invitations, offers, requests, and proposals dealing with potential or actual rejection,” in *Structures of Social Action*, eds J. M. Atkinson and J. Heritage (Cambridge: Cambridge University Press), 102–128. doi: 10.1017/cbo9780511665868.009
- De Stefani, E. (2021). Embodied responses to questions-in-progress: silent nods as affirmative answers. *Discourse Process.* 58, 353–371. doi: 10.1080/0163853X.2020.1836916
- Deppermann, A. (2013). Turn-design at turn-beginnings: multimodal resources to deal with tasks of turn-construction in German. *J. Pragmat.* 46, 91–121. doi: 10.1016/j.pragma.2012.07.010
- Dingemans, M., and Floyd, S. (2014). “Conversation across cultures,” in *Cambridge Handbook of Linguistic Anthropology*, eds N. J. Enfield, P. Kockelman, and J. Sidnell (Cambridge: Cambridge University Press), 447–480. doi: 10.1017/cbo9781139342872.021
- Dryer, M. S., and Haspelmath, M. (eds) (2013). *WALS Online*. Leipzig: Max Planck Institute for Evolutionary Anthropology.
- Du Bois, J. W. (2012). *Representing Discourse*. Santa Barbara, CA: University of California. (Fall 2012 Version).

- Floyd, S., Rossi, G., and Enfield, N. J. (eds) (2020). *Getting Others to Do Things: A Pragmatic Typology of Recruitments*. Berlin: Language Science Press.
- Goodwin, C. (1981). *Conversational Organization: Interaction between Speakers and Hearers*. New York: Academic.
- Goodwin, C. (2013). The co-operative, transformative organization of human action and knowledge. *J. Pragmat.* 46, 8–23. doi: 10.1016/j.pragma.2012.09.003
- Goodwin, C. (2017). *Co-Operative Action*. Cambridge: Cambridge University Press.
- 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
- Hayashi, M. (2005). Joint turn construction through language and the body: notes on embodiment in coordinated participation in situated activities. *Semiotica* 156, 21–53.
- Helmer, H., Reineke, S., and Deppermann, A. (2016). A range of uses of negative epistemic constructions in German: ICH WEIß NICHT as a resource for dispreferred actions. *J. Pragmat.* 106, 97–114. doi: 10.1016/j.pragma.2016.06.002
- Heritage, J. (1984). “A change-of-state token and aspects of its sequential placement,” in *Structures of Social Action: Studies in Conversation Analysis*, eds J. M. Atkinson and J. Heritage (Cambridge: Cambridge University Press), 299–345. doi: 10.1017/cbo9780511665868.020
- Heritage, J. (2010). “Questioning in medicine,” in “*Why Do You Ask?*” *The Function of Questions in Institutional Discourse*, eds A. F. Freed and S. Ehrlich (New York, NY: Oxford University Press), 42–68. doi: 10.1093/acprof:oso/9780195306897.003.0003
- Heritage, J. (2012). The epistemic engine: sequence organization and territories of knowledge. *Res. Lang. Soc. Interact.* 45, 30–52. doi: 10.1080/08351813.2012.646685
- Heritage, J. (2015). Well-prefaced turns in English conversation: a conversation analytic perspective. *J. Pragmat.* 88, 88–104. doi: 10.1016/j.pragma.2015.08.008
- Heritage, J., and Clayman, S. (2010). *Talk in Action: Interactions, Identities, and Institutions*. Hoboken, NJ: Wiley-Blackwell.
- Heritage, J., and Sorjonen, M. L. (eds) (2018). *Between Turn and Sequence: Turn-Initial Particles Across Languages*. Amsterdam: Benjamins.
- Hopper, P. J., and Traugott, E. C. (2003). *Grammaticalization*, 2nd Edn. Cambridge: Cambridge University Press.
- Huang, C.-T. J. (1989). “Pro-drop in Chinese,” in *The Null Subject Parameter*, eds O. Jaeggli and K. Safir (Dordrecht: Foris), 185–224.
- Hutchby, I. (2002). Resisting the incitement to talk in child counseling: aspects of the utterance “I don’t know”. *Discourse Stud.* 4, 147–168. doi: 10.1177/14614456020040020201
- Jefferson, G. (1984). “Transcript notation,” in *Structures of Social Action: Studies in Conversation Analysis*, eds J. M. Atkinson and J. Heritage (Cambridge: Cambridge University Press), 346–369.
- Kärkkäinen, E., and Thompson, S. A. (2018). Language and bodily resources: ‘response packages’ in response to polar questions in English. *J. Pragmat.* 123, 220–238. doi: 10.1016/j.pragma.2017.05.003
- Keevallik, L. (2003). *From Interaction to Grammar: Estonian Finite Verb Forms in Conversation*. Uppsala: Uppsala University.
- Keevallik, L. (2011). “The terms of not knowing,” in *The Morality of Knowledge in Conversation*, eds T. Stivers, L. Mondada, and J. Steensig (Cambridge: Cambridge University Press), 184–206. doi: 10.1017/cbo9780511921674.009
- Keevallik, L. (2013). The interdependence of bodily demonstrations and clausal syntax. *Res. Lang. Soc. Interact.* 46, 1–21. doi: 10.1080/08351813.2013.753710
- Kendon, A. (1967). Some functions of gaze-direction in social interaction. *Acta Psychol.* 26, 22–63. doi: 10.1016/0001-6918(67)90005-4
- Kendon, A. (2004). *Gesture: Visible Action as Utterance*. Cambridge: Cambridge University Press.
- Kendrick, K. H., and Holler, J. (2017). Gaze direction signals response preference in conversation. *Res. Lang. Soc. Interact.* 50, 12–32. doi: 10.1080/08351813.2017.1262120
- Kendrick, K. H., and Torreira, F. (2015). The timing and construction of preference: a quantitative study. *Discourse Process.* 52, 255–289. doi: 10.1080/0163853x.2014.955997
- Kidwell, M. (2006). “Calm down!” the role of gaze in the interactional management of hysteria by the police. *Discourse Stud.* 8, 745–770. doi: 10.1177/1461445606069328
- Korpus DIALOG 2.0 (2020). *Korpus DIALOG The DIALOG Corpus*. Prague: Czech Language Institute.
- Lerner, G. (2003). Selecting next speaker: the context sensitive operation of a context-free organization. *Lang. Soc.* 32, 177–201. doi: 10.1017/s004740450322202x
- Levinson, S. C., and Torreira, F. (2015). Timing in turn-taking and its implications for processing models of language. *Front. Psychol.* 6:731. doi: 10.3389/fpsyg.2015.00731
- Li, X. (2014). Leaning and recipient intervening questions in Mandarin conversation. *J. Pragmat.* 67, 34–60. doi: 10.1016/j.pragma.2014.03.011
- Li, X. (2016). Some interactional uses of syntactically incomplete turns in Mandarin conversation. *Chin. Lang. Discourse* 7, 237–271. doi: 10.1075/cld.7.2.03li
- Lindström, J., and Karlsson, S. (2016). Tensions in the epistemic domain and claims of no-knowledge. A study of Swedish medical interaction. *J. Pragmat.* 106, 129–147. doi: 10.1016/j.pragma.2016.07.003
- Linell, P., Hofvendahl, J., and Lindholm, C. (2003). Multi-unit questions in institutional interactions: sequential organizations and communicative functions. *Text* 23, 539–571.
- Maschler, Y. (2012). Emergent projecting constructions: the case of Hebrew *yada* (‘know’). *Stud. Lang.* 36, 785–847. doi: 10.1075/sl.36.4.03mas
- Maschler, Y. (2017). “The emergence of Hebrew *loydea/loydat* (‘I dunno masc/fem’) from interaction: blurring the boundaries between discourse marker, pragmatic marker, and modal particle,” in *Pragmatic Markers, Discourse Markers and Modal Particles: New Perspectives*, eds A. Sansò and C. Fedriani (Amsterdam: Benjamins), 37–69. doi: 10.1075/slcs.186.02mas
- Maynard, D. W. (1997). The news delivery sequence: bad news and good news in conversational interaction. *Res. Lang. Soc. Interact.* 30, 93–130. doi: 10.1207/s15327973rlsi3002_1
- Mondada, L. (2014). The local constitution of multimodal resources for social interaction. *J. Pragmat.* 65, 137–156. doi: 10.1016/j.pragma.2014.04.004
- Mondada, L. (2018). *Conventions for Multimodal Transcription*. Available online at: https://franzoesistik.philhist.unibas.ch/fileadmin/user_upload/franzoesistik/home/Personen/Mondada/Unterordner/Mondada_conv_multimodality.pdf (accessed March 15, 2020).
- Müllerová, O. (1996). Výskyt a funkce slova *no* v českých textech prostě sdělovacího stylu. *Stylistyka* 4, 222–229.
- Oloff, F., and Havlík, M. (2018). An initial description of syntactic extensions in spoken Czech. *Pragmatics* 28, 361–390. doi: 10.1075/prag.17003.olo
- Park, I., and Kline, J. (2020). Incomplete utterances as critical assessments. *Discourse Stud.* 22, 441–459. doi: 10.1177/1461445620914669
- Pekarek Doehler, S. (2021a, forthc). How grammar grows out of social interaction. From multi-unit to single-unit question. *Open Linguistics*.
- Pekarek Doehler, S. (2016). More than an epistemic hedge: French *je sais pas* ‘I don’t know’ as a resource for the sequential organization of turns and actions. *J. Pragmat.* 106, 148–162.
- Pekarek Doehler, S. (2019). At the interface of grammar and the body. *Chais pas* (‘dunno’) as a resource for dealing with lack of recipient response. *Res. Lang. Soc. Interact.* 52, 365–387. doi: 10.1080/08351813.2019.1657276
- Pekarek Doehler, S. (2021b, forthc). Word-order affects response latency: action projection and the timing of responses to question-word questions. *Discourse Process.* 58, 328–352. doi: 10.1080/0163853x.2020.1824443
- Pekarek Doehler, S., and Deppermann, A. (2021). Longitudinal CA: Introduction to the Special Issue. *Res. Lang. Soc. Interact.* 54, 127–141. doi: 10.1080/08351813.2021.1899707
- Persson, R. (2020). Taking issue with a question while answering it: prefatory particles and multiple sayings of polar response tokens in French. *Res. Lang. Soc. Interact.* 53, 380–403. doi: 10.1080/08351813.2020.1786977
- Pomerantz, A. (1984). “Agreeing and disagreeing with assessments: some features of preferred/dispreferred turn shapes,” in *Structures of Social Action*, eds J. M. Atkinson and J. Heritage (Cambridge: Cambridge University Press), 57–111. doi: 10.1017/cbo9780511665868.008
- Pomerantz, A. (1986). Extreme case formulations: a way of legitimizing claims. *Hum. Stud.* 9, 219–229. doi: 10.1007/bf00148128
- Pomerantz, A., and Heritage, J. (2013). “Preference,” in *Handbook of Conversation Analysis*, eds J. Sidnell and T. Stivers (Cambridge: Cambridge University Press), 210–228.

- Raymond, G. (2003). Grammar and social organization: yes/no type interrogatives and the structure of responding. *Am. Soc. Rev.* 68, 939–967. doi: 10.2307/1519752
- Roberts, F., Francis, A. L., and Morgan, M. (2006). The interaction of inter-turn silence with prosodic cues in listener perceptions of “trouble” in conversation. *Speech Commun.* 48, 1079–1093. doi: 10.1016/j.specom.2006.02.001
- Roberts, F., Margutti, P., and Takano, S. (2011). Judgments concerning the valence of inter-turn silence across speakers of American English, Italian, and Japanese. *Discourse Process.* 48, 331–354. doi: 10.1080/0163853X.2011.558002
- Robinson, J. D. (2020). One type of polar, information-seeking question and its stance of probability: implications for the preference for agreement. *Res. Lang. Soc. Interact.* 53, 425–442. doi: 10.1080/08351813.2020.1826759
- Rossano, F. (2012). *Gaze Behavior in Face-to-Face Interaction*. Ph.D. Dissertation. Nijmegen: Max Planck Institute for Psycholinguistics Series.
- Rossano, F., Brown, P., and Levinson, S. C. (2009). “Gaze, questioning, and culture,” in *Conversation Analysis: Comparative Perspectives*, ed. J. Sidnell (Cambridge: Cambridge University Press), 197–249.
- Sachs, H. (1987). “On the preferences for agreement and contiguity in sequences in conversation,” in *Talk and Social Organisation*, eds G. Button and J. R. E. Lee (Clevedon: Multilingual Matters), 54–69.
- Sacks, H. (1992). *Lectures on Conversation*. Oxford: Blackwell.
- Schegloff, E. A. (1987). “Recycled turn beginnings: a precise repair mechanism in conversation’s turn-taking organization,” in *Talk and Social Organisation*, eds G. Button and J. R. E. Lee (Clevedon: Multilingual Matters), 70–85.
- Schegloff, E. A. (1988). On an actual virtual servo-mechanism for guessing bad news: a single case conjecture. *Soc. Probl.* 35, 442–457. doi: 10.1525/sp.1988.35.4.03a00080
- Schegloff, E. A. (1996). “Turn organization: one intersection of grammar and interaction,” in *Interaction and Grammar*, eds E. Ochs, E. Schegloff, and S. A. Thompson (Cambridge: Cambridge University Press), 52–133. doi: 10.1017/cbo9780511620874.002
- Schegloff, E. A. (2007). *Sequence Organization in Interaction: A Primer in Conversation Analysis*. Cambridge: Cambridge University Press.
- Scheibman, J. (2000). *I dunno*: a usage-based account of the phonological reduction of *don’t* in American English conversation. *J. Pragmat.* 32, 105–124. doi: 10.1016/S0378-2166(99)00032-6
- Selting, M., Auer, P., Barth-Weingarten, D., Bergmann, J., Bergmann, P., and Birkner, K. (2011). A system for transcribing talk-in-interaction: GAT 2 translated and adapted for English by Elizabeth Couper-Kuhlen and Dagmar Barth-Weingarten. *Gesprächsforschung* 12, 1–51. doi: 10.1075/slsi.30.01sor
- Sidnell, J. (ed.) (2009). *Conversation Analysis: Comparative Perspectives*. New York: Cambridge University Press.
- Sidnell, J. (ed.) (2011). *Conversation Analysis*. Cambridge: Cambridge University Press.
- Skogmyr Marian, K., and Pekarek Doehler, S. (2021, forthc). Multimodal trajectories for indexing cognitive search: a longitudinal L2 study. *Soc. Interact. Video Based Stud. Hum. Social.*
- Stivers, T. (2010). An overview of the question–response system in American English conversation. *J. Pragmat.* 42, 2272–2281.
- Stivers, T., and Robinson, J. D. (2006). A preference for progressivity in interaction. *Lang. Soc.* 35, 367–392.
- Stivers, T., Enfield, N. J., Brown, P., Englert, C., Hayashi, M., Heinemann, T., et al. (2009). Universals and cultural variation in turn-taking in conversation. *Proc. Natl. Acad. Sci. U.S.A.* 106, 10587–10592. doi: 10.1073/pnas.0903616106
- Stivers, T., Mondada, L., and Steensig, J. (eds) (2011). *The Morality of Knowledge in Conversation*. Cambridge: Cambridge University Press.
- Streeck, J. (2009). *Gesturecraft. The Manu-Facture of Meaning*. Amsterdam: Benjamins.
- Thompson, S. A., Fox, B. A., and Couper-Kuhlen, E. (2015). *Grammar in Everyday Talk: Building Responsive Actions*. Cambridge: Cambridge University Press.
- Tsui, A. B. M. (1991). The pragmatic functions of *I don’t know*. *Text* 11, 607–622.
- Vatanen, A., Endo, T., and Yokomori, D. (2020). Cross-linguistic investigation of projection in overlapping agreements to assertions: stance-taking as a resource for projection. *Discourse Process.* 58, 1–20. doi: 10.1080/0163853X.2020.1801317
- Weatherall, A. (2011). *I don’t know* as a prepositioned epistemic hedge. *Res. Lang. Soc. Interact.* 44, 317–337. doi: 10.1080/08351813.2011.619310
- Whitehead, K. A. (2015). Everyday antiracism in action: preference organization in responses to racism. *J. Lang. Soc. Psychol.* 34, 373–388.

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: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Copyright © 2021 Pekarek Doehler, Polak-Yitzhaki, Li, Stoenica, Havlik and Keevallik. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.