

Polar responses in Russian across modalities and across interfaces*

Maria Esipova (University of Oslo)

May 26, 2021

Abstract

This paper investigates gestures and prosody in polar responses in Russian as part of a larger research program of studying meaning as it is expressed through various channels and constrained at various levels of representation and their interfaces. Based on the data on head nods and a gestural–intonational cluster used to question the rationale behind the antecedent speech act in Russian responses, it argues that gestures and intonational contours should be treated on a par with spoken words and their parts when it comes to fitting them into typologies of meaning-encoding expressions in spoken language. It also shows, based on the data on linear placement of gestural and spoken polarity markers in Russian as well as prosodic grouping in Russian (and English) polar responses, that studying gestural content and prosodic properties of utterances can help us reveal various interface constraints in natural language.

keywords: semantics, pragmatics, prosody, polar responses, response particles, gestures, Russian

1 Introduction

The question of how different types of polarity are encoded cross-linguistically in polar responses to questions and assertions, as in (1), has received much attention in semantics literature (Farkas and Bruce 2010; Krifka 2013; Roelofsen and Farkas 2015, etc.).

- (1) A: {Did Nina pass the exam?, Nina passed the exam.}
B: {Yes she did, No she didn't}¹

¹For feedback at various stages of this project (which have been very much scattered in time) I would like to thank Byron Ahn, Donka Farkas, Floris Roelofsen, Nicholas Rolle, Philippe Schlenker, the participants of the 'Mean & W(h)ine' group at NYU, and the reviewers and audience at FASL 28. Special thanks to Anna Szabolcsi for being the prime mover of this project.

However, natural language utterances are more than strings of spoken (or signed) words. Gestures (i.e., movements of hands, head, and other body parts, as well as facial expressions) and prosody contribute significantly and non-trivially to the meaning of utterances. Polar responses are an especially rich empirical ground for studying multichannel expression of meaning and how it is constrained at various levels of representation and their interfaces, because they are sensitive both to the utterance-internal and the utterance-external material, and because linguistic communities typically have conventionalized gestures that occur in polar responses. More recently, thus, researchers started incorporating fine-grained gestural and prosodic data into the picture (e.g., González-Fuente et al. 2015).

In this paper, I continue this effort by making novel empirical observations about gestures, intonational contours, linear placement of polarity markers, and prosodic grouping in Russian polar responses and by discussing their theoretical implications for how “secondary channel” expressions contribute meaning and for the grammar of polar responses. In particular, I argue that gestures and intonational contours should be treated as first-class citizens when it comes to fitting them into typologies of meaning-encoding expressions (section 3). I do so by showing that (i) gestural polarity markers, such as head nods and head shakes, fit into the existing typology of polarity markers in a predictable way, but lexicalize independently of spoken polarity markers, and (ii) some other meanings often encoded in responses can be expressed via various channels. I also raise the question of how polar responses are constrained at the interfaces (section 4). I conclude, based on observations about linear placement of polarity markers and prosodic grouping, that while it is not clear how exactly the two types of polarity encoded in polar responses (described in section 2) should be distinguished syntactically, if at all, pragmatics often forces us to structure our polar responses in a certain way.

Before proceeding, I would like to add a caveat that no intonation-labeling system based on assumptions of the autosegmental-metrical approach, akin to MAE-ToBI (Beckman and Ayers 1997), currently exists for Russian. Therefore, my discussion of Russian intonation will be mostly couched in pre-theoretical descriptive terms. Sound files with accompanying TextGrid files and pitch contour drawings as well as videos for selected examples discussed in this paper can be found at <https://osf.io/fx9gu/>.

2 Background: absolute and relative polarity in responses

Roelofsen and Farkas (2015) maintain that polar responses to questions and assertions

¹Throughout the paper, I omit all punctuation marks from the target responses so that they don't get interpreted as indicators of how these responses are most naturally prosodified.

are categorized by two polarity types: ABSOLUTE POLARITY, i.e., polarity of the response itself ([+] or [-]), and RELATIVE POLARITY, i.e., polarity with respect to the antecedent speech act ([AGREE] or [REVERSE]). Roelofsen and Farkas treat the objects in square brackets as morphosyntactic features; I adopt their terminology descriptively. Notation-wise, I show both features for each response and enclose the feature realized by the given particle instance (when it's clear what it is) in a box.

Languages have different inventories of polarity markers used in polar responses with respect to the features they can realize. For example, English *yes* and *no* can realize both types of polarity, which can be seen in responses to negative polar questions (under the low reading of negation, i.e., 'Is it the case that not *p*?') or negative assertions, where the two types of polarity come apart:

- (2) A: {Did Nina not pass the exam?, Nina didn't pass the exam.}
- B: (i) Yes she didn't [AGREE, -]
 (ii) No she didn't [AGREE, -]
 (iii) Yes she did² [REVERSE, +]
 (iv) No she did [REVERSE, +]

In both (2-i) and (2-ii), the absolute polarity is negative ([−]), as the PREJACENT (i.e., the part of the response without the polarity particle) contains sentential negation, but the relative polarity is positive ([AGREE]), since the response matches the antecedent in polarity. In (2-i), the polarity particle *yes*, thus, realizes relative polarity, while in (2-ii), the polarity particle *no* realizes absolute polarity. Conversely, in (2-iii) and (2-iv), the absolute polarity is positive ([+]), since the prejacent contains no negation, but the relative polarity is negative ([REVERSE]), since the response reverses the polarity of the antecedent. Accordingly, *yes* in (2-iii) realizes absolute polarity, while *no* in (2-iv) realizes relative polarity.

Many languages have a dedicated particle for [REVERSE, +] responses only. For example, French uses *si* exclusively in [REVERSE] responses to negative antecedents:

- (3) A: Nina n'a pas passé l'examen {?, .}
 Nina NEG-has NEG passed the-exam
 'Did Nina not pass the exam?, Nina didn't pass the exam.'
- B: {Si, *oui, ?non} elle l'a passé
 SI OUI NON she it-has passed³
 'Yes she passed it' [REVERSE, +]

²Not all English speakers find this response fully natural, especially when it's a reaction to a question rather than an assertion. Furthermore, prosodification is especially important in this case. Whatever the source of this degradedness, it's still different from the categorical nature of the corresponding Russian data.

- (4) A: {Est-ce que Nina a passé l'examen ?, Nina a passé l'examen.}
 is-it that Nina has passed the-exam Nina has passed the-exam
 '{Did Nina pass the exam?, Nina passed the exam.}'
- B: (i) {*Si, oui} elle l'a passé.
 SI OUI she it-has passed
 'Yes she passed it' [AGREE, +]
- (ii) {*Si, non} elle ne l'a pas passé.
 SI NON she not it-has not passed
 'Yes she passed it' [REVERSE, -]

Russian has a gap in the spoken polarity particle inventory, as shown in (5). Like English *no*, Russian *net* can realize both types of negative polarity, i.e., both [REVERSE] and [-]. However, unlike English *yes*, Russian *da* can only realize relative, but not absolute positive polarity, i.e., it can only realize [AGREE], but not [+].

- (5) A: Nina ne sdala ekzamen{?, .}
 Nina NEG passed exam
 '{Did Nina not pass the exam?, Nina didn't pass the exam.}'
- B: (i) Net ne sdala
 NET NEG passed
 'No she didn't' [AGREE, -]
- (ii) Da ne sdala
 DA NEG passed
 'Yes she didn't' [AGREE, -]
- (iii) Net sdala
 NET passed
 'No she did' [REVERSE, +]
- (iv) *Da sdala
 DA passed
 Intended: 'Yes she did' [REVERSE, +]

In subsection 3.1, I will show that this gap in Russian is filled with head nods, which can realize both [AGREE] and [+].

3 Gestures and intonational contours in polar responses

Many linguistic communities have conventionalized head gestures, such as head nods, head shakes, upward head movements,⁴ etc., that are used, among other things, in polar responses. Polar responses can also contain less obvious head and upper body movements that are more likely to fly under the radar of metalinguistic awareness, but

³According to Paloma Jeretić (p.c.), the most natural response here would be one with a reduplicated *si-si*.

⁴These are single abrupt head movements upwards used, for instance, in Bosnia, Bulgaria, Greece, etc., in negative responses. Contra to popular belief, they are distinct from nods used in positive responses and, in some cultures, are accompanied with a click sound. I thank Maria Kouneli and Dunja Veselinović (p.c.) for clarifying these facts to me.

also come with robust meaning associations; the same can be said about intonational contours in polar responses.

In this section, I look at how some of such gestures and intonational contours are used in Russian polar responses and discuss how they fit into existing typologies of meaning-encoding expressions. In particular, I show that head nods and head shakes follow the general typological rules for polarity markers, as described in Roelofsen and Farkas 2015 (contra González-Fuente et al. 2015), but they lexicalize independently of spoken polarity particles. I also discuss a gestural–intonational cluster consisting of a complex gesture whose most prominent feature is a head tilt and a distinct rising contour that can accompany Russian responses and is used to question the rationale behind the antecedent speech act. I argue that this cluster needs to be fit into the typology of rising declaratives, the cross-modal typology of expressions that integrate with spoken utterances they co-occur with prosodically, but not compositionally, and the cross-modal typology of expressions that question or comment on the premises of the antecedent speech act.

3.1 Head nods

In many cultures, head nods are used in polar responses.⁵ González-Fuente et al. (2015) adduce production data showing that Russian (and Catalan) speakers produce nods in both [AGREE] and [REVERSE] responses.⁶ González-Fuente et al. offer no formal analysis for gestures, but informally suggest that nods can be used both for confirming and rejecting an antecedent proposition. That would make them unlike any described spoken polarity particles and would run against the typological predictions in Roelofsen and Farkas 2015, which claim that it's impossible for a single item to be able to realize both [AGREE] and [REVERSE].

There are two main issues with the (interpretation of the) data in González-Fuente et al. 2015, however. First, when making their generalizations, they don't distinguish between what they call "strong" vs. "repeated" nods. Those "strong" nods or a subset thereof might be forward head movements that are used cross-linguistically to mark (contrastive) focus (e.g., House et al. 2001; Dohen et al. 2006), in particular, in [REVERSE] responses. These focus-marking head gestures are articulatorily and perceptually distinct from polarity-marking nods, but can be mistaken for them in a coarse-grained analysis. Throughout this paper I make sure that the nods I am talking about are repeated nods (which I gloss as NOD-NOD), even if the amplitude of the nods in a se-

⁵Of course, head nods have other functions, such as backchanneling, which I will not talk about.

⁶They use different terms, but the gist is the same.

quence drops dramatically after the first nod and regardless of how emphatic the first nod in the sequence is (which can perhaps sometimes be a fusion of a focus-marking and a polarity-marking head gesture). This should not be taken to mean, of course, that only repeated nods can mark polarity; for instance, a single large-amplitude nod can certainly be used as a standalone affirmative response.

Second, and most importantly, when making generalizations about responses used to reject antecedent propositions, González-Fuente et al. don't distinguish between [REVERSE, +] and [REVERSE, -] responses. Yet, I observe that, in Russian, nods are good in [AGREE] and [REVERSE, +], but not in [REVERSE, -] responses, as shown in (6) and (7).⁷

- (6) A: 'Did Nina pass the exam?'
 B: (i) $\overline{\text{Da sdala}}^{\text{NOD-NOD}}$ [AGREE, +]
 (ii) $*\overline{\text{Net ne sdala}}^{\text{NOD-NOD}}$ [REVERSE, -]
- (7) A: 'Did Nina not pass the exam?'
 B: (i) $\overline{\text{Net ne sdala}}^{\text{NOD-NOD}}$ [AGREE, -]
 (ii) $\overline{\text{Net sdala}}^{\text{NOD-NOD}}$ [REVERSE, +]

It should be noted that the nod sequence in (7-ii) is prosodically different from those in (6-i) and (7-i) (in particular, the former seems to be more rapid overall, and the first nod in the sequence seems to have a larger amplitude). However, like with other co-speech gestures, the prosody of NOD-NOD is to a large extent parasitic on the prosody of the speech string the gesture co-occurs with,⁸ and in (7-ii), said prosody is contrastive. To the extent that it is not parasitic on the vocal prosody, the prosody of the nod sequence likely is also contrastive. Thus, the difference between the nod sequences in the two cases with respect to their prosodic properties is expected. That said, in both cases, it is still identifiably a sequence of repeated nods. Proposing that in these two cases we are dealing with two different, independently lexicalized gestures would make as much sense as proposing that two instances of the same spoken morpheme produced with two different intonational contours are, in fact, instances of two different, independently lexicalized morphemes.

I, therefore, maintain that in Russian, nods can realize both [AGREE] and [+], which puts them in the same typological bin as English *yes*. They furthermore fill the gap in the Russian polarity marker inventory created by the lack of a spoken particle that could realize [+], showing that gestural polarity markers aren't simply gestural manifestations of spoken polarity particles within a given language, but lexicalize indepen-

⁷I write head gestures co-occurring with speech as superscripts and use overlining to indicate their approximate temporal alignment.

⁸See, e.g., Loehr 2004 for prosodic properties of co-speech hand gestures.

dently of those, while still fitting properly into the typology of polarity markers.

3.2 Head shakes

Head shakes (labeled *SHAKE*), which are also common cross-culturally, are used in Russian to mark both [REVERSE] and [-], as shown in (8). This makes them akin to English *no* and Russian *net*.

- (8) A: ‘Did Nina not pass the exam?’
 B: (i) $\overline{\text{Net ne sdala}}^{\text{SHAKE}}$ [AGREE, -]
 (ii) $\overline{\text{Net sdala}}^{\text{SHAKE}}$ [REVERSE, +]

However, (8-ii) is somewhat awkward to produce (but not impossible), as aligning a head shake with contrastive vocal prosody seems articulatorily challenging. Furthermore, there might exist a pressure to realize contrastive focus gesturally, which has to be suppressed in (8-ii), since, unlike nods, head shakes are articulatorily incompatible with the focus-marking head movements discussed in the previous subsection.

3.3 The “why would you even say that?” gestural–intonational cluster

Yet another gesture that often emerges in polar responses, but as far as I know, has not been discussed in detail, is a complex gesture consisting of a head tilt, an (optional?) lip purse, and an (optional) slight shrug. This gesture is articulatorily incompatible with repeated nods or head shakes and typically comes hand in hand with a distinct rising intonational contour. The latter is illustrated in Figure 1 for the response in (5-iii), where it is juxtaposed with a regular contrastive focus contour for the same response.

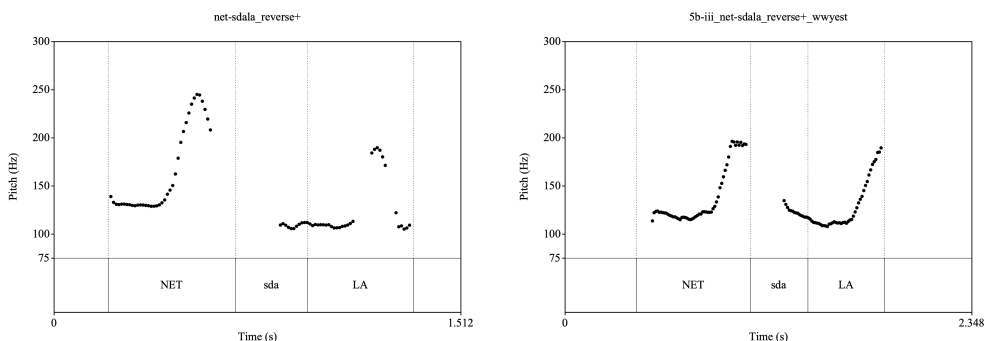


Figure 1: Intonational contours in regular [REVERSE, +] responses with contrastive focus (left) vs. “why would you even say that?” [REVERSE, +] responses (right) in Russian. Both examples consist of two prosodic phrases (one for the particle, one for the pre-jacent), and the contours are repeated twice in each instance.

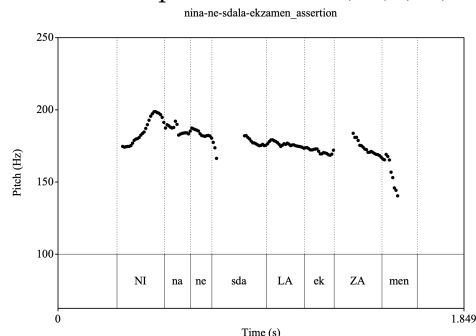
This gestural–intonational cluster can occur outside of polar responses as well, so it

doesn't mark polarity. I maintain that its role is to signal that the speaker is questioning the rationale behind the antecedent speech act. For example, it can be used in responses in which the speaker disagrees with the antecedent assertion or the bias in a biased question to create a "why would you even say that?" (WWYEST) effect. Or it can be used in response to unbiased questions, polar or constituent, if the speaker believes that the person who asked the question should already know the response, so they are questioning the reasons for asking the question in the first place. Further decomposition of the gestural and the intonational components of this cluster into smaller meaningful units is a potentially intriguing endeavour, but falls outside the scope of this paper.

If I were to use MAE-ToBI to label Russian intonation, I would label the WWYEST contour as $L^* H-L\%$,⁹ the contrastive contour on the left in Figure 1 as $L^*+H L-L\%$, and the regular, non-contrastive, non-questioning contours in Russian polar responses as $(L+)H^* L-L\%$ or $H+!H^* L-L\%$. Of course, using MAE-ToBI in this way would be inappropriate. That said, rising contours in English declaratives have been claimed to have various "questioning" functions as well (see Jeong 2018 for a recent discussion). As far as I know, questioning the rationale behind the antecedent speech act has not been explicitly listed among those. Note that this is not the same as questioning some part of the content of the antecedent utterance (including its presuppositions), which is what both English and Russian "incredulous" rising declaratives do:¹⁰

- (9) A: John went to the airport to pick up his sister.
 B: John has a sister? $L^* H-H\%$ (Jeong 2018, (1b))

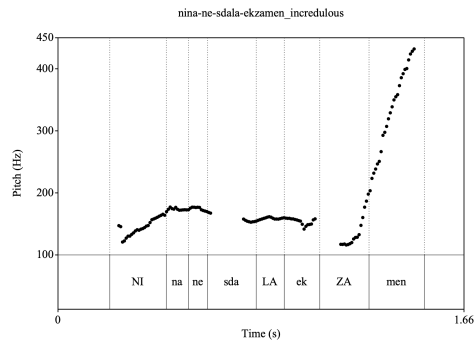
- (10) A: Nina ne sdala ekzamen.
 Nina NEG passed exam
 'Nina didn't pass the exam. $(H^*) (H^*) H^* L-L\%$ '



- B: Nina ne sdala ekzamen?
 Nina NEG passed exam
 'Nina didn't pass the exam? $L^* H-H\%$ '

⁹This is not very easy to see in Figure 1, but when there is more segmental material after the nuclear pitch accent, the tail of the WWYEST contour plateaus out, similarly to $H-L\%$ in English.

¹⁰See Gunlogson 2003 for an earlier discussion of these in English.



Apart from the robust phonetic difference between the two types of rising contours in Russian (a rise to a mid-to-high plateau vs. an extra-high rise), in “incredulous” rising declaratives, the rising contour compositionally combines with the propositional content of the utterance it is hosted by and questions it. However, in WWYEST responses in Russian, there is no questioning of the content of the response itself (for example, the response can contain adverbs like *konečno* ‘of course’); instead, what is questioned is the rationale behind the antecedent speech act. The WWYEST gestural–intonational cluster, thus, seems to form an independent speech act parasitizing on the segmental material it co-occurs with; in a way, it is hijacking the host utterance’s prosody for its own purposes. This would make WWYEST responses akin to “uptalk” rising declaratives from the typology of English rising declaratives in Jeong 2018, where the contour doesn’t compositionally comment on the content of the host utterance either, but is used to build rapport, signal a persona, etc.

I, thus, maintain that, on the one hand, WWYEST responses need to be fit into the typology of rising declaratives. On the other hand, they need to be fit into a cross-modal typology of expressions that integrate with the host utterance at some level(s) of representation (for example, in prosody), but either don’t integrate with it compositionally, or do so vacuously (the beginnings of such a typology can be found in Esipova 2020, which focuses primarily on expressive content). Having established this need, I leave its proper fulfillment for future research.

At this point, one might wonder if there exist languages that lexicalize the WWYEST meaning as a spoken morpheme. Russian adversative *da* (which is distinct from the polarity-marking *da*) in *da net* polar responses also seems to contribute a WWYEST-like flavor (*da net* responses are especially good in co-occurrence with the WWYEST gestural–intonational cluster). Similarly, Russian *že* particle (described, e.g., in McCoy 2003) can be used with a WWYEST-ish flavor in response to (polar or constituent) questions to signal that the person asking the question should have known the answer. French has *mais* {*oui, non, si*} polar responses, and while my French-speaking consul-

tants were not able to come up with precise generalizations as to when those are used, some of the uses of *mais* ('but') in polar responses seem to be of the WWYEST-like kind. I leave a proper investigation of WWYEST cross-linguistically and cross-modally for future research, but it is clear that WWYEST-like meanings can be expressed via multiple channels and should be brought together into a unified typology.

4 Absolute and relative polarity at the interfaces

While the existence of the two types of polarity, absolute and relative, is an empirical fact, it is not at all clear how they are operationalized at various levels of representation and their interfaces. For instance, how are the two types of polarity represented syntactically? And how does a cooperative speaker structure their polar responses, given the syntactic possibilities at hand? In this section, I will argue that there exist pragmatic constraints on the structure of polar responses that go beyond what is in principle possible in the syntax. Based on the Russian data on relative linear placement of gestural and spoken polarity markers within one utterance, I will argue for a pragmatic constraint that urges the speaker to disagree with the antecedent speech act first, before asserting what they believe to be the case (if they are going to do both). I will also adduce data on prosodic grouping in both Russian and English polar responses that suggest that there exists yet another pragmatic constraint that encourages the speaker to package their reaction to the antecedent speech act as its own speech act (when the grammar of a given language allows that).

4.1 Linear placement constraints on the two polarity types

It is uncontroversial that both types of polarity can be marked within one utterance (however this process is operationalized in the syntax/semantics interface). This seems to be the case with the previously mentioned French *si*, which is only licensed in responses that are both [REVERSE] and [+], and, even more obviously, with bi-morphemic particle clusters like the Romanian *ba nu* and German *ja doch*. These particle clusters have a fixed particle order (which is why I refer to them as "clusters"), so I will set them aside and will focus instead on cases of co-occurring polarity markers in which one could in principle expect more flexibility.¹¹

Generally speaking, having two standalone spoken particles realizing two different types of polarity, like in (11), is not impossible in either English or Russian.

(11) A: Did Nina not pass the exam?

¹¹I am also setting aside English responses like *yeah no*, *yeah no for sure*, etc.

B: Yes no she didn't [AGREE, -]¹²

At the very least, such responses seem to be frequently produced in non-sterile speech, even though they often come off as confused. Because of the overall entropic nature of such mixed responses, it is often hard to judge whether they exhibit further constraints on the relative linear placement of the polarity markers. It turns out, however, that such judgements are much clearer when one of the target polarity markers is gestural.

We have already seen that we can realize the two types of polarity simultaneously within one utterance by a gesture and a spoken particle in examples like (7), where *net* is realizing [-] or [REVERSE] while NOD is realizing [AGREE] or [+], respectively. The linear placement of the gestural marker is crucial, however. Pre-speech nods are OK in [AGREE], but not in [REVERSE] responses to negative questions or assertions:

- (12) A: '{Did Nina not pass the exam?, Nina didn't pass the exam.}'
 B: (i) NOD-NOD net ne sdala [AGREE, -]
 (ii) ??NOD-NOD net sdala [+ , REVERSE]

However, a similar contrast doesn't seem to obtain for head shakes:

- (13) A: '{Did Nina not pass the exam?, Nina didn't pass the exam.}'
 B: (i) SHAKE da ne sdala [- , AGREE]
 (ii) SHAKE $\overline{\text{sdala}}^{\text{NOD-NOD}}$ [REVERSE, +]

The generalization seems to be that it is possible to realize both polarity types within one utterance with a head gesture and a spoken particle, but in [REVERSE, +] responses, relative polarity should preferably come first linearly.

The nature of this constraint can't be syntactic. Utterances consisting solely of standalone spoken particles or gestures are in principle possible, therefore, a standalone polarity marker can in principle form an entire speech act, however it is derived (via ellipsis or not). Furthermore, utterances can contain several speech acts. Therefore, the syntax should in principle be able to derive any linear string of polarity markers. Furthermore, the contrast above cuts across the absolute vs. relative distinction in a way that singles out [REVERSE, +] responses, which make up a natural class from a pragmatic, but not obviously from a syntactic standpoint. [REVERSE, +] responses are always reactions to assertions or biased questions and, thus, always lead to a CONVERSATIONAL CRISIS in Roelofsen and Farkas's (2015) terms, i.e., they always signal incompatible commitments or biases of the speech act participants.

With that in mind, I am putting forward a pragmatic principle *Disagree First!*, which

¹²Here and in similar examples, the order of the features in [] reflects the linear order of the particles trying to realize them.

requires that if the speaker is going to both disagree with the content or biases of the antecedent speech act and assert what they believe to be the truth, they should disagree first and only then make their assertion.

Moving on, one might ask whether *Disagree First!* applies to responses with two spoken particles. As I said above, the entropy they create might be too high to have strong introspective judgements, more so than for gestures, which are easier to ignore and are often produced unconsciously. However, the gestural data make the contrasts clearer, so now we know what to look for in examples with two spoken particles.

4.2 Prosodic grouping in polar responses

Spoken polarity particles exhibit prosodic grouping preferences. In particular, in both English and Russian, relative-polarity-realizing particles tend to be in their own prosodic phrases (PrPs).¹³ This is especially obvious in the case of Russian *da*, which can only realize relative polarity and, thus, always prefers to be in its own PrP:

- (14) A: 'Did Nina pass the exam?'
 B: (i) $\{(PrP \text{ Net}) (PrP \text{ ne sdala}), (PrP \text{ Net ne sdala})\}$ [REVERSE] or [-]
 (ii) $\{(PrP \text{ Da}) (PrP \text{ sdala}), ??(PrP \text{ Da sdala})\}$ [AGREE]

We furthermore observe that in English [REVERSE, +] responses, *yes* prefers to be in the same PrP as the prejacent, but the same doesn't hold for *no* in [AGREE, -] responses:

- (15) A: Nina didn't pass the exam.
 B: (i) $\{(PrP \text{ Yes she did}), ??(PrP \text{ Yes}) (PrP \text{ she did})\}$ [REVERSE, +]
 (ii) $\{(PrP \text{ No she didn't}), (PrP \text{ No}) (PrP \text{ she didn't})\}$ [AGREE, -]

So, there seems to be a general tendency for relative-polarity-realizing markers to be prosodically independent across the board, on the one hand, and for absolute-polarity-realizing markers to be prosodically close to the prejacent in [REVERSE, +] responses, on the other. Once again, these tendencies seem to ultimately have pragmatic rather than syntactic roots. I take the prosodic grouping facts above to reflect the tendency for relative-polarity-realizing markers to be independent speech acts, which are packaged into their own PrPs. The dispreference for two PrPs in (15-i) then is due to the fact that in the "high stakes" [REVERSE, +] case, the speaker should probably avoid misleading the addressee, even briefly, into thinking that they are agreeing with them.

¹³In line with the caveat at the end of the Introduction, I remain ignorant about the specific prosodic grouping inventory in Russian, hence the vague term.

¹⁴Beware of the previously mentioned adversative *da*, though, which is always a clitic and has a completely different set of uses than the polarity-marking *da*. Here it would be signalling the speaker's annoyance.

Now, it was observed by Paloma Jeretič (p.c.) that placing *yes* after the preajcent in [REVERSE, $\boxed{+}$] responses doesn't help:

- (16) A: {Did Nina pass the exam?, ??Did Nina not pass the exam?}
 B: (_{PrP} She did) (_{PrP} yes)

This is puzzling. Such a response shouldn't be "misleading", since by the time the speaker gets to utter the polarity particle, they would have already asserted what they believe to be the case. Therefore, this string should be possible under the construal whereby *yes* is realizing [+] in its own speech act packaged into its own PrP. Donka Farkas (p.c.) suggested that in such cases realizing absolute polarity by a postposed polarity particle would be redundant, given that the preajcent-only speech act has already done so. This seems plausible, but I do think that more needs to be said about postposed polarity particles in general. This is yet another issue I leave for future research.

One final note is that the contrast in (15) also holds for the epistemic adverbs *sure* and *konečno* in English and Russian, respectively:¹⁵

- (17) A: I will not pass this exam.
 B: {??(_{PrP} Sure) (_{PrP} you will), (_{PrP} Sure you will)} [REVERSE, +]
- (18) A: 'I will not pass this exam.'
 B: {??(_{PrP} Konečno) (_{PrP} sdaš'), (_{PrP} Konečno sdaš')}
 sure pass.FUT.2SG
 'Sure you will' [REVERSE, +]

I take it that it is the same pragmatic process as in (15) that prevents the split into two speech acts (and, consequently, two PrPs) in the [REVERSE, +] responses in (17) and (18).

Moving on, it would be interesting to investigate how prosodic grouping considerations in polar responses apply to gestures, especially, considering that, unlike spoken morphemes, those typically have the option of linearizing as co-speech.

5 Conclusion

In this paper, I aimed to show that by studying natural language utterances multimodally we can gain a better understanding of how "secondary channel" expressions, such as gestures and intonational contours, contribute meaning, on the one hand, and of how our utterances are constrained more generally, on the other. To do so, I focused on gestural content and prosodic properties of polar responses in Russian.

¹⁵This was observed by Russian-speaking non-linguists in the comment thread under a Facebook post by Asya Pereltsvaig.

I have argued that gestures and intonational contours should be treated as bona fide linguistic objects across the board and, in particular, when it comes to fitting them into typologies of meaning-encoding expressions. I have used two case studies to support this view. By investigating head nods in Russian polar responses, I have shown that we can fit head gestures into the typology of polarity markers in a predictable way, and we can do so independently of spoken polarity particles within a given language. I have also argued that the gestural–intonational cluster that emerges in Russian responses (polar and otherwise) and is used to question the rationale behind the antecedent speech act needs to be fit into the typology of rising declaratives, the cross-modal typology of expressions that integrate with the utterances they are hosted by prosodically, but not compositionally, and the cross-modal typology of expressions that are used to question or comment on the rationale of the antecedent speech act.

I have also demonstrated that investigating gestural and prosodic data can reveal pragmatic constraints that might not be evident if we only look at spoken utterances as strings of words. To that effect, I have looked at relative linear placement of polarity markers, gestural and spoken, within a single response (in Russian), which has revealed the existence of a pragmatic constraint that urges cooperative speakers to disagree with the antecedent speech act before they assert what they believe to be the case (if they are going to do both). I have also shown that certain prosodic grouping properties of polar responses (in Russian and English) indicate that speakers tend to package relative-polarity-realizing markers into their own speech acts and tend to avoid doing so with absolute-polarity-realizing markers when the conversational stakes are high.

I hope that the multi-modal approach to meaning expression and interface constraints thereupon will be extended to other empirical domains in the future.

Bibliography

- Beckman, M. E. and Ayers, G. (1997). Guidelines for ToBI labelling. Version 3.0. The Ohio State University Research Foundation.
- Dohen, M., Lœvenbruck, H., and Harold, H. (2006). Visual correlates of prosodic contrastive focus in French: Description and inter-speaker variability. In Hoffmann, R. and Mixdorff, H., editors, *Proceedings of Speech Prosody*, pages 221–224.
- Esipova, M. (2020). Composure and composition. Ms., University of Oslo, in revision.
- Farkas, D. F. and Bruce, K. B. (2010). On reacting to assertions and polar questions. *Journal of Semantics*, 27(1):81–118.
- González-Fuente, S., Tubau, S., Espinal, M. T., and Prieto, P. (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.
- Gunlogson, C. (2003). *True to Form: Rising and Falling Declaratives as Questions in English*. Routledge, New York.
- House, D., Beskow, J., and Granström, B. (2001). Timing and interaction of visual cues for prominence in audiovisual speech perception. In *Proceedings of the European Conferences on Speech Communication and Technology (Eurospeech)*, pages 387–390.
- Jeong, S. (2018). Intonation and sentence type conventions: Two types of rising declaratives. *Journal of Semantics*, 35(2):305–356.
- Krifka, M. (2013). Response particles as propositional anaphors. In Snider, T., editor, *Proceedings of SALT 23*, pages 1–18.
- Loehr, D. P. (2004). *Gesture and intonation*. PhD thesis, Georgetown University.
- McCoy, S. (2003). Connecting information structure and discourse structure through “Kontrast”: The case of colloquial russian particles *-TO*, *že*, and *VED’*. *Journal of Logic, Language and Information*, 12(3):319–335.
- Roelofsen, F. and Farkas, D. F. (2015). Polarity particle responses as a window onto the interpretation of questions and assertions. *Language*, 91(2):359–414.
-