

Dialogue Act Annotation

Application of ISO 24617-2 annotation guidelines to Slovenian ROG
corpora and the SloBench ASR evaluation dataset

Darinka Verdonik

University of Maribor, Faculty of Electrical Engineering and Computer Science
Maribor, Slovenia · June 2026

Version 01

Vsebina

1 Introduction.....	4
2 Dialogue act annotation	5
3 Segmentation guidelines	8
3.1 Basic principle	8
3.2 Keep one main clause with its dependent clauses.....	8
3.3 Split coordinated independent clauses when each gives a new step or action	8
3.4 Keep tightly integrated coordination together.....	9
3.5 Split explanatory, causal, result, and purpose clauses when they are independent enough.....	9
3.6 Split discourse markers when they are prosodically independent	10
3.7 Split feedback tokens when they function independently	10
3.8 Segment short questions and question tags separately when they function independently	10
3.9 Treat elliptical and verbless units as possible segments	11
3.10 Split afterthoughts, appositions, and added specifications when they are prosodically separate.....	11
3.11 Do not split ordinary hesitations unless they have their own function	12
3.12 Segment repairs, abandoned starts, and self-corrections when the speaker restarts or withdraws material	12
3.13 Reporting frames and direct speech may be segmented separately	13
3.14 Nonverbal events are separate only when communicatively relevant.....	13
3.15 Discontinuous segments	13
3.16 Practical decision checklist.....	14
4 Annotating dimension	15
4.1 Task	15
4.2 Auto-Feedback.....	16
4.3 Allo-Feedback	16
4.4 Turn Management.....	16
4.5 Time Management	17
4.6 Discourse Structuring.....	17
4.7 Social Obligations Management	17
4.8 Own Communication Management	17
4.9 Partner Communication Management	17
4.10 Unspecified Dimension	18
4.11 Multidimensionality	18
4.12 Practical decision checklist.....	18

6 Annotating semantic content for Inform function	29
6.1 General rules	29
6.2 Semantic-content labels	29
6.2.1 reportedContent	29
6.2.2 emotionalExclamation	30
6.2.3 selfExperience	30
6.2.6 requirementConstraint	31
6.2.7 possibilityOption	31
6.2.8 conditionalScenario	31
6.2.9 causalExplanation	32
6.2.10 referentialSpecification	32
6.2.11 eventAction	32
6.2.12 stateDescription	33
8 Conclusion	34
9 References	35

1 Introduction

Dialogue act annotation describes what speakers do when they communicate. Speakers do not only produce words, phrases, or sentences; they perform communicative actions. They ask questions, provide information, agree, disagree, request something, give instructions, apologize, thank, show understanding, hesitate, repair their own speech, manage the flow of interaction.

The idea that speaking is a form of acting is rooted in speech act theory, especially in the work of Austin (1975) and Searle (1979). Dialogue act annotation, however, is not a direct application of classical speech act theory. The original speech act categories are often too abstract and idealized for corpus-based analysis, as authentic conversational data is highly context-dependent and the most frequent communicative functions do not necessarily align with theoretically derived classifications.

The most common purpose of dialogue act annotation is to make the communicative functions of utterances explicit, so that linguistic research can systematically examine everyday interaction: how speakers negotiate meaning, manage turn-taking, express actions such as opinions, requests, agreements or disagreements, and perform influence acts such as persuading, instructing, warning or prompting others to act. It is also widely used in technologies such as dialogue systems, conversational AI, tutoring systems, educational dialogue analysis, emotion-aware systems and human–machine interaction, because annotated dialogue acts help machines recognize user intentions, track interaction structure, and generate contextually appropriate responses.

This document presents practical guidelines for applying the ISO 24617-2 standard for dialogue act annotation to Slovenian spoken data. Specifically, it was used for the annotation of dialogue acts in the ROG training corpus of spoken Slovenian and in the SloBench ASR evaluation dataset for speech technologies.

The ISO 24617-2 standard has been first released in 2012 and updated in 2020. The present guidelines follow the version from 2012. Although the 2020 version of ISO 24617-2 was reviewed and is occasionally cited in this document, the present annotation scheme does not include the categories newly added in that version: the dimensions Task Management and Contact Management, the communicative function Test Question, or the Contact Management functions. These categories were excluded because the annotated data did not contain communication types that clearly required them.

2 Dialogue act annotation

Several annotation schemes have shaped current approaches to dialogue act annotation. The overview in Table 1 summarizes the main contributions of generic dialogue act annotation schemes discussed in the literature. Although a number of task-specific schemes also exist, the present overview focuses only on generic schemes.

Table 1: Dialogue act annotation schemes

Scheme	Main contribution	Practical note
Speech act theory (Austin 1975; Searle 1969)	Introduced the idea that utterances perform actions such as assertions, directives, commissives, or expressives.	Although it has been used in some cases for annotation of corpus data its categories are relatively limited, and certain types of actions performed in authentic interaction are not covered.
DAMSL (Allen, Core 1997)	Provided a multidimensional framework for dialogue act annotation.	The framework is theoretically rich, but it is strongly oriented toward task-oriented data and permits an excessive number of possible tag combinations.
SWBD-DAMSL (Jurafsky et al. 1997)	Simplified DAMSL for the large-scale annotation of telephone conversations in the Switchboard corpus and became highly influential in automatic dialogue act modelling due to the broad availability of annotated data.	The framework is practical and influential, but its tagset is uneven in granularity and its theoretical basis is less transparent.
AMI (AMI 2005)	Designed for annotation of multimodal meeting interaction in the AMI corpus.	The approach is intent-based and theoretically interesting, but its tagset is relatively sparse and the scheme is less widely recognized.
DIT++ (Bunt 2021)	Systematic and fine-grained, especially in its treatment of metadiscursive acts such as feedback, turn management and time management, while retaining the traditional dialogue act categories for task-performing acts.	Background for the ISO 24617-2 dialogue act annotation standard.
DART (Weisser 2020)	Offers pragmatic, corpus-oriented annotation with intuitive labels.	Provides a very large tagset, but the lack of higher-level categorization and sufficiently detailed documentation makes it difficult to apply outside its original environment.

Language into Act theory (Cresti 2020)	Corpus-based framework for dialogue act annotation that analyses spoken interaction through prosodically identifiable utterance units and classifies their illocutionary functions within a broader theory of speech activity.	Theoretically rich and empirically grounded approach, but it develops an extensive theoretical model, while practical annotation guidelines remain scarce and not publicly available.
ISO 24617-2 (2012; 2020)	Provides a standardized multidimensional framework, DiAML representation, and segmentation principles.	Used here as the annotation framework.

The ISO 24617-2 (2012) standard for dialogue act annotation, used here as the annotation framework, does not treat the utterance as the default unit of annotation, unlike many previous schemes. Dialogue acts are annotated on functional segments: the smallest stretches of communicative behaviour that perform one or more communicative functions (ISO 24617-2 2020: 3). A speaker turn may contain several functional segments, and one functional segment may be discontinuous, embedded, overlapping, or spread across turns.

Each functional segment may be annotated at several levels; however, these levels are hierarchically interconnected. For example, each dimension has its own set of communicative functions, and only inform function has several additional semantic content labels, which helps avoid an excessive number of tag combinations. Table 2 gives an overview of the annotation levels that can be assigned to each functional segment.

Table 2: Overview of ISO 24617-2 levels of annotation

Level	Description	Use in this approach
Sender and addressee	Usually implicit in two-person dialogue; important in multiparty data.	
Dimension	The aspect of communication addressed by the dialogue act, typically topic under discussion, but also the management of understanding, attention, turn-taking, timing, discourse structure, social relations, etc.	Annotated.
Communicative function	What the speaker is doing with the segment, e.g., asking, informing, answering, confirming, requesting, suggesting, thanking, apologizing, taking the turn, stalling, etc.	Annotated.
Semantic content	What the dialogue act is about, e.g., a piece of information, a situation, an action, an event, an object, a property, or a relation.	Annotated.
Dependence relations	Show that one dialogue act depends on an earlier act or segment for its interpretation. For example, an answer can be linked to the question it answers.	Not annotated.
Qualifiers	Qualifiers certainty, conditionality, and sentiment can be added as additional information to a communicative function.	Sentiment is annotated for each functional segment or a stretch of segments.

Rhetorical relations	Show how one act is semantically or pragmatically connected to another, for example relations between dialogue acts, such as cause, explanation, justification, elaboration, contrast, or example.	Not annotated.
----------------------	--	----------------

Practical annotation workflow:

1. Identify the smallest functional segment.
2. Decide which dimension the segment mainly belongs to.
3. Assign the communicative function in context.
4. If the communicative function is inform, assign one semantic-content tag.
5. If the segment seems multifunctional, try to split it. If it cannot be split, annotate only the dominant function in context.
6. Do not annotate mechanically on the basis of the transcription or word form alone. Always listen to the speech and make annotation decisions based on acoustic impression and context.

3 Segmentation guidelines

3.1 Basic principle

In the present guidelines, speech is segmented into communicative units, with C-units used as the basic/default segmentation unit. A C-unit is defined as an independent main clause together with all subordinate clauses that depend on it (Biber et al. 1999).

This syntactically based principle is not part of the ISO 24617-2 segmentation procedure. ISO 24617-2 defines annotation units semantically and pragmatically, as functional segments, that is, minimal stretches of communicative behaviour that express one or more communicative functions. Applying this definition directly, however, requires the annotator to first identify the communicative function and only then determine the segment to which this function applies. Such a procedure requires highly qualified and experienced annotators.

Based on our annotation experience, C-units provide a useful syntactic basis for identifying likely boundaries between functional segments. In spontaneous speech, however, syntax alone is not sufficient. Prosody also provides important segmentation cues. Ultimately, the proposed boundaries are evaluated against the functional definition of the unit. Common cases in which clauses and non-clausal material frequently combine to form a single functional unit or split to form two units are defined in Sections 3.2 to 3.14 below.

3.2 Keep one main clause with its dependent clauses

Keep a main clause together with subordinate clauses that complete, specify, or restrict it.

Examples:

- Me zanima, kakšna je njihova kebab pica.
[‘I wonder what their kebab pizza is like.’]
- Saj veš, da če samo dosti počasi scrollaš, pol ne opaziš.
[‘You know that if you scroll slowly enough, then you do not notice.’]
- Če bo pa nam kaka karta ostala, bomo pa dali.
[‘If we have any ticket left, we will give it away.’]

Do not split subordinate clauses mechanically. Split them only when they function as a separate communicative unit.

3.3 Split coordinated independent clauses when each gives a new step or action

When coordinated clauses each express a separate action, step, or piece of information, segment them separately.

Examples:

- Tam greš na ono spletno stran od Ryanaira,
[‘There you go to that Ryanair webpage,’]

- pogledaš datume,
[‘you check the dates,’]
- pogledaš cene
[‘you check the prices’]
- in pol si izbereš.
[‘and then you choose.’]

This is especially common in instructions, narratives, and lists of actions.

3.4 Keep tightly integrated coordination together

Do not split coordination when the first part is incomplete without the second, especially in alternatives, contrasts, or fixed comparative structures.

Examples:

- Ali boš šla samo ti ali bo šla samo Nadja ali samo Mario.
[‘Will only you go, or only Nadja, or only Mario.’]
- Ne samo one velike vreče po petindvajset kil, imajo tudi majhna pakiranja.
[‘They do not only have those large twenty-five-kilo bags, they also have small packages.’]

If the coordinated parts form one choice, comparison, or combined meaning, keep them together.

3.5 Split explanatory, causal, result, and purpose clauses when they are independent enough

Subordinate clauses that express a different communicative function from the main clause to which they are connected should be segmented as separate units. This is especially common with clauses that add an explanation, reason, consequence, or purpose to previously provided information.

Examples:

- TikTok nimam,
[‘I do not have TikTok,’]
- ker vidiš, kake pišejo.
[‘because you see what they write.’]

Another example:

- Ne bo odprl nobene spletne strani,
[‘It will not open any webpage,’]
- tako da ga sploh ne moreš apdejtati.
[‘so you cannot update it at all.’]

Do not split causal clauses automatically. If the clause is tightly integrated into the C-unit and does not form its own communicative move, keep it together.

3.6 Split discourse markers when they are prosodically independent

Discourse markers are segmented separately when they are prosodically separated, emphasized, turn-initial, or used to organize the discourse.

Examples:

- Kaj zdaj,
[‘What now,’]
- si kupila karte za na Sardinijo?
[‘did you buy tickets for Sardinia?’]

Another example:

- No,
[‘Well,’]
- ma tam si naloži tisto aplikacijo od Ryanair
[‘there you download that Ryanair app’]

Common discourse markers include:

ja, no, aja, glej, ej, zdaj, dobro, tako da.

Keep the marker in the same segment if it is prosodically fused with the rest and does not perform a separate discourse function.

Example:

- Ja ne vem.
[‘Well I don’t know.’]

3.7 Split feedback tokens when they function independently

Short forms such as mhm, aha, aja, ja, okej, and ne are separate segments when they provide feedback, confirmation, agreement, or response on their own.

Examples:

- Aha.
[‘Aha.’]
- Ja, ja.
[‘Yes, yes.’]
- Mhm.
[‘Mhm.’]

3.8 Segment short questions and question tags separately when they function independently

Short question-like elements are separate segments when they ask for confirmation, check understanding, or invite response.

Examples:

- Ne?
[‘Right?’]
- A ne?
[‘Right?’]
- Veš?
[‘You know?’]
- A imaš?
[‘Do you have it?’]

If the question tag is syntactically and prosodically integrated into the preceding clause, it may remain in the same segment.

3.9 Treat elliptical and verbless units as possible segments

Spoken language often contains fragments without a finite verb. Segment them separately when they perform their own communicative function, such as answering, specifying, evaluating, identifying, or adding information.

Examples:

- Na pregled.
[‘For a check-up.’]
- Bančno?
[‘Bank one?’]
- Z zajtrkom?
[‘With breakfast?’]
- Polno plaž.
[‘Lots of beaches.’]
- Ja, super.
[‘Yes, great.’]

Fragments should not be dismissed as incomplete. If they communicate something on their own, they are valid segments.

3.10 Split afterthoughts, appositions, and added specifications when they are prosodically separate

If a speaker adds extra information after the main unit and it sounds like an afterthought, apposition, clarification, or specification, segment it separately.

Examples:

- Tisti jajc je bil tako.
[‘That egg was like that.’]

- Malo brezveze.
[‘A bit pointless.’]

Another example:

- Jaz sem vzel pico.
[‘I took a pizza.’]
- s slanino, šunko.
[‘with bacon, ham.’]

Keep the added material in the same segment if it is syntactically integrated and does not sound independent.

3.11 Do not split ordinary hesitations unless they have their own function

Most hesitations, lengthenings, repetitions, and filled pauses remain inside the surrounding segment.

Example:

- Lahko eee z doplačilom,
[‘Possibly uh with an extra payment,’]

However, segment filled pauses separately when they are salient, isolated, or clearly used for stalling.

Examples:

- eee
[‘uh’]
- Zdaj eem ...
[‘Now erm ...’]

A silent pause alone is not normally annotated as a separate segment. Detailed annotation of hesitation phenomena is left to disfluency annotation.

3.12 Segment repairs, abandoned starts, and self-corrections when the speaker restarts or withdraws material

If the speaker abandons a structure or starts again, the abandoned or corrected material may form a separate segment.

Examples:

- Koliko cajta ...
[‘How much time ...’]
- koliko cajta bi bila tam?
[‘how much time would you be there?’]

Another example:

- Pa kaj b()
[‘And what w()’]
- kje boste pa parkirali?
[‘where will you park?’]

Do not segment every small self-repair, but only when there is a real abandoned form, correction, restart, or withdrawn formulation. Detailed annotation of self-repair phenomena is left to disfluency annotation.

3.13 Reporting frames and direct speech may be segmented separately

When a speaker introduces quoted or reported speech, the reporting frame may be a separate segment, especially before direct speech.

Examples:

- Je rekla:
[‘She said:’]
- Sem rekel:
[‘I said:’]
- In grem hitro atiju:
[‘And I quickly say to dad:’]

The quoted material is then segmented according to its own communicative structure.

3.14 Nonverbal events are separate only when communicatively relevant

Events such as laughter, voice noises, or other nonverbal material may be separate segments if they carry communicative value.

Examples:

- [smeh]
[‘[laughter]’]
- [glas]
[‘[voice/sound]’]

If the nonverbal event only accompanies the surrounding speech and does not function independently, keep it with the relevant speech segment or do not treat it as a separate dialogue-act segment.

3.15 Discontinuous segments

If one functional segment is interrupted by another functional segment, use plus signs to mark the interrupted segment. The first part ends with +, and the continuation begins with +.

ja kaj pa+ | eee | +hotel? ['yes what about+ | uh | +the hotel?']

- ja kaj pa+ -> segmentID:0001, dimension:Task, communicativeFunction:setQuestion
- eee -> segmentID:0002, dimension:TimeManagement, communicativeFunction:stalling
- +hotel? -> segmentID:0001, dimension:Task, communicativeFunction:setQuestion

3.16 Practical decision checklist

For each possible boundary, ask:

1. Does the material form one main clause with its dependent clauses?
→ Keep together.
2. Does the next part perform a new action, answer, question, evaluation, instruction, or discourse move?
→ Split.
3. Is it an independent discourse marker, feedback token, hesitation, repair, or question tag?
→ Split if it has its own function.
4. Is it only a hesitation or small disfluency inside a larger unit?
→ Keep inside the larger segment.
5. Is it an afterthought, apposition, short specification, or fragment with its own meaning?
→ Split.
6. Is the boundary unclear?
→ Prefer the C-unit, unless prosody or communicative function clearly supports a split.

4 Annotating dimension

The dimension level says what kind of thing the speaker is doing in the conversation. For example, is the speaker: talking about the main topic/task, showing understanding, asking whether the other person understood, taking or giving the turn, buying time, managing social routines?

4.1 Task

Use the Task dimension when the segment contributes to the main conversational content, activity, or practical purpose of the dialogue.

Examples:

- Bolj poceni je. ['it is cheaper.']
- Piše na aplikaciji od Ryanaira. ['it says on the Ryanair app.']
- A različne so? ['Are they different?']

Tip 1: If the speaker explicitly informs the hearer about the ongoing management of the conversation, the segment should not be annotated in the Task dimension, even if it has the communicative function inform. Instead, annotate it as Discourse Structuring, because the content concerns the organization of the dialogue rather than the topic itself.

Examples:

- Bom ti to pozneje povedal. ['I will tell you that later.']
- Dosti o tem. ['Enough about that.']
- Zdaj ne me spraševati. ['Now do not ask me.']
- Saj nima veze. ['It does not matter.']
- Zdaj sva malo razlagala bolj take koncerte, ne, rokerske. ['Now we have been explaining more those kinds of concerts, right, rock ones.']
- Drugače pa, ko si me vprašala, kaka je moda. ['Otherwise, when you asked me what fashion is like.']

Tip 2: Although an utterance may have the form of a question, it can in fact express an opinion or stance, for example surprise or disbelief and should be annotated as Inform communicative function.

Examples:

- A res? ['Really?'] – expressing surprise
- Komu se da. ['Who would bother.']
- Kdo bi dal toliko denarja za telefon, k ima šestdeset hercev. ['Who would pay that much money for a phone that has sixty hertz.']

Tip 3: Some question-like utterances function as metadiscourse and should therefore be annotated in the Discourse Structuring dimension. These utterances do not primarily ask for task-related information. Instead, they help organize the conversation: they introduce something the speaker is

about to say, return to a previous topic, search for a word, check the previous formulation, or manage the flow of discourse.

Examples:

- Veš, kaj ti moram povedati? ['Do you know what I have to tell you?']
- Kje sva ostala? ['Where were we?']
- Kako je že beseda? ['What's the word again?']
- Kako si rekel? ['How did you say it?']

4.2 Auto-Feedback

Use when the speaker signals their own processing of what the other participant said: hearing, understanding, accepting, or registering the previous contribution.

Examples:

- Mhm. ['Mhm.']
- Aha. ['Aha.']
- Aja. ['Oh, I see.']

Tip: Short responses are Auto-Feedback only if they mean “I hear/understand/register what you said”.

4.3 Allo-Feedback

Use when the speaker checks or comments on the addressee's processing.

Examples:

- ne? ['right?']
- veš? ['you know?']
- saj veš ['you know already']

Tip: Auto-Feedback = “I understand you”; Allo-Feedback = “Do you follow me?”

4.4 Turn Management

Use when the segment regulates the speaker role: taking, accepting, keeping, releasing, assigning, or grabbing the turn.

Examples:

- Ja, ['Yes,'] when taking or accepting the turn
- No, ['Well,'] when starting a turn
- Čakaj, ['Wait,'] when grabbing or holding the floor

Tip: Turn functions are position-dependent.

4.5 Time Management

Use when the segment helps the speaker gain time or suspend speaking briefly.

Examples:

- eee ['uh']
- mislim, ['I mean,']
- čakaj malo, ['wait a moment,']

Tip: Use Time Management for hesitation or planning, not for every discourse marker.

4.6 Discourse Structuring

Use when the segment organizes the conversation: opens, shifts, resumes, closes, or transitions.

Examples:

- Kaj zdaj, ['So what now,']
- No, ['Well,']
- Ja, glej. ['Yes, look.']
- Veš, kaj se je še meni zgodilo? ['Do you know what else happened to me?']

Tip: Use this when the form organizes the discourse rather than providing feedback or topic content.

4.7 Social Obligations Management

Use for conventional social acts such as greetings, goodbyes, thanking, apologies, and responses to these acts.

Examples:

- Hej. ['Hey.']
- hvala ti. ['thank you.']
- ni za kaj. ['you are welcome.']

Tip: Use only when the social act is actually performed.

4.8 Own Communication Management

Use when the speaker manages problems in their own speech: abandoned beginnings, retractions, errors, and self-corrections.

Examples:

- Na l() eee na ['On l() uh on']
- ne, ['no,'] when correcting one's own previous formulation

Tip: If the speaker is only gaining time, use Time Management.

4.9 Partner Communication Management

Use when the speaker completes or corrects the other participant's speech.

Examples:

- Etiketo. ['The label.']
- Poganjke. ['The shoots.']
- Filme. ['Films.']

Tip: Do not use this for ordinary overlap or supportive feedback.

4.10 Unspecified Dimension

Use only when no specific dimension can be assigned, especially when the segment is unintelligible or abandoned before its purpose can be recognized.

Examples:

- No ... ['Well ...']
- pa ... ['and ...']
- ne pa ... ['not but ...']

Tip: This is a last-resort label.

4.11 Multidimensionality

ISO 24617-2 supports multidimensional annotation. In the present practical annotation approach, however, multiple dimensions and multiple communicative functions are not assigned to the same segment.

If a segment appears to belong to more than one dimension or to express more than one function, first check whether it can be divided into smaller functional segments. If different parts perform different functions, annotate the parts separately. If the same words could perform more than one function at the same time, assign only the single dimension and function that best represent the main communicative role in context.

Do not annotate implied, secondary, or mechanically inferable functions. For example, if eee 'uhm' both buys time and helps the speaker keep the turn, annotate the function that is most central in context: Time Management if the main role is stalling, Turn Management if the main role is keeping or taking the floor.

4.12 Practical decision checklist

Main role of the segment	Dimension
Contributes to the main conversational content, activity, or practical purpose of the dialogue	Task
Shows the speaker's own understanding or uptake	Auto-Feedback
Checks or comments on the addressee's understanding	Allo-Feedback
Manages who speaks	Turn Management
Gains time or briefly suspends speaking	Time Management
Organizes topic, transition, or discourse flow	Discourse Structuring

Performs greeting, thanking, apology, goodbye, or similar social act	Social Obligations Management
Repairs or manages the speaker's own speech	Own Communication Management
Completes or corrects the partner's speech	Partner Communication Management
Cannot be classified reliably	Unspecified Dimension

5 Annotating communicative functions

After choosing the dimension, assign the communicative function. The communicative function says what the speaker is doing with the segment. Always annotate the function in context, not from the word form alone. In this simplified annotation scheme, each segment receives only one dimension and one communicative function. If a segment seems to express more than one function, split it if possible; otherwise choose the function that is most central in context.

5.1 Task functions

Use Task functions when the segment contributes to the conversational content, activity, or practical purpose. General-purpose functions are listed here with the Task dimension because this is how they are used in the present scheme.

Dimension / function	Use when	Examples
Task / question	Use for a Task question when a more specific question type cannot be assigned.	A nisi imela kaj bolj pametnega za barati? ['Didn't you have anything smarter to ask?']
Task / propositionalQuestion	Use for yes/no questions where the expected answer is yes or no.	Ste imeli kakšne igre? ['Did you have any games?']
Task / setQuestion	Use for open questions asking for a value, person, place, reason, manner, quantity, or other content.	Kaj se lahko tudi zdaj sadi? ['What can also be planted now?']
Task / checkQuestion	Use when the speaker checks whether an expected proposition is correct.	Ti si bil na karavli, ne? ['You were at the border post, right?']
Task / choiceQuestion	Use when the speaker offers alternatives.	Pixel eight ali Pixel nine? ['Pixel eight or Pixel nine?']

Questions about speech production, hearing, understanding, or discourse management should first be checked against the relevant non-Task dimension, for example Own Communication Management, Auto-Feedback, or Discourse Structuring.

Dimension / function	Use when	Examples
Task / inform	Use when the speaker provides topic-related information that is not the first direct response to a question. This is the most frequent tag and covers a wide range of content, including facts, opinions, arguments, conditions, explanations, emotional states, and similar information. Add Semantic Content tag to annotate in more details.	Piše na aplikaciji od Ryanaira. ['It says on the Ryanair app.']
Task / answer	Use only for the first segment that directly responds to an information-seeking question. Longer follow-up explanations are usually annotated as Inform.	Za prvi maj še ne. ['Not yet for the first of May.']
Task / confirm	Use when the speaker confirms an expected proposition, usually after a check question.	Ja, točno. ['Yes, exactly.']
Task / disconfirm	Use when the speaker rejects an expected proposition or gives a negative confirmation response.	Ne, tisto je črno. ['No, that one is black.']
Task / agreement	Use when the speaker aligns with a previous statement, opinion, or evaluation.	To je res. ['That is true.']
Task / disagreement	Use when the speaker does not align with a previous statement, opinion, or evaluation.	To ni fajn. ['That is not good.']
Task / correction	Use when the speaker corrects previous information.	iPad OS. ['iPad OS.']

If the segment mainly expresses what the speaker intends or needs to do, consider whether a commissive function is more appropriate than inform: Bom šla čez vikend pol. ['I will go during the weekend then.']

If it mainly tells the addressee what they should, could, or need to do, consider whether a directive function is more appropriate: In tam greste na izlet z ladjo. ['And there you go on a boat trip.']

Dimension / function	Use when	Examples
Task / request	Use when the speaker asks the addressee to do something.	Daj, please. ['Come on, please.']
Task / instruct	Use when the speaker tells the addressee how to do something, often as a step in a procedure.	In pol izbereš karto. ['And then you choose the ticket.']
Task / suggest	Use when the speaker proposes a possible action, usually leaving the addressee a choice.	Drugače pa lahko še korenček seješ. ['Otherwise you can also sow carrots.']
Task / addressOffer	Use for a response that deals with an offer but does not clearly accept or decline it.	A res lahko? ['Can you really?']
Task / acceptOffer	Use when the speaker accepts an offer.	V redu. ['All right.']
Task / declineOffer	Use when the speaker declines an offer.	Raje ne. ['I'd rather not.']
Dimension / function	Use when	Examples
Task / offer	Use when the speaker offers to do something for the addressee.	Lahko ti jo pošljem. ['I can send it to you.']
Task / promise	Use when the speaker commits themselves to doing something.	Bomo. ['We will.']
Task / addressRequest	Use for a response that deals with a request but does not clearly accept or decline it.	Bom pogledala. ['I will check.']
Task / acceptRequest	Use when the speaker accepts a request.	Ja. ['Yes.']
Task / declineRequest	Use when the speaker refuses a request.	Ne, tega ne morem. ['No, I cannot do that.']
Task / addressSuggest	Use for a response that deals with a suggestion but does not clearly accept or decline it.	Bom še premislila. ['I will think about it.']
Task / acceptSuggest	Use when the speaker accepts a suggestion.	Ni problema. ['No problem.']
Task / declineSuggest	Use when the speaker rejects a suggestion.	Ne, to ne morem. ['No, I cannot do that.']

5.2 Feedback functions

Feedback functions describe how participants process previous dialogue. Auto-Feedback concerns the speaker's own processing; Allo-Feedback concerns the addressee's processing.

Dimension / function	Use when	Examples
Auto-Feedback / autoPositive	Use when the speaker signals that they heard, understood, accepted, or registered the previous contribution.	Aha. ['Aha.']/ Mhm. ['Mhm.']
Auto-Feedback / autoNegative	Use when the speaker signals a problem in their own hearing or understanding.	Kaj? ['What?']
Allo-Feedback / alloPositive	Use when the speaker indicates that the addressee has understood or processed the previous contribution correctly.	Ja, to sem mislila. ['Yes, that is what I meant.']
Allo-Feedback / alloNegative	Use when the speaker indicates that the addressee has not understood or processed the previous contribution correctly.	Ne, narobe si razumel. ['No, you understood it wrongly.']
Allo-Feedback / feedbackElicitation	Use when the speaker checks whether the addressee follows, understands, or agrees enough to continue.	A ne? ['Right?']/ Saj veš. ['You know.']

Short forms such as mhm, aha, ja, and okej are feedback only when they acknowledge or check processing. If they answer a question, confirm information, or accept an offer, use a Task function instead.

5.3 Turn Management functions

Use Turn Management when the segment regulates who has the speaker role.

Dimension / function	Use when	Examples
Turn Management / turnTake	Use when the speaker starts speaking at a normal transition point.	Ja, dobro. ['Yes, all right.']
Turn Management / turnAccept	Use when the speaker accepts the turn after it has been assigned or made available.	Aja, ['Oh,']
Turn Management / turnGrab	Use when the speaker takes the floor before the previous speaker has clearly finished.	No, ['Well,']
Turn Management / turnAssign	Use when the speaker explicitly gives the turn to someone else.	Ti povej. ['You tell it.']
Turn Management / turnKeep	Use when the speaker signals that they have not finished and wants to keep the floor.	In zdaj, ['And now,']
Turn Management / turnRelease	Use when the speaker makes the floor available to the other participant.	Tako da ... ['So ...']

5.4 Time Management functions

Use Time Management when the speaker gains time or briefly suspends the dialogue.

Dimension / function	Use when	Examples
Time Management / stalling	Use when the speaker buys time while planning what to say.	Eee. ['Uh.'] / Eem. ['Erm.']
Time Management / pausing	Use when the speaker explicitly suspends speaking for a short time.	Čakaj malo. ['Wait a moment.']

5.5 Discourse Structuring functions

Use Discourse Structuring when the segment organizes the flow of the conversation rather than contributing directly to the topic.

Dimension / function	Use when	Examples
Discourse Structuring / interactionStructuring	Use when the speaker shifts, resumes, closes, summarizes, or marks a transition in the conversation.	Dosti o tem. ['Enough about that.'] / Drugače pa ... ['Otherwise ...']
Discourse Structuring / opening	Use when the speaker opens an interactional move or attracts attention at the beginning of a move.	Glej, ['Look,'] / Ej, ['Hey,']

If an inform segment states what will happen, has happened, or should happen within the conversation itself, annotate it as Discourse Structuring, not Task: Bom ti to pozneje povedal. ['I will tell you that later.']

5.6 Social Obligations Management functions

Use Social Obligations Management for conventional social acts and responses to them.

Dimension / function	Use when	Examples
Social Obligations Management / initGreeting	Use when the speaker initiates a greeting.	Hej. ['Hey.']
Social Obligations Management / returnGreeting	Use when the speaker returns a greeting.	Bok, bok. ['Hi/bye, hi/bye.']
Social Obligations Management / initSelfIntroduction	Use when the speaker introduces themselves.	Jaz sem Ana. ['I am Ana.']
Social Obligations Management / returnSelfIntroduction	Use when the speaker responds with their own self-introduction.	Jaz pa Maja. ['And I am Maja.']
Social Obligations Management / apology	Use when the speaker apologizes.	Sorry. ['Sorry.']
Social Obligations Management / acceptApology	Use when the speaker accepts or downplays an apology.	Nič hudega. ['No problem.']
Social Obligations Management / thanking	Use when the speaker thanks someone.	Hvala ti. ['Thank you.']

Dimension / function	Use when	Examples
Social Obligations Management / acceptThanking	Use when the speaker responds to thanks.	Ni za kaj. ['You're welcome.']
Social Obligations Management / initGoodbye	Use when the speaker initiates a goodbye.	Adijo. ['Goodbye.']
Social Obligations Management / returnGoodbye	Use when the speaker returns a goodbye.	Bok. ['Bye.']

Use these functions only when the social act is actually performed. A quoted or ironic hvala or sorry should be annotated according to its function in context.

5.7 Own Communication Management functions

Use Own Communication Management when the speaker manages problems in their own speech.

Dimension / function	Use when	Examples
Own Communication Management / selfError	Use for an error or problematic production that is marked as such in the speaker's own speech.	Pol p(), ['Then p(),']
Own Communication Management / retraction	Use when the speaker abandons or withdraws part of their own speech.	Koliko cajta ... ['How much time ...']
Own Communication Management / selfCorrection	Use when the speaker corrects their own previous formulation.	Eee oziroma ne letos, lani. ['Uh, or rather not this year, last year.']

If the speaker is only gaining time, use Time Management. Use Own Communication Management only when there is an error, abandoned formulation, retraction, or correction.

5.8 Partner Communication Management functions

Use Partner Communication Management when the speaker completes or corrects the other participant's speech.

Dimension / function	Use when	Examples
Partner Communication Management / completion	Use when the speaker supplies a missing or unfinished part of the partner's utterance.	Etiketo. ['The label.']
Partner Communication Management / correctMisspeaking	Use when the speaker corrects the partner's spoken formulation.	Skencalali. ['Cancelled.']

5.9 Unspecified function

Dimension / function	Use when	Examples
Unspecified Dimension / unspecifiedFunction	Use only when no specific dimension or communicative function can be assigned, usually because the segment is unintelligible, incomplete, or cannot be interpreted reliably.	No ... ['Well ...']

Before using this label, check whether the segment may instead be stalling, retraction, interactionStructuring, autoPositive, or another contextually recoverable function.

6 Annotating semantic content for Inform function

In ISO 24617-2, semantic content is what the dialogue act is about: information, an action, an event, a situation, an object, a property, or a relation. In these guidelines, semantic content is annotated only for segments with the communicative function inform. These labels are semantic-content types, not additional communicative functions.

6.1 General rules

- Apply one semantic-content label only to segments already annotated as inform.
- Assign exactly one label per segment.
- If one segment contains two clearly separable content types, split it if possible.
- If the segment cannot be split, choose the first clearly applicable label in the priority order.
- Uncertainty, approximation, weak recall, or guessing are not separate semantic-content labels in this scheme.

Overview of labels:

1. reportedContent
2. emotionalExclamation
3. selfExperience
4. evaluationStance
5. humourIrony
6. requirementConstraint
7. possibilityOption
8. conditionalScenario
9. causalExplanation
10. referentialSpecification
11. eventAction
12. stateDescription

6.2 Semantic-content labels

6.2.1 reportedContent

Use when the segment reports, quotes, paraphrases, or refers to what someone said, asked, thought, claimed, wrote, or generally says.

Ask: Whose speech or thought is being reported?

Examples:

- Je rekla: ['She said:']
- Sem rekel: ['I said:']
- Je rekla hčera, da naj ji to prišparam. ['The daughter said that I should save this for her.']
- Pač tako pravijo. ['That is what they say.']

Boundary rule: Use `reportedContent` whenever the segment is part of reported speech or reported thought, even if the reported content itself contains an event, opinion, condition, requirement, or explanation. This applies to all segments that belong to the reported content.

6.2.2 `emotionalExclamation`

Use when the segment is mainly a short emotional or expressive reaction, such as surprise, disbelief, frustration, admiration, annoyance, shock, or excitement.

Ask: Is the speaker mainly reacting emotionally rather than describing, explaining, or evaluating something in detail?

Examples:

- Jezus. ['Jesus.']
- Pa nena reči. ['Do not say that.']
- O fak. ['Oh fuck.']
- Pff. ['Pff.']
- Ma mati božja. ['Oh Mother of God.']

Boundary rule: Use `emotionalExclamation` for brief expressive outbursts. Use `evaluationStance` if the segment mainly evaluates something, and `selfExperience` if it describes the speaker's own feeling or state.

6.2.3 `selfExperience`

Use when the segment informs about a participant's own knowledge, memory, perception, emotion, preference, interest, bodily state, ability, or personal experience.

Ask: What does the participant know, remember, feel, like, perceive, experience, or not know?

Examples:

- Ne vem. ['I do not know.']
- Se niti ne spomnim. ['I do not even remember.']
- Take španske so mi všeč. ['I like those Spanish ones.']
- Komaj čakam, da jo vidim. ['I can hardly wait to see her.']

Boundary rule: Use `evaluationStance` when the speaker mainly evaluates an external person, object, event, or situation.

6.2.4 `evaluationStance`

Use when the segment expresses a direct evaluation, opinion, judgment, stance, attitude, or subjective assessment of something.

Ask: How is the speaker evaluating something?

Examples:

- Ja, super. ['Yes, great.']
- Je pač res grozno. ['It is really terrible.']
- Najboljši so bili pingvini. ['The penguins were the best.']

Boundary rule: Use selfExperience for the participant's own feeling, memory, perception, or bodily condition. Use humourIrony when the main point is humorous, ironic, teasing, or joking rather than a direct evaluation.

6.2.5 humourIrony

Use when the segment mainly expresses humour, irony, joking, teasing, playful exaggeration, sarcasm, or a non-literal humorous stance.

Ask: Is the speaker mainly joking, teasing, being ironic, or making a humorous comment?

Examples:

- Somebody check on her. ['Somebody check on her.']

Boundary rule: Use evaluationStance when the segment expresses a straightforward opinion or judgment. Use humourIrony only when the humorous, ironic, teasing, or non-literal stance is central in context.

6.2.6 requirementConstraint

Use when the segment informs about a necessity, obligation, rule, restriction, requirement, constraint, or required step.

Ask: What must be done, what is required, or what is not allowed?

Examples:

- Moraš imeti aplikacijo. ['You have to have the application.']
- Pa mora biti prav napisano. ['And it has to be written correctly.']
- Morajo biti vse črke prav. ['All letters have to be correct.']

Boundary rule: Use possibilityOption for can/may/available-choice meanings.

6.2.7 possibilityOption

Use when the segment informs about an available option, possibility, permission, choice, or something that can be done.

Ask: What is possible, available, allowed, or optional?

Examples:

- In tam lahko najameš avto. ['And there you can rent a car.']
- Pijačo si lahko kupiš tudi na letalu. ['You can also buy a drink on the plane.']
- Imaš na izbiro oni bungalov na plaži. ['You have that bungalow on the beach as an option.']

Boundary rule: Use conditionalScenario when the main point is an if/then or imagined scenario.

6.2.8 conditionalScenario

Use when the segment describes a conditional, hypothetical, generic, imagined, or non-actual scenario.

Ask: What would happen if something were the case?

Examples:

- Če se izgubi. ['If it gets lost.']
- Če daš sladkor, prej kipne. ['If you add sugar, it boils over sooner.']
- Če bi bil Windows, bi se povezal. ['If it were Windows, it would connect.']

Boundary rule: Use possibilityOption if the main point is simply that something is possible; use requirementConstraint if the main point is what must be done.

6.2.9 causalExplanation

Use when the segment informs about a cause, reason, explanation, justification, purpose, or consequence.

Ask: Why? Because of what? For what purpose? With what result?

Examples:

- Ker smo imeli tam mi tudi problem. ['Because we also had a problem there.']
- Zato morajo biti standardne mere. ['That is why they have to be standard measurements.']
- Tako da ga sploh ne moreš apdejtati. ['So you cannot update it at all.']

Boundary rule: Use causalExplanation only when the reason, purpose, explanation, or consequence is the main contribution.

6.2.10 referentialSpecification

Use when the segment mainly names, identifies, classifies, defines, points to, or specifies a person, object, place, event, concept, time, number, quantity, measure, price, duration, order, or location.

Ask: What exactly is being identified or specified?

Examples:

- Mercedes CVS. ['Mercedes CVS.']
- Magnetno resonanco. ['Magnetic resonance imaging.']
- To je bil Almost Heaven. ['That was Almost Heaven.']
- V Zagrebu, v Domu sportova. ['In Zagreb, in Dom sportova.']

Boundary rule: Use this mainly for short fragments or segments whose main role is identification or specification. If the segment reports an action, use eventAction; if it describes a property or situation, use stateDescription.

6.2.11 eventAction

Use when the segment informs about an event, action, happening, activity, movement, visit, purchase, trip, or something that happened, is happening, or will happen.

Ask: What happened, is happening, or will happen?

Examples:

- Ful smo se vozili z vlaki. ['We travelled a lot by train.']
- Zdaj sem pa kupil jih še pet. ['Now I bought five more of them.']
- Ona je takrat dobila karte. ['She got the tickets then.']

- Pol so nas nagnali. [‘Then they chased us away.’]

Boundary rule: Use stateDescription for static properties, circumstances, possession, availability, or descriptions.

6.2.12 stateDescription

Use when the segment informs about a state, situation, circumstance, property, possession, availability, description, or how something is or was.

Ask: What is or was the case? What is or was something like?

Examples:

- Je bilo kar dosti ljudi. [‘There were quite a lot of people.’]
- Tudi nismo imeli najboljše pozicije. [‘We also did not have the best position.’]
- Air pomeni, da nima ventilatorja. [‘Air means that it does not have a fan.’]
- Polno plaž. [‘Lots of beaches.’]

Boundary rule: Use stateDescription as the final default factual-content tag when none of the more specific labels applies.

8 Conclusion

These guidelines provide a practical framework for annotating dialogue acts in Slovenian spoken interaction. They adapt ISO 24617-2 to the needs of the ROG corpus and the SloBench ASR evaluation dataset by combining a function-based view of dialogue acts with clear rules for segmentation, dimension selection, communicative-function annotation, and semantic-content annotation.

The annotation is multilayered, but the layers are organized in a top-down way. First, the annotator selects the dimension, which predefines the set of possible communicative functions. Then, one communicative function is selected within that dimension. Only one communicative function, *inform*, is further subdivided into semantic-content labels, which make this very broad category more informative while keeping the overall scheme manageable.

The main principle throughout the annotation process is to interpret each segment in context. Annotators should not rely only on word form, punctuation, or grammatical structure alone, but should also consider prosody, sequential position, speaker intention, and the role of the segment in the interaction. When a segment appears multifunctional, it should first be split if possible; if not, only the most central dimension and communicative function should be assigned.

This simplified approach preserves the main strengths of ISO 24617-2 while making annotation practical and consistent for everyday spoken dialogue. It supports the annotation of topic-related content, feedback, turn and time management, discourse structuring, repairs, social acts, and fine-grained semantic distinctions within *inform*. The resulting data can be used both for linguistic analysis and for the development of automatic dialogue-act annotation models.

9 References

- Allen, J., & Core, M. (1997). Draft of DAMSL: Dialog Act Markup in Several Layers.
- AMI. (2005). AMI Dialogue Act Annotation Scheme.
- Austin, J. L. (1975). *How to Do Things with Words*. Harvard University Press.
- Biber, D., Johansson, S., Leech, G., Conrad, S., & Finegan, E. (1999). *Longman Grammar of Spoken and Written English*. Longman.
- Bunt, H. (2021). DIT++ Taxonomy of Dialogue Acts, Annotation Scheme, and DiAML Markup Language, Release 5.2. Available at: <https://dit.uvt.nl/>
- Cresti, E. (2020). The pragmatic analysis of speech and its illocutionary classification according to the Language into Act Theory. In S. Izre'el, H. Mello, A. Panunzi, & T. Raso (Eds.), *In Search of Basic Units of Spoken Language: A Corpus-Driven Approach* (pp. 181–220). John Benjamins.
- ISO. (2012). ISO 24617-2:2012: Language resource management — Semantic annotation framework (SemAF) — Part 2: Dialogue acts. International Organization for Standardization.
- ISO. (2020). ISO 24617-2:2020: Language resource management — Semantic annotation framework (SemAF) — Part 2: Dialogue acts. International Organization for Standardization.
- Jurafsky, D., Shriberg, E., & Biasca, D. (1997). *Switchboard SWBD-DAMSL Shallow-Discourse-Function Annotation Coders Manual*.
- Searle, J. R. (1969). *Speech Acts: An Essay in the Philosophy of Language*. Cambridge University Press.
- Searle, J. R. (1979). *Expression and Meaning: Studies in the Theory of Speech Acts*. Cambridge University Press.
- Weisser, M. (2020). The DART Scheme. Available at: https://martinweisser.org/DART_scheme.html