Enhancing ISO 24617-2: Formalizing Apology and Thanking Acts for Spoken Russian Dialogue Annotation

Ksenia Klokova
MIPT
Moscow, Russia
klokova.ks@mipt.ru

Anton Bankov
Independent researcher
Moscow, Russia
ant.s.bankov@gmail.com

Nikolay Ignatiev
HSE, MIPT
Moscow, Russia
ignatievnickolay@gmail.com

Abstract

This paper refines ISO 24617-2's Social Obligations Management dimension by formalizing apology and thanking acts for Russian dialogue annotation. Addressing gaps in formal definitions and limited response strategies, we propose culture-neutral semantic cores using Wierzbicka's universal primes and update semantics. We introduce three response functions: address (minimal acknowledgment), downplay (mitigation), and decline (reinforcement). Validated through qualitative analysis, this framework captures empirical strategies—including non-response, formulaic minimization, and strategic obligation maintenance—unaddressed in the current standard. Our approach maintains ISO compatibility while eliminating unsubstantiated elements like obligatory response pressure, enhancing annotation accuracy for Russian dialogue.

1 Introduction

Natural dialogue involves nuanced negotiation of social obligations that extends beyond binary frameworks. While ISO 24617-2:2020 provides a comprehensive framework for Social Obligations Management (SOM), its treatment of core functions like *apology* and *thanking* reveals opportunities for refinement. The standard currently lacks formal definitions for these functions and offers limited coverage of response strategies beyond acceptance — a gap particularly evident when applied to casual spoken dialogues.

Building on Wierzbicka's universal primes and empirical observations, we propose a formalization of the *apology/thanking* and the hierarchy of respective response functions. This approach accommodates cross-linguistic variation while addressing empirical observations from Russian dialogue data, where conventional response taxonomy fail to capture strategies like non-committal addressing, downplaying, or explicit declination. By providing formal definitions and expanded response taxonomies for apology and thanking functions, this work offers a framework that can be utilized in implementing more nuanced social reasoning in automated systems.

Our primary contributions are:

- Formal definitions for refined *apology* and *thanking* using update semantics compatible with ISO 24617-2.
- Extended response taxonomy introducing *addressApology/Thanking*, *downplayApology/Thanking*, and *declineApology/Thanking* functions.
- Validation framework demonstrating applicability to Russian through qualitative analysis of movie dialogues.

This paper is structured as follows: Section 2 reviews ISO 24617-2's SOM dimension and prior work; Section 3 establishes our theoretical foundation in politeness and semantic primitives; Section 4 describes the Russian Multimedia Politeness Corpus and annotation methodology; Sections 5 and 6 present formalizations and case studies for apologies and thankings, respectively; Section 7 discusses implications for dialogue annotation standards.

2 Related Work

The ISO 24617-2 standard (Bunt et al., 2012, 2020; ISO Central Secretary, 2020) is a multidimensional dialogue act annotation scheme based on the Dynamic Interpretation Theory (DIT) (Bunt, 2000b) and the DIT++ taxonomy (Bunt, 2009). It provides the semantic framework for the analysis of the communicative behaviour of dialogue participants and

Proceedings of the 21st Joint ACL - ISO Workshop on Interoperable Semantic Annotation (ISA-21), pages 53–62

has been sucessfully applied to the dialogue act annotation in a number of languages (Petukhova et al., 2014; Yoshino et al., 2018; Ngo et al., 2018; Roccabruna et al., 2021; Oleksy et al., 2022; Hwaszcz et al., 2023).

Communicative functions in the Standard are mapped to 10 different dimensions across 5 contexts (Bunt, 2000a). These dimensions are responsible for the information about the task or activity which motivates the dialogue, including dealing with social obligations. The corresponding Social Obligations Management dimension includes functions related to greeting, introduction, apology, thanking, leave-taking, etc. The extended taxonomy suggested by Gilmartin et al. (2017, 2018) includes functions to account for the common social intentions such as politeness questions.

Communicative functions are defined by their update semantics: updates they impose onto addressee's context (Bunt, 2011; Petukhova, 2011). These updates consist of basic semantic concepts called semantic primitives. There are both general-purpose semantic primitives, which can be used in forming updates for communicative functions in any of the dimensions, and dimension-specific primitives. Although the ISO Standard does not provide formal definitions of communicative functions in terms of these updates, they were used during the process of creating the Standard (Bunt et al., 2010). To our knowledge, the specified formal definitions of the communicative functions (Bunt, 2012, 2014) did not include the definitions for *apology*, *thanking* and their response functions.

3 Pragmatic Perspectives on Apology and Thanking

While apologies universally function as remedial acts addressing normative violations through expressions of regret and responsibility acceptance, their specific linguistic realizations and contextual applications are culture-bound. While the politeness autonomy-based framework (Brown and Levinson, 1987) and cross-linguistic preconditions (speaker involvement, recognized breach, perceived harm) (Blum-Kulka and Olshtain, 1984) offer valuable analytical tools, they do not provide a framework for cross-cultural generalization.

Similarly, thanking acts manifest culture-specific expressions of benefit acknowledgment and debt management (Coulmas, 1981), operating as positive politeness strategies within the universal politeness framework (Brown and Levinson, 1987). Both speech acts fundamentally negotiate social valence—apologies repairing negative equilibrium, thankings reinforcing positive bonds — yet their concrete realizations vary cross-culturally.

Wierzbicka (1991) proposed that universal primes provide the most suitable framework for capturing this variation, reducing apology/thanking to:

> I did/didn't do something (to/for you).
> I feel something bad because of this.
> (Wierzbicka, 1991, p. 126)

These culture-neutral explications avoid ethnocentric presuppositions inherent in other models.

However, any operationalization grounded in real language necessitates culture-linked correlates. When formalizing definitions and annotating examples, naming associated feelings is unavoidable. In the proposed formalization (Sections 5.2 and 6.2):

- Apology operationalized through the semantic primitive *regret* (Petukhova, 2011, p. 158), which encapsulates Wierzbicka's definition. A corresponding emotional correlate that dominates Russian interactions concerning breaches of social norms is *guilt*.
- Similarly, in thanking we operate through the primitive that describes the state of being *grateful* (Petukhova, 2011, p. 158) and its associated correlate *indebtedness*.

While *regret/guilt* and *grateful/indebtedness* may serve as operational correlates in other languages, we acknowledge that these specific emotional mappings reflect Russian cultural patterns and should not be assumed universal. Our methodological contribution lies in proposing that annotation of apology- and thanking-related acts in any language should identify and employ the specific core emotions that culturally drive these speech act realizations.

4 Data and Annotation

This study leverages dialogues from the Russian Multimedia Politeness Corpus (Klokova et al., 2023), a resource designed to model politeness phenomena in contemporary spoken Russian. The RMPC comprises manually transcribed excerpts from modern Russian films, capturing five core

politeness scenarios: greetings, acquaintance rituals, apologies, thankings, and leave-takings. Each transcript includes punctuation annotation and is enriched with paralinguistic features — specifically non-verbal markers (e.g., gestures, facial expressions) and sociopragmatic variables grounded in foundational sociolinguistic frameworks (Brown and Levinson, 1987; Helfrich, 1979; Holmes, 1995; Mills, 2003).

For the purpose of this study, we initially applied the ISO 24617-2:2020 annotation framework to apology and thanking sequences within the corpus. The discovered limitations in the current primary (*apology, thanking*) and response (*acceptApology/Thanking*) functions necessitated an iterative extension of the taxonomy. The extended scheme was then tested through targeted annotation of salient sequences — rather than full dialogue application — to validate its descriptive adequacy.

4.1 Annotation Procedure

We selected 110 dialogues containing apology-relevant frames and 105 dialogues containing thanking-relevant frames from the corpus. Each dialogue was independently annotated by two annotators, with a third annotator performing final reconciliation of disagreements.

Our annotation approach focused exclusively on segments that instantiated the core emotional and pragmatic functions of apologizing and thanking, conceptualized through the emotional correlates of 'regret' (guilt) and 'gratitude' (indebtedness), respectively. These states were primarily inferred through linguistic cues (formulae and explicit admissions of such feelings). The corresponding (un)resolved states were inferred through subsequent conversational context, the speaker's intonation, and non-verbal signals.

This framework required excluding supportive strategies that, while co-occurring within apology or thanking sequences, serve different pragmatic functions. For apologies, we excluded strategies such as offers of repair or accounts that lack essential expressions of regret or guilt acknowledgment. Similarly, for thankings, we excluded strategies like exclamations of surprise or compliments that do not signal indebtedness. Although such strategies might substitute for nuclear expressions, their reliable annotation presents considerable challenges due to dependence on multiple contextual variables, ranging from the scale of the wrong- or right-doing to the speaker's actual psychological state.

This principled exclusion maintains theoretical coherence by distinguishing core speech acts from accompanying strategies. Conflating these would obscure the distinct pragmatic mechanisms underlying different politeness phenomena and prevent accurate identification of linguistic realizations of regret and gratitude.

Following this filtering, our final dataset comprised 103 dialogues with apology acts and 92 with thanking acts. Inter-annotator agreement (Krippendorff's alpha) reached 0.92 for all segments and 0.98 for matching segments, indicating high reliability. Representative examples are analyzed in Sections 5.3 and 6.3.

5 Apology

5.1 Current ISO 24617-2 Definition

Current ISO 24617-2 offers two communicative functions for annotating apology interactions: *apology* and *acceptApology*, which refers to the downplay response strategy. Their definitions are given as follows:

> **/apology/:** Communicative function of a dialogue act performed by the sender, S, in order to signal that he/she wants the addressee, A, to know that S regrets something; S puts pressure on A to acknowledge this.
>
> **/acceptApology/:** Communicative function of a dialogue act performed by the sender, S, in order to mitigate the feelings of regret that the addressee, A, has expressed.

Empirical evidence from our data reveals frequent absence of responses to apologies (Section 5.3). This demonstrates that apologies do not universally impose response pressure on the addressee, contrary to the current specification. We therefore refine the definition of the *apology* function by eliminating the pressure-to-respond component.

5.2 Formalization

The exact formalized preconditions of the apology-related communicative functions are provided in the Table 1.

Apology consists of the elementary update functions which inform the addressee of the *regret* (Section 3), which the speaker experiences, and their

Comm. Function	Update Semantics	Explanation
Apology	$\mathbf{Regret}(S, \mu)$ $\mathbf{Want}(S, \mathbf{Bel}(A, \mathbf{Regret}(S, \mu)))$	Sender, S, informs the addressee, A, that S regrets some action or information, μ
Address Apology	$\mathbf{Bel}(S, \mathbf{Regret}(A, \mu))$ $\mathbf{Bel}(S, \mathbf{Want}(A, \mathbf{Bel}(S,$ $\mathbf{Regret}(A, \mu))))$	Sender, S, acknowledges addressee's, A, regret for some action or information, μ
Downplay Apology	$\neg\mathbf{Want}(S, \mathbf{Regret}(A, \mu))$ $\mathbf{Bel}(S, \mathbf{Regret}(A, \mu))$ $\mathbf{Bel}(S, \mathbf{Want}(A, \mathbf{Bel}(S,$ $\mathbf{Regret}(A, \mu))))$	Sender, S, acknowledges addressee's, A, regret for some action or information, μ, and wants to mitigate the A's feelings
Decline Apology	$\mathbf{Want}(S, \mathbf{Regret}(A, \mu))$ $\mathbf{Bel}(S, \mathbf{Regret}(A, \mu))$ $\mathbf{Bel}(S, \mathbf{Want}(A, \mathbf{Bel}(S,$ $\mathbf{Regret}(A, \mu))))$	Sender, S, acknowledges addressee's, A, regret for some action or information, μ, and wants to reinforce the A's feelings

Table 1: Formalized preconditions for the proposed communicative functions of apology and response to apology (S = sender, A = addressee, μ = some action or information)

desire to communicate this *regret* to the addressee. Given that an apology does not imply pressure to respond, the elementary update function responsible for the pressure was omitted.

In order to account for possible response strategies, we propose the following hierarchy of *apology*-related response functions: *addressApology* is the broader concept and *downplayApology* and *declineApology* are its conceptual domain (see Figure 1).

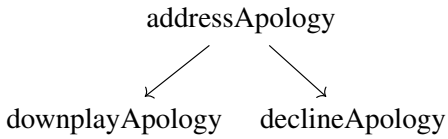

Figure 1: Apology responses

The *addressApology* function is responsible for the more general and non-committal responses. Thus, its formalized preconditions consist only of the elementary update functions corresponding to the acknowledgement of the *regret*, expressed by the addressee with the *apology* function previously in the dialogue. The *downplayApology* function corresponds to the *acceptApology* in ISO 24617-2 and *Apology-downplay* in DIT++. In this function, the sender wants to mitigate the addressee's feeling of *regret*, i.e. the sender communicates that they do not want the addressee to experience *regret*; thus, $\neg\mathbf{Want}(S, \mathbf{Regret}(A, \mu))$ elementary update function is present. On the contrary, in *declineApology* the sender wants to reinforce the addressee's feeling of *regret*; thus, $\mathbf{Want}(S, \mathbf{Regret}(A, \mu))$ elementary update function is used.

5.3 Response Strategies to Apology

Our annotation identified 140 segments as apology acts. The majority of apologies (99 instances, 71%) received no explicit response from the recipient. Among the responses that did occur, downplaying emerged as the most frequent strategy, accounting for 17% (24 instances) of all apologies, followed by declining at 8% (11 instances) and addressing at 6% (9 instances). The remaining three cases involved reciprocal apologies, where recipients responded with their own apology acts.

5.3.1 Apology without Response

Our annotation reveals contexts where apologies elicit no response, particularly when functioning as discourse organizers rather than debt-negotiation acts (1). Other possible contexts involve apologizing for misspeaking, at the end of an interaction (2), and preemptive apologies.

(1) [Rus] Дмитрий: *Извините*... Мне бы Галаганову Нину Сергеевну найти? Девушка: В зале она, молодой человек.

[Translation] Dmitry: *Excuse me*... I'm looking for Nina Sergeevna Galaganova. [SOM:apology]

Young woman: She's in the hall, young man.

(2) [Rus] Олег: Всё, ухожу, ухожу. *Извини, пожалуйста*. *уходит*

[Translation] Oleg: Alright, I'm going, I'm going. *Forgive me, please*. *leaves* [SOM:apology]

Example (2) demonstrates that pressure within the *apology* function is not merely ignored by the addressee but is absent from the speaker's communicative intent. In this instance, the speaker concludes a conflictual exchange for which they bear responsibility. While the speaker's utterance stems from guilt and constitutes an apology act, the speaker simultaneously terminates the dialogue — a move that would be incompatible with exerting pressure on the addressee to elicit a response to the apology.

5.3.2 Addressing Apologies

Addressing constitutes non-committal acknowledgment without guilt mitigation, typically realized through minimal tokens (3) or clarification requests (4) that maintain rather than resolve guilt conditions or avoid evaluating the offense's validity.

(3) [Rus] Маша: *Извини*, я не могла раньше, честное слово.
Костя: *Угу*.

[Translation] Masha: *Sorry*, I couldn't make it earlier, I swear. [SOM:apology]
Kostya: *Uh-huh*. [SOM:addressApology]

(4) [Rus] Маша: Виктор Сергеевич, *простите меня*.
Виктор: *За что*?

[Translation] Masha: Viktor Sergeyevich, *forgive me*. [SOM:apology]
Viktor: *For what*? [SOM:addressApology]

5.3.3 Downplaying Apologies

Downplaying responses actively mitigate the apologizer's guilt burden, characterized by an intention to reduce perceived offense severity. Common formulaic apology minimizers could include ничего страшного (*no worries*) or все в порядке (*it's ok*) as in Example (5).

(5) [Rus] Фима: *Извините, что поздно звоню*. Вам говорить удобно?
Вадим: Да, Фима. *Ничего, ничего*.

[Translation] Fima: *Sorry for calling so late*. Can you talk? [SOM:apology]
Vadim: Yes, Fima. *It's alright, no problem*. [SOM:apologyDownplay]

Two other strategies are shown in Example (6): the speaker employs apology minimization through jocular disbelief followed by imperative termination, systematically dismantling the addressee's guilt assertion.

(6) [Rus] Иван: *Прости нас с матерью*, сынок.
Дмитрий: *Да ты чё*, бать? (...)
Иван: *Всё, что с тобой случилось, это наша вина*.
Дмитрий: *Перестань*, бать.

[Translation] Ivan: *Forgive me and your mother*, son. [SOM:apology]
Dmitry: *What're you on about*, Dad? [SOM:apologyDownplay] (...)
Ivan: *All that happened to you it's our fault*. [SOM:apology]
Dmitry: *Stop it*, Dad. [SOM:apologyDownplay]

5.3.4 Declining Apologies

The *decline* response type maintains or reinforces guilt rather than alleviates it. The most straightforward examples of such utterances would be я тебя не прощаю (*I don't forgive you*) or мне не нужны твои извинения (*I don't need your sorry*). Other possible strategies could involve highlighting negative consequences or costs incurred, and establishing avoidable fault by specifying an alternative action (as in (7)). In Example (8) the speaker declines an apology by postponing resolution.

(7) [Rus] Маша: Максим, *извините*, вы, наверное, меня не дождались?
Максим: *Я час ждал*.
Маша: Ой, а меня на работе задержали.
Максим: *Могли бы позвонить*.

[Translation] Masha: Maxim, *I'm sorry*, you probably didn't wait for me, did you? [SOM:apology]
Maxim: *I waited for an hour*. [SOM:declineApology]

Masha: Oh, they kept me late at work.
Maxim: *You could have called* [SOM:declineApology].

(8) [Rus] Олег: Кать, *прости меня*.
Катя: Олег, *давай всё потом, пожалуйста*? Я тебя очень прошу, ну я очень хочу спать.

[Translation] Oleg: Katya, *forgive me*. [SOM:apology]
Katya: Oleg, *let's talk later, please*? I'm begging you, come on... I really want to sleep. [SOM:declineApology]

6 Thanking

6.1 Current ISO 24617-2 Definition

Similar to apology, ISO 24617-2 offers two communicative functions for annotating thanking interactions: *thanking* and *acceptThanking*, which refers to the downplay response strategy. Their definitions are given as follows:

> **/thanking/:** Communicative function of a dialogue act performed by the sender, S, in order to inform the addressee, A, that S is grateful for some action performed by A; S puts pressure on A to acknowledge this.
>
> **/acceptThanking/:** Communicative function of a dialogue act performed by the sender, S, in order to mitigate the feelings of gratitude which the addressee, A, has expressed.

However, once again, our data shows that non-committal and even declining response strategies to thanking are possible and that thanking does not necessarily pressure the addressee to respond. For illustrative examples, refer to Section 6.3.

6.2 Formalization

The exact formalized conditions of the thanking-related communicative functions are provided in the Table 2.

Similar to the *apology* function, *thanking* function consists of the elementary update functions, informing the addressee of the *gratitude*, which the sender experiences, and their desire to communicate this *gratitude*. It does not include the elementary update function responsible for the pressure to acknowledge addressee's *gratitude* in response.

The hierarchy of response communicative functions to *thanking* is similar to that of *apology*: *addressThanking* is the broader concept, while *downplayThanking* and *declineThanking* are its conceptual domain (see Figure 2).

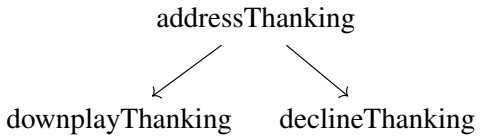

Figure 2: Thanking responses

The *addressThanking* function occurs in more general and non-committal responses and its formalized conditions consist only of the elementary update functions corresponding to the acknowledgement of the *gratitude*, expressed by the addressee. In *downplayThanking* (*acceptThanking* in ISO 24617-2 and *Thanking-downplay* in DIT++), the sender wants to mitigate the addressee's feeling of *gratitude*; thus, $\neg\mathbf{Want}(S, \mathbf{Grateful}(A, T, \mu))$ elementary update function is present. Contrary to downplaying, in *declineThanking* the sender wants to reinforce the addressee's feeling of *gratitude* (for example, expecting favor in the future); thus, $\mathbf{Want}(S, \mathbf{Grateful}(A, T, \mu))$ elementary update function is used.

6.3 Response Strategies to Thanking

Our annotation resulted in 118 segments classified as thanking acts. Most expressions of gratitude (82 instances, 70%) went without explicit acknowledgment. When responses were present, addressing was the predominant strategy at 16% (19 instances) of all thankings, while downplaying occurred in 5% (6 instances) and declining in 3% (3 instances). Similar to apologies, the remaining eight cases featured reciprocal thanking.

6.3.1 Thanking Without Reply

Similar to apologies, our annotation of thanking sequences reveals contexts where no reply is pragmatically required. Particularly, this is the case in casual interactions or when the gratitude expression terminates the conversational exchange (e.g., service encounters or farewells). This absence of response may be conditioned by: 1) conversational position – terminal thanking acts often lack replies;

Comm. Function	Update Semantics	Explanation
Thanking	$\mathbf{Grateful}(S, T, \mu)$ $\mathbf{Want}(S, \mathbf{Bel}(A, \mathbf{Grateful}(S, T, \mu)))$	Sender, S, informs the addressee, A, that S is grateful to some person(s), T, for some action or information, μ (most often T coincides with A)
Address Thanking	$\mathbf{Bel}(S, \mathbf{Grateful}(A, T, \mu))$ $\mathbf{Bel}(S, \mathbf{Want}(A, \mathbf{Bel}(S,$ $\mathbf{Grateful}(A, T, \mu))))$	Sender, S, acknowledges addressee's, A, gratitude to some person(s), T, for some action or information, μ
Downplay Thanking	$\neg\mathbf{Want}(S, \mathbf{Grateful}(A, T, \mu))$ $\mathbf{Bel}(S, \mathbf{Grateful}(A, T, \mu))$ $\mathbf{Bel}(S, \mathbf{Want}(A, \mathbf{Bel}(S,$ $\mathbf{Grateful}(A, T, \mu))))$	Sender, S, acknowledges addressee's, A, gratitude to some person(s), T, for some action or information, μ, and wants to mitigate the A's feelings
Decline Thanking	$\mathbf{Want}(S, \mathbf{Grateful}(A, T, \mu))$ $\mathbf{Bel}(S, \mathbf{Grateful}(A, T, \mu))$ $\mathbf{Bel}(S, \mathbf{Want}(A, \mathbf{Bel}(S,$ $\mathbf{Grateful}(A, T, \mu))))$	Sender, S, acknowledges addressee's, A, gratitude to some person(s), T, for some action or information, μ, and wants to reinforce the A's feelings

Table 2: Formalized conditions for the proposed communicative functions of thanking and response to thanking (S = sender, A = addressee, T = some person(s), μ = some action or information)

2) socio-relational factors – familiarity (short social distance), age and hierarchical disparity may override the expectation of a response. Example (9) demonstrates this pattern in a child-adult interaction with a clear status asymmetry.

(9) [Rus] Девочка: Дядя, а можно, пожалуйста, мячик?
Игорь: Конечно.
Девочка: *Спасибо*.

[Translation] Little girl: Mister, can I have the ball please?
Igor: Of course.
Little girl: *Thank you*. [SOM:thanking]

6.3.2 Addressing Thanking

Addressing responses function as a minimal non-committal response (*gratitude* is recognized without an attempt to alleviate or reinforce the speaker's expressed obligation). They often involve formulaic markers пожалуйста (*welcome*), interjections угу (*uh-huh*) (10).

(10) [Rus] Егор: Я понимаю, что достал. В общем, *спасибо, что помогаешь мне*.
Катя: *Угу*. Давай, пока.

[Translation] Yegor: I know I'm being annoying. Anyway, *thanks for helping me out*. [SOM:thanking]
Katya: *Uh-huh*. Alright, see ya. [SOM:addressThanking]

Example (11) illustrates strategic addressing in conflict discourse. Here, the response *I see* acknowledges the expression of *gratitude* and avoids mitigation.

(11) [Rus] Алина: *Я очень благодарна твоей маме*.
Борис: *Я вижу*.

[Translation] Alina: *I'm really grateful to your mom*. [SOM:thanking]
Boris: *I see*. [SOM:addressThanking]

6.3.3 Downplaying Thanking

We operationalize *thankingDownplay* as a response only when the speaker actively minimizes the addressee's debt acknowledgment. This could be achieved with intentional mitigation via explicit verbal cues (e.g., rhetorical questions, minimizers) that negate the need for *gratitude* (12). Common minimizers could include phrases like без проблем (*no problem*) or забей (*forget it*).

(12) [Rus] Отец Кати: (...) *Спасибо, что доехали до меня*.
Катя: Пап, *ну ты шутишь что ли*? *Ну как же бы мы не доехали*?

[Translation] Katya's father: And I love you too. *Thank you for coming to see me*. [SOM:thanking]
Katya: Dad, *are you kidding*? [SOM:thankingDownplay]

How could we not come? [SOM:thankingDownplay]

In contrast to formulaic expressions, a downplay can also be characterized by pragmatic markedness. In (13) there is a three-part thanking sequence which emphasizes the debt acknowledgment, followed by a blatant rejection of the debt frame.

(13) [Rus] Алина: *Спасибо*, дорогой мой. Золотце моё, *ты меня очень сильно выручил*. *Я твоя должница*.
Юра: *Да иди ты*, *что ты говоришь*? Всё, давай, целую, до понедельника.

[Translation] Alina: *Thank you* [SOM:thanking], my dear. My darling one, *you really helped me out big time* [SOM:thanking]. *I owe you one*. [SOM:thanking]
Yura: *Screw you* [SOM:thankingDownplay], *what are you talking about* [SOM:thankingDownplay]? Alright then, kisses, see you Monday.

6.3.4 Declining Thanking

Declining responses actively maintain or intensify debt obligations. Possible strategies include the affirmation of the outstanding obligation (и правильно (*rightly so*)) or explicit debt reminders (не забудь, кто тебе помог – *don't forget who helped you*). Example (14) illustrates the strategic avoidance of debt closure. The speaker's topic shift *Shall we go?* occurs as a response to preemptive thanking before request fulfillment. By doing so, she acknowledges the gratitude and creates conditional obligation (debt remains pending).

(14) [Rus] Борис: Всё-таки, пожалуйста, не кричи на неё. Мы давно живём, я хорошо знаю, когда ты громко говоришь, когда кричишь.
Алина: *молчит*
Борис: *Спасибо*.
Алина: *Поехали*?

[Translation] Boris: Still, please don't yell at her. We've lived together long enough - I know well when you're just speaking loudly and when you're actually shouting.
Alina: *remains silent*
Boris: *Thank you*. [SOM:thanking]
Alina: *Shall we go*? [SOM:declineThanking]

7 Conclusion

In this paper we proposed the refinement to the ISO 24617-2's Social Obligations Management dimension through formal extensions for apology- and thanking-related functions. By grounding definitions in Wierzbicka's universal primes, we established culture-neutral semantic cores for apology and thanking, while accommodating Russian-specific realizations through correlates like *guilt* and *indebtedness*.

Our key contributions — the *address*, *downplay*, and *decline* response functions — resolve empirical gaps in the current Standard, enabling precise annotation of strategies observed in Russian dialogues: non-response (e.g. in the opening positions), formulaic downplaying, and strategic declination. The proposed formalizations maintain ISO compatibility through update semantics while omitting unsubstantiated elements like automatic "pressure to respond".

Future work involves extending our research to other Social Obligations Management functions, such as leave-taking and greetings, using the same methodology of culture-neutral formalization grounded in empirical data. The second objective lies in complete annotation of the communicative functions in the dialogues from the Russian Multimedia Politeness Corpus. Ultimately, this research aims to contribute to the development of evaluation frameworks that can assess how well conversational agents, large language models, and other dialogue systems understand and deploy appropriate politeness strategies within specific cultural and contextual parameters.

References

Shoshana Blum-Kulka and Elite Olshtain. 1984. Requests and apologies: A cross-cultural study of speech act realization patterns (ccsarp). *Applied Linguistics*, 5(3):196–213.

Penelope Brown and Stephen C. Levinson. 1987. *Politeness: Some Universals in Language Usage*. Studies in Interactional Sociolinguistics. Cambridge University Press.

Harry Bunt. 2000a. Dialogue pragmatics and context specification. In Harry Bunt and William Black, editors, *Abduction, Belief and Context in Dialogue: Studies in computational pragmatics*, pages 81–149. John Benjamins Publishing Company, Amsterdam.

Harry Bunt. 2000b. Dynamic interpretation and dialogue theory. In M. Martin Taylor, Françoise Néel, and Don Bouwhuis, editors, *The Structure of Multimodal Dialogue*, pages 139–188. John Benjamins Publishing Company, Amsterdam.

Harry Bunt. 2011. The semantics of dialogue acts. In *Proceedings of the Ninth International Conference on Computational Semantics (IWCS 2011)*.

Harry Bunt. 2012. The semantics of feedback. In *Proceedings of the 16th Workshop on the Semantics and Pragmatics of Dialogue (SEMDIAL 2012), Paris, France*, pages 118–127. University Paris-Diderot, Paris Sorbonne-Cite.

Harry Bunt. 2014. A context-change semantics for dialogue acts. In Harry Bunt, Johan Bos, and Stephen Pulman, editors, *Computing Meaning: Volume 4*, pages 177–201. Springer Netherlands, Dordrecht.

Harry Bunt, Jan Alexandersson, Jean Carletta, Jae-Woong Choe, Alex Chengyu Fang, Koiti Hasida, Kiyong Lee, Volha Petukhova, Andrei Popescu-Belis, Laurent Romary, Claudia Soria, and David Traum. 2010. Towards an ISO standard for dialogue act annotation. In *Proceedings of the Seventh International Conference on Language Resources and Evaluation (LREC'10)*, Valletta, Malta. European Language Resources Association (ELRA).

Harry Bunt, Jan Alexandersson, Jae-Woong Choe, Alex Chengyu Fang, Koiti Hasida, Volha Petukhova, Andrei Popescu-Belis, and David Traum. 2012. ISO 24617-2: A semantically-based standard for dialogue annotation. In *Proceedings of the Eighth International Conference on Language Resources and Evaluation (LREC'12)*, pages 430–437, Istanbul, Turkey. European Language Resources Association (ELRA).

Harry Bunt, Volha Petukhova, Emer Gilmartin, Catherine Pelachaud, Alex Fang, Simon Keizer, and Laurent Prévot. 2020. The ISO standard for dialogue act annotation, second edition. In *Proceedings of the 12th Language Resources and Evaluation Conference*, pages 549–558, Marseille, France. European Language Resources Association.

H.C. Bunt. 2009. The dit++ taxanomy for functional dialogue markup. In *Proceedings of EDAML@AAMAS, Workshop Towards a Standard Markup Language for Embodied Dialogue Acts*, pages 13–24. International Foundation for Autonomous Agents and Multi-agent Systems.

Florian Coulmas. 1981. *"Poison to Your Soul" Thanks and Apologies Contrastively Viewed*, pages 69–92. De Gruyter Mouton, Berlin, New York.

Emer Gilmartin, Christian Saam, Brendan Spillane, Maria O'Reilly, Ketong Su, Arturo Calvo, Loredana Cerrato, Killian Levacher, Nick Campbell, and Vincent Wade. 2018. The ADELE corpus of dyadic social text conversations:dialog act annotation with ISO 24617-2. In *Proceedings of the Eleventh International Conference on Language Resources and Evaluation (LREC 2018)*, Miyazaki, Japan. European Language Resources Association (ELRA).

Emer Gilmartin, Brendan Spillane, Maria O'Reilly, Christian Saam, Ketong Su, Benjamin R. Cowan, Killian Levacher, Arturo Calvo Devesa, Lodana Cerrato, Nick Campbell, and Vincent Wade. 2017. Annotation of greeting, introduction, and leavetaking in dialogues. In *Proceedings of the 13th Joint ISO-ACL Workshop on Interoperable Semantic Annotation (ISA-13)*.

Hede Helfrich. 1979. Age markers in speech. In Klaus R. Scherer and Howard Giles, editors, *Social markers in speech*, chapter 3, page 63–107. Cambridge University Press, Cambridge, UK.

Janet Holmes. 1995. *Women, Men and Politeness*. Routledge, London, UK.

Krzysztof Hwaszcz, Marcin Oleksy, Aleksandra Domogała, and Jan Wieczorek. 2023. ISO 24617-2 on a cusp of languages. In *Proceedings of the 19th Joint ACL-ISO Workshop on Interoperable Semantics (ISA-19)*, pages 40–46, Nancy, France. Association for Computational Linguistics.

ISO Central Secretary. 2020. Language resource management — semantic annotation framework (semaf) part 2: Dialogue acts. Standard ISO 24617-2:2020, International Organization for Standardization, Geneva, Switzerland.

Ksenia Klokova, Maxim Krongauz, Valery Shulginov, and Tatiana Yudina. 2023. Towards a russian multimedia politeness corpus. In *Computational Linguistics and Intellectual Technologies: Proceedings of the International Conference "Dialogue 2023"*, pages 233–244.

Sara Mills. 2003. *Gender and politeness*. Cambridge University Press, Cambridge, UK.

Thi-Lan Ngo, Pham Khac Linh, and Hideaki Takeda. 2018. A Vietnamese dialog act corpus based on ISO 24617-2 standard. In *Proceedings of the Eleventh International Conference on Language Resources and Evaluation (LREC 2018)*, Miyazaki, Japan. European Language Resources Association (ELRA).

Marcin Oleksy, Jan Wieczorek, Dorota Drużyłowska, Julia Klyus, Aleksandra Domogała, Krzysztof Hwaszcz, Hanna Kędzierska, Daria Mikoś, and Anita Wróż. 2022. DiaBiz.Kom - towards a Polish dialogue act corpus based on ISO 24617-2 standard. In *Proceedings of the 29th International Conference on Computational Linguistics*, pages 3631–3638, Gyeongju, Republic of Korea. International Committee on Computational Linguistics.

Volha Petukhova. 2011. *Multidimensional dialogue modelling*. Phd thesis, Tilburg University, Tilburg, Netherlands.

Volha Petukhova, Martin Gropp, Dietrich Klakow, Gregor Eigner, Mario Topf, Stefan Srb, Petr Motlicek, Blaise Potard, John Dines, Olivier Deroo, Ronny Egeler, Uwe Meinz, Steffen Liersch, and Anna Schmidt. 2014. The DBOX corpus collection of spoken human-human and human-machine dialogues. In *Proceedings of the Ninth International Conference on Language Resources and Evaluation (LREC'14)*, pages 252–258, Reykjavik, Iceland. European Language Resources Association (ELRA).

Gabriel Roccabruna, Alessandra Cervone, and Giuseppe Riccardi. 2021. Multifunctional iso standard dialogue act tagging in italian. In *Proceedings of the Seventh Italian Conference on Computational Linguistics CLiC-it (2020)*, pages 365–372, Bologna, Italy. Accademia University Press.

Anna Wierzbicka. 1991. *Cross-cultural pragmatics: The semantics of human interaction.* Mouton de Gruyter, Berlin.

Koichiro Yoshino, Hiroki Tanaka, Kyoshiro Sugiyama, Makoto Kondo, and Satoshi Nakamura. 2018. Japanese dialogue corpus of information navigation and attentive listening annotated with extended ISO-24617-2 dialogue act tags. In *Proceedings of the Eleventh International Conference on Language Resources and Evaluation (LREC 2018)*, Miyazaki, Japan. European Language Resources Association (ELRA).