
DOES MY QUESTION ANSWER YOUR ANSWER?

BIBLIOGRAPHIC REPORT

UE 705 - SUPERVISED PROJECT 2019-2020

Authors

Alena YAKAVETS
Axel DIDIER

Supervisors

Maxime AMBLARD
Maria BORITCHEV

December 20, 2019

Contents

1	Introduction to the subject	1
1.1	Dialogues	1
1.2	Questions	2
1.3	Answers	3
2	Context	4
2.1	Semantics of discourse	4
2.2	Semantics of dialogues	5
2.3	Challenges in question-answer analysis	5
2.3.1	Implicatures	6
2.3.2	The role of context in a dialogue	7
2.3.3	Non-sentential utterances	8
2.4	Machine learning	10
3	Dialogue annotation: project background	11
3.1	Coffee or Tea? Yes.	11
3.2	What's the answer: dialogue annotation	11
3.2.1	Annotation schema: questions	12
3.2.2	Annotation schema: features	12
3.2.3	Annotation schema: answers	13
3.2.4	Annotation experiments	13
3.2.5	Callouts	14
3.3	Where's the answer: dialogue annotation	14
3.3.1	Separating form from function in classification	15
3.3.2	Compatibility and mismatch in question-answer pairs	16
3.3.3	Ambiguities in answers classification	17
3.3.4	Text segmentation algorithm	17
3.3.5	Callouts	17
4	Future work prospects	19
	References	20

- *Do you want coffee or tea?*
- *Yes!*
- *Motivate your answer*
- *Go, answer! Go!*

1 Introduction to the subject

Dialogues are one of the core forms of human communication. They have for a long time attracted the minds of linguistic researchers, and have become one of the focuses in the field of natural language processing and artificial intelligence. Formalising spontaneous dialogues is a complex task, touching upon multiple scientific domains, including semantics, pragmatics, as well as logic and cognitive science.

This bibliography report is a pre-phase to a project, that is a logical continuation of two other yearly student works ("What's the Answer: Dialogue Annotation" [1] and "Where's the Answer: Dialogue Annotation" [2]), carried out at Université de Lorraine, and supervised by Maxime Amblard and Maria Boritchev. In a bigger frame, it fits within their more global researches of dynamic discourse modeling and SLAM project (Schizophréniee et Langage, Analyse et Modélisation) developed by the Sémagramme group at LORIA whose main goal is to analyze the conversations between schizophrenic patients and identify inconsistencies in their dialogues. All these works address the issue of semantic and pragmatic inconsistencies in a human dialogue, and attempts to find a way to detect and quantify this phenomenon (Boritchev and Amblard, 2018 [3]).

Our project involves tackling questions and answers annotation for spontaneous dialogues, incomprehension phenomena and machine learning experiments – focusing the notion of the answer. In this report we'll first talk about some base concepts for dialogues, questions, answers and the challenges of their interactions in the frame of natural language processing. We'll then move to a brief overview of some relevant scientific studies in this domain, and lastly give more details on the background of this project and our prospects for the work to be done during the next semester.

1.1 Dialogues

To have a common ground when talking about dialogues, we looked at how various dictionaries (Merriam-Webster, Apple Dictionary, Wikipedia, Oxford, Collins, Cambridge English dictionaries) define the term. The main shared idea is that a dialogue is a conversation between two or more people or an exchange between a person and something else (e.g. computer). If we focus more specifically on the purpose, a dialogue can be defined as a discussion between different parties in order to: negotiate, reach an agreement, explore a particular subject, find a resolution of a problem, exchange ideas or opinions, etc.

There exist various ways to approach dialogue definition, which can vary depending on the context. In his article "The Four Types of Conversations: Debate, Dialogue, Discourse, and Diatribe", David W. Angel suggests to distinguish a dialogue from other types of conversations, based on the direction and tone of the exchange: 1-way vs. 2-way, cooperative vs. competitive (Angel, 2016 [4]). Dialogues are described as 2-way cooperative conversations, as opposed to the three other types, as shown in the Figure 1.

Indeed, one can assume that in a regular dialogic conversation the agents cooperate with each other. Though there exist many scenarios, especially outside of task-oriented dialogues, where due to conflicting preferences, the participants can make certain moves that are not content-cooperative (e.g. lying, misleading implicatures, limited informativeness, not answering) (Lascares, 2012 [5]).

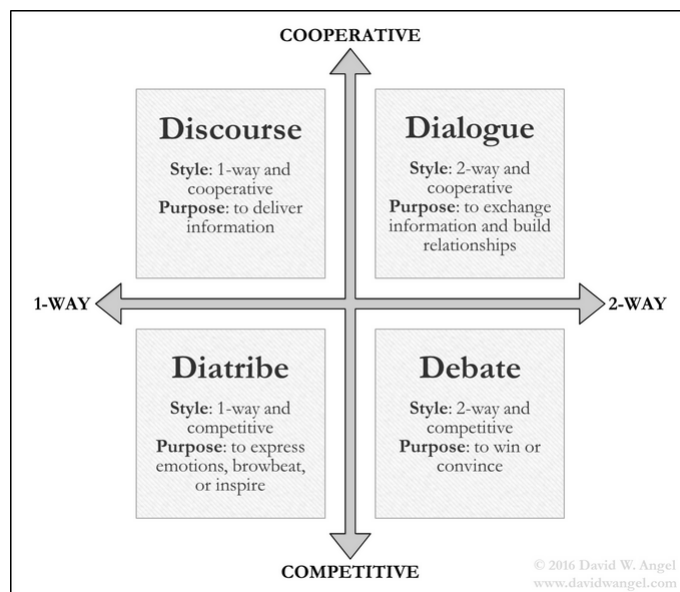


Figure 1: The Four Types of Conversations by David W. Angel

However, it is important to point out that a dialogue is not a succession of unrelated utterances, but still a cooperative effort, that can be defined by a common purpose or mutually accepted direction (which, in its turn, has flexibility to change over the course of the interaction) (Grice, 1975 [6]).

Spontaneous dialogues between humans are very vast in their nature. In addition to commonly observed features, one should not neglect to take into account conversational specifics that can occur due to personality type, cultural context, age-related capabilities and different mental health conditions of participants. Therefore it would be reasonable to assume that we should look at the broadest possible scales of definitions, without limiting ourselves only to cooperative conversations.

Regardless of different approaches in defining the scope or nature of a dialogue, one of the most observed phenomenon in this domain is the exchange of information, facilitated by the presence of questions and answers. To better understand what it involves, let's have a look at their respective definitions.

1.2 Questions

*"The man who asks a question is a fool for a minute,
the man who does not ask is a fool for life."
— Confucius*

Similarly to dialogues, to approach questions, we first referred to their dictionary definitions. A question is usually described as a sentence or phrase expressed in an interrogative form, addressed to someone in order to elicit information (in the form of an answer), or as a matter requiring resolution or discussion. If we look at the speech acts theory, a question can be understood as an **illocutionary** act in the field of pragmatics, or as a special kind of proposition in frameworks of formal semantics.

In Austin's framework, **locution** is what was said and meant, **illocution** is what was done, and **perlocution** is what happened as a result. When somebody says "Is there any salt?" at the dinner table, the illocutionary act is a request: "please give me some salt", even though the locutionary act (the literal sentence) was to ask a question about the presence of salt. The perlocutionary act (the actual effect), might be to cause somebody to pass the salt (Austin, 1975 [7]).

Questions play a fundamental role in a number of fields: linguistically, they make information exchange possible, allowing speakers to raise issues and to steer the conversation towards certain goals. Cognitively, questions drive the process of inquiry (Friedman, 2013 [8]), allowing us to pursue and achieve knowledge; additionally, they have been argued to play a key role in human reasoning. The role of questions in logic is perhaps less evident, but not less crucial (Ciardelli, 2017 [9]).

It is important to point out that in a dialogue not all questions receive a dedicated response, even if it's a properly formed question. On the other hand, certain types of questions are not supposed to receive an answer altogether: rhetorical questions are interrogative in form, but because they are not expected to be answered, one may say that they should not be considered true questions. In a way, certain non-interrogative grammatical structures may be viewed as questions too, like in the example of the following imperative sentence: "Tell me your address".

To avoid interpretation ambiguities with the named difficult cases, especially in context of machine learning and when prosodic information is not available, it would be safer to consider any interrogative sentence, ending with a question mark, as a question (even if it's an elliptical sentence).

1.3 Answers

*"Questions don't have to make sense, Vincent," said Miss Susan. "But answers do."
— Terry Pratchett, Thief of Time*

If you browse through dictionary definitions of answers, you would see them described as a statement (either spoken or written) that is made to reply or response to a question, request or accusation. If a question presupposes a solution, the answer is meant to be a solution to this problem (e.g. correct answer for a test or quiz). An answer can also be an action or behavior in return to a request, such as "Could you pass the salt, please?".

Judging by the above definitions we can already sense the challenge of identifying what and where exactly the answer is, and whether it is possible to objectively define whether an answer can be considered satisfactory. This task would seem much easier if we could set more precise boundaries.

H. Paul Grice in his work "Logic and Conversation" [6] introduces some maxims, that an answer should follow. **Quantity**: an answer should have high information density, i.e. be as informative as required and concise. **Quality**: an answer should not contain any false information or that with lack of evidence. **Relation**: an answer should be relevant, on point. **Manner**: an answer should not be ambiguous, obscure, but rather stay perspicuous. The author however admits that he has stated his maxims "as if this purpose were a maximally effective exchange of information; this specification is, of course, too narrow, and the scheme needs to be generalized to allow for such general purposes as influencing or directing the actions of others."

In many linguistic works an answer is often viewed in light of the question it belongs to, and it is expected to be coherent with it. Therefore they are often analysed within the frame of question-answer pairs. Vijayendra Mohanty, storyteller, looks at their interaction as follows: "An answer is anything that satisfies the conditions set by the question. However, this does not necessarily mean that an answer is something that will satisfy the questioner." (Vijayendra, 2014 [10]).

2 Context

Before going to the practical part of our supervised project next semester, we wanted to acquire some scientific culture on the study of the answers, and more generally on the acts of dialogue and conversational analysis.

Analysing the state of art allowed us to get familiar with different points of view of the scientists and researches on our subject and to orient us for the rest of our work.

2.1 Semantics of discourse

Discourse semantics emerged in an attempt to solve certain problems in formal theories of meaning for single sentences, which were mostly related to the interpretation of pronouns and other anaphoric elements. One could track a dependency between the meaning of an individual sentence and the information given by previous sentences in the discourse. Discourse semantics developed a formal analysis of a discourse context and of the interaction between the meaning of a sentence and the context in which it needs to be interpreted. The essential idea of discourse semantics is that the meaning of a sentence is a relation between contexts (Asher, 1998 [11]). We will present different points of view about the subject.

Discourse approaches actually started with certain types of sentences introducing ambiguous quantifications and anaphors, such as donkey sentences: "Every farmer who owns a donkey beats it.". Two readings are possible: only one donkey is beaten or there is one donkey for each farmer.

Richard Montague shows how English could be treated as a formal language. Montague develops the idea that a language follows a logical structure, and he links semantics, language and logic together (Montague, 1973 [12]). However, Montague's approach mishandles some linguistic phenomena, that are specific to discourse. In the semantics of traditional natural languages, only individual sentences are examined, but context also plays an important role in its meaning, such as:

(1) "Jerry entered the kitchen. Georges followed him. He opened the fridge".

"He" refers to individual constants introduced previously in order to obtain their meaning.

Montague's contribution to linguistic semantics, called Montague grammar, was later the basis of several studies. However Montague grammar did not cover such important notion as dynamics of language, some future studies lead to the resolution of this issue.

The first was the Discourse Representative Theory (DRT), developed in 1981 by Hans Kamp. It introduced an update of the context of interpretation by adding dynamical aspects (Kamp, 1981 [13]). The main idea is that DRT includes a level of abstract mental representations (discourse representation structures, DRS) within its formalism, which gives it a natural ability to handle meaning across sentence boundaries. There are two critical components to DRS: a set of discourse referents to represent entities under discussion, and a set of DRS conditions to represent information that has been given about discourse referents. However, it does not account the existence of discourse relations underlying the notion of discourse cohesion.

To retrieve this issue, Nicolas Asher and Alex Lascarides started developing and set up the first Segmented Discourse Representation Theory (SDRT) (Asher and Lascarides, 2003 [14]). The idea is to take into consideration rhetorical relations between segments of discourse. The SDRT is defined as an extension of DRS and is particularly suitable for the analysis of discourse, because it is fully formalized and it allows for computational modelling of language.

Discourses are dynamic things - new information gets communicated, affecting the state of the conversation and the states of minds of the conversational participants. de Groote enriched Montague semantics with a notion of context, that allows discourse dynamics to be tackled. He solves dynamicity-related issues with the idea of compositional meaning (de Groote, 2006 [15]). The Con-

tinuation Based Dynamic Semantics (CBDS, also called Type Theoretical Dynamic Logic (TTDL) in previous literature), is based on Church's simply typed λ -calculus (Church, 1940 [16]). His solution lies in the idea of solving dynamical issues by defining the remaining context in a specific way.

2.2 Semantics of dialogues

The semantics of dialogue is a fundamental subject that has been extensively studied from numerous points of view. We will introduce different works on this subject.

For Jonathan Ginzburg, a dialogue is the primary medium of language use (Ginzburg, 2016 [17]). Moreover, studying dialogue means a careful study of the nature of context. The context has a role to play in determining what one can or should say at a given point and also how to say it. For Ginzburg, there are two fundamental problems in the dialogue: conversational relevance and conversational meaning.

For David Schlangen, in dialogue, we use language to mean things for our interlocutors. When we mean something, the important component is the meaning of what we say. However, in different contexts, we can mean different things with the same words. He investigates the role that compositional semantics can play in theories of what interlocutors mean (Schlangen, 2015 [18]). Schlangen explains that syntax is concerned with the analysis of the composition of sentences, semantics with specification of the meaning of sentences, and pragmatics with determination of the meaning of utterances.

He makes a deliberate distinction between sentences which possess a literal meaning and utterances which carry the speaker meaning, such as this example :

(2) "It is raining".

"It is raining here"; "It is raining at the place we are looking at in a live transmission"; etc. Otherwise, David Schlangen assumes that a dialogue can be a monological discourse, and in this point of view dialogue utterances are utterances of sentences.

We now have defined that a dialogue is a type of discourse where several stakeholders participate in a cooperative conversation. When the discourse (monologue) becomes a dialogue with the introduction of another speaker in the process, the interpretation of the discourse is necessary to take into account for the mutual acceptance of the discursive segments to stay a cooperative conversation. In contrast, we can have non-cooperative dialogues where people (e.g: player) adopt conflicting intentions in their attempt to win a game (e.g: Settlers of Catan [5]). Finally, we understand that discourse relations are crucial to account for the interaction between questions and answers during an exchange.

To conclude, these two works have almost the same understanding of the semantics content and they focus both on the context, which has a role to play in determining the meaning of the sentences. We will focus on the context and give it great importance to our future work. Moreover, we are interested in the Question and Answer pair in dialogue. Now, we will explain the relationship between question and answer in dialogue.

2.3 Challenges in question-answer analysis

In an unsupervised spontaneous conversation, there's a lot of freedom in how an answer (as well as a question) can behave. But even with this freedom, there are certain constraints where an utterance or an action, that follow a question, would no longer be recognised as a fitting answer. In this case we'll be talking about semantic or pragmatic inconsistencies in human language. Indeed, if

we manage to define non-correspondence criteria between questions and answers in a dialogue, we could get a better grasp of incomprehension phenomenon formalization, which could be especially interesting for automation (Boritchev and Amblard, 2018 [3]).

More formally, the mismatch between a question and an answer correlates with the following idea of Searl: In a conversation, each illocutionary act accepts a limited set of appropriate illocutionary acts as replies. Sometimes the such reply is very tightly constrained by the act that precedes it, as in question and answer sequences; and sometimes it is more open, as in casual conversations that move from one topic to another. The principle is somewhat similar to what you can encounter in a game, where one move creates and restricts the range of possible appropriate countermoves (Searl, 1985 [19]).

To address this issue, as well the one of determining the answer span and various speaker behavior in a dialogue negotiation span, we'll rely on last year's project "Where's the Answer: Dialogue Annotation" [2] and the suggested Negotiation Phase algorithm (see section 3.3.4 for more details).

There are certain scenarios, where an answer or even a question may seem irrelevant to an uninformed observer, however it may be still satisfactory for the interlocutor. Dialogue context, shared experience or background, assumptions, implicatures, anaphoras – all add complexity to the question-answer analysis, especially within a spontaneous dialogue.

2.3.1 Implicatures

Implicatures in conversations is a broad subject, that touches upon shared cultural background, common experience, missed logical steps, strive to minimize speech effort, incomplete utterances, *etc.* Implicatures can be described either as the act of meaning or implying one thing by saying something else (i.e. not using a literal meaning), or the object of that act. Some classic examples of this phenomenon are: metaphors, figures of speech, irony, memes, *etc.* Implicatures can be conventional, in different senses, or unconventional. They serve a variety of goals, depending on intention and conventions: communication, maintaining good social relations, misleading without lying, style, verbal efficiency. Common forms of implicature are usually learned along with language acquisition. Historical linguistics traces the evolution of conversational implicatures into idioms (Davis, 2016 [20]).

Conversational implicatures have become one of the principal subjects of pragmatics, with a challenge to distinguish senses and entailments from generalized conversational implicatures. At the source of the implicatures theory we can see H. P. Grice [6], who tried to explain and predict conversational implicatures, describe how they arise and are understood.

Let's refer to an example, where interlocutors A and B are talking about a mutual friend C, who is now working in a bank. Person A asks how C is getting on in his new job, and B replies: "Oh quite well, I think; he likes his colleagues, and he hasn't been to prison yet". Such reply may make A inquire, what B was implying, what he was suggesting, or even what he meant by saying C had not yet been to prison. The answer might be one of the following things: C is the sort of person likely to yield to the temptation provided by his occupation, that C's colleagues are really very unpleasant people, and so on. The exact understanding will depend on the shared knowledge or opinion about C in between A and B. It is clear though, that whatever B implied, is distinct from what B said (that C hasn't been to prison yet).

Let's take another example: "He is in a grip of a vice". Given the knowledge of the English language, but no knowledge of the circumstances of the utterance, we can deduct that "he" was (1)either unable to stay free from a bad character trait or (2)that some part of "him" was caught in a certain tool or instrument. But to be able to fully identify what the speaker meant, one would need to know the identity of "he", the time of the utterance, the meaning of the phrase in this particular occasion, *etc.*

Based on the above we see that implicatures, especially the non-conventional ones, could become one of the biggest challenges in the question-answer analysis and machine learning experiments.

2.3.2 The role of context in a dialogue

In discourse and conversation analysis there are many studies and discussions of context. In traditional linguistics, context is limited to the verbal context surrounding some word or sentence and helps to determine the meaning of words. In conversation, context refers to the sentences and words that surround any part of a discourse and the social situation in which spoken or written verbal utterances that are being used can change the meaning. Listeners and speakers must speak cooperatively and mutually accept one another to be understood in a particular way. The cooperative principle describes how effective communication in conversation is achieved in common social situations.

In the book "Inquisitive Semantics" (Ciardelli, Groenendijk, Roelofsen, 2017 [21]), they tackle some general issues in question. The issues expressed by a question in a natural language is rarely completely determined by grammar alone; it depends on the conversational context in various ways. We will see different questions to illustrate some relevant contextual factors:

- (3) a. Which students passed the exam?
- b. Who comes to the concert?
- c. Where is Mary?

A first important contextual parameter is the **domain of quantification**. As example, the issue in (3a) depends on the set of students which are relevant in a particular context. The contextual parameter in question issue can also need a **method of identification**, where the person who asks the question wants to ask a precise description (e.g: description of a game card). With the who questions, the question can get either **mention-some** or **mention-all** interpretation. In (3b) the speaker rather wants just one or the complete set of people who come to the concert. Finally, the issue in (3c) depend of the **level of granularity** (Ginzburg, 1995 [22] [23]), such as:

- d. Mary is at home.
- e. Mary is in the bathroom.

In certain context, the answer (d) can resolve the question (3c) but in another context, the location of Mary needs to be determined more precisely, for instance with the answer (e). More issues can exist, but lets move on the some examples of issues in light of context-dependency.

In 1997, Paul Piwek tackle some general grounds for rejecting answerhood, which does not take context-dependence seriously (Piwek, 1997 [24]). Not only answers can be restricted by the dialogue context, but questions too: defining if a question makes sense to be asked, depends on the background. There seems to be no point in asking a question, where the answer is already part of the common background (keeping rhetorical, ironic and other types of phatic questions aside). For instance, if the interlocutors share the information that nobody has seen Mary, then the question "Who has seen Mary?" is inappropriate.

Some answers may convey a valid feature to fulfill the question only in a certain context. For example, it's part of common knowledge that if May is at home, then her car is in the garage. We can then imagine a situation where John asks "Where is Mary's car?" and an answer "Mary is at home" would make sense. However on it's own this answer does not comply with the rules for the answer format, set by the wh-constituent "where".

An interesting way to answer a wh-question, without providing the exact missing information, is to rule out some options (some sort of a negative answer). For example, the question "Where is Mary?" can be answered "She is not at home". Such response is not exhaustive, but gives an important update to the shared context, where an answer "At home" is no longer a possible .

Lastly, there are situations where a question raises new questions, answering which will bring the questioner closer to an answer for the original question, i.e. the answer to the precision question implies the answer to the original one and/or resolves it.

In all these perspective questions and answers are apt to change in the different contexts. However, there is still an issue for the interpretation and understanding of the answer when he doesn't fulfil the requirement of the question.

2.3.3 Non-sentential utterances

Acquiring a language by a child is often viewed in correlation with an increase in the mean length of utterance. What's notable for adults is the ability to use short utterances whenever the context allows. This side of linguistic competence can be observed in the usage of **non-sentential utterances (NSUs)**, a common phenomenon in spoken dialogue. NSUs are fragmentary utterances that do not have the form of a full sentence according to most traditional grammars (they lack an overt verbal more generally predicative constituent), but that nevertheless convey a complete clausal meaning (Ginzburg, 2007 [25]). Based on the context, it is usually possible to reconstruct their assumed full form and meaning.

The following example from London-Lund Corpus illustrates some short answers and reprise utterances used for acknowledge or request clarification:

A: Wasn't he refused the chair in Oxford?

B: Who ?

A: Skeat. Wasn't he refused

B: That's Meak.

A: Oh Meak, yes.

In their paper "Answers without questions: The emergence of fragments in child language" (Ginzburg and Kolliakou, 2009 [26]), Jonathan Ginzburg and Dimitra Kolliakou present the results of a corpus study of the emergence of certain classes of NSUs, based on the data gathered from English- and Greek-speaking children. The authors explore adult NSUs distribution in a dialogue and share their principal finding, that the main classes of **non-sentential queries (NSQs)** are acquired much later than **non-sentential answers (NSAs)**, which they call the "**late short query effect**". These results have significant implications for the grammatical analysis of NSUs and, more generally, for how context needs to be integrated with grammar.

A well developed taxonomy for NSUs, based on the British National Corpus, was introduced by Fernández and Ginzburg (Fernández and Ginzburg, 2002 [27]). The split they observed for each NSU class in a dedicated sub-corpus, with examples, is shown in the Table 1.

Different NSU classes are typically related to different resolution constraints, in order to resolve NSUs appropriately it is crucial to identify the intended kind of NSU (Ginzburg, 2007 [25]). The study on the emergence of NSUs was focused on five major sub-classes:

1. Short answers: this NSU class refers to typical responses to (possibly embedded) wh-questions
2. Affirmative answers: plain and repeated, a signal that a previous declarative utterance was understood and/or accepted, can repeat a part of the antecedent utterance
3. Rejection answers: plain and helpful, the latter providing an appropriate alternative or correction to the previous utterance

NSU class	Example	Total
Plain Acknowledgement	<i>A: ... B: mmh</i>	599
Short Answer	<i>A: Who left? B: Bo</i>	188
Affirmative Answer	<i>A: Did Bo leave? B: Yes</i>	105
Repeated Acknowledgement	<i>A: Did Bo leave? B: Bo, hmm.</i>	86
Clarification Ellipsis	<i>A: Did Bo leave? B: Bo?</i>	79
Rejection	<i>A: Did Bo leave? B: No.</i>	49
Factive Modifier	<i>A: Bo left. B: Great!</i>	27
Repeated Affirmative Answer	<i>A: Did Bo leave? B: Bo, yes.</i>	26
Helpful Rejection	<i>A: Did Bo leave? B: No, Max.</i>	24
Sluice	<i>A: Someone left. B: Who?</i>	24
Check Question	<i>A: Bo isn't here. Okay?</i>	22
Filler	<i>A: Did Bo ...B: leave?</i>	18
Bare Modifier Phrase	<i>A: Max left. B: Yesterday.</i>	15
Propositional Modifier	<i>A: Did Bo leave? B: Maybe.</i>	11
Conjunction + Fragment	<i>A: Bo left. B: And Max.</i>	10
Total dataset		1283

Table 1: NSUs in a sub-corpus of British National Corpus

4. Clarification ellipsis: reprise fragments used to clarify an utterance that has not been fully comprehended
5. Sluicing: all wh-question NSUs, can be reprise and direct

The approach to incorporate NSUs into grammatical analysis, would depend on whether NSUs need to be assimilated with some other grammatical phenomenon (phonological reduction, or anaphora) or they are significantly unique. We'll be looking, respectively, at **unitarian** or **constructionist** theories. The former associate ellipsis resolution with a single, typically extra-grammatical mechanism. In the latter NSUs are incorporated into grammar as distinct constructions, with certain contextual characteristics to govern their use. In their research Ginzburg and Kolliakou argue that unitarian approaches cannot be used to formulate a theory of NSUs acquisition. In contrast, they advocate for following a **dialogue-oriented constructionism** approach.

Ginzburg and Kolliakou analyse the idea that the late short query effect can be explained either pragmatically or mechanistically (Ginzburg and Kolliakou, 2009 [26]). In the pragmatic approach the absence of NSQs are explained by either **contextual unavailability** of prerequisites for NSQs and/or by the **semantic undesirability** of the contents that NSQs may express in the interaction. A mechanistic approach explains the lack of NSQs by the absence of appropriate linguistic or conversational competence.

To show that pragmatic explanation is not sufficient, the team examined the following situations: whether the children got appropriate contexts for NSQs, and how likely they will produce an NSQ in those cases (contextual availability); whether NSQs encode the meanings that the children would want to convey in a given situation (semantic undesirability).

At a point when a child can productively use sentential queries, elliptical declaratives and polar lexemes **yes** and **no**, non-sentential questions are almost absent in their speech, despite being present in the child's caregivers speech and despite the context that would allow the usage of such NSUs.

The evidence showed that children would not exploit the triggering context for NSQs with a significantly higher likelihood compared to adult speakers, and would formulate their clarification requests with non-elliptical constructions. The data also indicates that the nature of the corpus didn't bias the children against using NSQs. This implies that the absence of NSQs in their speech is determined by insufficiently developed competence, which supports the mechanistic explanation.

In their work the authors also provided an account of the late short query effect using a dialogue-oriented construction grammar, based on two components. The first one being propositional NSUs and sentential interrogatives without NSUs: the specification of a grammar and contextual system in which short answers and sentential interrogatives co-exist without NSQs. The second one covered the order of acquisition certain of classes of NSU constructions, explaining the emergence of NSAs before NSQs in relation to such notions as semantic composition complexity and accessibility to contextual parameters. They have shown how the late short query effect can be modelled in an approach to grammar, which integrates phonological/syntactic/semantic information with detailed specification of dialogue context. NSUs are treated as constructions, which, among other things, carry a specification of the contextual conditions that regulate their resolution.

2.4 Machine learning

"The next big step for Machine Learning is natural language understanding, which aims to give machines the power to understand not just individual words but entire sentences and paragraphs."
— Yann LeCun

In recent years many successful machine learning applications have been developed. Machine learning is a field of study focused on the development of algorithms that can learn from data. The purpose of such algorithms is to extrapolate patterns from large sets or make predictions based on input provided to the system. This is transformed into a model which can be used to make predictions on new data. The properties of machine learning make it a suitable tool to use when developing a classifier.

Classification belongs to the category of supervised learning. There are two types of learners in classification: lazy learners (e.g: k-nearest neighbor) and eager learners (e.g: Decision Tree, Naive Bayes, Artificial Neural Networks). Eager learners construct a classification model based on the given training data before receiving data classification. It takes a long time to train and less time to predict due to the model construction.

There are different ways to experiment with machine learning, when working with dialogue classification schemes. A previous work related to our subject [1] tried two approaches: the use of a traditional statistical classification algorithm and the use of neural networks. We will introduce the first approach with the Decision Tree for the next section.

The **Decision Tree** is a commonly used algorithm in Natural Language Processing (NLP) research. It produces good results for classification tasks and has been successfully applied to domains like question classification (Zhang, 2003 [28]) and coreference resolution (Soon, 2001 [29]). Decision tree builds classification or regression models in the form of a tree structure. It uses an if-then rule set which is mutually exclusive and exhaustive for classification, the rules are learned sequentially using the training data one at a time. However, a decision tree can be over-fitted. This phenomenon occurs when the tree is designed to perfectly fit all samples in the training data set. An over-fitted model has a very poor performance on the unseen data, even though it gives an impressive performance on training data.

3 Dialogue annotation: project background

As mentioned in the introduction, our project carries over the work of the previous studies supervised by Maxime Amblard and Maria Boritchev, namely: "What's the Answer: Dialogue Annotation" [1] and "Where's the Answer: Dialogue Annotation" [2]. We will introduce their main ideas and conclusions, as they outline the base for our future work.

3.1 Coffee or Tea? Yes.

One of the possible ways to define incomprehension in a dialogue is when speech acts follow each other in a usual way, but their combination - on the contrary - doesn't make any sense. Defining the non-correspondence of an answer to a question is not always an easy task, especially when we talk about trying to automate it. Among other things, one needs to consider identifying the start of the answer and its span, which may leave quite a lot of room for interpretation. In the scope of this project the authors chose to put certain difficulties aside and for now only focused on logical incoherences in the combination of speech acts. However in further work it is envisioned to broaden their coverage. The aim of this project is thus to build a compositional logical model for dialogues, which would allow to quantify the phenomenon of logical inconsistencies, and also to automatically identify the situations in a conversation, when speakers don't understand each other.

Using Type Theoretical Dynamic Logic model (de Groote, 2006 [15]) and Abstract Categorical Grammar toolkit (Pogodalla, 2016 [30]) allows to produce logical representation of sentences in natural language. It should be possible to do the same with speech acts - the parsing is similar, and the methods developed for general discourse can also be applied to dialogues. To account for logical representations of speech acts not being as straightforward, dialogues can be divided into smaller parts, the so called **negotiation phases**. The outcome intuitively corresponds to regrouping the dialogue into self-contained sub-dialogues according to the set of topics that are brought up. Logical modeling of questions and answers is at the core of the current work, with an idea to later implement compositional mapping from Natural Language to Inquisitive Logic.

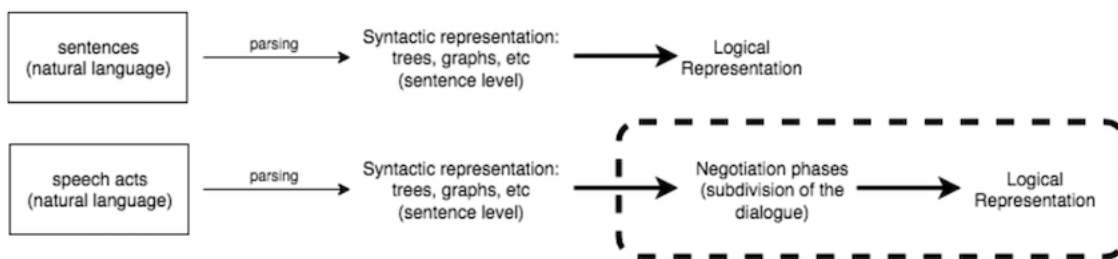


Figure 2: Dialogue modelling: architecture of the process

The Figure 2 shows the architecture of the research process. The upper part has been implemented. The focus of the current work lies within the area, surrounded by the dotted line. To be applied on real-life, uncontrolled human dialogues [3].

3.2 What's the answer: dialogue annotation

This supervised project [1] was delivered in 2018 by María Andrea Cruz Blandón, Gosse Minnema and Aria Nourbakhsh within the Cognitive Science master program at Université de Lorraine, supervised by Maxime Amblard and Maria Boritchev. Focusing on identification of classification of question-answer pairs in dialogues, it pursued three main objectives:

1. Design a classification schema for questions (3.3.1), features (3.3.2) and answers (3.3.3);
2. Create an annotation guide and apply it to manually annotate several dialogues;
3. Explore machine learning approaches to automate annotation.

3.2.1 Annotation schema: questions

In their first attempt to create an annotation schema, the team developed their strategy based on the related work of Freed (Freed, 1994 [31]), who approached questions categorization within an **information continuum**, ranging them from purely factual information requests to the ones that convey social information. Within this continuum questions were divided into classes based on the combination of functional and formal criteria. Another inspiration was Stolcke (Stolcke, 2000 [32]), whose annotation schema for dialogue acts included a rather developed set of question types: **yes/no questions**, **wh-questions**, **rhetorical questions**, *etc.*

The team used corpus-driven approach, starting with a simple model of the iconic Y/N and Wh-questions, and expanding based on the cases encountered in their corpora of natural conversations. Trying to associate each question type with a clear set of syntactic, semantic and pragmatic characteristics, the team identified prototypical questions as well as some of their deviant variations. Three additional question types were distinguished: completion suggestion questions (where you complete the utterance of the previous speaker), disjunctive questions (where you have a set options to choose from for your reply) and phatic questions (where social function is more important: show that you are still paying attention, you express surprise or it's rhetorical). The schema defines the total of **five question types**:

1. YN: Yes/No question - asks the other participant to confirm or deny a proposition;
2. CS: Completion suggestion - proposes an expression to the other participant to be used to finish their statement;
3. DQ: Disjunctive question - asks the hearer to choose between one of several options;
4. WH: Wh-question - contains a wh-constituent, introduced by a wh-word like what, who, where, when, which, how;
5. PQ: Phatic question - has purely communicative purpose, does not expect an informative answer.

Additional characteristics: The guide also distinguished in between simple and complex (multiple), and also quoted questions.

3.2.2 Annotation schema: features

Depending on the question type, their form and function, the team used a set of supplementary tags in their annotation schema, to denote the type of the missing information, its semantic role. This precision for the content of the question is expected to be matching with the received answer. This type of annotation was adapted from the features set by Boritchev (Boritchev, 2017 [33]):

1. TMP: Temporality (When?)
2. LOC: Location (Where?)
3. AG: Agent (Who?)

4. CH: Characteristic (How?)
5. OW: Owner (Whose?)
6. RE: Reason (Why?)
7. TH: Theme (Whom?)

3.2.3 Annotation schema: answers

The core idea behind the annotation schema for answers within a dialogue is that their shape is restricted by the preceding question type and features, listed above. It's expected to have correlation in between these parameters, as described respectively in the Table 2.

Additionally, in a conversation certain question are answered, some are not. But even when a question is answered, the reply may not be directly providing the requested information, certain logical steps may be skipped. To account for such cases, where the reply is not strictly consistent with the question, but still sufficiently responds to it, the tag **unrelated topic** (UT) was introduced.

Tags	Name	Question Type
PA	Positive Answer	YN, CS
NA	Negative Answer	YN, CS
FA	Feature Answer	DA, WH
PHA	Phatic Answer	YN, CS, DQ, WH, PQ
UA	Uncertainty Answers	YN, CS, DQ, WH, PQ
UT	Unrelated Topic	YN, CS, DQ, WH, PQ
DA	Deny the Assumption	YN, CS, DQ, WH, PQ

Table 2: Answer and their correlation to a question type

3.2.4 Annotation experiments

Based on the above, the team created guidelines for manual annotation. They were put in practice and, following the evaluation of the results, updated to introduce a list of improvements.

In order to minimize disagreements in classification, an additional measure was used: a precedence order for question types. Going from more specific question types, with easily identifiable characteristics, to more generic ones, the precedence order for dialogue annotation looks as follows: (1) Wh-questions, (2) Disjunctive questions, (3) Yes/No questions, (4) Completion suggestion, and (5) Phatic questions.

The team conducted some experiments with machine learning. The most promising results were obtained with the application of a decision tree algorithm, based on a set of hand-designed features, that rely on formal characteristics of a question and not on discourse context. Full list of such features can be found in the Table 3. Despite being quite simple, the algorithm performed with a good accuracy score = 0.73 and and F1 score = 0.58, significantly outperforming the majority-class baseline algorithm (Acc = 0.47, F1 = 0.31). Most of the mistakes were due to the length of question parameter, where questions above certain length were mistakenly identified as phatic. Similarly to manual annotation, there has also been some confusion with phatic questions that contain wh-words.

Feature	Description	Value
has_wh	The question contains a wh-constituent	True, False
has_or	The question has contains the word "or"	True, False
has_inversion	The question has inverted structure	True, False
has_tag	The question is a tag question	True, False
has_utt_similar	The last utterance has at least 50% similar words	True, False
last_utt_incomplete	The last utterance is incomplete	True, False
has_cliche	The question contains a cliché (e.g. "really?", "you know?")	True, False
length	The length of the question (number of words)	Numerical value

Table 3: Extracted features for the classification task

To check what input parameters are the most informative, the team conducted some experiments with neural architectures: RNN (recurrent neural network) and BOW (bag-of-words). In their results the team observed that with their relatively small dataset, these models suffered from overfitting and were outperformed by the decision tree model:

- RNN: Acc = 0.76, F1 = 0.44
- BOW: Acc = 0.54, F1 = 0.24

3.2.5 Callouts

This is a good start to create a classification and annotation schema for questions and answers within a dialogue. The project covered both manual annotation experiments and improvements, as well as machine learning with different models (decision tree and neural networks). The features set used for classification is not very extensive, and was easy to interpret in the decision tree experiments. However it mixes form and function features of a question and doesn't seem to allow enough agility, that is needed for spontaneous dialogues in real-life situation. In some cases, the annotation guidelines were a bit ambiguous, some broad definitions that were present there left room for interpretation and resulted in annotator disagreement.

Automatic annotation process only classifies questions. It needs to be extended to account for answers. Question-Answer span was called out as one of the impediments in tracking, also identifying exact answer for a particular question can be problematic. No tracking between questions and answers was proposed, to see what belongs where. Bigger corpus may be more interesting for experiments with neural networks.

3.3 Where's the answer: dialogue annotation

This supervised project [2] was delivered in 2019 by Marta Carletti, Lea Dieudonat and Yiting Tsai within the Natural Language Processing master program at Université de Lorraine, supervised by Maxime Amblard and Maria Boritchev. It focused on taxonomy of questions and answers within spontaneous non-cooperative dialogues, and introduced a significantly updated annotation schema and corresponding guidelines, reworking the previous year approach. The main purpose of this project was defined as contributing to a better understanding of the nature of questions and answers, and proposing a fine-grained annotation schema that could account for their complex structure.

3.3.1 Separating form from function in classification

The report introduced the need to analyze the questions and answers in isolation and as a pair. It also stressed the importance to classify and annotate their forms separately from the functions, so that they could coexist in a more agile frame and would allow to show various combinations of matching and mismatching, that existing in real life dialogues. The form would be defined by syntactic requirements (syntactic inversions, included lexical items, etc.), and the function - by the semantic ones (the intention of a speaker).

To expand on the possible question forms, related on the Table 4, the team suggested differentiating disjunctive questions into inclusive (accepting both suggested options) and exclusive (only one choice is possible). They also define auxiliary-deontic questions (expressing duty, obligation or permission) vs. auxiliary-epistemic (knowledge acquisition).

Name	Tag	Description	Examples
Yes no	YN	Contains do support, inversion	Are you fine?
Wh	WH	Contains WH Require: FEATURE	What time is it?
Disjunctive inclusive	DQ_I	Contains "or", inclusive	Are you a citizen of European Union or Switzerland? If yes, click here
Disjunctive exclusive	DQ_E	Contains "or", exclusive	Do you want tea or coffee?
Auxiliary deontic	AUX_D	Contains an auxiliary deontic	Can you open the window for me?
Auxiliary epistemic	AUX_E	Contains an auxiliary epistemic	Can you survive all this?

Table 4: Reworked question forms

Answers have their own form too. Based on the corpus analysis, the team settled on a list of answer forms, as described in Table 5. It does not have an exhaustive set of examples for each tag, but rather intuitively corresponds to the usual meaning each form conveys.

Name	Tag	Examples
Yes no	YN	Yes, yeah, yep, jeez, sure, of course, absolutely../ No, nope, no way, not at all, nah..
Wh	WH	I go home tomorrow, When I., Because I..
Uncertain	UNC	I'm not sure, maybe, still don't know, could be..
Unknown	UNK	I don't know, dunno, have no clue..

Table 5: Reworked answer forms

Name	Tag	Description
Completion suggestion	CS	The speaker completes the turn of another speaker
Phatic	PHA	Phatic function
Ask confirmation	ASK_CONF	The speaker asks the truth value of a proposition or the hearer's engagement to an action
Ask feature	ASK_FEAT	The speaker asks for a feature
Ask to perform	ASK_PERF	The speaker asks to perform an act
Reported speech	RS	The speaker report someone's else question

Table 6: Reworked question functions

Compared to last year's work, where form and function attributes were mixed, this reports explicitly extracts the question and answer functions as described in the Table 6 and the Table 7, keeping all the form tags aside.

Name	Tag	Examples
Yes no	YN	Yes, yeah, yep, jeez, sure, of course, absolutely../ No, nope, no way, not at all, nah..
Wh	WH	I go home tomorrow, When I., Because I..
Uncertain	UNC	I'm not sure, maybe, still don't know, could be..
Unknown	UNK	I don't know, dunno, have no clue..

Table 7: Reworked answer functions

The list of Wh-features stayed the same: temporality, location, agent, theme, owner, reason, characteristic. No new tags were introduced.

3.3.2 Compatibility and mismatch in question-answer pairs

To better understand how the questions and answers interact with each other, and why it can result in comprehension or lack of thereof, we can relay to the concepts of symmetry and mismatch. Consequently, in a dialogue, one can observe full symmetry of form and function, and, on the other hand, asymmetry either in form or in function, defined by the nature of a given question. It's important to note that asymmetry in these parameters does not necessarily trigger a mismatch or would result in incomprehension. The respective compatibility tables, as suggested in this work, can be found in the Table 8 and Table 9.

Question Forms	Expected answer forms
YN	F _{o1} < YN,UNC,UNK >
WH	F _{o2} < WH,UNC,UNK >
DQ_I	F _{o3} < YN,UNC,UNK >
DQ_E	F _{o4} < WH,UNC,UNK >
AUX_D	F _{o5} < YN, NONE, PERF >
AUX_E	F _{o6} < YN,UNC,UNK >

Table 8: Compatibility by form

Question Function	Expected answer function
CS	F _{u1} < REFUSE, ACCEPT, PHA, GIVE_CONF, REPORT >
PHA	F _{u2} < REFUSE, PHA, GIVE_CONF, REPORT, NONE >
ASK_CONF	F _{u3} < REFUSE, ACCEPT, GIVE_UNC, GIVE_UNK, GIVE_CONF >
ASK_FEAT	F _{u4} < GIVE_FEAT, GIVE_UNC, GIVE_UNK >
ASK_PERF	F _{u5} < GIVE_PERF, NONE, GIVE_UNK, GIVE_UNC, YN >
RS	F _{u6} < PHA,REPORT,NONE >

Table 9: Compatibility by function

3.3.3 Ambiguities in answers classification

In their mismatch analysis, the team pointed out the existence of direct and indirect answers, based on whether answer fulfills the semantic requirements of the question. To avoid going too much into details, the cases of indirectness were put into two big categories: dialogical functions (speaker makes a comment about what was said before) and conversational implicatures (when the meaning depends on the conversational context). It's worth to mention that detection of indirectness is a very broad topic, that may need a research dedicated solely to it.

The team has also outlined an algorithm to resolve complex exchanges containing multiple questions and answers within the same negotiation phase, that involves accounting for the answers that directly follow the questions, and the ones that arrive within a bigger span due to intermediate questions or remarks. In total, we can observe 5 tags to categorize questions:

1. Question form
2. Question function
3. Expected answer form
4. Expected answer function
5. Feature

3.3.4 Text segmentation algorithm

Spontaneous dialogues rarely consist of only one-to-one question and answer pairs. In a natural conversation, a negotiation phase can consist of several questions and answers, with a question in reply to a question or remarks in between answers, unanswered questions, etc. Consider an example from the Saarbrücken Corpus of Spoken English:

A: No I don't suppose I got invited
B: Did you? Good ha ha ha good
A: Do you want me to go?
B: Yes

Cases like this one are complex to treat, so taking into account different possible combinations between two speakers in an embedded exchange, the team set up a graph that represents the paths of their interaction.

One of the main issues during corpus annotation was deciding whether to consider multiple questions in the same negotiation phase or not. This algorithm proved to be a very efficient way to solve the problem and improve the agreement between annotators.

3.3.5 Callouts

The new model seems to have more precise classification approach for annotating questions and answers in a dialogue. However one may argue that it has become too granular. The team has achieved better results in some areas, compared to last year, however we can still notice ambiguities in certain classification areas, which results in insufficient observed annotator agreement. We wondered if it's possible to distinguish an exact reason (and therefore solution) for not having enough agreement. Is it the classification that needs to better cover spontaneous dialogues examples? Is it due to ambiguities in instructions interpretation? Is it a human factor, that makes us perceive the information transmitted in both questions and answers, especially when contextual or logical implications are present?

Legend
 1 = Speaker 1
 2 = Speaker 2
 Qrel = Question related
 Qunrel = Question unrelated
 R = Response (It can be S or A)
 S = Statement
 A = Answer
 EB = End of NP before the segment
 EA = End of NP after the segment
 Numbers from 1 to 20: paths for the reader (references to smooth the explanation of the examples)

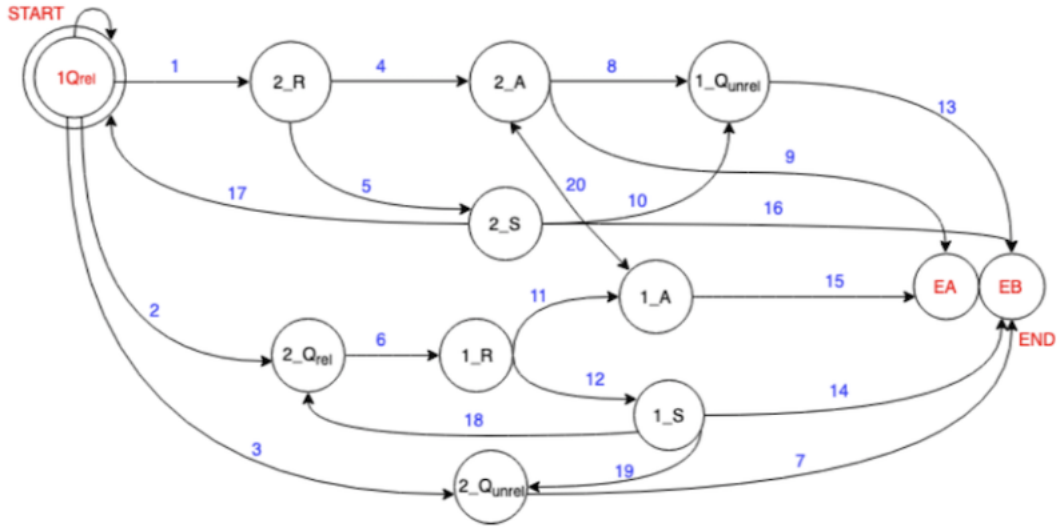


Figure 3: Negotiation phase algorithm

There have been no machine learning experiments based on this model. All the annotation experiments and evaluations have been done only for the manual part. Classification requirements have to be specified for machine learning, various approaches need to be experimented upon.

4 Future work prospects

Question-answer classification schemes within a dialogue have been on linguists' agenda for several years. Development of technology naturally provoked high interest in the field of automation for this domain, which raised new issues and highlighted the importance of question-answer taxonomies. Machine learning seems the right technology to address this challenge.

For our future work we have identified several paths. The most natural one would be in creating an algorithm for machine learning, based on the improvements, introduced in last year's project "Where's the Answer: Dialogue Annotation". Similarly, we would like to do some experiments with decision-tree algorithms and with neural networks. We would also like to explore some gaps, called out in the two previous projects, namely in the domain of classifying indirect answers.

Before proceeding to algorithmisation, we'll need to lock down the classification schema, ideally trying to find a compromise between the two approaches, suggested in the previous supervised projects. Based on the presence of the updates, we may need another round for manual annotations to test our optimised solution and corresponding updates in the encoding system. We will use the Saarbrücken Corpus of Spoken English, similar to last years projects, and the DinG (Dialogues in Games) corpus, if it is ready at that time. After that we can proceed to developing machine learning specifications and respective tests.

To better analyse question-answer pairs, we would like to introduce a way to mark what answer belongs to what question. This will allow us to study the answers behavior: average span, situations where questions remain unanswered and why, if an answer appears far from the question - is there anything particular about them, etc.

Based on the above, we'll be able to see if it's feasible and profitable to do a deep dive into the domain of mismatched answers and questions, and add more details into their classification, or it is better to keep them generalized into bigger groups for the time being. Implicatures are definitely fascinating, but may be hard to tackle in their full complexity at this point. And such ambiguous notions as irony, for example, we'll clearly remain outside of scope for our study.

We initially considered participating in the DinG corpus creation as well. However, taking into account a significant scope of work, estimated for the machine learning part, we decided not to strongly commit to it. This option remains open, if the time allows.

We also considered increasing the list of languages for annotation experiments with Russian or Belarusian, based on the linguistic competence of one of our group members. Belarusian is not a very wide-spoken language. Its currently available corpus is rather small, and mostly consist of scientific articles and literature publications, there's no dedicated sub-corpus for dialogues. As for the Russian, we did not manage to find a corpus for spontaneous dialogues that would be publicly available. After some considerations we realised that it would have been too much work for the next semester to also try to create own own corpora for these languages, so we opted out of this idea as well.

In the end we decided to focus on optimising the annotation schema, algorithmisation and machine learning experiments. We hope that our findings would be helpful for understanding the specifics of semantic and pragmatic inconsistencies in a human dialogue, within natural language processing paradigm.

References

- [1] Maria-Andrea Cruz-Blandón, Gosse Minnema, Aria Nourbakhsh, Maria Boritchev, and Maxime Amblard. Toward dialogue modeling: A semantic annotation scheme for questions and answers. *arXiv preprint arXiv:1908.09921*, 2019.
- [2] Marta Carletti, Lea Dieudonat, and Yiting Tsai. Where’s the answer: Dialogue annotation. 2019.
- [3] Maria Boritchev and Maxime Amblard. Coffee or tea? Yes. In *The 22nd workshop on the Semantics and Pragmatics of Dialogue*, Aix-en-Provence, France, November 2018. Laurent Prévot, Magalie Ochs and Benoît Favre.
- [4] David W. Angel. The four types of conversations: Debate, dialogue, discourse, and diatribe. <https://medium.com/@DavidWAngel/the-four-types-of-conversations-debate-dialogue-discourse-and-diatribes-898d19eccc0a>, 2016. Accessed: 2019-12-15.
- [5] Markus Guhe and Alex Lascarides. Trading in a multiplayer board game: Towards an analysis of non-cooperative dialogue. In *Proceedings of the Annual Meeting of the Cognitive Science Society*, volume 34, 2012.
- [6] H Paul Grice. Logic and conversation, syntax and semantics. *Speech Acts*, 3:41–58, 1975.
- [7] John Langshaw Austin. *How to do things with words*. Oxford university press, 1975.
- [8] Jane Friedman. Question-directed attitudes. *Philosophical Perspectives*, 27(1):145–174, 2013.
- [9] Ivano Ciardelli. Question meaning= resolution conditions. *Logic and Logical Philosophy*, 26(3):383–416, 2017.
- [10] Mohanty Vijayendra. What is an answer? <https://www.quora.com/What-is-an-answer>, 2014. Accessed: 2019-12-15.
- [11] Nicholas Asher. Discourse semantics. 1998.
- [12] Richard Montague. The proper treatment of quantification in ordinary english. In *Approaches to natural language*, pages 221–242. Springer, 1973.
- [13] Hans Kamp. A theory of truth and semantic representation. *Formal semantics-the essential readings*, pages 189–222, 1981.
- [14] Nicholas Asher, Nicholas Michael Asher, and Alex Lascarides. *Logics of conversation*. Cambridge University Press, 2003.
- [15] Philippe De Groote. Towards a montagovian account of dynamics. In *Semantics and Linguistic Theory*, volume 16, pages 1–16, 2006.
- [16] Alonzo Church. A formulation of the simple theory of types. *The journal of symbolic logic*, 5(2):56–68, 1940.
- [17] Jonathan Ginzburg. The semantics of dialogue, 2016.
- [18] David Schlangen. Semantics and dialogue. *The Handbook of Contemporary Semantic Theory*, pages 687–713, 2015.
- [19] John R Searle, John Rogers Searle Searle, S Willis, Daniel Vanderveken, et al. *Foundations of illocutionary logic*. CUP Archive, 1985.

- [20] Wayne A Davis. Implicature. In *Irregular Negatives, Implicatures, and Idioms*, pages 51–84. Springer, 2016.
- [21] Ivano Ciardelli, Jeroen Groenendijk, and Floris Roelofsen. Inquisitive semantics. *NASSLLI lecture notes*, 187, 2012.
- [22] Jonathan Ginzburg. Resolving questions, i. *Linguistics and philosophy*, 18(5):459–527, 1995.
- [23] Jonathan Ginzburg. Resolving questions, ii. *Linguistics and Philosophy*, pages 567–609, 1995.
- [24] Paul Piwek. What is an answer? that is the question, 1997.
- [25] Raquel Fernández, Jonathan Ginzburg, and Shalom Lappin. Classifying non-sentential utterances in dialogue: A machine learning approach. *Computational Linguistics*, 33(3):397–427, 2007.
- [26] Jonathan Ginzburg and Dimitra Kolliakou. Answers without questions: The emergence of fragments in child language. *Journal of Linguistics*, 45(3):641–673, 2009.
- [27] Raquel Fernández and Jonathan Ginzburg. Non-sentential utterances: Grammar and dialogue dynamics in corpus annotation. In *Proceedings of the 19th international conference on Computational linguistics-Volume 1*, pages 1–7. Association for Computational Linguistics, 2002.
- [28] Dell Zhang and Wee Sun Lee. Question classification using support vector machines. In *Proceedings of the 26th annual international ACM SIGIR conference on Research and development in information retrieval*, pages 26–32. ACM, 2003.
- [29] Wee Meng Soon, Hwee Tou Ng, and Daniel Chung Yong Lim. A machine learning approach to coreference resolution of noun phrases. *Computational linguistics*, 27(4):521–544, 2001.
- [30] Sylvain Pogodalla. Aegtk: un outil de développement et de test pour les grammaires catégorielles abstraites. 2016.
- [31] Alice F Freed. The form and function of questions in informal dyadic conversation. *Journal of Pragmatics*, 21(6):621–644, 1994.
- [32] Andreas Stolcke, Klaus Ries, Noah Coccaro, Elizabeth Shriberg, Rebecca Bates, Daniel Jurafsky, Paul Taylor, Rachel Martin, Carol Van Ess-Dykema, and Marie Meteer. Dialogue act modeling for automatic tagging and recognition of conversational speech. *Computational linguistics*, 26(3):339–373, 2000.
- [33] Maria Boritchev. Approaching dialogue modeling in a dynamic framework. 2018.