Toward a linguistically grounded dialog model for chatbot design

Anna Dall'Acqua*
Università di Bologna & Injenia S.r.l.

Fabio Tamburini**
Università di Bologna

The increasing interest in various types of conversational interfaces has been supported by a progressive standardization of the technological frameworks used to build them. However, the landscape of available methodological frameworks for designing conversations is much more fragmented. We propose a highly generalizable methodology for designing conversational flows rooted in a functionalist-pragmatics perspective, with an explicit adherence to a conversationalist approach. In parallel, we elaborate a practical-procedural workflow for undertaking chatbots projects in which we situate the theoretical starting point. At last, we elaborate a general case-study on which we transpose the identified approach in Italian language and using one of the most authoritative NLU platforms.

1. Introduction

One of the most exciting innovations that we are experiencing in the last decade is the massive widespread of conversational interfaces, such as chatbots or virtual assistants (Tsvetkova et al. 2017; Chaves et al. 2019; Dale 2016). The various attempts that have been made to classify these technologies (Radziwill and Benton 2017; Følstad, Skjuve, and Brandtzaeg 2019; Hussain, Sianaki, and Ababneh 2019; Mathur and Singh 2018) and the absence of an unequivocal taxonomy (Braun and Matthes 2019) surely contribute to the lack of a methodological approach for designing conversational agents. They are perceived as something in between humans and web search engines characterised by a conversational way of expression and the capability of managing input and output in natural language (Dale 2016; Braun and Matthes 2019).

We are witnessing a flourishing literature about technologies, techniques and applications for building conversational interfaces (Ahmad et al. 2018; Adamopoulou and Moussiades 2020). Unfortunately, we cannot say the same for the elaboration of methodological guidelines that can be pursued for the designing of conversational interfaces, especially from a linguistic point of view.

This is the context in which our research comes to light. We think that a solid anchor in linguistics and therefore a scientific knowledge of what human conversation is may be the key for identifying a generalizable methodological approach for designing conversational agents.

* Dept. of Classical Philology and Italian Studies, University of Bologna, Italy and Injenia S.r.l., Bologna, Italy. E-mail: anna.dallacqua2@unibo.it
** Dept. of Classical Philology and Italian Studies, University of Bologna, Italy. E-mail: fabio.tamburini@unibo.it. Corresponding author.

© 2021 Associazione Italiana di Linguistica Computazionale

1.1 A methodological lack in conversation design

A relevant set of studies have been produced on this topic. The first agents were implemented with simple *pattern-matching* techniques and *template-based* responses (Weizenbaum 1966; Colby, Weber, and Hilf 1971). They could support a continuative concatenation of utterances but they were still far away from today's state of the art. The same pattern recognition model shaped the architecture of A.L.I.C.E., a chatbot annotated with AIML, a mark-up language derived from the metalanguage XML (Wallace 2003; Shawar and Atwell 2007). After 2016 there has been a growing interest for chatbots in various areas and applications, mostly because they were considered new productive and entertaining objects not reducible as mere assistants, but capable of a way of interacting that brings them closer to users (Dale 2016).

Today, most of them rely on machine learning algorithms and Natural Language Understanding modules, but still even the more conversational of the agents can just vaguely simulate conversational exchanges enriched by mutual understanding as we know it as humans.

Adamopolou and Moussiades (2020) distinguish between two ways of developing chatbots: using any programming language like Java, Clojure, Python, C++, PH, Ruby and Lisp or using state-of-the-art platforms. At this time, the leading NLU cloud platforms supported by machine learning are: Google's Dialogflow in both versions ES and CX, Facebook's wit.ai, Microsoft LUIS, IBM Watson Conversation and Amazon Lex. These platforms share a common *information-retrieval* approach based on what Moore and Arar (2019) identify as *Intent-Entity-Context-Response (IECR)* paradigm. An *intent* "[...] represents a mapping between what the user says and what action should be taken by the chatbot" (Adamopoulou and Moussiades 2020, 377). Intents recognize the conversational action a user is performing, while *entities* are tools used to extract particular details and parameters values from natural language inputs. They can be either system-defined or customized by the developer. *Contexts* are "[...] strings that store the context of the object the user is referring to" (Adamopoulou and Moussiades 2020, 378) used to capture the context of the current topic. *Responses* consist of what the chatbot actually answers in chat. This approach aims to extract contextual and detailed information from users' inputs and respond accordingly to the users' intention, extracting domain-specific entities and associating the corresponding intent, which means that machine learning algorithms are used for intent identification and entities extraction tasks, but responses are typically pre-authored by a designer (Moore and Arar 2019).

Generative models capable of automatically generating answers considering current and previous user messages are also in production, but there are still difficulties in building and training them and they are not available in the major commercial platforms (Adamopoulou and Moussiades 2020). Studies on the evaluation of these platforms show similar performances in terms of combined *f-score* (Liu et al. 2019; Braun et al. 2017), with slight differences in intent identification task (Canonico and De Russis 2018), especially with longer utterances (Zubania et al. 2020).

Although the technological progresses and the wide technical landscape here outlined, we are facing today a lack in designing domain specific conversational interfaces. The current state of conversational interfaces is limited in terms of established user interface design patterns: it is still unclear when chatbots should be text-based or button-based, or which are the best practices in designing a chatbot conversations. The question about how to structure the interaction with this new medium for creating efficient conversational experiences is still opened. Schiavo and Fadhil (2020) investigate the available scientific literature about interaction patterns and design principles in health-

care and identify four common theoretical themes in which the specific features are categorized: bot-user interaction, bot-response, bot-development, user experience. Since our work has a predominant linguistic focus, we mainly concentrate our attention on linguistic features, such as tone of voice, flexibility of responses, conversation length and user engagement in general. Schiavo and Fadhil (2020) treat each feature separately, offering relevant suggestions, but no univocal applicable design methodology in that sense.

Some studies embrace sociolinguistic theories: Chaves et al. (2019) apply to the design of a specific use case the *register theory*, Bennett (2018) and Dippold et al. (2020) identify Interactional Sociolinguistics as key to express chatbot's personality through language in responses and prompts design.

There are multiple works concerning the users' perceptions while texting with chatbots and what they would expect from a satisfying conversation with them: Hill et al. (2015) demonstrate that users hold long conversations also with conversational agents, adapting to chatbots language without overlooking that they were actually chatting with robots. Svenningsson and Faraon (2019) identify the factors of perceived humanness in chatbots' responses and underline their possible applications in terms of design guidelines. Jain et al. (2018) focus on new chatbots' users identifying guidelines more related to flow buildings. Kvale et al. (2019) draw practical and theoretical implications from a manual analysis of chatbots conversations, such as the value of cross-disciplinary teams and the need of diligence in chatbots training. Although these studies end with practical advice, they are far too generic for laying the foundations for a methodology.

There is consistent number of systematic guidelines on how to design conversational interfaces with a practical-computational procedure on how to approach chatbots' projects, reported also in McTear (2020). Some of them do have a commercial vocation (Hall 2018), other focus on technical issues (Shevat 2017; Dasgupta 2018) but even the more linguistically or cognitive oriented ones do not display a clear affiliation to a complete framework of analysis rooted in linguistics (Pearl 2016; Cohen, Giangola, and Balogh 2004).

We think that it is fundamental for creating effective conversational agents that should actually converse with humans and whose aim is to simulate the mechanism of human interaction to refer to a solid linguistic framework. Since pragmatics is the area of linguistics that primarily focuses on language in use also in interactional contexts, we agree with Bianchini et al. (2017) on the importance of pragmatics in developing new chatbots examples. Furthermore, we also agree with Bennett (2018) with the identification of Conversation Analysis as a methodological key to design better conversational flows.

2. Theoretical analysis of *dialogue* and *conversation*

2.1 A pragmatic *perspective*

The term *pragmatics* is conventionally credited to Charles W. Morris (1938) who first introduced a "pragmatic dimension" in the context of relations between signs, interpreters and objects (Bazzanella 2008). Influenced by Charles P. Peirce (1932) and in agreement with Carnap (1938), he distinguished three "dimensions of semiosis" (Morris 1938, 21), in which pragmatics addressed the relations between signs and who use and interpret them (Horn and Ward 2006). Since pragmatics is an interdisciplinary "hardly a well-integrated field of research" (van Dijk 2009, 13), it is preferable to speak about

a pragmatic *perspective* towards language instead of a pragmatic theory (Bazzanella 2008; Bublitz and Norrick 2011). The adoption of a pragmatic perspective also allows to embrace the contributions from philosophy, psychology, sociology, linguistics and the multiple definitions it has (Levinson 1983; Leech 1983; Katz and Fodor 1963; Ariel 2010; Turner 1999).

The philosophers of language Austin (1962), Grice (1975) and Searle (1969) influenced a notion of pragmatics in contrast with the chomskyan analysis of language as an abstract instrument independent by the context of use. Reflections in this direction are a consistent part of the most common handbooks of pragmatics, such as Levinson (1983), Leech (1983), Mey (1993), Yule (1996), and Verschueren (1999). It thus seems reasonable "[...] to claim that the ensuing pragmatic turn was most notably induced by J.L. Austin, J.R. Searle and H.P. Grice, who were interested in utterance meaning rather than sentence or word meaning, i.e. in studying unique historical events created by actual speakers to perform linguistic acts in actual situational contexts in order to accomplish specific goals" (Bublitz and Norrick 2011, 2).

The approach adopted in this work is *functionalist*, that is, "[...] that it attempts to explain facets of linguistic structure by reference to non-linguistic pressures and causes" (Levinson 1983, 7). This perspective aims to explain linguistic phenomena relying on pragmatics principles (Givon 1979; Hymes 1962) and opens to different developable possibilities such as an ethnomethodological method rooted in sociology (Garfinkel 1996; Goffman 1983; Sacks, Schegloff, and Jefferson 1992) and a psycholinguistic approach such as the alignment model (Pickering and Garrod 2004; Branigan, Pickering, and Cleland 2000; Szmrecsanyi 2005).

In both cases, the study of pragmatics is connected to the use of language in communication. Since communication inevitably involves at least two parties, the primary focus of pragmatics are "language use and language users in interaction" (Bublitz and Norrick 2011, 4). Towards this intersubjective dimension, Kerbrat-Orecchioni (2001) speaks of *pragmatique interactionniste*, whose main objects of research are manifestations of verbal interactions, such as dialogue and conversation.

Dialogue could be taken as "[...] the elementary and universal form of human communication" (Luckmann 1990, 58), whose basic principles are most salient in conversations and authentic discourses (Linell 2001). Levinson (1983, 284) defines conversation as "[...] the predominant kind of talk in which two or more participants freely alternate in speaking, which generally occurs outside specific institutional settings". In the broadest sense, it includes both face-to-face social communications and technology-mediated forms of interactions: all these different manifestations can be classified according to different criteria. A common distinction bases on the final scope of the exchange: social interactions' aim is building and maintaing rapports, while transactional interactions mainly fulfill practical goals (Brown and Yule 1983; Clark et al. 2019). Hakulinen's classification (2009) takes into account the degree of institutionality, the activity type or genre, the channel and participation framework; Linell and Luckmann (1991) consider the degree of asymmetry between the interlocutors. According to Schegloff, the *ordinary conversation* is the most general and flexible type of conversation from which the other types are adapted for particular purposes. It is defined as "[...] the basic medium of 'interactional exchange' [...] in whatever practices it is embodied in those settings" (Schegloff 1999, 413). Moore and Arar (2019) identify *service*, *teaching* and *counseling conversations* as derived typologies from the ordinary conversations. They are all charachterized by roles's fixedness and influenced by their settings. Since Moore and Arar (2019) embrace a strictly conversationalist point of view, their focus is on identi-

fying the underlying structure of conversation, which is suitable for slight adaptations according to the settings and contexts real conversations may occur in.

2.2 Pragmatic frameworks of analysis

There are several useful frameworks to analyze dialogical and conversational interactions. In the field of pragmatics, Haugh (2012) discerns two key trends to place conversational interactions: at the level of meaning and abstract principles referring to the works of language philosophers such as Grice (1989) and Searle (1969); at the level of the performance the analysis of authentic data referring to Conversation Analysis (Sacks, Schegloff, and Jefferson 1974) and Interactional Sociolinguistics (Gumperz 1982). In this section, we summarise the key points of each and the possible adaptation to a human-computer interaction.

The first trend is situated on a cognitive level and understands conversation as a *"joint activity"* whose progression is determined by the concatenation of "joint actions" (Clark 1996, 30). They are the result of the coordination of individual actions on two levels: "There is coordination of both content, what the participants intend to do, and processes, the physical and mental systems they recruit in carrying out those intentions." (Clark 1996, 59). Regarding language use, "[...] a central problem is coordinating what speakers mean" (1996, 73). The idea of conversation as action determined by an undercurrent of communicative intention between the participants was formulated by Austin (1962) and Searle (1969, 1983). These studies have been enormously influential in the pragmatic approach of conversation: they allow to "[...] formalise rules and principles by which speakers mean (and to a lesser extent do) things" in conversation, abstracting from the conversation itself (Haugh 2012, 251). For example, the austinian notion of performativity frames new perspectives in human-computer-conversations, such as the collaborative action of *"We" Human-and-Technology* (Cho and Yoon 2013; Cho 2015) and the methodology of the *Performative Experience Design* (Spence 2016).

Intentionality and delivery of implicit meanings are Grice's main objects of inquiry (1975). He can be considered one of the pioneers of *inferential pragmatics* (Ariel 2012). Most of all the cooperation principle and the conversational maxims proposed by Grice (1975) and later updated by Sperber and Wilson (1995) are a consistent part of the study of pragmatics today and involved in the implementation of dialogue systems from a methodological point of view: Jacquet et al. (2018; 2019; 2019b) evaluate the violation of the gricean maxims in textual online conversations; Saygin and Cicekli (2002) propose an empirical study study of human-computer interactions within the context of the Loebner Prize Contest.

Lakoff's theory of politeness (1973) is an attempt of expansion of Grice's conversational maxims. This theory has been extensively criticized because it is hardly generalizable (Al-Duleimi, Rashid, and Abdullah 2016) and the key terms used in it are culturally determined and they therefore need to be clearly defined (Brown 1976; Tannen 1984).

Brown and Levinson (1987) propose instead an expansion of the studies on politeness made by Goffman (1967) introducing the concepts of *positive face* and *negative face*, which are respectively the need to be approved by the others and the need of autonomy. The importance of politeness in the realization of conversational interfaces is attested among the others by Følstad et al. (2018) and Nordheim, Følstad, and Bjørkli (2019), who list politeness as a factor perceived to affect trust in chatbots for customer service. As far as practical chatbots realizations are concerned, Hall (2018) includes it as an element to take into account during the conversation design, while De Jong, Theune, and Hofs

(2008) adapt the model of politeness strategy elaborated by Brown and Levinson (1987) and of the linguistic alignment (Pickering and Garrod 2004) to a virtual museum guide.

In relation to the second trend, the dominant perspectives are *Interactional Sociolinguistics* (IS) and *Conversation Analysis* (CA).

The core idea of Interactional Sociolinguistics is that what happens in a sequence of talk can be analyzed in its social contexts and that humans in talk accomplish social goals. A central concept of this approach are *contextualization cues*, "[...] by which speakers signal and listeners interpret what the activity is, how the semantic content is to be understood and how each sentence relates to what precedes or follows" (Gumperz 1982). Feine et al. (2020) offer an overview of the implementation of social cues in different kinds of conversational agents, while Bennet (2018) translates conversational cues in the realm of text-based chats, arguing that a strategic manipulation of orthography to convey conversation cues could help the design of chatbots personality and could situate them on different levels of *enthusiasm* or *considerateness*. Dippold et al. (2020) attest how a microlevel of design linguistic analysis based on Interactional Sociolinguistics can be useful for chatbot designers for creating engaging interactions and provide specific guidelines. Relevant studies in the field of Computer-Mediated Communication (CMC) describe the modification on the different levels of language that apply on digital mediated communication: Crystal (2001) and Herring (2012) are pioneers for the English language, while for Italian the work from Pistolesi (2018) is certainly a relevant reference point.

Another approach to dialogical interaction is Conversation Analysis. Levinson (1983) includes in his textbooks on pragmatics a chapter entitled *Conversational structure*, in which he compares *Conversation Analysis* and *Discourse Analysis* as two opposite methodological frameworks with a preference for the first one, characterized by an empirical and inductive vocation.

CA is a subfield of sociology whose origin is influenced by ethnomethodological studies (Garfinkel 1996), but the publications of the firsts and more influential CA contributions in 1974 *A simplest systematics* in the flagship *Journal of the Linguistic Society* and of *The preference of self-correction* in 1977 on *Language* both by Schegloff, Sacks and Jefferson established a relationship between CA and Linguistics from its origins (Fox et al. 2018).

The aim of this discipline is providing a *systematic description* of oral language *practices* (Schegloff 1992, 120) between humans, in order to formalise it into key structural elements that occur in the variety of contexts in which conversations may take place. Unlike ethnomethodology, whose observations are based on memory and intuition (Pallotti 2007), the methodological approach of CA is fundamentally *empirical* and its focus was a description of language as a tool used by social actors in interactions. In this respect, the object of study is an *interactive activity contextually situated*, where sentences are "[...] produced by someone, for someone else, at a certain time, in a certain way" (Hoey and Kendrick 2017). In the words of Schegloff and Sacks (1992, 70), CA is a "[...] naturalistic observational discipline that could deal with the detail of social action(s) rigorously, empirically and formally".

The systematic descriptions of such practices leads to discovering the *machinery* (Sacks 1984, 84) underlying conversations, a *mechanics* of how people naturally talk in a variety of settings (Sacks, Schegloff, and Jefferson 1974; Schegloff 2007) made of some key concepts that regulate interactions. This approach supports the existence of some structural patterns that occur in conversational exchanges without consideration of the delivered content and with slight adaptations according to the settings and contexts of

realization: it is therefore clear the enormous potential it may have in Human-Computer interaction studies.

We will describe four elements of the descriptive apparatus for analyzing interactional structures, adhering to the analysis made by Pallotti (2007), Moore and Arar (2019) and Hoey and Kendrick (2017), which are today "[...] common stock for everyone doing CA" (Pallotti 2007, 7).

2.3 Key concepts of the conversation *machinery*

The strategy people use to manage the conversational traffic in interaction and the distribution of talk among the parties is *turn-taking*, known to be the feature that makes conversations orderly without significant clashes, overlappings or long pauses. Sacks et al. (1974, 702) describe it as a *simplest systematics* composed of two components and a coordination of the ending of the turn with the start of the next. The *turn-constructional units (TCU)* consist of linguistic unit-types such as sentential, clausal, phrasal or lexical constructions that form a "[...] recognizably complete utterance in a given context" (Hoey and Kendrick 2017). These bound units are defined in functional terms, being understood that usually "[...] a complete (linguistic) action corresponds to a complete syntactic unit, so that the TCU boundary turns out to coincide with the clause boundary" (Pallotti 2007, 8). Once a turn is perceived as completed, occur a turn-transfer using turn-allocation techniques such as self-selection and other selection in specific *transition-relevance place (TRP)*. A hierarchically organized set of rules governs the turn construction and coordinates the transition so as to minimize the gaps and overlappings (Sacks, Schegloff, and Jefferson 1974).

Turns do not occur haphazardly, but are *sequentially organized* into coherent courses of actions (Schegloff 2007). The minimal unit of sequential organization is a two-move sequence, the *adjacency pairs*, in which the connection between the parts depends on the *conditional relevance*: the occurrence of a first pair sets up the relevance of the second part to follow (Schegloff 1968). Schegloff and Sacks (1973) give some examples of adjacency pairs, such as question-answer, greeting-greeting, offer-acceptance/refusal. Moore and Arar (2019, 65) also include farewell-farewell, assessment-assessment, inquiry-answer, request-grant/deny, invitation-acceptance/decline, accusing-admitting/denying. Since the first part sets up an *expectation*, the absence of the accepted second part is noticeable and the participants may require explanations or justifications for not having answered or for not having chosen the *preferred* option (Schegloff, Jefferson, and Sacks 1977). Even if the first part is usually directly followed by the second part, it could happen that for satisfying various requirements the completion of the first part has to be suspended for one or more turns. Sequences are therefore inherently expandable through additional turns over and above the two basic units of sequences. Expansions are allowed before the first pair part *(pre-expansion)*, between the first and the second pair part *(inserted expansion)* and after the second pair part *(post-expansion)* (Schegloff 2007, 26). Another kind of sequence are the *storytelling sequences* (Jefferson 1978), used to express stories, anecdotes, or instructions whose content have to be distributed on multiple turns. They are often introduced by pre-announcement, namely the story preface (Schegloff 2007, 41). Sequences are an instrument for organizing utterances produced by the participants, while sequences themselves are organized into *activities* that define the "overall structural conversation" (Schegloff and Sacks 1973, 71), such as conversation openings and closings, instruction giving or troubleshooting.

Another relevant element in CA is *turn design*, which deals with how speakers build their turns to achieve some goals and to deliver contents for a specific audience (Drew

2020). Turn design principles concern the speakers' orientation to contiguity and their tendency to display connections between what they are saying and what the other said in prior turns; the specific lexico-syntactic adopted by the speakers to pursue the desired action and *the recipient design principle*, which is "[...] a multitude of respects in which the talk by a party in a conversation is constructed or designed in ways which display an orientation and sensitivity to the [...] co-participants [...] with regard to word selection, topic selection, admissibility and ordering of sequences, options and obligations for starting and terminating conversations [...]" (Sacks, Schegloff, and Jefferson 1974, 727). It implies the consideration of the relationship between the interlocutors, their mutual knowledge and *common ground* (Clark 1996). The other principle to take into account in turn design is *minimization*, known as the speakers' tendency to deliver a message or to complete an action without using more words than necessary, but still being recognisable from the recipient the conversation is tailored to (Sacks and Schegloff 2007). In other words, the recipient's design is prioritarian over minimization: the speaker has to be efficient, using as few words as possible without preventing the interlocutor to understand.

The last element of conversation machinery are the *repair practices* spontaneously accomplished by the speakers in case of troubles in speaking, hearing or understanding (Schegloff, Jefferson, and Sacks 1977). Hoey and Kendrick (2017) describe the three basic components of a repair procedure: a trouble source, a repair initiation (i. e. a signal that begins the repair procedure) and the repair solution (i. e. the actual repair, for example paraphrasing or repeating a word part of the prior turns). Both the speaker *(self)* and the recipient *(other)* can initiate a repair procedure and/or accomplish a repair solution. Repair mechanisms are thus distinguished relying on who initiates the repair (*self-initiated* or *other-initiated*) and on who effectively accomplishes the repair (*self-repaired* or *other-repaired*).

Some attempts have been made adopting CA as a theoretical framework for implementing conversational agents: Luff et al. (1990) early imagined the potentialities of CA both as analytical tool in HCI applications and as inspiration for design methodology. Wooffit et al. (1997) adopt CA as a sociological perspective for studying human-computer dialogues. More recently, Lotze (2016) includes CA in the theoretical approaches used to analyze human-chatbot corpora. Hirst (2001) reviews Luff, Gilbert and Frohlich (1990) focusing on the different conceptual perspectives from which CA and Discourse Analysis are shaped, defending the necessity of considering CA studies in the field of Natural Language Understanding (NLU) technologies and applications. At that time technologies were not mature enough to adhere to such a complex theoretical framework, but at present time more and more studies go in this direction, such as the works from Gervits et al. (2020) and Michael and Möller (2020).

2.4 The *Natural Conversation Framework* (NCF) as a promising starting point for a design methodology

The work by Moore and Arar (2019) on NCF represents an innovative proposal in the scientific-industrial landscape. Moore and Arar worked together at IBM-Research in designing prototypes for novel forms of interactions for conversational interfaces and Moore is currently developing a conversational methodology founded on the qualitative models from the field of CA. This is not the first work in this direction: Moore (2013) and Moore et al. (2018) collect interesting contributions on specific design issues generalizable to various use-cases, such as Bickmore et al. (2018) and Candello and Pinhanez (2018) and lay the groundwork for the practical guide published in 2019.

Moore and Arar (2018) especially introduce the lack of a methodology for designing conversations and invite us to embrace the complexities of human dialogues in order to create machines we can interact with in a natural way.

Moore and Arar (2019) can be considerd a potential starting point for drawing a complete conversation methodology for different reasons. First of all, they argue that among the possible natural-language interaction styles the *conversation-centric style* is the future of AI interfaces[1], since it aims to reproduce a real conversation-first way of interaction. Therefore, they situate themselves on a higher level than simply offering a versatile practical procedure: they embrace a specific linguistic approach and translate it in the realization of human-machine dialogues. The sub-field of research of the *User Experience Design* they steer is rooted in CA, whose methodological principles and key elements are described in previous section.

The NCF traces the basics mechanics of conversational patterns documented in CA and consists of four parts: an underlying interaction model based on expandable sequences, a distinctive content format based on the interaction model, a reusable pattern language for common conversational activities, a general method for navigating conversational applications.

The interaction model is based on the sequential structure of conversations, in which sequences are "[...] general patterns that [...] can be used and reused in all kinds of different situations and settings, for all kinds of different purposes" (Moore and Arar 2019, 65). Like adjacency pair sequences and the storytelling sequences in CA, this model should support sequence expansions, as "[...] natural indicators of the participants' state of understanding of a turn-by-turn basis" (Moore and Arar 2019). This is a more natural and interactive pattern than the simple two-turn sequence model of the majority of chatbots and virtual assistants currently available.

In order to apply the interaction model accurately, a particular format is required for the content of the conversational application. (Moore and Arar 2019, 70) express it translating the principle of minimization (Sacks and Schegloff 2007) into three guiding principles: "limit agent utterances to a single sentence or less", "break paragraphs down into their parts" and "let users control the level of detail". These criteria enable designers to break-up document-formatted content into bite-sized intents, which can be requested by the users through simple queries.

Nevertheless NCF does not provide a library of industry-specific content in the form of intents and entities, but a systematic set of dialogue patterns that constitute various aspects of conversational competence, enable a variety of social activities and can be configured to a wide range of use cases (Moore and Arar 2019). The catalogue proposed by Moore and Arar (2019) is made of 15 types of patterns and 100 subpatterns[2] and is directly inspired from the examination of naturally occurring observations.

The patterns can be divided into three categories: *conversational activities*, *sequence level management* and *conversation-level management* (see Figure 1). Conversational activities include patterns for managing content inside the boundaries of the conversation, such as inquiries, requests or extended tellings. The remaining two categories help users and agents to manage the interaction itself and occur on two levels: management

1 Other interaction-styles that imply the recognition of natural language inputs are the *system-centric style*, the *content-centric style* and the *visual-centric style*. They differ from the conversation-centric style because they do recognize and produce strings in natural language, but they do not exhibit the conversational actions distinctive of the human way of communicating (Moore and Arar 2019).

2 The list is not exhaustive. Here https://ibm.biz/BdzwQU are some new patterns IBM researchers are experimenting with.

patterns on a sequence level help managing successful sequential interactions through repair procedures, management patterns on a conversation level help managing conversations' openings, closings and disengaging. Finally, the navigation method consists of six simple conversational actions the users can take to get around the conversational space.

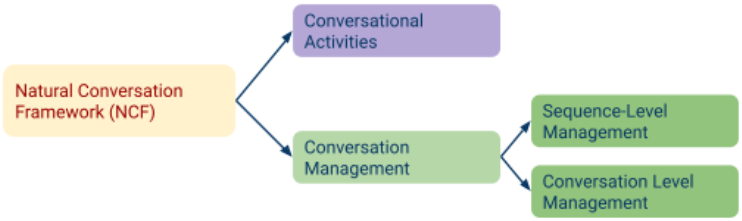

Figure 1
Graphical representation of the *Natural Conversation Framework* patterns.

The so defined set of patterns is much more similar to a *"language"*than a *"library"*: these patterns are designed to be used together and are particularly useful in service encounters conversational interactions, which are fundamentally practical dialogues characterized by a goal-seeking nature (Allen et al. 2001).

To show how the patterns are presented, we provide an example of sequence closer patterns, from the category of the sequence management patterns (Moore and Arar 2019, 142). Sequence closers enable users to communicate that the previous action is completed and that they are ready to move on:

```
U: INQUIRY/REQUEST
A: RESPONSE
U: ACKNOWLEDGMENT/ASSESSMENT/LAUGHTER
<OPTIONAL PAUSE>
A: LAST TOPIC CHECK

U: what's your favourite movie?
A: I guess I like movies with a strong AI lead.
U: ok
(2.0)
A: Anything else?
```

Each pattern is expandable with other subpatterns or may present different options according to its conversational structure. Patterns display the same format: an abstract model of the pattern with generic social actions or intents in capital letters and an adaptation of the pattern to a particular use-case. Here the user does the first part of an adjacency pair (an inquiry or a request) successfully granted or answered by the agent. If the agent did not succeed, the third position would not have been filled with an acknowledgement but it would have been expanded with a repair initiator, like "what did you say?" or "define AI". In this case, the agent does not need to say anything more after the acknowledgement and, after 2-3 seconds pause in a vocal interaction or 20-30 seconds pause in a textual interface, the agent checks if the user has other requests or topics.

The adherence to an Intent-Entity-Context-Response paradigm allow to apply the language pattern of the NCF to various platforms and use-cases. Specifically, the examples shown in the textbook are all realized on Watson, but we think that the same patterns could be realized also on other platforms with the same paradigm, like Google's Dialogflow. Furthermore Moore and Arar (2019) methodology is primarily focused on the realization of vocal conversational interaction in English, but we argue that the same patterns with slight adaptation could be implemented also in textual conversational interfaces and in other languages, like in Italian.

3. A roadmap towards the implementation

3.1 A possible proposal: a summary roadmap

In this section, we provide a practical-procedural workflow for approaching chatbot projects based on our work experience in the field and on the most influential textbook guidelines on this topic (Cohen, Giangola, and Balogh 2004; Pearl 2016; Shevat 2017; Hall 2018; Dasgupta 2018; McTear 2020). This workflow is made of macro and micro-levels integrated together and should involve stakeholders from different departments, considering at least the joint effort of marketing, linguistic design and the technical development orchestrated by a flexible methodology that opens with the project's requirements definition up to the implementation.

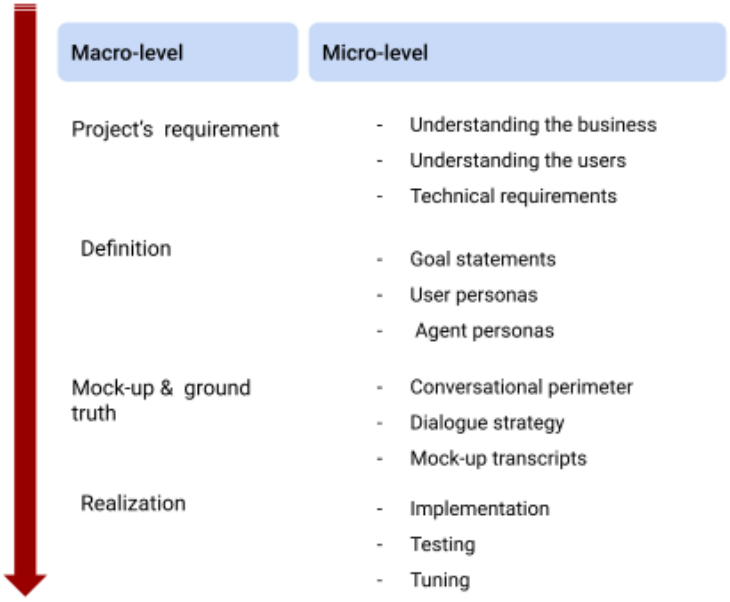

Figure 2
Schema for the proposed roadmap.

The stage of **project's requirements** definition produces as output the acknowledgement of the feasibility of the conversational agent's project, the identification of the potential target users and of the technical requirements the application should possess. We substantially agree with the procedure proposed in Cohen et al. (2004, 46) that covers three micro-levels: *understanding the business*, *understanding the user*,

technical requirements. There are multiple business issues that need to be understood and questions to be answered. Answering these questions can provide significant guidance to also write down metrics for the overall evaluation of the project.

The second element of the requirement phase is the understanding of the population of expected target users, which needs to be understood both in terms of *characteristics/needs* and of *usage modality*. Cohen et al. (2004, 49) list various elements for understanding and taking into account the final user's needs and characteristics.

The second perspective regards how, when, where and why the final users will use the application (Cohen, Giangola, and Balogh 2004, 50). They suggest also two practical ways to get these information: a preliminary overview of the various touchpoints of the company focusing on offered functionalities and transmitted feels, and organizing meeting with the company. The key attitude to develop to gain this information coincides with the first stage of the *Design Thinking* process assumed in Moore and Arar (2019), *empathize*. Empathy towards the interlocutors to get an understanding of the business and the users can be obtained through observations of how the final users currently interact with the industry or resolve the task that will be supplied by the chatbots, or through the engagement of the people directly involved. From a linguistic point of view, examples of useful materials to collect and analyze may be emails, call transcripts, or messagges from the final users to the company.

The third requirement is getting an understanding of the application, focusing on the *technical requirements* of it. This stage is a prerogative of the technical department and its main point is getting an understanding of the application from a technical point of view, evaluating the feasibility itself and the strategies that need to be elaborated for solving tasks and subtasks.

The second macro-level point is **definition**, whose goal is to draw conclusions from the preliminary analysis of the project's requirements in order to define user needs more formally. We identify three elements that have to be defined: goal statements, user personas and agent personas. With *goal statements* we mean the definition of key design criteria learned from the analysis of requirements. They involve the definition of user goals grouped into broader statements, defined also with respect to the technical, financial and organizational constraints of the specific project.

The *user personas* is a fictional representation of the target user: a systematization of the collected user's qualities to define a prototype of who will typically interact with the final product. The user personas should also reflect the users' pain points, in order to improve the user journey. From our personal experience in the field, an efficient way to systematize the users' pain points is to write them down synthetically and associate them with benefits that the chatbot could provide in relation to them. An example is shown in Table 1.

The last element of the second section is the definition of some characteristics of the *agent personas* (Hall 2018; Cohen, Giangola, and Balogh 2004; Pearl 2016). In our experience on the field, we do agree with the synthetical sketch proposed in Moore and Arar (2019) that splits the agent personas design into three components: *agent job description*, *agent personality* and *agent self-knowledge*. A starting point for describing the qualities and the language of an agent, is imagining a job it is supposed to do. What role is the agent supposed to substitute or replace? Trying to list down duties and activities the chatbot is intended to assolve and the expected experiences and qualifications, as if it was a real job candidature, can help to be consistent also in the design itself and throughout the project. This procedure helps to identify jargon and recurrent technical terms the chatbot is supposed to manage. The *agent personality* needs to be characterized

Table 1
Pain points in the user journey associated with possible chatbot's benefits.

Pain Points	Chatbot benefits
Long waiting time on the phone	Instant answer or escalation to an operator
Difficult information retrieval	Personalised user journey to the retrieved information starting with an initial disambiguation
Hard understanding of complex and long documents	Systematisation of information in small slots and simplified language

in terms of communicative style, its level of formality and generally the tone of voice the target users expect to find in the conversational agent they are talking to. Two other core issues of the agent personality are the assignment of a gender to a chatbot, with the cultural and social implications that this may bring with, and the opportunity of humor in it. From our experience on the field, strongly anthropomorphized chatbots usually do belong to a gender, which most of the time is female (West, Kraut, and Chew 2019). There are also cases of neutral chatbots that reproduce an animal, a vegetable or a fantasy character. It depends on what kind of character or conception of gender we aim to reproduce in a virtual reality, being aware of the risk of reproducing virtually biases or prejudices belonging to the real word (Strengers and Kennedy 2020).

Humor can be an efficient strategy to build trust, especially if used in secondary responses that do not cover the main topics of the chatbot. Since the users expect a chatbot to be productive and efficient (Brandtzæg and Følstad 2017, 2018; Piccolo, Mensio, and Alani 2018; Zamora 2017), humoristic responses rather than informative ones can be counterproductive and may indeed frustrate the user. On the other hand, receiving a humoristic answer in an unexpected context such as online conversation with a machine, can increase the surprise effect and it may induce the users to continue the conversation (Jain et al. 2018).

Another correlated aspect involved in sketching a chatbot personality is the definition of some conversational paths that do not constitute the core topics of the chatbots but they are in some way related to it and can entertain the user. For example, a customer-service chatbot of an online motorbike clothing may provide an answer for a question like "What is your favourite motorbike brand?".[3] Another way to reinforce the users' confidence towards the agent is working on the agent self-knowledge (Przegalinska et al. 2020; Følstad and Brandtzaeg 2020; Følstad, Nordheim, and Bjørkli 2018). A conversational agent can not have real perception of itself, but providing conversational paths that may help the users to navigate the conversational space created by the chatbot and the chatbot itself can be an efficient way for helping the user understanding what the chatbot can actually do and say and asking him more pertinent questions. Questions like this may regard the chatbot itself ("What are you?" "Are you a human?" "What is a

3 Business and commercial constraints have to be considered as well. In this case, the chatbot may not be able to provide an answer citing a specific brand, and therefore indirectly supporting a brand, but it can answer with a generic: "In my lonely virtual world I can only ride the wings of fantasy". Furthermore, some popular NLU platforms like Dialogflow contain pre-built conversational agents enriched with small-talks conversational paths covering generic topics such as weather, hour or day of the week.

chabot?") or the competences of the chatbot ("What can you do?" "What can I ask you?" "What do you know'").

The third step in the proposed methodology is called **mock-up & ground truth**. Once the target users and the goal statements have been sketched out, we can move on to the drafting of the *conversational perimeter*. This term identifies a sort of table that holds together the groups of topics managed by the agent, structured in a way that reflects the Intent-Entity-Context-Response paradigm of the most common NLU platforms. In this phase, it is important to define the intents coherently and functionally to the goals of the agent. If conversational data collected during the preliminary phase such as call phone transcripts or emails are available, we suggest to group them following a *bottom-up labelling approach*. This approach consists of first grouping collected utterances into wide general categories, like "questions", "problems", "getting information" and then proceeds refactoring and splitting the so funded categories into more specific ones, considering for example the topic of the question, the action required to satisfy it and so on. Through this procedure we can both identify intents and train them with authentic linguistic material that constitute the so-called training phrases. A schema of how it works is shown in Figure 3 representing a schematisation of the reasonings behind the identification of the topics that will be handled by the agent, but it is not proper conversational perimeter as we mean it.

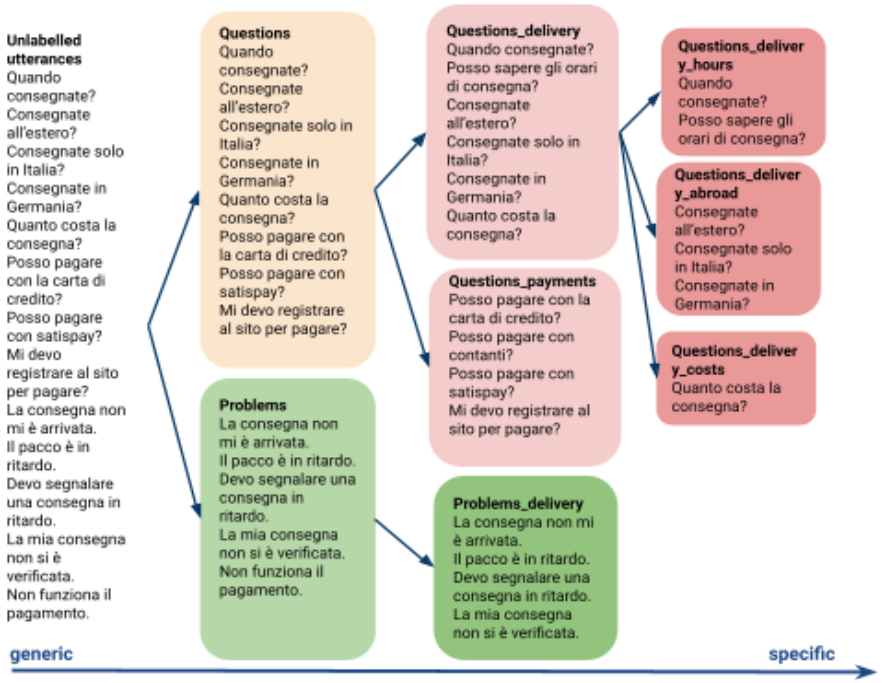

Figure 3
Schema of the bottom-up labelling approach for the perimeter design.

The conversational perimeter is the final elaboration of this reasoning. It also includes a section dedicated to the responses, that should coincide with the actual responses that the agent gives in chat and has to be updated with new information

or modifications. Furthermore, it also contains some examples of the *training phrases*, utterances that final users can potentially say to formulate a request. Training phrases are categorized in *intents*, in turn grouped according to the topic they refer to. A conversational perimeter offers a systematization of the topics handled by the conversational agent, organized in intents like the platform it will be developed on; and it shows how simple question-answer exchanges will be managed directly in chat, with some examples of utterances the final users may produce and the answers they will be given. Table 2 shows the different section of a conversational perimeter, re-adapting the utterances from Figure 3 to the final elaboration of the perimeter.

Table 2
A section of a conversational perimeter.

Topic	Training phrases	Intent	Response
Consegna	Quando consegnate? Posso sapere gli orari di consegna? Quando fate la consegna?	Consegna_orari	Consegniamo tutti i giorni della settimana, compresi i festivi.
Consegna	Consegnate all'estero? Consegnate solo in Italia?	Consegna_estero	Consegniamo solo in Italia, non all'estero.
Consegna	Quanto costa la consegna?	Consegna_costi	Il costo della consegna è gratuito per gli utenti registrati e di 4 euro per gli ospiti.
Consegna	La consegna non mi è arrivata. Il pacco è in ritardo. Devo segnalare una consegna in ritardo. La mia consegna non si è verificata.	Consegna_ritardo	Se ritieni che ci sia un ritardo di consegna, posso inoltrare una segnalazione al servizio clienti, mi basta solo qualche tuo dato.
Pagamenti	Posso pagare con carta di credito? Posso pagare con satispay? Posso pagare con contanti?	Pagamenti_metodi	Puoi pagare con paypal, satispay e carta di credito registrata.
Pagamenti	Mi devo registrare al sito per pagare?	Pagamenti_account	Puoi comprare dal nostro sito sia come ospite sia come utente registrato, a te la scelta!

We need to consider that a complete conversational perimeter is usually much bigger than this: it depends on the specific project, but in our experience it can contain between fifty and one hundred intents. Furthermore, it contains only the simplest form of interaction: the responses are static, which means that they do not change dynamically taking information from external sources, but they are always the same, even though on the most common NLU platforms they can be randomized, that is to a specific intent can be associated with one or more responses randomly picked-up by the agent to create conversational variety. The form of the responses should respect the basic language-specific pragmatic norms, as shown in section 4.

The second micro-level step of the mock-up and ground truth mapping is dedicated to the *dialog strategy*, concerned with the effective building of the dialogue. It answers the question: how will the back and forth between the agent and the final users be? Will generally the agent start the conversation or the user? If the conversation flow is supposed to be more complicated than a simple question-answer and it requires the following of a specific path, we recommend the use of graphical tools for visualising

the steps of the path we have imagined and possible variations, such as Google Draw or XMind.

Mock-up transcripts is thought especially for more complex flows, that need more than one conversational turn to be developed and that may also be represented on one of the graphical tools mentioned above. In this phase there are two alternatives that can be pursued. From one side we can simply write down the dialogues we would like to reproduce with the agent (Cohen et al. (2004) and Pearl (2016) call them *sample dialogues*) and read it aloud to see if they sound human before implementing it, using either programming languages or commercial platforms that allow to concatenate conversational flows reproducing the characteristics identified by CA mentioned before. Alternatively, instead of creating the flows from scratch and implementing them, our proposal considers the Natural Conversation Framework and especially the sequential patterns proposed by Moore and Arar (2019). The latter approach appears more innovative and more structured, and it is the one we would like to adopt.

The final macro-level step of the roadmap is the concrete **realization** of the agent and consequently the effective implementation of the defined flows and selected sequential patterns on a NLU platform, on specific chatbot tools or using the most common programming languages.

After the implementation, a fundamental step before the roll out is *testing*. Like McTear (2020) claims, there is still no unified and univocal testing approach, but it depends on the implementation method behind the agent (Deriu et al. 2020) and on the project's requirements that need to be evaluated. Testing has a double function: to evaluate the efficiency of the developed application and, if possible, to improve the actual functionalities with a tuning activity. From our experience on the field, we suggest various test steps before the final roll-out. This allows developers and designers to evaluate the results and, if necessary, to tune some aspects of the application before the final version. A *dialog transversal test* (Cohen, Giangola, and Balogh 2004; Pearl 2016) for evaluating the behaviour of the system in every dialogue state and in every condition seems very important, trying out some out-of-perimeter utterances to verify the proper response of the system also in such cases. A fundamental aspect that needs to be considered in this context, is that a conversational agent is a constantly-in-progress creature: even though it is finished and especially if it supports a NLU tool, it needs to be updated with new linguistic materials in order to make the performances better and better. It has to be seen as an alive creature and project, that is nurtured by language and, in some way, produces language: therefore, training and maintenance are continuative activities. There are multiple elements that can be tuned or updated to make the performance of the chatbot better: the training phrases in natural language, the responses in case of changings in the information to deliver, new conversational paths. Table 3 shows a possible evaluation framework taken from our experience on the field and resulted from the combination of two variables: *in/out of perimeter*, which is referred to the coherence of the utterance produced by the user in relation to the conversational perimeter of the conversational agent; and *correct/wrong*, which is the effective evaluation of the response given by the conversational agent in that specific context. From every possible combination of these variables we provide possible improvements that can be undertaken.

4. Adaptation of the methodology and implementation

After the definition of a theoretical approach rooted in pragmatics and a practical-computational operating procedure towards the implementation, in this section we

Table 3
Evaluation grid.

	In perimeter	**Out of perimeter**
Correct	The user's utterance is in perimeter and the chatbot answers with the expected associated response: no improvements needed.	The user's utterance is out of perimeter and the chatbot answers properly activating the expected fallback intent: no improvements needed.
Wrong	The user's utterance is in perimeter and the chatbot answers wrongly with a response associated with another intent. Improvements may be: updating the training phrases of the missed intent, verifying the training phrases of other intents that may cause miss-match, or modify the response with more information.	The user's utterance is out of perimeter and the chatbot doesn't activate the fallback intent,but a response associated with another intent. Improvements may be: adding more training phrases to a fallback intent, adding training phrases to the missed intent, verifying training phrases that may have caused miss-match.

expand and enrich the procedural workflow with the implementation of a selection of patterns on one of the most authoritative NLU commercial platforms also for the Italian language, namely Google's Dialogflow (Zubania et al. 2020).

The selection of the patterns has been made according to two criteria: (a) the adaptation to a text-based modality of interaction; (b) the suitability to a customer-service context (Szymanski and Moore 2018).

We elaborate a general customer-service case study on which we transpose a selection of patterns in Italian. This practical section on the implementation aims to demonstrate the high generalizability of the approach, still considering the language-specific pragmatic implications in prompt design, and its suitability also to business-oriented contexts of use.

4.1 General case-study description

The structure of conversation belonging to a customer-service domain is similar to the more general category of service conversations, i.e. dialogic exchanges in which a person (in the role of a customer or a citizen) requests services or information and another person on behalf of an organization or an institution, provides services or information.

Since the focus of our work is not to build a comprehensive conversational agent but to demonstrate the effectiveness of our methodology for the design of sequential flows, we do not dwell on the details of the conversational perimeter of the case-study.

It suffices to say that it is a customer-service chatbot, whose goal is to provide repetitive information to support customers on the e-commerce of a chain of shops. It can provide information about typical online-shopping requests, such as deliveries, expeditions, payments and returns and it manages issues and problems related to the state of the orders. Customers may own a *fidelity card*, a card they collect points on to gain special discounts. Furthermore, customers can register on the website and activate an online profile with all the details on their customer's situation. Even though the main goal of the conversational agent is not to perform complex task, it should be able to support the final users through the registration process on the website. The chatbot

handles it through some questions focused on the extraction of data. We can distinguish between two kinds of data to extract: (a) user's name and user's email address are necessary data to complete the registration process. Without the collection of these data, the registration can not be performed; (b) the fidelity card number is an optional data. If the user decides to not provide it or the user does not possess a fidelity card, the registration process it is not compromised.

The registration process is the interaction that we are going to transpose in Italian using the NCF pattern and implement on Google's Dialogflow, in order to demonstrate the validity and the high generalizability of the presented theoretical and methodological approach.

4.2 Pattern selection and transposition

The registration process on the website can be realised basically associating three patterns of the NCF described in Table 4, namely the **pattern A2.6 Open Request Summary** the **pattern A2.7 Warrant Request & Refusal** and the **pattern A2.11 Open Request Repairs**.They all belong the the first group of pattern, *conversational activities*, and therefore help to manage what happens *inside* the boundaries of the conversation itself.

As shown in the section 2.4, each pattern is made of an abstract model of social actions in capital letters and an adaptation of the pattern to an hypothetical dialogue.

The pattern *A2.6 Open Request Summmary* manages requests that need a progressive level of detail to be satisfied, like a registration process, in which a set of data need to be collected in order to perform an action or to fill-in a module. The peculiarity of this pattern is the final summary at line (8), that should allow a user's confirmation or modification of the inserted information. The modification can be performed using the pattern *A2.11 Open Request Repairs*, which handles repair procedures. The pattern *A2.7 Warrant Request & Refusal* manages user's questions concerning a detail request and the possible refusal of giving a specific data. As described in 4.1, we have both compulsory and optional data that need to be handled in different way.

In the next section, we explain the pragmatic aspect we take into account for the transposition into Italian and we then implement them on Google's Dialogflow.

4.3 Pragmatic aspects for the transposition into Italian

The pragmatic-functionalist perspective adopted in this work emerges both in the turn-taking sequential structure of the patterns of NCF and in the transposition of the chosen sequential patterns into the Italian language. The observations made by Cohen et al. (2004) about prompts design are still actual and rich in ideas, but they are concerned exclusively with the English language.

In our work these aspects become relevant for the distribution of content in the Italian transposition of the selected patterns. We are going to transpose patterns from a language to another, namely from English to Italian, respecting the predominant tendency of positioning on the left side of the utterance the topic/given information and the comment/new information on the rest of the sentence. This norm can be violated for obtaining particular sense-effects and for enlightening a part of an utterance that in a normal order of the component would instead be positioned at the end of the sentence. In Italian the most used strategies for producing this effect on a written channel of communication are the inversion of the normal order of the syntactic components through *focalizations* such as *dislocations* and *cleft sentences*.

Table 4
Selected patterns for the transposition into Italian.

Pattern A2.6 Open Request Summary	
1 U: PARTIAL REQUEST	U: I need to register
2 A: DETAIL REQUEST	A: Okay. What's your name?
3 U: DETAIL	U: Dave Bowman.
4 A: DETAIL REQUEST	A: What's your job title?
5 U: DETAIL	U: astronaut
6 A: DETAIL REQUEST	A: What is your organizational affiliation?
7 U: DETAIL	U: NASA
8 A: REQUEST SUMMARY	A: So Dave Bowman, Astronaut at Nasa. Is that correct?
9 U: SUMMARY CONFIRM	U: yes
10 A: GRANT	A: Ok.
11 U: SEQUENCE CLOSER	U: thank you
12 A: RECEIPT	A: You're welcome!
Pattern A2.7 Warrant Request & Refusal	
1 A: DETAIL REQUEST	A: What's your job title?
2 U: WARRANT REQUEST	U: why do you need to know?
3 A: WARRANT	A: Job title enables us to know the range of professions of our participants.
4 U: REFUSAL	U: none of your business.
5 A: ACKNOWLEDGMENT	A: No problem. What's your organizational affiliation?
Pattern A2.11 Open Request Repairs	
1 U: FULL REQUEST	U: Can you recommend a nearby Mexican restaurant?
2 A: GRANT	A: Mario's located at Beach and Main.
3 U: REPAIR INITIATOR	U: What do you mean?
4 A: REPAIR	A: The Mexican Restaurant, Mario's, is located at the intersection of Beach Boulevard and Main Street.
5 U: SEQUENCE CLOSER	U: thanks

Another element on which Cohen et al. (2004) dwell on are *discourse markers* and their use in prompts's content design for conversational agents in English, organized into various classes of use and considered an efficient strategy for enhancing the natural flow of conversational technology. Without going deeper into the classification of discourse markers in Italian (among the others Bazzanella (1995) and Berretta (1984)), it suffice to say that we will take them into consideration for their pragmatic impact on sentences organizations and structures during the patterns' transposition. In Italian, they are indeed significantly important for their interactional function on both sides of the speakers, their meta-textual function as demarcations, focalizations and riformulations and finally for their cognitive functions (Bazzanella 2008). As Cohen et al. claim, it

can be that discourse markers are perceived as informal or slang, even though they are classified as a *functional* category with no explicit formality degree.

What instead has to do with the level of formality of an utterance and a piece of discourse in general (Clark 1996), are *register* and *consistency*. Even though the definition of register is controversial (Bazzanella 2008), we can say it involves the psychological and social rapports between the speakers, the circumstances in which the communication takes place and the adopted channel (Halliday 1994). This is one of the dimensions of variations of language. Other dimensions of variations are determined by the spatial origins and the geographical distribution of the speakers, by socio-cultural elements such as level of instruction, age and competences and channel of communication. Especially in technical or highly specialized work context, the use of jargon is an important issue to take into account. It is indeed acceptable if all the speakers do share similar background and analogous competencies in relations to the main content expressed by the conversational agent we are working on.

The channel of communication is also an issue with significant implications on the distribution of content and therefore on relation to the drafting of the responses. Without deepening into the characterization of the language variations adopted in digital contexts of communication, we have to consider that the variant of language adopted with a text-based conversational agent is an intermediate solution between the two opposite poles of written and oral language. This variation takes place in a written form but it shares important elements with speech (Pistolesi 2018), since it is a type of writing that considers more the acoustic effect than the visual one.

All these sociolinguistic and pragmatic aspects have to be considered as requirements for transposing the selected patterns into Italian. In the next section, we are going to outline a generic customer-service use-case and in followings there is the effective realization and transposition of the patterns on Google Dialogflow.

4.4 Implementation on Google's Dialogflow

We realize the selected NCF patterns on Google's Dialogflow ES.

The three patterns have been combined together in order to cover multiple scenarios we may face during a registration process. For doing that, we implemented fourteen intents on Dialogflow

01_reg_00_registrazione_generico	*01_reg_03_registrazione_riepilogo*
01_reg_01_giustificazione_nome	*01_reg_03_rifiuto_email*
01_reg_01_registrazione_carta	*01_reg_03_spiegazione_email*
01_reg_01_rifiuto_nome	*01_reg_04_modifica_dati*
01_reg_02_registrazione_email	*01_reg_04_riepilogo_corretto*
01_reg_02_spiegazione_tessera	*01_reg_04_riepilogo_negativo*
01_reg_02_tessera_rifiuto	*01_reg_05_ringraziamenti*

connected together by input and output contexts manually set and we use three system entities for the parameter extraction: (a) *sys.person* to extract and memorise the user's name, (b) *sys.number* to extract and memorise the fidelity card's number and (c) *sys.email* to extract and memorise the user's email.

The opening intent *01_reg_00_registrazione_generico* is activated by training phrases that express the intention of starting a registration process on the website. As Figure 4 displays, the contents' distribution of the agent's response follows the topic/comment order described in 4.3: the first part introduces the registration process already known

by the user, since he/she asked for it, and prepares the speaker to provide a set of data. In this case, the user provides the requested data and the user's name is thus memorised through the annotation of the training phrases with the *sys.person* entity and the extraction of the corresponding parameter.

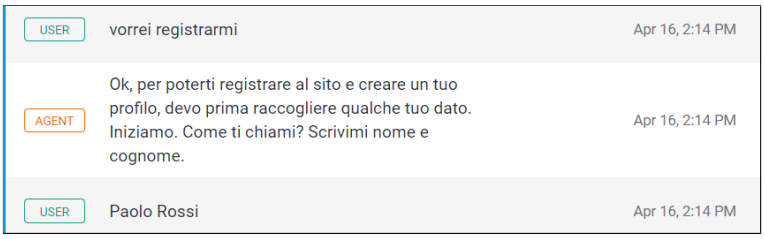

Figure 4
Realization of the pattern *Open Request Summary*. (TRANS. U:I would like to register/A:Ok, in order to register on the site and create your own profile, I must first collect some of your data. Let's begin. What's your name? Write your name and surname/U: Paolo Rossi).

As the pattern *A2.7 Warrant Request & Refusal* demonstrates, repetitive patterns in a process based on the data extraction, such as the outlined registration, are users' requests of warranty and explanation concerning a specific data. We provided two intents for managing questions about the reasons of requesting a name in this context and the refusal of providing it, respectively the intent *01_reg_01_giustificazione_nome* and the intent *01_reg_01_rifiuto_nome*, both conncted to the hierarchical intent *01_reg_00_registrazione_generico* by contexts. The intent *01_reg_01_rifiuto_nome* is also subordinated to the *01_reg_01_giustificazione_nome*, which means that after an explanation's request by the user the chatbot is able to handle also a refusal of providing the requested data (see Figure 5).

The chatbot's responses are more extended then the examples provided by Moore and Arar (2019). This is because we are specifically working on a text-based interface, which allows us to insert also a bit longer prompt than with only vocal conversational assistant. As we claimed in Section 4.1, the user's name is a required information without which the registration process cannot be performed. The conversational agent offers an alternative to the user, *("Se vuoi, puoi registrarti anche dal vivo, ti basta andare alla cassa [...]")* because we are in a business-oriented context, where the customer satisfaction is the most important requirement (Gnewuch, Morana, and Maedche 2017). If after the activation of the intent *01_reg_01_rifiuto_nome* the user at the end provides the name, the registration process can proceed again.

The second data to extract is the fidelity card number, which is not compulsory and it can be skipped without compromising the registration. Also in this case we have a *good path* (Pearl 2016) where the number is immediately provided, and two expanding intents used to manage clarification's requests and refusal related to this topic, namely the intent *01_reg_02_spiegazione_tessera* and the intent *01_reg_02_tessera_rifiuto*. Since the fidelity card number is not a sensibile data, instead of a warranty we provide further explanation about the number (where it is, what it is, why it is requested) and possibility of managing refusal, especially if determined by incomprehension, impossibility to find the card or absence of the card (see Figure 6).

In any case, if the fidelity's card has been provided or not, the registration conversational flows proceeds to the third and last data, the user's email, which is, like the name, a compulsory data to complete the registration, and it is therefore handled by the intents *01_reg_03_rifiuto_email* and *01_reg_03_spiegazione_email* with training phrases

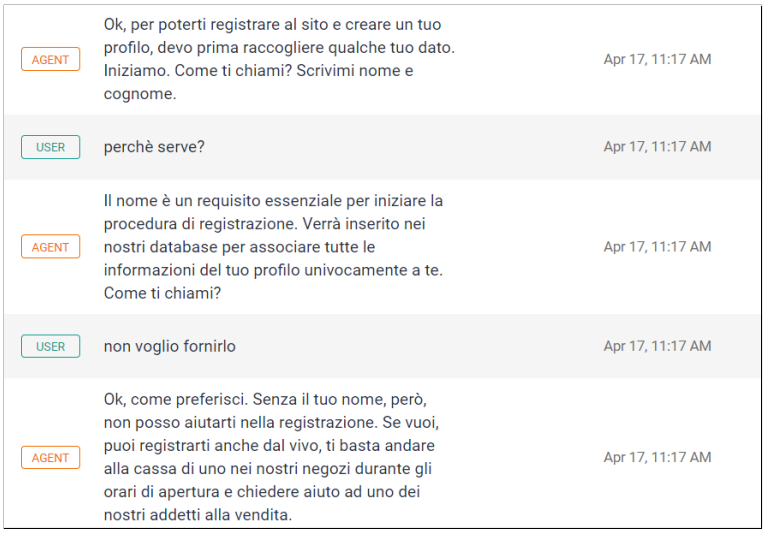

Figure 5
Realization of the pattern *Warrant Request & Refusal.* (TRANS. A:Ok,...*same as before*...surname/U:Why is it needed?/A:The name is an essential requirement to start the registration process. It will be entered in our databases to associate all the information in your profile uniquely with you. What's your name?/U:I don't want to provide it/A:Ok, as you like. Without your name, however, I cannot help you register. If you want, you can also register live, just go to the cashier of one of our stores during opening hours and ask for help from one of our sales staff).

similar to ones dedicated to the name, with the difference that they are annotated with the *sys.email* entity.

As displayed in the pattern *A2.6 Open Request Summary*, after the collection of data we have a summary of all the provided information memorised through contexts and reported in the responses of the intent *01_reg_03_registrazione_riepilogo* through coding some specific variables through a specific Dialogflow syntax, as Figure 7 shows. These are the responses that the conversational agent displays once the *01_reg_03_registrazione_riepilogo* is activated by the insertion of an email address. They contain parameters associated to the data extracted with entities and stored in the conversational flow through contexts, according to the form: *#context-name.parameter-name*. The two responses are chosen randomly from Dialogflow: it means that there is no certainty that if the user did not provide the fidelity card's number in the flow the conversational agent will display the response with *"carta fedeltà non fornita"*. Our tests reveal a tendency to choose the appropriate response according to the collected data, but for more precise result we recommend to customize the response with some code. In Figure 8 we show how they appear if filled with the corresponding value.

The summary ends with a agent's request of confirmation before validating the inserted data. If the user confirms, the data are validated and the sequence can be closed with a sequence closer and its reception. If the users would like to modify one of the inserted data, a repair procedure can be started. The repair procedure can pass through the intermediate intent *01_reg_04_riepilogo_negativo*, where the user simply acknowledges that the proposed summary is not the correct, or it can directly activate the intent *01_reg_04_modifica_dati*, the core-intent of the repair. Figure 9 shows a

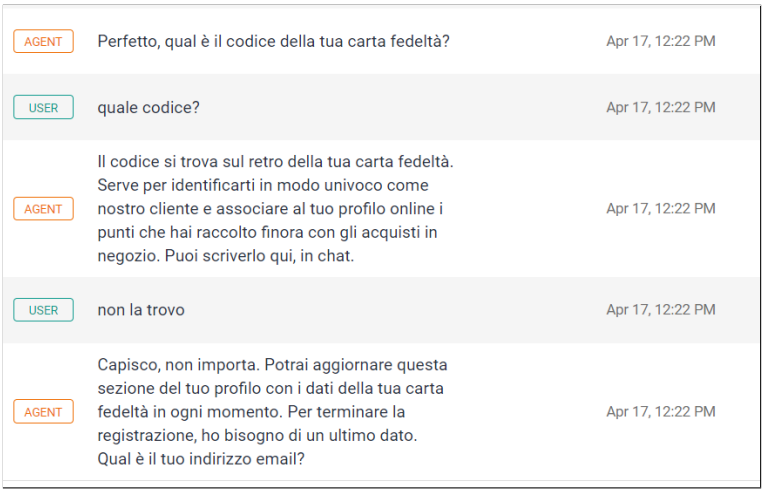

Figure 6
Realization of a slightly modified version of the pattern *Warrant Request & Refusal* according to the type of data to extract. (TRANS. A:Perfect, what is your loyalty card code?/U:Which code?/A:The code can be found on the back of your loyalty card. It is used to uniquely identify you as our customer and associate the points you have collected so far with in-store purchases to your online profile. You can write it here in chat/U:I cannot find it/A:I see, it doesn't matter. You can update this section of your profile with your loyalty card details at any time. To end the registration, I need one more data. What is your e-mail address?).

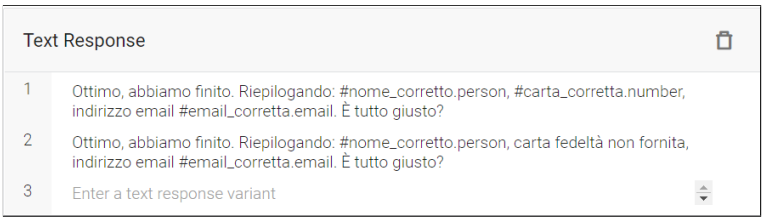

Figure 7
Dialogflow's syntax for the final recap in the pattern *Open Request Summary*.

repair procedure with the activation of both the intents *01_reg_04_riepilogo_negativo* and *01_reg_04_modifica_dati*.

The repair procedure is handled by the intent *01_reg_04_modifica_dati*, connected to the previous ones by contexts and whose training phrases are annotated with the system entities used before. The difference here is that the parameters associated to the entities and extracted from the user's sentences are saved with another value. The new extracted value is used in the response with the same `#context-name.parameter-name` syntax, in order to overwrite the previous value with the new one. The next figures show the parameter section in Dialogflow and the responses with the extracted parameters.

The prompts are enriched with discourse markers that functionally structure the progression of the conversational flow (*"Va bene"*, *"Ok, quindi"*, *"Ok"*, *"Perfetto"*,*"Ottimo, abbiamo finito"*).

This demonstration does not expect to exhaust all the possibilities and expansions that a comprehensive registration process may present. The repair procedure handles for instance the modification of a single data and we did not focus on the insertion of

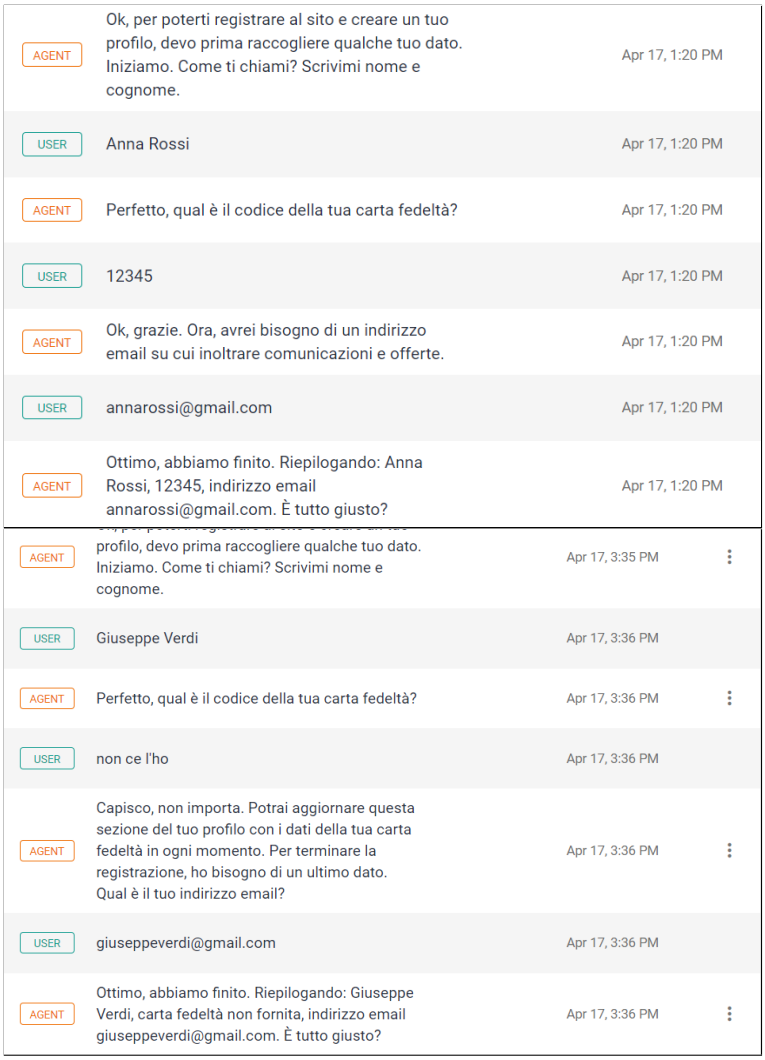

Figure 8
Realization of the pattern *Open Request Summary* with the two scenarios: in the first conversation the user provides all the requested data while in the second the users does not provide the fidelity card number. (TRANS. *First Part*: A:Ok,...*as in Figure 5*...surname/U:Anna Rossi/A:Perfect, what is your loyalty card code?/U:12345/A:OK thanks. Now I need an email address to forward communications and offers to/U:annarossi@gmail.com/A:Good, we have finished. Summarising: ... Is it correct? *Second Part*: ...*same as before*... U:/I do not have it/A:I see, it doesn't matter. You can update this section of your profile with your loyalty card details at any time. To end the registration, I need one more data. What is your e-mail address?/U:giuseppeverdi@gmail.com/A: ...*same as before*...).

unexpected or wrong-spelled data. The goal of our implementation is to demonstrate the effectiveness and the validity of an approach, which is supported by an high-level theoretical method and which is easily transposable into other languages and using various NLU commercial platforms.

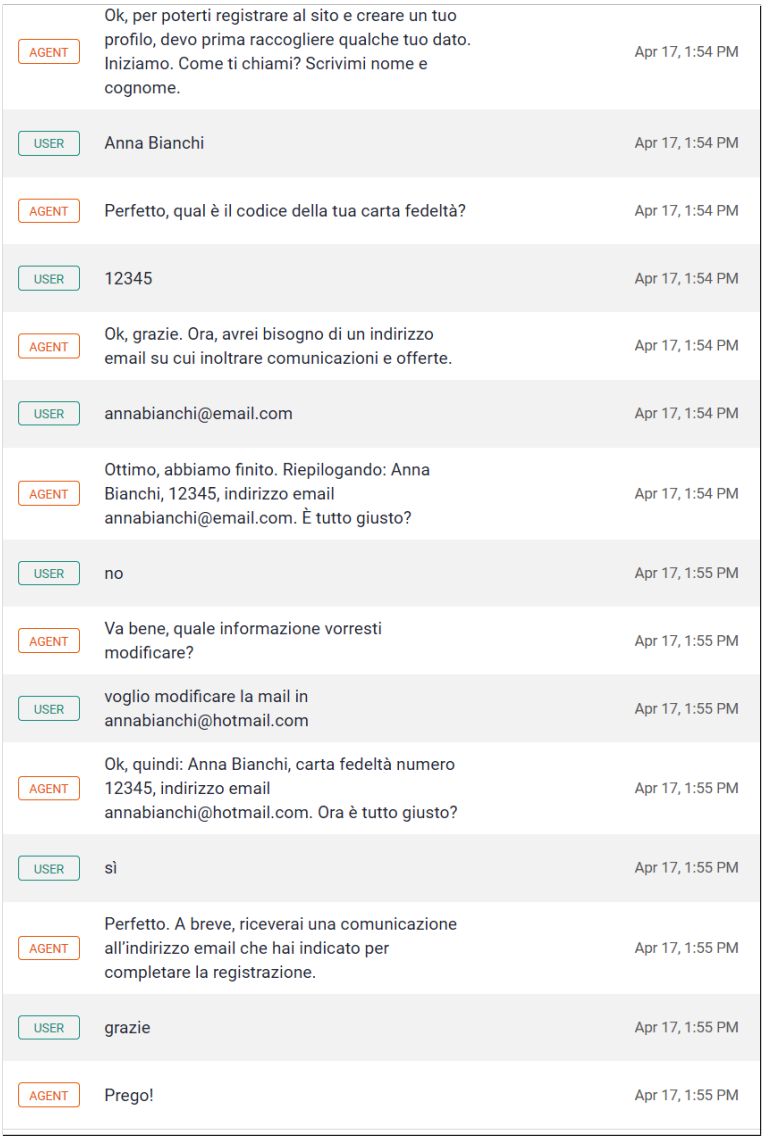

Figure 9
Realization of the pattern *Open Request Summary* and *Open Request Repairs*. (TRANS. *The first 7 steps as in previous figures*/U:No/A:Ok, which information would you like to change?/U:I would like to change the email in annabianchi@hotmail.com/A:Ok,...*same as before*.../U: Yes/A:Perfect. Soon, you will receive a communication to the email address you indicated to complete the registration/U:Thank you/A:You are welcome!).

5. Conclusions and future directions

We have therefore obtained a formal, highly generalisable procedure for building conversational agents. Our procedure is explicitly anchored to Conversation Analysis as authoritative methodological approach for realising sequential exchanges and reproducing typical features of human conversational interactions. We additionally transposed

Figure 10
Dialogflow's syntax and parameters for the realization of the pattern *Open Request Repairs*.

the patterns in Italian considering the pragmatic implications of the chosen language and selected a flexible case-study that allows to easily recreate the patterns in multiple contexts and situations.

The generalisability of the selected patterns is thus subordinated to two main factors: the cultural implications that may occur in the transposition of the patterns into another language and the level of complexity of the conversational flows to implement. The problem of the transposition of the pattern deals with the pragmatics of the language and the cultural expectations related to the customer experience that may influence the dialogue design as well. As Brandtzaeg and Følstad (2017) claim the search of productivity in chatbot use is explicitly anchored to Western culture. This aspect conditions the entire dialogue design, for example in positioning the chatbot scope within the very first conversational turns, in order to not waste time. In our experience, the adaptation of the pattern have been made between two rather culturally similar language, but it might not have been the same with, for example, an oriental culture and its language. Beyond the strictly linguistic aspects, also the project requirements and the agent personas definitions are also culturally defined (Ruane, Birhane, and Ventresque 2019; Cardinal, Gonzales, and Rose 2020). The generalisability of the patterns is also determined by the scope of the specific conversational agent that need to be implemented and by the level of granularity of the conversational flows it should have: the patterns can been seen as a base model that can be simplified or complicated depending on the individual needs.

As future directions of our research, we are going to further investigating the applicability of this approach on the Dialogflow CX version and possibly on other commercial NLU platforms.

Acknowledgements

We would like to thank Injenia S.r.l. for supporting this research.

CRediT author statement; ADA: Conceptualization, Methodology, Validation, Formal Analysis, Investigation, Writing (Original Draft), Writing (Review & Editing); FT: Conceptualization, Supervision, Project Administration, Writing (Review & Editing), Funding Acquisition.

References

Adamopoulou, Eleni and Lefteris Moussiades. 2020. An overview of chatbot technologies. *Artificial Intelligence Applications and Innovations*, 584:373–383.

Ahmad, Nahdatul Akma, Mohamad Hafiz Che Hamid, Azaliza Zainal, Muhammad Fairuz Abd Rauf, and Zuraidy Adnan. 2018. Review of chatbots design techniques. *International Journal of Computer Applications*, 181(8):7–10.

Al-Duleimi, Hutheifa Y., Sabariah Md. Rashid, and Ain Nadzimah Abdullah. 2016. A critical review of prominent theories of politeness. *Advances in Language and Literary Studies*, 7(6):262–270.

Allen, James F., Donna K. Byron, Myroslava Dzikovska, George Ferguson, Lucian Galescu, and Amanda Stent. 2001. Toward conversational human-computer interaction. *AI Magazine*, 22(4):27–38.

Ariel, Mira. 2010. *Defining Pragmatics*. Cambridge, Cambridge University Press.

Ariel, Mira. 2012. Research paradigms in pragmatics. In K. Allan and K. M. Jaszczolt K. M. Jaszczolt, editors, *The Cambridge Handbook of Pragmatics*. Cambridge, Cambridge University Press, pages 23–46.

Austin, John Langshaw. 1962. *How to Do Things with Words*. Oxford, Clarendon Press.

Bazzanella, Carla. 1995. I segnali discorsivi. In L. Renzi, G. Salvi, and A. Cardinaletti, editors, *Grande grammatica italiana di consultazione, vol. 3 (Tipi di frase, deissi, formazione delle parole)*. il Mulino, Bologna, pages 225–257.

Bazzanella, Carla. 2008. *Linguistica e pragmatica del linguaggio. Un'introduzione*. Laterza.

Bennett, Greg. 2018. Conversational style: Beyond the nuts and bolts of conversation. In R. J. Moore, R. Arar, H. Szymanski M., and G. Ren G. Ren, editors, *Studies in Conversational UX Design*. Springer, International Publishing, pages 161–180.

Berretta, Monica. 1984. Connettivi testuali in italiano e pianificazione del discorso. In *Linguistica testuale. Atti del XV congresso internazionale della Società di Linguistica Italiana*. pages 237–254.

Bianchini, Alessia, Francesco Tarasconi, Raffaella Ventaglio, and Mariafrancesca Guadalupi. 2017. "gimme the usual" - how handling of pragmatics improves chatbots. In *Proceedings of the Fourth Italian Conference on Computational Linguistics (CLiC-it 2017)*, pages 30–35, Roma, Italy, December.

Bickmore, Timothy W., Ha Trinh, Reza Asadi, and Stefán Ólafsson. 2018. Safety first: Conversational agents for health care. In Robert J. Moore, Raphael Arar, Margaret H. Szymanski, and Guang-Jie Ren, editors, *Studies in Conversational UX Design*. editors, Springer International Publishing, pages 161–180.

Brandtzæg, Petter Bae and Asbjørn Følstad. 2017. Why people use chatbots. In I. Kompatsiaris, J. Cave, A. Satsiou, G. Carle, E. Kontopoulos, S. Diplaris, and D. McMillan and, editors, *Internet Science. INSCI 2017 (Lecture Notes in Computer Science)*. volume 10673, Springer, pages 377–392.

Brandtzæg, Petter Bae and Asbjørn Følstad. 2018. Chatbots: changing user needs and motivations. *Interactions*, 25(5):38–43.

Branigan, Holly P., Martin J. Pickering, and Alexandra A. Cleland. 2000. Syntactic coordination in dialogue. *Cognition*, 75(2):B13–25.

Braun, Daniel, Adrian Hernandez Mendez, Florian Matthes, and Manfred Langen. 2017. Evaluating natural language understanding services for conversational question answering systems. In *Proceedings of the 18th Annual SIGdial Meeting on Discourse and Dialogue*, pages 174–185, Saarbrücken, Germany, August.

Braun, Daniel and Florian Matthes. 2019. Towards a framework for classifying chatbots. In *Proceedings of the 21st International Conference on Enterprise Information Systems (ICEIS 2019)*, pages 484–489, Heraklion, Greece, May.

Brown, Gillian and George Yule. 1983. *Discourse Analysis*. Cambridge University Press.

Brown, Penelope. 1976. Women and politeness: A new perspective on language and society. *Reviews in Anthropology*, 3(3):240–249.

Brown, Penelope and Stephen C. Levinson. 1987. *Politeness - Some universals in language usage*. Cambridge University Press.

Bublitz, Wolfram and Neal R. Norrick. 2011. Introduction: The burgeoning field of pragmatics. In W. Bublitz and N. R. Norrick, editors, *Foundations of Pragmatics*. vol. 1 of Handbooks of Pragmatics, Mouton de Gruyter, pages 1–20.

Candello, Heloisa and Claudio Pinhanez. 2018. Recovering from dialogue failures using multiple agents in wealth management advice. In R. J. Moore, M. H. Szymanski, R. Arar, and G. Ren G. Ren, editors, *Studies in Conversational UX Design*. Springer, International Publishing, pages 139–160.

Canonico, Massimo and Luigi De Russis. 2018. A comparison and critique of natural language understanding tools. In *CLOUD COMPUTING: The Ninth International Conference on Cloud Computing, GRIDs, and Virtualization*, pages 110–115, Barcelona, Spain, February.

Cardinal, Alison, Laura Gonzales, and Emma J. Rose. 2020. Language as participation: Multilingual user experience design. In *Proceedings of the 38th ACM International Conference on Design of Communication (SIGDOC '20)*, Denton, TX, October.

Carnap, Rudolf. 1938. Foundations of logic and mathematics. In O. Neurath, R. Carnap, and C. W. Morris and, editors, *International Encyclopedia of Unified Science, vol. I*. University of Chicago Press, pages 139–214.

Chaves, A. P., E. Doerry, J. Egbert, and M. Gerosa. 2019. It's how you say it: Identifying appropriate register for chatbot language design. In *Proceedings of the 7th International Conference on Human-Agent Interaction*, pages 102–109, Kyoto, Japan, May.

Cho, HyunKyoung. 2015. Toward a new design philosophy: Politics and the aesthetic of "we" human-and-technology in interaction design. In C. Stephanidis, editor, *HCI International 2015 - Posters' Extended Abstracts. HCI 2015. Communications in Computer and Information Science*. vol 528, Springer, Cham, pages 13–18.

Cho, Hyunkyoung and Joonsung Yoon. 2013. Toward a new design philosophy of hci: Knowledge of collaborative action of "we" human-and-technology. In M. Kurosu, editor, *Human-Computer Interaction. Human-Centred Design Approaches, Methods, Tools, and Environments. HCI 2013. Lecture Notes in Computer Science*. vol 8004, Springer, Berlin, pages 32–40.

Clark, Herbert H. 1996. *Using Language*. Cambridge University Press.

Clark, Leigh, Nadia Pantidi, Orla Cooney, Philip Doyle, Diego Garaialde, Justin Edwards, Brendan Spillane, Emer Gilmartin, Christine Murad, Cosmin Munteanu, Vincent Wade, and Benjamin R. Cowan. 2019. What makes a good conversation? challenges in designing truly conversational agents. In *Proceedings of CHI Conference on Human Factors in Computing Systems*, pages 1–12, Glasgow, Scotland, May.

Cohen, Michael, James P. Giangola, and Jennifer Balogh. 2004. *Voice User Interface Design*. Addison Wesley.

Colby, Kenneth Mark, Sylvia Weber, and Franklin Dennis Hilf. 1971. Artificial paranoia. *Artificial Intelligence*, 2(1):1–25.

Crystal, David. 2001. *Language and the Internet*. Cambridge University Press.

Dale, Robert. 2016. Industry watch the return of chatbots. *Natural Language Engineering*, 22(5):811–817.

Dasgupta, Ritwik. 2018. *Voice User Interface Design. Moving from GUI to Mixed Modal Interaction*. Apress.

de Jong, Markus, Mariët Theune, and Dennis Hofs. 2008. Politeness and alignment in dialogues with a virtual guide. In *Proceedings of the 7th International Conference on Autonomous Agents and Multiagent Systems*, Estoril, Portugal, May.

Deriu, Jan, Alvaro Rodrigo, Arantxa Otegi, Guillermo Echegoyen, Sophie Rosset, Eneko Agirre, and Mark Cieliebak. 2020. Survey on evaluation methods for dialogue systems. *CoRR*, abs/1905.04071.

Dippold, Doris, Jenny Lynden, Rob Shrubsall, and Rich Ingram. 2020. A turn to language: How interactional sociolinguistics informs the redesign of prompt:response chatbot turns. *Discourse, Context & Media*, 37(10043):2.

Drew, Paul. 2020. Turn design. In J. Sidnell and T. Stivers, editors, *The handbook of conversation analysis*. Wiley-Blackwell, Springer, Cham, pages 131–149.

Feine, Jasper, Ulrich Gnewuch, Stefan Morana, and Alexander Maedche. 2020. Gender bias in chatbot design. In A. Følstad, T. Araujo, S. Papadopoulos, E. L. Law, O. Granmo, E. Luger, and P. B. Brandtzaeg, editors, *Chatbot Research and Design. CONVERSATIONS 2019. Lecture Notes in Computer Science*. vol 11970, Springer, Cham, pages 79–93.

Fox, Barbara A., Sandra A. Thompson, Cecilia E. Ford, and Elizabeth Couper-Kuhlen. 2018. Conversation analysis and linguistics. In J. Sidnell and T. Stivers, editors, *Handbook of Conversation Analysis*. Oxford: Wiley-Blackwell, pages 726–740.

Følstad, Asbjørn and Petter Bae Brandtzaeg. 2020. Users' experiences with chatbots: findings from a questionnaire study. *Qual User Exp*, 5:3.

Følstad, Asbjørn, Cecilie Bertinussen Nordheim, and Cato Alexander Bjørkli. 2018. What makes users trust a chatbot for customer service? an exploratory interview study. In S. Bodrunova, editor, *Internet Science. INSCI 2018. Lecture Notes in Computer Science*. vol 11193, pages 194–208.

Følstad, Asbjørn, Marita Skjuve, and Petter Brandtzaeg. 2019. Different chatbots for different purposes: Towards a typology of chatbots to understand interaction design. In S. Bodrunova, O. Koltsova, A. Følstad, H. Halpin, P. Kolozaridi, L. Yuldashev, A. Smoliarova, and H. Niedermayer, editors, *Internet Science. INSCI 2018. Lecture Notes in Computer Science*. vol 11551, Springer, Cham, pages 145–156.

Garfinkel, Harold. 1996. *Studies in ethnomethodology*. Polity Press.

Gervits, Felix, Ravenna Thielstrom, Antonio Roque, and Matthias Scheutz. 2020. It's about time: Turn-entry timing for situated human-robot dialogue. In *Proceedings of the 21th Annual Meeting of the Special Interest Group on Discourse and Dialogue*, pages 86–96, virtual meeting, July.

Givon, Talmy. 1979. *Syntax and Semantics 12: Discourse and Syntax*. Academic Press.

Gnewuch, Ulrich, Stefan Morana, and Alexander Maedche. 2017. Towards designing cooperative and social conversational agents for customer service. In *Proceedings of the International Conference on Information Systems (ICIS)*, Seoul, South Korea, December.

Goffman, Erving. 1967. *Interaction Ritual: Essays on Face-to-Face Behavior*. Aldine Publishing Company, New York.

Goffman, Erving. 1983. The interaction order. *American Sociological Review*, 48:1–17.

Grice, H. Paul. 1975. Logic and conversation. In P. Cole and J. L. Morgan, editors, *Syntax and Semantics*. Speech Acts. Academic Press, 3.

Grice, H. Paul. 1989. *Studies in the Way of Words*. Harvard University Press.

Gumperz, John J. 1982. *Discourse Strategies*. Cambridge University Press.

Hakulinen, Auli. 2009. Conversation types. In S. D'hondt, J. Ostman, and J. Verschueren, editors, *The Pragmatics of Interaction*. John Benjamins, pages 55–65.

Hall, Erika. 2018. *Conversational design*. A Book Apart.

Halliday, Michael A. K. 1994. *An Introduction to Functional Grammar*. Edward Arnold.

Haugh, Michael. 2012. Conversational interaction. In K. Allan and K. M. Jaszczolt, editors, *The Cambridge Handbook of Pragmatics*. Cambridge University Press, pages 251–274.

Herring, Susan C. 2012. Grammar and electronic communication. In C. Chapelle, editor, *Encyclopedia of applied linguistics*. Wiley-Blackwell.

Hill, Jennifer, W. Randolph Ford, and Ingrid G. Farreras. 2015. Real conversations with artificial intelligence: A comparison between human–human online conversations and human–chatbot conversations. *Computers in Human Behavior*, 49:245–250.

Hirst, Graeme. 2001. Does conversation analysis have a role in computational linguistics? *Computational Linguistics*, 17(2):211–272.

Hoey, Elliott M. and Kobin H. Kendrick. 2017. Conversation analysis. In A. M. B. de Groot and P. Hagoort, editors, *Research Methods in Psycholinguistics: A Practical Guide*. Wiley, Blackwell, pages 151–173.

Horn, Laurence R. and Gregory Ward. 2006. Introduction. In L. R. Horn and G. Ward, editors, *The Handbook of Pragmatics*. Blackwell, pages xi–xix.

Hussain, Shafquat, Omid Sianaki, and Nedal Ababneh. 2019. A survey on conversational agents/chatbots classification and design techniques. In L. Barolli, M. Takizawa, F. Xhafa, and T. Enokido, editors, *Web, Artificial Intelligence and Network Applications. WAINA 2019. Advances in Intelligent Systems and Computing*. vol 927. Springer, Cham.

Hymes, Dell. 1962. The ethnography of speaking. In T. Gladwin and W. C. Sturtevant, editors, *Anthropology and Human Behavior*. Anthropological Society, pages 13–53.

Jacquet, Baptiste, Jean Baratgin, and Frank Jamet. 2018. The gricean maxims of quantity and of relation in the turing test. In *Proceedings of the 11th International Conference on Human System Interaction - HSI 2018*, pages 332–338, Gdańsk, Poland, June.

Jacquet, Baptiste, Jean Baratgin, and Frank Jamet. 2019b. Cooperation in online conversations: The response times as a window into the cognition of language processing. *Frontiers in Psychology*, 10(727):1–15.

Jacquet, Baptiste, Alexandre Hullin, Jean Baratgin, and Frank Jamet. 2019. The Impact of the Gricean Maxims of Quality, Quantity and Manner in Chatbots. In *Proceedings of the International Conference on Information and Digital Technologies*, pages 180–189, Zilina, Slovakia, June.
Jain, Mohit, Pratyush Kumar, Ramachandra Kota, and Shwetak N. Patel. 2018. Evaluating and informing the design of chatbots. In *Proceedings of the Designing Interactive Systems Conference*, pages 895–906, Hong Kong, June.
Jefferson, Gail D. 1978. Sequential aspects of storytelling in conversation. In J. Schenkein, editor, *Studies in the Organization of Conversational Interaction*. Academic Press, pages 219–48.
Katz, Jerrold and Jerry Fodor. 1963. The structure of a semantic theory. *Language*, 39:67–80.
Kerbrat-Orecchioni, Catherine. 2001. *Les actes de langage dans le discours. Théorie et fonctionnement*. Nathan.
Kvale, Knut, Olav Sell, Stig Hodnebrog, and Asbjørn Følstad. 2019. Improving conversations: Lessons learnt from manual analysis of chatbot dialogues. In A. Følstad, T. Araujo, S. Papadopoulos, E. L. Law, O. Granmo, E. Luger, and P. B. Brandtzaeg, editors, *Chatbot Research and Design. CONVERSATIONS 2019. Lecture Notes in Computer Science*. vol 11970, Springer, pages 187–200.
Leech, Geoffrey. 1983. *Principles of Pragmatics*. Longman.
Levinson, Stephen C. 1983. *Pragmatics*. Cambridge University Press.
Linell, Per. 2001. *Approaching Dialogue: Talk, interaction and contexts in dialogical perspectives.* John Benjamins Publishing Company.
Linell, Per and Thomas Luckmann. 1991. Asymmetries in dialogue: some conceptual preliminaries. In I. Marková and K. Foppa, editors, *Asymmetries in Dialogue*. Harvester Wheatsheaf, pages 1–20.
Liu, Xingkun, Arash Eshghi, Pawel Swietojanski, and Verena Rieser. 2019. Benchmarking natural language understanding services for building conversational agents. *arXiv 1903.05566v3*.
Lotze, Netaya. 2016. Chatbots. eine linguistische analyse. In J. Runkehl, P. Schlobinski, and T. Siever, editors, *Sprache – Medien – Innovationen*. vol. 9. Peter Lang.
Luckmann, Thomas. 1990. Social communication, dialogue and conversation. In I. Markovà and K. Foppa, editors, *The Dynamics of Dialogue*. Harvester Wheatsheaf, pages 45–61.
Luff, Paul, David Frohlich, and Nigel Gilbert. 1990. *Computers and Conversation*. Academic Press.
Mathur, Vinayak and Arpit Singh. 2018. The rapidly changing landscape of conversational agents. *arXiv 1803.08419v2*, pages 1–14.
McTear, Michael. 2020. *Conversational AI: Dialogue Systems, Conversational Agents, and Chatbots (Synthesis Lectures on Human Language Technologies)*. Morgan & Claypool.
Mey, Jacob L. 1993. *Pragmatics. An introduction*. Blackwell.
Michael, Thilo and Sebastian Möller. 2020. Simulating turn-taking in conversations with delayed transmission. In *Proceedings of the 21th Annual Meeting of the Special Interest Group on Discourse and Dialogue*, page 157–161, virtual meeting, July.
Moore, Robert J. 2013. Ethnomethodology and conversation analysis: empirical approaches to the study of digital technology in action. In S. Price, C. Jewitt, and B. Brown, editors, *The SAGE handbook of digital technology research*. SAGE Publications Ltd, pages 217–235.
Moore, Robert J. and Raphael Arar. 2018. Conversational ux design: An introduction. In R. J. Moore, M. H. Szymanski, R. Arar, and G. Ren, editors, *Studies in Conversational UX Design*. Springer, International Publishing, pages 1–16.
Moore, Robert J. and Raphael Arar. 2019. *Conversational UX Design. A Practitioner's Guide to the Natural Conversation Framework*. Association for Computing Machinery.
Moore, Robert J., Raphael Arar, and Guang-Jie Ren. 2018. *Studies in Conversational UX Design*. Springer, International Publishing.
Morris, Charles W. 1938. *Foundations of the Theory of Signs, Chicago University Press; reprinted in Writings on the General Theory of Sign*. Mouton.
Nordheim, Cecilie Bertinussen, Åsbjørn Følstad, and Cato Alexander Bjørkli. 2019. An initial model of trust in chatbots for customer service—findings from a questionnaire study. *Interacting with Computers*, 31(3):317–335.
Pallotti, Gabriele. 2007. Conversation analysis: Methodology, machinery and application to specific settings. In H. Bowles and P. Seedhouse, editors, *Conversation Analysis and Language for Specific Purposes*. Peter Lang, pages 37–68.
Pearl, Cathy. 2016. *Designing Voice User Interfaces. Principles of Conversational Experiences*. O'Reilly.
Peirce, Charles Sanders. 1932. *Collected Papers*. Cambridge University Press.

Piccolo, Lara, Martino Mensio, and Harith Alani. 2018. Chasing the chatbots: Directions for interaction and design research. In S. Bodrunova, editor, *Internet Science, INSCI*. Lecture Notes in Computer Science, Springer, page 157–169.
Pickering, Martin and Simon Garrod. 2004. Towards a mechanistic psychology of dialogue. *Behavioural and Brain sciences*, 27:169–225.
Pistolesi, Elena. 2018. Storia, lingua e varietà della comunicazione mediata dal computer. In G. Patota and F. Rossi F. Rossi, editors, *L'italiano e la rete, le reti per l'italiano*. Accademia della Crusca – GoWare, pages 16–34.
Przegalinska, Aleksandra, Leon Ciechanowski, Anna Stroz, Peter Gloor, and Grzegorz Mazurek. 2020. In bot we trust: A new methodology of chatbot performance measures. *Business Horizons*, 62:785–797.
Radziwill, Nicole M. and Morgan C. Benton. 2017. Evaluating quality of chatbots and intelligent conversational agents. *arXiv 1704.04579*.
Ruane, Elayne, Abeba Birhane, and Anthony Ventresque. 2019. Conversational AI: Social and ethical considerations. In *Proceedings for the 27th AIAI Irish Conference on Artificial Intelligence and Cognitive Science*, Galway, Ireland, December.
Sacks, Harvey. 1984. Notes on methodology. In J. M. Atkinson and J. C. Heritage, editors, *Structures of Social Action: Studies in Conversation Analysis*. Cambridge University Press, pages 21–27.
Sacks, Harvey and Emanuel A. Schegloff. 2007. Two preferences in the organization of reference to persons in conversation and their interaction. In N. J. Enfield and T. Stivers, editors, *Person Reference in Interaction: Linguistic, Cultural and Social Perspectives, Language Culture and Cognition*. Cambridge University Press, pages 23–28.
Sacks, Harvey, Emanuel A. Schegloff, and Gail Jefferson. 1974. A simplest systematics for organization of turn-taking for conversation. *Language*, 50:696–735.
Sacks, Harvey, Emanuel A. Schegloff, and Gail Jefferson. 1992. *Lectures on Conversation*, volume I, II. Blackwell.
Saygin, Ayse Pinar and Ilyas Cicekli. 2002. Pragmatics in human-computer conversations. *Journal of Pragmatics*, 34:227–258.
Schegloff, Emanuel A. 1968. Sequencing in conversational openings. *American Anthropologist*, 70(6):1075–1095.
Schegloff, Emanuel A. 1992. To searle on conversation: a note in return. In J. Searle, H. Parret, and J. Verschueren, editors, *(On) Searle on Conversation*. Benjamins, pages 113–128.
Schegloff, Emanuel A. 1999. Discourse, pragmatics, conversation, analysis. *Discourse Studies*, 1:405–435.
Schegloff, Emanuel A. 2007. *Sequence Organization in Interaction: A Primer in Conversation Analysis*, volume I. Cambridge University Press.
Schegloff, Emanuel A., Gail Jefferson, and Harvey Sacks. 1977. The preference for self-correction in the organization of repair in conversation. *Language*, 53(2):361–382.
Schegloff, Emanuel A. and Harvey Sacks. 1973. Opening up closings. *Semiotica*, 8:289–327.
Schiavo, Gianluca and Ahmed Fadhil. 2020. Designing for health chabots. *arXiv 1902.09022*.
Searle, John R. 1969. *Speech Acts. An Essay in the Philosophy of Language*. Cambridge University Press.
Searle, John R. 1983. *Intentionality: An Essay in the Philosophy of Mind*. Cambridge University Press.
Shawar, Bayan Abu and Eric Atwell. 2007. Chatbots: Are they really useful? *LDV-Forum*, 22(1):29–49.
Shevat, Amir. 2017. *Designing Bots: Creating Conversational Experiences.* O'Reilly.
Spence, Jocelyn. 2016. *Performative Experience Design*. Springer, Cham.
Sperber, Dan and Deirdre Wilson. 1995. *Relevance: Communication and Cognition*. Blackwell.
Strengers, Yolande and Jenny Kennedy. 2020. *Why Siri, Alexa, and Other Smart Home Devices Need a Feminist Reboot*. MIT Press.
Svenningsson, Nina and Montathar Faraon. 2019. Artificial intelligence in conversational agents: A study of factors related to perceived humanness in chatbots. In *Proceedings of the 2nd Artificial Intelligence and Cloud Computing Conference*, pages 151–161, Kobe, Japan, December.
Szmrecsanyi, Benedikt. 2005. Language users as creatures of habit: a corpus linguistics analysis of persistence in spoken language. *Corpus Linguistics and Linguistic Theory*, 1(1):113–150.
Szymanski, Margaret H. and Robert J. Moore. 2018. Adapting to customer initiative: Insights from human service encounters. In R. J. Moore, M. H. Szymanski, R. Arar, and G. Ren G. Ren,

editors, *Studies in Conversational UX Design*. Springer, International Publishing, pages 19–32.
Tannen, Deborah. 1984. *Conversational Style: analyzing talk among friends*. Ablex.
Tsvetkova, Milena, Ruth García-Gavilanes, Luciano Floridi, and Taha Yasseri. 2017. Even good bots fight: The case of wikipedia. *PLOS ONE*, 12(2):1–27.
Turner, Ken, editor. 1999. *The Semantics/Pragmatics Interface from Different Points of View*. Elsevier.
van Dijk, Teun A. 2009. *Society and Discourse: How Social Contexts Influence Text and Talk*. Cambridge University Press.
Verschueren, Jef. 1999. *Understanding Pragmatics*. Arnold.
Wallace, Richard S. 2003. *The Elements of AIML Style*. A.L.I.C.E. Artificial Intelligence Foundation Inc.
Weizenbaum, Joseph. 1966. ELIZA – A computer program for the study of natural language communication between man and machine. *Communications of the ACM*, 9(1):36–45.
West, Mark, Rebecca Kraut, and Han Ei Chew. 2019. *I'd Blush if I Could. Closing Gender Divides in Digital Skills through Education*. UNESCO.
Wooffitt, Robin, Norman Fraser, Nigel Gilbert, and Scott McGlashan. 1997. *Humans, Computers and Wizards. Analyzing human (simulated) computer interaction*. Routledge.
Yule, George. 1996. *Pragmatics*. Oxford University Press.
Zamora, Jennifer. 2017. I'm sorry, Dave, I'm afraid I can't do that: Chatbot perception and expectations. In *Proceedings of the 5th International Conference on Human Agent Interaction*, HAI '17, page 253–260, New York, NY, USA. Association for Computing Machinery.
Zubania, Matteo, Luca Sigalini, Ivan Serina, and Alfonso Emilio Gerevini. 2020. Evaluating different natural language understanding services in a real business case for the italian language. In *Proceedings of the 24th International Conference of Knowledge-Based and Intelligent Information & Engineering Systems*, pages 995–1004, Verona, Italy, September.