Beyond *yes* and *no*: Multimodal multilexical polar responses in German[1]

Cornelia LOOS — *University of Hamburg*
Sophie REPP — *University of Cologne*

Abstract. Response particles like *yes* and *no* fulfil two functions: they may affirm or reject a previous utterance, or they may indicate the response polarity. Languages differ with regard to the function that is given priority, which is reflected in the use of particles in responses to negative antecedents. Quantitative evidence shows that priorities are gradient and that there also is interindividual variation. Response strategies other than particle responses are less well studied but it is well known that languages use for instance echoing strategies, truth-related verbal and adjectival predicates, non-vocal gestural strategies, as well as sequential and simultaneous combinations of these. We present results from an open production experiment looking at the entire range of multilexical and multimodal response strategies in German. German uses a wide range of response elements (RE) including the ambiguous particles *ja* 'yes' and *nein* 'no', unambiguous RE like the adverb *genau* 'exactly' and the particle *doch* 'rejecting a negative antecedent', as well as stand-alone head nods and headshakes. We show inter alia that particles are preferred over non-particles, that stand-alone head movements are ambiguous like *ja* and *nein*, that co-speech headshakes may be used to help disambiguation whereas this is not the case for co-speech head nods, and that sequential combinations in principle are such that a second RE may disambiguate the first, although they typically are collocations that are used independently of actual ambiguity avoidance. We discuss the results in the context of models of response elements making use of presuppositional polarity features.

Keywords: response particle, polar response, German, co-speech gesture, head movements.

1. Introduction

Response particles like German *ja* 'yes' and *nein* 'no' can fulfill two functions when responding to an assertion. On the one hand, they affirm (*ja*) or reject (*nein*) the truth of the antecedent assertion. On the other hand, they signal the polarity of the response clause as either positive (*ja*) or negative (*nein*). The two functions coincide in responses to positive assertions:

(1) *Judge*: *Herr Beller, Sie waren während des Banküberfalls zu Hause.* antecedent
'Mr. Beller, you were at home during the bank robbery.'
Mr. Beller:
a. *Ja, war ich*. 'Yes, I was.' affirmation, positive polarity response
b. *Nein, war ich nicht*. 'No, I wasn't.' rejection, negative polarity response

In responses to negative assertions, this changes: In (2), Mr. Beller would affirm the negative assertion uttered by the judge with a negative response clause, preceded either by *ja* or *nein*

[1] This work was funded by the German Science Foundation DFG, project RE 2606/2-2 *Affirmative and rejective responses to negative assertions and questions*. We thank Franziska Busche and Sven Weber for acting in the video stimuli (narrator and dialogue partner, resp.), Franziska Busche for editing the video stimuli, Lukas Kurzeja and Nele Schillberg for helping with the recordings, and Isabeau Lettow, Lilli Kivelitz and Klara Voerkel for the annotation of the data in ELAN.

(2a). He would reject it with a positive response clause preceded by the particle *doch* (2b), which is the dedicated particle for rejecting negative assertions in German. The availability of *ja* and *nein* in the affirmation in (2a) shows that either particle can in principle be used to encode the intended meaning.

(2) *Judge*: *Herr Beller, Sie waren während des Banküberfalls* ***nicht*** *zu Hause.*
'Mr. Beller, you were **not** at home during the bank robbery.'
Mr. Beller:
a. *Ja/Nein, war ich nicht*. 'Yes, I wasn't.' *ja* = affirmation, *nein* = negative polarity
b. *Doch, war ich*. 'Doch, I was.' *doch* = rejection + positive polarity

Across languages, we find different preferences for the use of response particles to encode affirmation/rejection (traditionally called *truth-based languages*) vs. positive/negative polarity (*polarity-based languages*) (Pope, 1976; Jones, 1999). Like German, most languages do not follow one strategy exclusively but exhibit a graded preference for one response strategy over another (Roelofsen and Farkas, 2015). Experimental evidence from different languages confirms that response preferences more often than not are gradient and also exhibit substantial inter-individual variation, to the extent that speakers of the same language show opposite preferences (e.g., Brasoveanu et al. 2013; González-Fuente et al. 2015; Meijer et al. 2015; Goodhue and Wagner, 2015, 2018 Li et al. 2016; Claus et al.2017; Repp et al. 2019; Loos et al.2020, 2024; Geist and Repp, 2023).

Response particles are not the only conventionalized means to realize a polar response. Some languages regularly use the repetition of the lexical verb of the antecedent clause, i.e., a so-called *echoic verb* (e.g., Jones, 1999; González-Fuente et al. 2015) or an auxiliary (e.g., Finnish, see Holmberg, 2001). There are also gestural means to express a polar response, which can be used on their own, like head movements; head movements they can also serve as co-speech gestures (González-Fuente et al.2015; Li et al.2016; Esipova, 2021). Following Loos, Steinbach and Repp (2024), we are using the term *response element (RE*) here to refer to any vocal or gestural element that is used as a stand-alone polar response to assertions or questions. There are sequential combinations of RE, although there has only been a targeted experimental investigation of such combinations for German Sign Language (DGS) (Loos et al.2020, 2024).

In this paper, we present evidence from a study that looked at the entire range of multilexical and multimodal response strategies in German: particle responses, non-particle lexical RE, non-lexical vocalizations (like English *uh-uh*), stand-alone gestural RE (e.g., head movements), as well as sequential combinations of RE (both vocal and gestural), and co-speech gesture. We are focusing here on individual RE, sequential RE combinations and co-speech head movements without reporting on other co-speech gestures. The data that we present are from an open production study that used a discourse-completion task in a dialogue setting.

The goal of the paper is two-fold. On the one hand, we describe the preference patterns for German RE (combinations) as they arose in our experimental investigation. On the other hand, we discuss our data in relation to a prominent model of response particles, the so-called *feature model* (Roelofsen and Farkas, 2015; Farkas and Roelofsen, 2019), which was recently extended to non-particle response strategies in DGS by Loos et al. (2024). This discussion also allows us a close comparison of the response strategies employed in German vs. in DGS.
The paper is organized as follows. Section 2 introduces the feature model as it was developed for particle responses to capture the gradient preference patterns mentioned at the beginning of

this section, illustrated for German. Section 3 discusses non-particle responses and presents the adapted feature model, as it was proposed for DGS. The section also discusses the issue of sequential and simultaneous RE combinations in DGS. Section 4 presents the experiment that we conducted and discusses the results. Section 5 summarizes and offers a more general discussion with a view to theoretical modelling.

2. The feature model for particle responses

Particle responses have been analyzed in a range of theoretical frameworks, including anaphora accounts (Krifka, 2013; Roelofsen and Farkas, 2015) and ellipsis accounts (Kramer and Rawlins, 2011; Holmberg, 2013, 2015). For the purposes of this paper, we are adopting the *feature model* developed in Roelofsen and Farkas (2015) and Farkas and Roelofsen (2019), because it captures the intuitive difference between the two functions that response particles have (to signal truth vs. polarity), while at the same time accounting for the gradient preferences for the use of the particles in many languages. The feature model treats response particles as propositional anaphors, which can be used when certain presuppositions are met. The presuppositions are assumed to be triggered by presuppositional polarity features, which are hosted by a polarity head (Pol) in the left periphery of the clause. Pol selects as its complement the so-called *prejacent* (essentially the response clause), which may also be omitted. There are two polarity features in Pol, which differ in type. Which feature is present for each type depends on the discourse in which the response is uttered. *Absolute polarity features* presuppose that the polarity of the proposition denoted by the prejacent is positive or negative. So the absolute feature in the Pol head is either [+] or [–]. *Relative polarity features* presuppose that there is a unique salient proposition in the preceding discourse (i.e., the antecedent) whose polarity is identical with, or the opposite of, the polarity of the prejacent. So the relative feature in the Pol head is either [AGREE] or [REVERSE]. Combinatorics delivers four possible combinations of relative and absolute polarity features in the Pol head. For instance, [REVERSE,+] is the feature combination that occurs in the Pol head of responses that reject a negative antecedent with a positive response clause.

Cross-linguistic variation and gradient preferences in the use of response particles result from different feature mappings and different realization preferences that are modeled with (linear) optimality theoretic constraints (Keller, 2000). For instance, a language may map only [+] onto its YES-type particle, while another maps [AGREE] and [+] onto its YES particle. When there is a choice between several particles – as is often the case after negative antecedents because of 'multifunctional' mapping – optimality-theoretic constraints regulate preferences.

We illustrate the interaction of feature mapping and constraint ranking for German as it has been proposed to account for acceptability judgements reported in Claus et al. (2017) (Farkas and Roelofsen, 2019). (3) shows which features map onto which German response particles and which optimality-theoretic constraint rankings apply. The definition of the constraints is given in (4). For ambiguous response contexts, i.e., when responding to a negative assertion, Claus et al. report that the majority of German speakers prefer *ja* over *nein* in affirmations (discourse type [AGREE, –]), and *doch* over *nein* in rejections (discourse type [REVERSE,+]). *Doch* is unacceptable in affirmations, and *ja* in rejections. These findings can be captured as follows. The constraint EXPRESSIVENESS prioritizes the realization of both presuppositional features in Pol by only one particle. The high weight of this constraint explains why *doch* is preferred over *nein* in rejections: *doch* encodes both [REVERSE] and [+] whereas *nein* only

encodes [REVERSE]. The constraint MAXIMIZE RELATIVE prioritizes the realization of relative polarity features over the realization of absolute features. So when a relative and an absolute feature in Pol vie for expression, as is the case in affirmations of negative assertions (discourse [AGREE, –]), the relative feature [AGREE] is encoded preferentially. This explains the preference for *ja* in such discourses. Lastly, MAXIMIZE MARKED plays a role in rejections of negative assertions (discourse [REVERSE, +]). Here, *nein* (which signals the marked discourse move of rejection by realizing [REVERSE]) is acceptable, even if it is dispreferred in comparison to *doch*, whereas *ja* (which would realize unmarked positive polarity) is not acceptable.

(3) *German: Feature mapping* (Roelofsen and Farkas, 2015) *and ranking* (Claus et al. 2017)

Mapping	[+] and [AGREE]	→ *ja*
	[–] and [REVERSE]	→ *nein*
	[+, REVERSE]	→ *doch*
Ranking	EXPRESSIVENESS >> MAXIMIZE RELATIVE >> MAXIMIZE MARKED	

(4) *Optimality-theoretic constraints in the feature model* (Farkas and Roelofsen, 2019)

EXPRESSIVENESS:	Maximize the expression of feature content
MAXIMIZE RELATIVE:	Maximize the realization of relative polarity features
MAXIMIZE MARKED:	Maximize the realization of marked polarity features or feature combinations

In its original version, the feature model was intended to account for particle responses. As we laid out in the introduction, this is only a part of the response system of a language. There are also non-particle responses and vocalizations, as well as gestural responses like head nods and headshakes. Neither does the feature model account for RE combinations. In the next section, we discuss an extension of the feature model that was suggested by Loos, Steinbach, and Repp (2020, 2024) for DGS to also account for non-particle and non-manual responses. DGS has conventionalized the use of both lexical and non-manual RE.

3. Beyond (single) particle responses: Evidence from DGS

3.1. Non-particle responses

Similar to German and other spoken languages, DGS uses lexical elements corresponding to English *yes* and *no*: the manual response particles JA 'yes' and NEIN 'no'. In addition, DGS signers use a range of other RE. There are what we might call *response verbs*, i.e. verbal predicates that directly comment on the truth of the antecedent, e.g., STIMMT-neg '(is) not-right'; there also are response adjectives, e.g., FALSCH 'false'. Furthermore, signers use stand-alone head nods and headshakes, as well as mouthings adapted from German like /doch/. Different RE may be combined simultaneously (e.g., JA + head nod) or sequentially (e.g., JA STIMMT 'yes, (that's) right'); responses without a RE also occur. Loos et al. (2024) show that the realization preferences for ambiguous *yes/no*-type RE in DGS are the same across manual and non-manual RE: JA, NEIN, head nod and headshake can express relative and absolute polarity features but preferentially realize the relative polarity feature, i.e. affirm or reject an antecedent. Loos et al. argue that head nods and headshakes are bona fide RE in DGS, whose distribution has to be accounted for in a theory of polar responses. To accommodate non-manual and non-

particle responses, Loos et al. extend the feature model by the constraints MAXIMIZE MANUAL, DIRECT MAPPING, and AVOID AMBIGUITY, defined in (5). The definitions of MAXIMIZE MANUAL and AVOID AMBIGUITY are self-explanatory. DIRECT MAPPING prioritizes particle responses over non-particle responses: Particles are direct realizations of the features in the Pol head because particles are merged in the syntactic structure in Pol (vocabulary insertion). For other RE, the lexical semantics must match the discourse type encoded by the polarity features under Pol.

(5) MAXIMIZE MANUAL: Maximize feature realization via manual signs
DIRECT MAPPING: Maximize realizing Pol head features directly
AVOID AMBIGUITY: Maximize feature encoding via unambiguous RE (combinations)

Table 1 shows the OT tableau that Loos et al. propose to capture the distribution of the clause-initial, i.e., first RE in their data set. Like Farkas and Roelofsen (2019), they assume that constraints are weighted, that non-optimal candidates can be acceptable, and that violations accumulate. Turning to the constraints and their relative weight, the high weight of MAXIMIZE MANUAL captures the finding that head movements do occur as stand-alone RE but are far outnumbered by manual RE. DIRECT MAPPING captures the finding that response particles are overall preferred over response verbs or response adjectives, which cannot express the polarity features in the Pol head by direct vocabulary insertion. AVOID AMBIGUITY accounts for the observation that in ambiguous response contexts with negative antecedents STIMMT 'is.right' is preferred over JA: Loos et al. argue that STIMMT, like JA, is a response particle, but since STIMMT (unlike JA) is unambiguous – it only realizes [AGREE] – STIMMT is preferred. Since STIMMT-neg is not a particle, the greater weight of DIRECT MAPPING than AVOID AMBIGUITY accounts for the preference of NEIN over STIMMT-neg in rejections of negative antecedents. As mentioned above, DGS signers prefer truth-based responses over polarity-based ones, which is due to MAXIMIZE RELATIVE having greater weight than MAXIMIZE MARKED: the realization of unmarked [AGREE] is preferred over the realization of marked [–] in affirmations of negative antecedents.

Table 1. OT tableau: Feature mapping and OT constraint weighting for the three most frequent RE in DGS (adapted from Loos et al. 2024: 454).

	Discourse	Candidate	Features realized in discourse	MAXIMIZE MANUAL	DIRECT MAPPING	AVOID AMBIGUITY	MAXIMIZE RELATIVE	MAXIMIZE MARKED
positive antecedent	[AGREE, +]	**JA**	[AGREE] [+]					
		STIMMT	[AGREE]					
		head nod	[AGREE] [+]	*				
	[REVERSE, –]	**NEIN**	[REVERSE] [–]					
		STIMMT-neg	[REVERSE]		*			
		headshake	[REVERSE] [–]	*				
negative antecedent	[AGREE, –]	**STIMMT**	[AGREE]					*
		JA	[AGREE]			*		*
		NEIN	[–]			*	*	
	[REVERSE, +]	**NEIN**	[REVERSE]			*		
		STIMMT-neg	[REVERSE]		*			
		JA	[+]			*	*	

3.2. RE combinations

We already mentioned that a theory of polar responses should also account for RE combinations because both spoken and signed languages can combine RE simultaneously and sequentially. For DGS, Loos et al. (2024) show that head movements accompanying a manual RE are generally isofunctional with that RE, i.e. JA occurs almost exclusively with head nods and NEIN with headshakes. This is different from co-speech head movements, which do not necessarily align in function with a vocal response particle. In experimental data from Mandarin Chinese (Li et al. 2016), Catalan, and Russian (González-Fuente et al. 2015), more than half of the rejecting responses to a negative antecedent were accompanied by a head nod (marking [+]). Notably, at least in Mandarin Chinese and Russian, the vocal RE that was accompanied by the head nod encoded [REVERSE]: a *no*-type particle in Mandarin Chinese, an echoic verb in Russian. Vocal and co-speech gestural response elements may thus take on different functions in spoken languages and thus may potentially be used for disambiguation.

Looking at sequential combinations, both signed and spoken languages allow reduplicating a given particle (e.g. *sí, sí* 'yes, yes' in Catalan), combining two particles (e.g. JA STIMMT 'yes, (that's) right' in DGS), combining a particle and a response verb or adjective (e.g. NEIN, FALSCH 'no, wrong' in DGS), or combining a particle with an echoic verb (e.g. Russian *net, prikhodil* 'no, he has come', González-Fuente et al. 2015). In principle, then, sequential combinations provide room for disambiguating potentially ambiguous responses. We know that in DGS, this potential remains untapped: The most frequent RE combinations involving a potentially ambiguous JA or NEIN are collocations (JA STIMMT 'yes, (that's) right', NEIN STIMMT-neg 'no, (that's) not right') that occur equally often in ambiguous and unambiguous response contexts. Still, the existence of these collocations might be taken as evidence for the desire to express unambiguous agreement or disagreement across discourse types.

4. Multilexical and multimodal response strategies in German: Experiment

In the present paper, we explore what RE (combinations) German speakers use in open production. The goal of our study is to arrive at an overview of the multilexical and multimodal response strategies employed in German, and to address several issues that arise from previous observations for other languages. We explored whether speakers use specific RE – such as unambiguous response verbs – and/or RE combinations, to counter the ambiguity problem in responses to negative antecedents. Recall that for DGS, Loos et al. (2024) observed that this was only partly the case. Furthermore, we investigated the combination of verbal RE with co-speech head movements exploring inter alia the issue of isofunctionality. As laid out above, previous research has shown that DGS and the spoken languages that to date have been investigated differ in the occurrence of non-isofunctional simultaneous combinations, specifically in the use of the head nod. It is an interesting question if the ambient language German patterns with DGS, which has conventionalized head movements as RE, or with other spoken languages. We also looked at head movements beyond nods and shakes, to open the window more widely into the world of multimodal responses, which for reasons of space we nevertheless can only touch upon here. From a theoretical perspective, we discuss some of our findings in view of the feature model but highlight that a multivariate response (system) is subject to a complex set of constraints that must be captured in a more complex probabilistic model that is to be developed in future research.

4.1. Method

To address our research questions, we conducted an open production experiment, where participants engaged in a pseudo-dialogue with a video-recorded interlocutor. Similar to Loos et al. (2024), we expected that open production encourages participants to use a wide range of response strategies even in a laboratory setting. The chosen method was intended to strike a balance between ecological validity and the availability of a balanced data set that contained multiple data points for different discourse types for individual speakers. The participants' task in the experiment was dialogue completion. They gave a visual-oral response to positive and negative assertions in a series of dialogues with a pre-video-recorded speaker named Peter, while being video-recorded themselves. The participants took the role of Peter's interlocutor, named Alex (= gender-neutral name). Each dialogue between Peter and Alex was preceded by a short scene-setting passage, told by a pre-video-recorded narrator, who described an episode in Peter and Alex's lives. The dialogue was about this episode, specifically about whether some action had been carried out by a third person. Both the narration and Peter's utterance were in colloquial style with natural co-speech gesture and some props for the actual dialogue scene (coffee mugs, paper folders, a flower pot, scarfs, hats etc.), to create a relaxed atmosphere. (6) is a sample experimental item. The experiment had a 2×2 design with the factors ANTECEDENT POLARITY (positive/negative) and SPEECH ACT (affirmation/rejection). ANTECEDENT POLARITY was implemented by manipulating the sentence polarity of Peter's statement, which Alex affirmed or rejected in their target utterance. The choice of SPEECH ACT was guided by information about the true state of affairs (SoA) in the scene-setting passage: When the SoA was DONE, the participants were expected to affirm a positive antecedent and reject a negative antecedent, respectively. When the SoA was NOT DONE, the expectation was the opposite.

(6) Narrator: *Peter und Alex woll'n sich die vier Jahreszeiten als Bilder an die Wand hängen. Sie ham für die Gemälde 'ne Künstlerin engagiert. Und mit der hat Alex heute Vormittag telefoniert. Sie sagt…,* 'Peter and Alex want to put paintings of the four seasons on their wall. They got an artist to do the pictures for them. Alex talked with her this morning. She says …'

… *Frühling und Sommergemälde sind fertig; morgen fängt sie mit dem Herbst an.* '…spring and summer are almost done, and next week she'll get started on the fall painting.' — SOA DONE

… *Herbst und Winter sind fast fertig und nächste Woche fängt sie mit dem Sommergemälde an.* '…fall and winter are almost done, and next week she'll get started on the summer painting.' — SOA NOT DONE

Nachmittags trinken Peter und Alex Kaffee und unterhalten sich über die Künstlerin und ihre Bilder. 'In the afternoon, Peter and Alex are having coffee, and they are talking about the artist and her paintings.'

Peter: *Du, die Künstlerin hat das Sommergemälde schon gemalt* — POSITIVE ANTECEDENT
you the artist.FEM has the summer painting already painted
'Hey, the artist has already painted the summer painting.'

Du, die Künstlerin hat das Sommergemälde noch nicht gemalt — NEGATIVE ANTECEDENT
you the artist.FEM has the summer painting not yet painted
'Hey, the artist hasn't painted the summer painting yet.

There were 48 lexicalizations, i.e., episodes. Antecedent polarity was a between-item factor: In half of the lexicalizations, Peter made a positive statement, in the other half he made a negative statement (e.g., (6) was only presented with a negative antecedent). The items were distributed over two experimental lists so that each participant saw each lexicalization only once: either in the version eliciting an affirmation, or in the version eliciting a rejection.

28 native speakers of German (14f/14m) took part in the experiment, which was conducted at the XLinC Laboratory at the University of Cologne. Recordings were conducted in a quiet room. Participants were standing in front of a large video screen: the persons visible on the screen (narrator, Peter), were of the same size as the participants; the height of the screen was adjusted to the participants' height so that the participant would face the person on screen at eye level. Before the experiment started, participants were informed about their task by a video recording of the first author. They were instructed to react truthfully to Peter's statement on the basis of what the narrator had said, and were given a few sample responses with different response elements and a response clause, which were accompanied by co-speech gesture in the video recording. In each trial, the narrator appeared first, introducing the episode. Then Peter appeared, uttered the statement, at the end of which the screen froze. Participants responded to Peter's statement with a response of their choice. Then they clicked a button to move on to the next episode. Participants moved through the experiment at their own pace.

4.2. Results and discussion

The dataset consists of 1344 recordings. 4.9% had to be discarded because of a programming error (28 items, 2%) or because participants gave the wrong answer as evidenced by the response clause they uttered (39 items, 2.9%). This left 1277 observations for analysis. All data were annotated in ELAN (Wittenburg et al. 2006). We transcribed the entire response and annotated all vocal and gestural RE as well as the co-speech gestures accompanying the RE. The gestures came from nine articulators (head, hands, brows, eyes, mouth, shoulders, feet, torso, gaze). In this paper we only consider head movements.

In the statistical analysis, we fitted generalized mixed effects models as well as multinomial models, depending on the particular analysis. In this paper, we are focusing on utterance-initial single RE and RE combinations, and the head movements accompanying them.

4.3. The first response element (RE1)

1119 responses contained at least one RE (88%), see Table 2 for the mean proportion by condition. The statistical analysis revealed that there are more RE1 in affirmations than in rejections (b = -0.81, SE = 0.28, z = -2.9, $p < 0.01$), and that this difference is larger in responses to positive antecedents (b = 0.75, SE = 0.22, z = 3.5, $p < 0.001$).

Table 2. Mean proportion of responses with at least one RE by condition.

Antecedent	Speech Act	Mean proportion	SD
positive	affirmation	0.96	0.20
	rejection	0.77	0.42
negative	affirmation	0.91	0.28
	rejection	0.85	0.35

As Table 2 shows, rejections of positive antecedents without a RE are quite frequent (23%). It seems that participants avoid a straight *nein* 'no' in this discourse type. Instead, they use hedging devices and/or motivate their rejection, as illustrated in (7), which is the English translation of a sample response from the dataset.

(7) Peter: Malte has already completed his English homework.
Alex: He told me something different.

The frequent avoidance of rejection-signaling RE suggests that speakers aim to soften face-threatening speech acts. In his analysis of particle responses, Krifka (2013) already suggested that a constraint he termed *DISAGR is relevant for the choice of a particle. *DISAGR penalizes disagreement with the interlocutor. The avoidance of a straight rejection and the mentioning of some conflicting evidence, as in (7), instead, helps the speaker avoid conflict. The avoidance of conflict is a general communicative principle, as is also highlighted by Farkas and Bruce (2010) in the discussion of their Table model. We will see more evidence for speakers following this principle further below.

1066 RE1 are vocal RE. Figure 1 shows the distribution of vocal and non-vocal REs in first position (normalized for lexical expression, i.e., variants like *nee* and *nein* (both 'no') are pooled as *NEIN*).

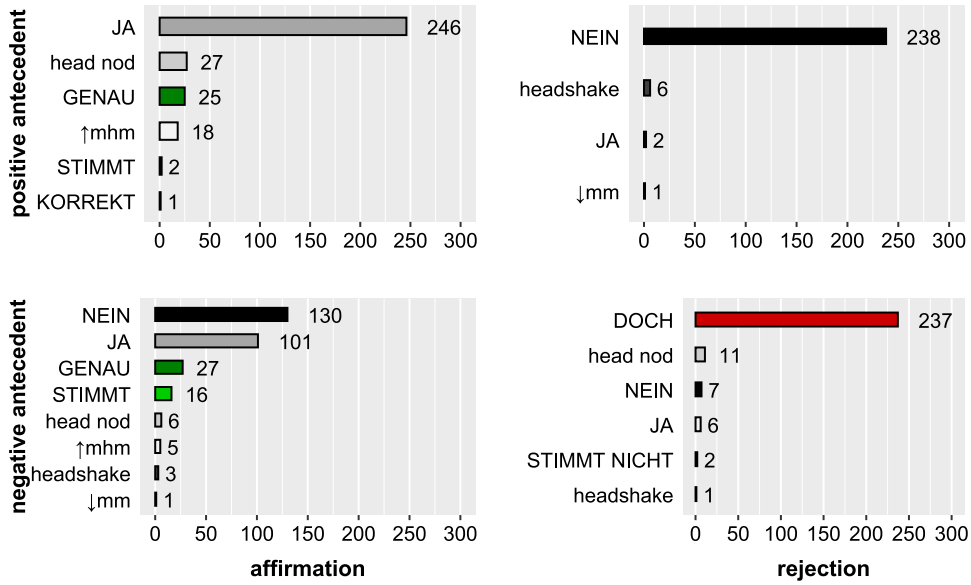

Figure 1. Distribution of vocal and non-vocal RE1 by condition. White, grey and black shades mark ambiguous RE; Green and red shades mark unambiguous RE that only occur in affirmations (green) or only in rejections (red).

The RE ↑mhm is a bisyllabic bilabial nasal vocalization with nasal [h] in the onset of the second syllable, which has a higher pitch than the first syllable. The RE ↓mm is a bisyllabic bilabial nasal vocalization with a glottal stop in the onset of the second syllable, which has a lower pitch than the first syllable.

As expected, *ja* 'yes' and *nein* 'no' are ambiguous. In affirmations of negative antecedents *nein* is more frequent than *ja* in this sample of the population (contrary to samples reported in

previous research, see Section 2). This suggests that the constraint MAXIMIZE MARKED has a high weight in this population sample: After negative antecedents, *nein* [REVERSE, –] realizes marked [–] whereas *ja* [AGREE, +] realizes unmarked [AGREE]. The stand-alone head nod and headshake are also ambiguous as they occur both in affirmations and in rejections of negative antecedents. Their frequency is too low to draw any conclusions regarding preferences. The low frequency of stand-alone non-vocal RE suggests a high ranking of a constraint that parallels the constraint MAXIMIZE MANUAL in DGS. Recall that in DGS manual RE were more frequent than non-manuals, suggesting that the main channel for conveying lexical information is prioritized. We suggest that in German, there is a spoken-language equivalent: MAXIMIZE VOCAL. Also recall, however, that for DGS it was possible to determine the preference patterns for stand-alone head movements, as these were more frequent than in German, which suggests a different conventionalization status for head RE in DGS.

Unambiguous agreeing RE, which realize [AGREE], are the adverb *genau* 'precisely' (a response adverb), the finite verb *stimmt* 'is correct' (a response verb), and the vocalization ↑mhm. The adverb *genau* is the most frequent of these RE. Unambiguous rejecting RE are the particle *doch*, which predominates, and the negative response verb *stimmt nicht* 'is not correct'. The particle *doch* realizes [REVERSE,+]: it does not occur after positive antecedents. *Stimmt nicht* is not restricted in this way, as we have verified by additional corpus searches. Thus, *stimmt nicht* realizes [REVERSE]. The vocalization ↓mm in this sample only realizes [–], i.e. it does not occur in rejections of negative antecedents. This may simply be due to the overall low number of non-lexical RE or to the great dominance of *doch*. This dominance suggests a high weight of EXPRESSIVENESS, as is expected for German (Claus et al. 2017).

Overall, the frequency distribution of particle and non-particle responses suggests that the constraint DIRECT MAPPING has great weight in German (as it does in DGS): particle responses are much more frequent than response verbs or adverbs, i.e., than RE that do not directly realize a feature in the Pol head.

4.4. Co-speech head movements simultaneous with RE1

Figure *2* shows the distribution of co-speech head movements accompanying a vocal RE1 (i.e., head movements that were not stand-alone gestures) ($n = 1065$). The majority of the vocal RE1 came with a head movement ($n = 811$, 76.2%). There was an interaction of the two experimental factors ($b = 0.39, 0.08$, $z = 4.85$, $p < 0.001$). After positive antecedents, there were more co-speech head movements in affirmations (86.3%) than in rejections (68.9%). After negative antecedents, it was the other way round (70.0% vs. 78.2%). Regarding head movement types, Figure *2* indicates that their overall frequency and distribution by condition differed, also see Table 3. Head nods were most frequent, outnumbering utterances with a vocal RE1 without accompanying head movement. Right and left turns (turns 'away') were the second-most frequent head movement, closely followed by forward pushes and forward and backward tilts.

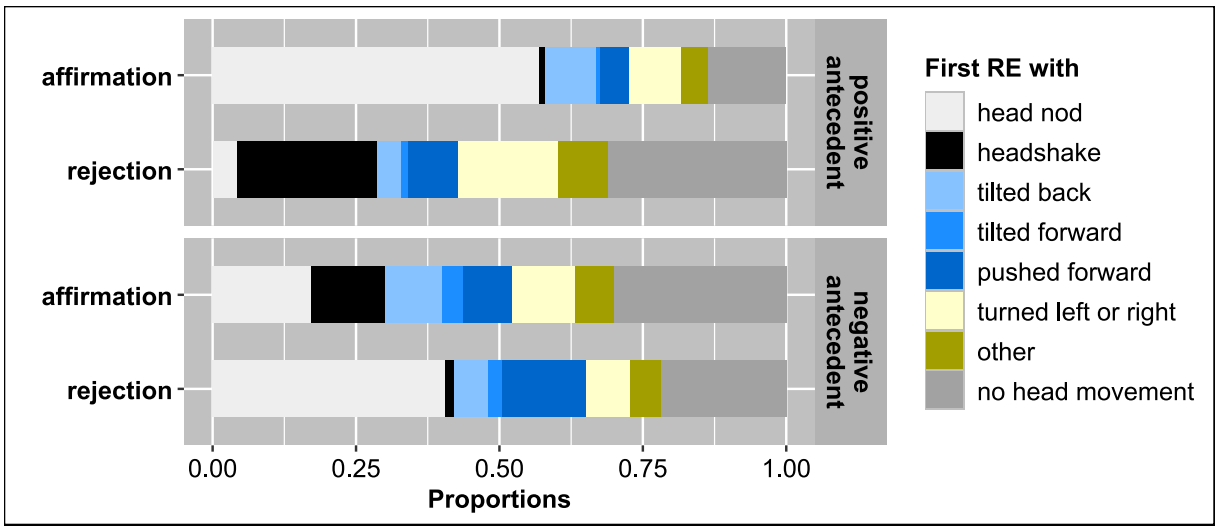

Figure 2. Co-speech head movements on RE1 in utterances with at least one RE (n = 1065). Proportions by condition.

Table 3. Distribution of head movement types: Counts and proportions by condition.

Head movement	**Total *n***	**Speech act**	**Positive antecedent**		**Negative antecedent**	
			n	proportion within condition	*n*	proportion within condition
no head movement	*254*	*affirmation*	*40*	*13.7*	*84*	*30.0*
		rejection	*75*	*31.1*	*55*	*21.8*
head nod	326	affirmation	166	56.8	48	17.1
		rejection	10	4.1	102	40.5
headshake	102	affirmation	3	1.0	36	12.9
		rejection	59	24.5	4	1.6
forward/backward tilt	97	affirmation	28	9.6	38	13.6
		rejection	13	5.4	21	8.3
pushed forward	100	affirmation	15	5.1	24	8.6
		rejection	21	8.7	37	14.7
turned left or right	118	affirmation	26	8.9	31	11.1
		rejection	42	17.4	19	7.5
other	68	affirmation	14	4.8	19	6.8
		rejection	21	8.7	14	5.6

We analyzed the effects of the experimental factors on the distribution of head movement type by fitting a multinomial mixed effects model (R package `mclogit`, Elff 2024), with 'no movement' as the baseline. Recall that the presence (or absence) of a head movement showed an interaction effect itself, with fewer head movements in responses with negative polarity.

The analysis revealed significant effects for several of the head movement types, see Table 4. For head nods, there was an interaction of the experimental factors suggesting that in relation to the baseline, there were more nods in affirmations than in rejections of positive antecedents and vice versa for responses to negative antecedents, the differences being very large. Thus, co-speech head nods align with positive response polarity [+]. The largest number of head nods

occurred in affirmations of positive antecedents, so there seems to be a preference for marking this particular discourse move by a nod. This is different from the spoken languages that we discussed, where the nod was more clearly associated with [+] only. This is also different from DGS, where co-speech nods are much more restricted to affirmations. In our data for German, the nod also regularly occurs in rejections.

For co-speech headshakes, there also is an interaction of the experimental factors in relation to the baseline. Headshakes are greatly reduced in responses with positive polarity, i.e., in affirmations of positive antecedents and in rejections of negative antecedents. So, similar to the baseline (no head movement), headshakes align with negative response polarity [–], albeit with larger differences. The alignment with [–] is different from DGS, where the headshake is more closely associated with rejections.

Regarding the other head movement types, Table 4 lists effects for pushing the head forward (interaction), for pooled forward/backward tilts (interaction), and for pooled right/left turns (antecedent effect) in relation to the no-head-movement baseline. Forward pushes and tilts do not align with response polarity, unlike the baseline. Forward pushes are more frequent in rejections than in affirmations, which may be interpreted as making the rejection more forceful or emphatic. This needs closer scrutiny in future research because it would go against the assumption that speakers avoid face-threatening acts. Forward/backward tilts are more frequent in affirmations than in rejections, which might be related to them being 'half nods'. Turns show a similar alignment with negative response polarity as the baseline (but show a small antecedent effect). Since headshakes also align with negative response polarity one might speculate that turns (away) share this characteristic because they share part of the movement with headshakes.

Table 4. Model estimates for head movement types. Effects are given for negative antecedent and for rejection. The asterisks indicate significance levels.

Head movement type	***Effect of***	***Estimate***	***SE***	***z***	***p***	
head nod	ANTECEDENT	-0.16	0.11	-1.43	0.15	
	SPA	-0.56	0.11	-4.99	< 0.001	***
	ANTECEDENT × SPA	-1.15	0.11	-10.19	< 0.001	***
headshake	ANTECEDENT	0.16	0.21	0.76	0.44	
	SPA	0.14	0.21	0.69	0.49	
	ANTECEDENT × SPA	1.03	0.21	4.94	< 0.001	***
pushed forward	ANTECEDENT	-0.15	0.13	-1.21	0.23	
	SPA	0.14	0.13	1.13	0.26	
	ANTECEDENT × SPA	-0.29	0.13	-2.29	< 0.05	*
forward/backward	ANTECEDENT	-0.09	0.13	-0.70	0.49	
tilt	SPA	-0.39	0.13	-3.10	< 0.01	**
	ANTECEDENT × SPA	-0.31	0.13	-2.43	< 0.05	*
turned left or right	ANTECEDENT	0.26	0.12	2.26	< 0.05	*
	SPA	-0.05	0.12	-0.46	0.64	
	ANTECEDENT × SPA	-0.02	0.12	-0.18	0.86	
other	ANTECEDENT	0.13	0.14	0.95	0.34	
	SPA	-0.03	0.14	-0.19	0.85	
	ANTECEDENT × SPA	-0.09	0.14	-0.61	0.54	

As for the potentially disambiguating function of co-speech head nods and headshakes in responses to negative antecedents, our data point to head nods not being used for disambiguation. All 102 occurrences of nods in rejections of negative antecedents accompany *doch*, which is not ambiguous itself. The 48 nods in affirmations of negative antecedents accompany (a) an unambiguously agreeing element (*n* = 9; *genau*, *stimmt*), thus not serving disambiguation, (b) *ja* (25), (c) ↑mhm (4), and (d) *nein* (10). Regarding *nein*, one might interpret the simultaneous head nod as a non-isofunctional cue for affirmation but note that nods also occur in a few rejections of positive antecedents. In these rejections, the nod might serve a different function, for instance acknowledgement of the fact that something was said. Overall, the co-speech head nod really is ambiguous between marking polarity and speech act.

The headshake might be seen as more indicative of preferring one function. The four occurrences in rejections of negative antecedents accompany (a) *doch* (*n* = 2), which is not ambiguous, and (b) *nein* (2). In affirmations of negative antecedents, the majority of the 36 headshakes accompany *nein* (30), which also marks negative polarity in this discourse move, two occurrences are with unambiguous elements (*genau, stimmt*), and four are with *ja*, which might be taken to suggest that the shake more closely associates with negative polarity.

4.5. Sequential combinations of (at least) two response elements

299 responses had (at least) a second RE following the first RE (26.7% of responses with at least one RE). Table 5 shows the distribution by condition. The statistical analysis revealed a main effect of speech act ($b = –0.73$, $SE = 0.10$, $z = –7.49$, $p < .001$) and an interaction of the two experimental factors ($b = –0.69$, SE = 0.09767, $z = –7.036$, $p < .001$). There were fewer RE2 in rejections than in affirmations, and this effect was larger after positive antecedents. The low frequency of RE2 in rejections of positive antecedents parallels the low frequency of RE1 in the total set of responses (Section 4.3). In the case of RE2, the number is particularly small. As before, we assume that the avoidance of the (cumulative) use of RE serves adhering to a constraint like Krifka's (2013) *DISAGR. However, as we will see below, the effect for RE2 might also be related to the use of certain lexical collocations.

Table 5. Mean proportions of responses with one RE that also contained a second RE.

Antecedent	Speech act	mean	SD
positive	affirmation	0.41	0.49
	rejection	0.06	0.23
negative	affirmation	0.29	0.45
	rejection	0.27	0.44

Figure 3 shows the overall frequency of the most frequent RE1-RE2 combinations ($n > 5$). There is one combination that is particularly frequent, namely *ja genau* 'yes precisely'. The adverb *genau* is unambiguously affirming and thus has the potential to disambiguate ambiguous *ja* after negative antecedents. The same holds for the response verb *stimmt* 'is correct' in the combination *ja stimmt*. However, these two *ja* AGREE sequences occur more often after positive than after negative antecedents, see Figure 4 in the next section. Thus, speakers do not add the unambiguously affirming RE2 for disambiguation.

The combination *nein* AGREE (*nein genau* and *nein stimmt*) occurs only in affirmations of negative antecedents, as expected. The addition of the unambiguously affirming RE2 here clearly serves disambiguation. Unambiguously rejecting *doch* sometimes follows the ambiguous head nod, where *doch* clearly disambiguates: these combinations only occur in rejections of negative antecedents. Another combination where there is potential for disambiguation after negative antecedents is *nein stimmt nicht* 'no, is not correct' but these occurrences are very rare. Finally, the doubling of *doch* has no functional advantage in terms of disambiguation. Native speakers report that for them *doch doch* is less face-threatening than a single *doch*.

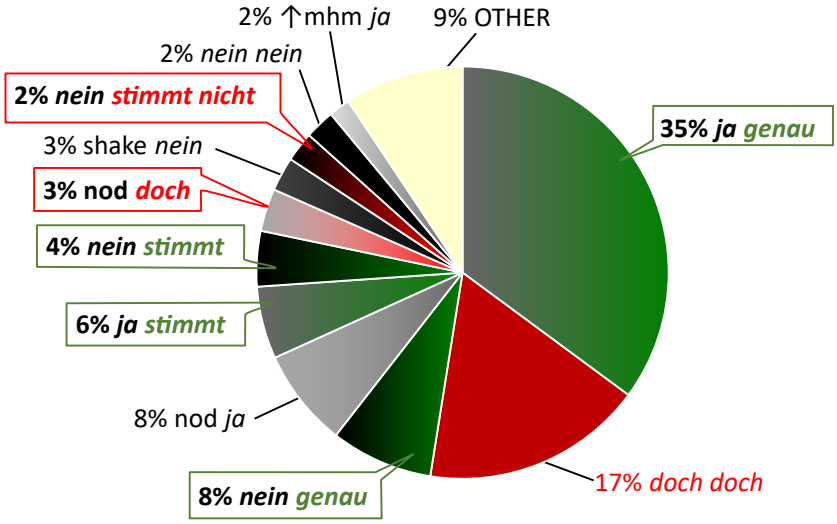

Figure 3. RE1 & RE2 combinations.

5. Towards a theory of response strategies

In our discussion of the results in the previous section, we highlighted the relevance of various optimality-theoretic constraints that were proposed for the choice of RE including simultaneous and sequential combinations (for feature model: Roelofsen and Farkas, 2015; Farkas and Roelofsen, 2019; Loos et al. 2024; for other theoretic work: Krifka, 2013). In this section, we are offering a general discussion of our findings in relation to these constraints, concentrating on non-simultaneous response strategies i.e. single RE responses and RE1-RE2 combinations, with only a few comments on simultaneous head movements. A thorough discussion of all the gestural components is beyond the scope of this paper. Ultimately, a theoretical model of response strategies will have to have a different format than the OT accounts proposed in the earlier literature, requiring a probabilistic approach that includes the many variables affecting the way we express meaning in responses. For ease of exposition in this Section, Figure *4* summarizes the distribution of single RE1 and RE1-RE2 sequences by condition for $n > 10$, sorted by descending frequency.

The most straightforward finding of our open production study in German is that speakers use a RE more often than they do not – i.e., bare response clauses are used less frequently – and when they use a RE, they use vocal means, typically a lexical item, specifically a particle. These findings parallel those for DGS. We suggest that the use of available lexical items in the main articulatory channel of the language (MAXIMIZE MANUAL/VOCAL) plausibly reflects the

unmarked way to interact. The preference for particles over non-particle responses suggests that the DIRECT MAPPING of the presuppositional polarity features in the Pol head is preferred, again reflecting unmarked language use: if there is a functionally dedicated lexical item, use it! This perspective also accounts for the use of a RE vs. no RE. Not using a dedicated item must be pragmatically motivated (see below).

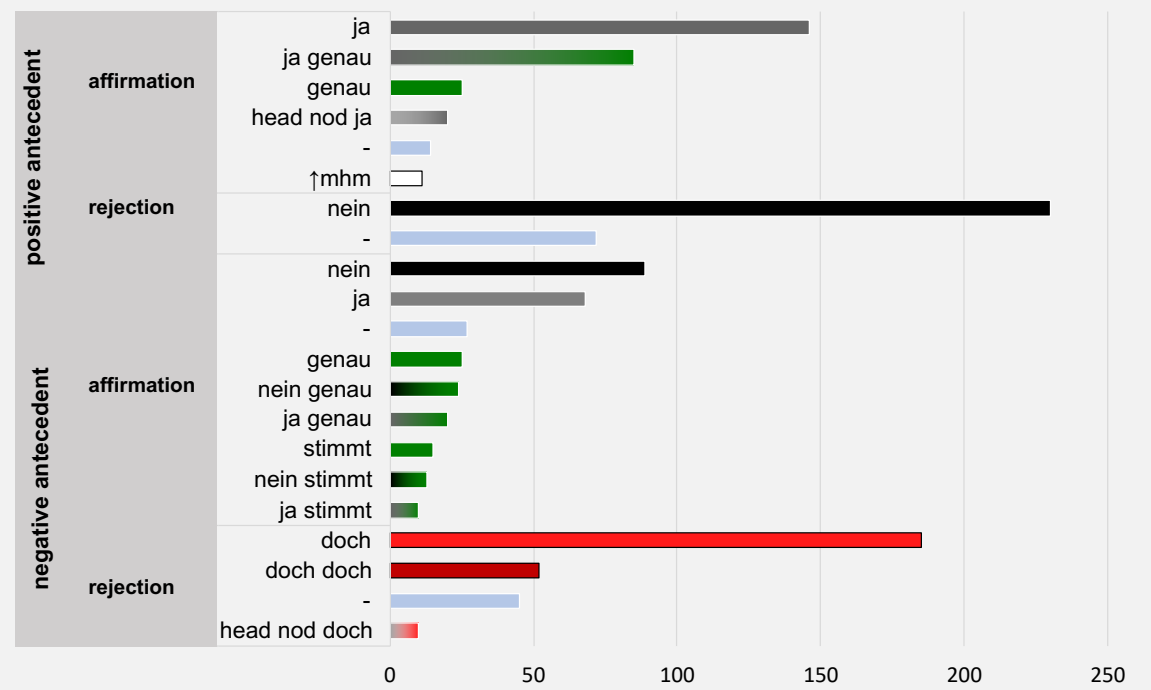

Figure 4. Single RE and RE sequences with a frequency n > 10 by condition.

Considering the three most frequent response strategies in affirmations of positive antecedents (see Figure *4*), we can say that the constraint DIRECT MAPPING is responsible for a preference for *ja* over *genau* but also for *ja genau* over *genau*: *ja genau* contains a particle, so there is direct mapping. However, DIRECT MAPPING cannot account for the preference of *ja* over *ja genau*. When considering all the discourse types we find that the use of a single RE overall is more frequent than the use of RE1-RE2 combinations. This suggests that there is a preference for brevity, which follows directly from Grice's (1989) maxim of manner, and which can be formulated as BE BRIEF. BE BRIEF would then also distinguish between *ja* and *ja genau*.

Regarding the polarity-specific constraints MAXIMIZE MARKED and MAXIMIZE RELATIVE, we mentioned in Section 4.3 that in our sample of the population, MAXIMIZE RELATIVE has a lower weight than in the majority group identified by Claus et al. (2017), as is evidenced by the lack of a preference for *ja* over *nein* in affirmations of negative antecedents. In fact, *nein* is more frequent than *ja* in this discourse type. Crucially such an outcome is not unexpected: We know that there is interindividual variation regarding this constraint, so a different group of individuals will show a different aggregated pattern.

Turning to *doch*, we highlighted in Section **Error! Reference source not found.** that the preference for this particle is guided by the constraint EXPRESSIVENESS, which favors the realization of two features by one particle. The constraint can be viewed as an instance of the more general principle *Maximize Presupposition!*. Loos et al. (2024) highlight that EXPRESSIVENESS is not relevant in DGS because DGS has not grammaticalized a RE like

German *doch* (although there are some speakers that use /doch/ as a mouthing accompanying the particle JA). In our production study we observed that *doch* (or a reduplication of *doch*) is essentially the only lexical means that speakers use in rejections of negative antecedents. The particle *nein* occurred seven times in this discourse type, which is a very low number, *ja* occurred six times. In the acceptability studies conducted by Claus et al. (2017), *doch* clearly received the highest ratings but *nein* was not unacceptable, unlike *ja*. We propose that this difference between studies is due to the different experimental tasks. In our study, participants chose the response they deemed most appropriate. In rejections of negative assertions, this clearly was *doch*. EXPRESSIVENESS is a very powerful constraint. In acceptability studies, however, participants try to comprehend what they are presented with. Since *nein* realizes the feature [REVERSE], which has a high realization need because it is marked, *nein* can be considered as serving the purpose of rejection. The six instances of *ja* in the current study prima facie are surprising: *ja* was judged unacceptable in the acceptability studies. However, we assume that *ja* can be used for signaling that the speaker has understood that their interlocutor has said something, which may occur in open oral production but is a less likely interpretation in written acceptability studies.

Regarding the reduplication *doch doch*, BE BRIEF favours single *doch*, as is evidenced by its higher frequency. Thus, although the avoidance of a face-threatening act seems desirable – recall that reduplication of *doch* feels less face-threatening for native speakers – BE BRIEF trumps *DISAGR. Interestingly, *doch doch* is similarly frequent as the omission of a RE in rejections of negative antecedents (see Figure *4*). Further recall that in rejections of positive antecedents, omission also is a frequently used strategy. As we suggested above, RE omission can be interpreted as a way to avoid conflict. Considering that RE omission is superior when it comes to brevity, it might be surprising that *doch doch* and RE omission are equally frequent. Yet surprise is unwarranted because *doch doch* obeys DIRECT MAPPING, which is not the case for RE omission. Thus, we see that there is an intricate interaction of constraints for the choice of RE (combinations).

Briefly turning to co-speech head movements, we observed that the head nod is not subject to the same constraints as its vocal counterpart *ja*, which is different from DGS. Co-speech head nods are ubiquitous in our data and only are infrequent in rejections of positive antecedents. The fact that they regularly occur both in affirmations and in rejections of negative antecedents, which is different from head nods in other spoken languages, makes them inept for disambiguation. Headshakes align regularly with negative response polarity, which connects them with *nein*, whose [REVERSE] function in rejections of positive antecedents is blocked by *doch*. Head gestures that are not used as stand-alone RE show patterns that are related to the speech act (forward pushes and tilts) or to polarity (turns). However, it seems clear that they can be only one cue of many that contribute to the meaning of a response as they occur in only a small proportion of utterances. This issue must be investigated further when other gestures are taken into account, too.

In conclusion, we have shown that German speakers employ a wide range of multilexical and multimodal response strategies but clearly prefer lexical items that are dedicated to responding (*ja, nein, doch*). Speakers almost always combine lexical and non-lexical means in a response, which confirms and emphasizes the multimodal nature of responding, and of course of communication in general. We have shown the relevance of several constraints influencing the choice of RE in German, and we have argued that most constraints are not response-specific but instantiate general communicative principles (MAXIMIZE VOCAL, DIRECT MAPPING, BE

BRIEF, *DISAGR, EXPESSIVENESS (*Maximize presupposition*)). Cross-linguistic differences can be explained by different repertoires of lexical items, different degrees of conventionalization (e.g., for gesture) and response-specific constraints pertaining to presuppositional polarity features.

References

Brasoveanu, A., D. Farkas, and F. Roelofsen (2013). N-words and sentential negation: evidence from polarity particles and VP-ellipsis. *Semantics and Pragmatics* 6(7), 1–33. DOI: https://doi.org/10.3765/sp.6.7

Claus, B.A., M. Meijer, S. Repp, and M. Krifka (2017). Puzzling response particles: an experimental study on the German answering system. *Semantics and Pragmatics* 10(19), 1–52. DOI: https://doi.org/10.3765/sp.10.19

Elff, M. (2024). `mclogit`. Multinomial logit models, with or without random effects or overdispersion. Version: 0.9.9. R package.

Esipova, M. (2021). Polar responses in Russian across modalities and across interfaces. *Journal of Slavic Linguistics* 29 (FASL 28 extra issue), 1–11.
Online: http://ojs.ung.si/index.php/JSL/article/view/151

Farkas, D. F., and K. B. Bruce. (2010). On reacting to assertions and polar questions. *Journal of Semantics*, *27*(1), 81-118. DOI: https://doi.org/10.1093/jos/ffp010

Farkas, D. and F. Roelofsen (2019). Polarity particles revisited. *Semantics and Pragmatics* 12(15), 1–16. DOI: https://doi.org/10.3765/sp.12.15

Geist, L. and S. Repp (2023). Responding to negative biased questions in Russian. In P. Biskup, M.Börner, O. Mueller-Reichau, and I. Shcherbina (eds.), *Advances in Formal Slavic Linguistics 2021, 145–177.* Berlin: Language Science Press.
Online: http://langsci-press.org/catalog/book/411

González-Fuente, S., S. Tubau, M. Espinal, and P. Prieto (2015). Is there a universal answering strategy for rejecting negative propositions? Typological evidence on the use of prosody and gesture. *Frontiers in Psychology* 6(899), 1–16.
DOI: https://doi.org/10.3389/fpsyg.2015.00899

Goodhue, D. and M. Wagner (2015). It's not just what you say, it's how you say it: intonation, *yes* and *no*. *North East Linguistic Society (NELS)* 45, 227–240.
Online: https://semanticsarchive.net/Archive/TRjYmNiN/GoodhueWagner_2015.pdf

Goodhue, D. and M. Wagner (2018). Intonation, *yes* and *no*. *Glossa*, 3(1):5, 1–45. DOI: https://doi.org/10.5334/gjgl.210

Grice, P. (1989). *Studies in the way of words*. Cambridge, MA.: Harvard University Press.

Holmberg, A. (2001). The syntax of *yes* and *no* in Finnish. *Studia Linguistica*, 55(2), 141–175. DOI: https://doi.org/10.1111/1467-9582.00077

Holmberg, A. (2013). The syntax of answers to polar questions in English and Swedish. *Lingua* 128, 31–50. DOI: https://doi.org/10.1016/j.lingua.2012.10.018

Holmberg, A. (2015). *The Syntax of yes and no*. Oxford: Oxford University Press.

Jones, B.M. (1999). *The Welsh Answering System*. Berlin: De Gruyter.

Keller, F. (2000). *Gradience in Grammar: Experimental and Computational Aspects of Degrees of Grammaticality*. Ph. D. thesis, University of Edinburgh.
Online: https://homepages.inf.ed.ac.uk/keller/publications/phd.pdf

Kramer, R. and K. Rawlins. (2011). Polarity particles: An ellipsis account. *North East Linguistic Society (NELS)* 39, 479–492.

Online: https://faculty.georgetown.edu/rtk8/NELS%20paper.pdf
Krifka, M. (2013). Response particles as propositional anaphors. In T. Snider (Eds.), *Proceedings of the 23rd Semantics and Linguistic Theory Conference*, 1–18. University of California, Santa Cruz. DOI: https://doi.org/10.3765/salt.v23i0.2676
Li, F., S. González-Fuente, P. Prieto, and M.T. Espinal (2016). Is Mandarin Chinese a truth-based language? Rejecting responses to negative assertions and questions. *Frontiers in Psychology* 7(1967), 1–10. DOI: https://doi.org/10.3389/fpsyg.2016.01967
Loos, C., M. Steinbach, and S. Repp (2020). Affirming and rejecting assertions in German Sign Language (DGS). In M. Franke, N. Kompa, M. Liu, J.L. Mueller, and J. Schwab (eds.), *Proceedings of Sinn und Bedeutung* 24(2), 1–19. Osnabrück University and Humboldt University of Berlin. DOI: https://doi.org/10.18148/sub/2020.v24i2.883
Loos, C., M. Steinbach, and S. Repp (2024). Polar response strategies across modalities: Evidence from German Sign Language (DGS). *Language* 100(3), 433–467. DOI: https://doi.org/10.1353/lan.2024.a937185
Meijer, A.M., B. Claus, S. Repp, and M. Krifka (2015). Particle responses to negated assertions: Preference patterns for German *ja* and *nein*. In T. Brochhagen, F. Roelofsen, and N. Theiler (eds.), *Proceedings of the 20th Amsterdam Colloquium*, 286–295. Universiteit van Amsterdam.
Pope, E. (1976). *Questions and Answers in English*. The Hague: Mouton.
Repp, S., M. Meijer, and N. Scherf (2019). Responding to negative assertions in Germanic: on *yes* and *no* in English, Dutch and Swedish. In M.T. Espinal, E. Castroviejo, M. Leonetti, L. McNally, and C. Real-Puigdollers (eds.), *Proceedings of Sinn und Bedeutung 23*(2), 267–285. Universitat Autònoma de Barcelona, Ikerbasque & UPV/EHU, Universidad de Alcalá, and Universitat Pompeu Fabra. DOI: https://doi.org/10.18148/sub/2019.v23i2.611
Roelofsen, F. and D. Farkas (2015). Polarity particle responses as a window onto the interpretation of questions and assertions. *Language* 91(2), 359–414. DOI: https://doi.org/10.1353/lan.2015.0017
Wittenburg, P., H. Brugman, A. Russel, A. Klassmann, and H. Sloetjes (2006). ELAN: A professional framework for multimodality research. In Proceedings of LREC. Genoa, Italy. European Language Resources Association (ELRA).