

# Dialog Acts in Greeting and Leavetaking in Social Talk

Emer Gilmartin  
Trinity College Dublin  
Ireland  
gilmare@tcd.ie

Brendan Spillane  
ADAPT, Trinity College Dublin  
Ireland  
brendan.spillane@adaptcentre.ie

Maria O'Reilly  
Trinity College Dublin  
Ireland  
moreil12@tcd.ie

Ketong Su  
ADAPT, Trinity College Dublin  
Ireland  
ketong.su@adaptcentre.ie

Christian Saam  
ADAPT, Trinity College Dublin  
Ireland  
saamc@adaptcentre.ie

Benjamin R. Cowan  
University College Dublin  
Ireland  
benjamin.cowan@ucd.ie

Nick Campbell  
Trinity College Dublin  
Ireland  
nick@tcd.ie

Vincent Wade  
ADAPT, Trinity College Dublin  
Ireland  
vincent.wade@adaptcentre.ie

## ABSTRACT

Conversation proceeds through dialogue moves or acts, and dialog act annotation can aid the design of artificial dialog. While many dialogs are task-based or instrumental, with clear goals, as in the case of a service encounter or business meeting, many are more interactional in nature, as in friendly chats or longer casual conversations. Early research on dialogue acts focussed on transactional or task-based dialogue but work is now expanding to social aspects of interaction. We review how dialog annotation schemes treat non-task elements of dialog – greeting and leave-taking sequences in particular. We describe the collection and annotation, using the ISO Standard 24617-2 Semantic annotation framework, Part 2: Dialogue acts, of a corpus of 187 text dialogues and study the dialog acts used in greeting and leave-taking.

## CCS CONCEPTS

• **Human-centered computing** → **Natural language interfaces; Interaction design theory, concepts and paradigms;** • **Computing methodologies** → **Discourse, dialogue and pragmatics;**

## KEYWORDS

Dialog acts, casual conversation, intelligent agents

### ACM Reference Format:

Emer Gilmartin, Brendan Spillane, Maria O'Reilly, Ketong Su, Christian Saam, Benjamin R. Cowan, Nick Campbell, and Vincent Wade. 2017. Dialog Acts in Greeting and Leavetaking in Social Talk. In *Proceedings of 1st ACM SIGCHI International Workshop on Investigating Social Interactions with Artificial Agents (ISIAA'17)*. ACM, New York, NY, USA, 2 pages. <https://doi.org/10.1145/3139491.3139493>

Permission to make digital or hard copies of part or all of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for third-party components of this work must be honored. For all other uses, contact the owner/author(s).  
ISIAA'17, November 13, 2017, Glasgow, UK  
© 2017 Copyright held by the owner/author(s).  
ACM ISBN 978-1-4503-5558-2/17/11.  
<https://doi.org/10.1145/3139491.3139493>

## 1 INSTRUMENTAL AND INTERACTIONAL DIALOGUE

Text has become a medium for practically synchronous interaction. Historically, written messages were asynchronous and did not approach the fine-grained interaction and collaboration of talk. Dialog systems model spoken or written synchronous/near-synchronous interactions, often to fulfill a task but increasingly to create the illusion of social interaction. With live text exchange a part of everyday life we have seen an explosion of casual writing, performed not for a formal purpose but to fulfill social goals.

## 2 ANNOTATION OF INTERACTIONAL TALK

Existing dialog act annotation schemes are often quite task-based. There have been several annotation schemes developed, often in conjunction with particular corpora or experiments [1, 2, 4, 5]. The ISO standard is very useful as it (i) amalgamates contributions from pre-existing schemes, and (ii) is multifunctional and multi-dimensional - several acts can apply to stretches within the same contribution.

Most schemes include a number of social obligation management functions. In a survey of 14 schemes, Petukova found that 10 included greeting functions, 4 included introductions, 6 had good-byes, 5 included apology type functions, and 5 contained thanking. The Social Obligations Management dimension of the ISO standard contains nine communicative functions.

## 3 DATA COLLECTION AND ANNOTATION

A corpus of 187 dialogs was collected and annotated with the ISO standard to provide training data for the ADELE project, a personalized intelligent companion. The dialogs were text-based and dyadic via a web-based interface. Each participant was given a persona with information on home, relationships, nationality, job, hobbies and interests, and instructed to discover this information about the interlocutor and also to discover any facts or interests in common. The corpus contained examples of greeting and leave-taking and casual talk for practically all of the conversations gathered. There

were 37 participants (26M/11F, age 18-43), either native English speakers or meeting IELTS level 6.5.

During a pilot annotation, annotators noted components in extended greeting/introductions and leavetaking (henceforth GIL) sequences which could not easily be annotated, as in the following example:

- (1) A: Hi
- (2) B: Hello, I'm Ann. I'm from Mexico City. Yourself?
- (3) A: Hi Ann, nice to meet you. I'm John.
- (4) B: Hey John, nice to meet you too. How are you today?
- (5) A: Good, good. You? I'm from Paris, living in London now.
- (6) B: I'm in good form!.

In the fragment there are four instances of **hello**, **hi**, or **hey**. The first two can be accounted for by the ISO but not the latter as there is no 'generic' greet tag, but only initialGreet and returnGreet. The expression **nice to meet you** and response **nice to meet you too** in lines 3 and 4 are clearly formulaic greetings but it is unclear how to annotate them. A similar situation obtains with the **How are you today? – Good, good and You?** (ellipsis of **How are you?**) – **I'm in good form!** in lines 4-6. If the first part of these adjacency pairs are annotated as setQuestions and the second parts as informs or answers, these tags could be placed in the SOM dimension. However, these composite treatments are clumsy to implement during annotation and do not reflect the illocutionary force of the expressions as clearly as existing SOM tags (initialGreet, returnGreet) do for the **Hi** and **Hello** in lines 1 and 2.

Additional acts were created to more easily mark such sequences and similarly problematic sequences in leave-taking. Table 1 shows the new acts.

**Table 1: Acts introduced for the ADELE annotation**

| Act      | Common Examples             | Functional Area |
|----------|-----------------------------|-----------------|
| ntmy     | Nice to meet you            | Greeting        |
| repNtmy  | Nice to meet you too        | Greeting        |
| hay      | How are you?                | Greeting        |
| repHay   | Fine                        | Greeting        |
| greet    | Hello                       | Greeting        |
| wntmy    | It was lovely to meet you   | Leavetaking     |
| repWntmy | It was nice to meet you too | Leavetaking     |

The GIL sequences in 187 conversations were annotated and then analysed.

## 4 GIL SEQUENCES

Greeting sections were marked as beginning with the first utterance of the conversation, and ending with the last production of a formulaic greeting/introduction or greeting/introduction response. Leave-taking sequences were marked from the first attempt to close the conversation to the final utterance of the conversation. The data contained 9231 turns or 'utterances' where a turn was defined as the text entered before a user pressed return. The vast bulk of utterances were tagged with a single label (7811, 84.7%), 1209 (13%) had two tags, 181 (2%) had three tags, while 26 (0.3%) and 3 utterances had four and five tags.

Of 10889 dialog act tags, 2336 or 21.5% were included in GIL sequences. 1329 tags related to greeting and 1007 to leave-taking. GIL sequences sometimes contained other acts unrelated to greeting, introduction, or leave-taking, as in the above example where **I'm from Mexico City. Yourself?** in line 2 are an inform and setQuestion related to the task. The question is answered on line 5 near the end of the greeting/introduction sequence. The number of dialog acts directly involved in GIL sequences was calculated by disregarding such 'interloping' acts. Greeting/introduction alone accounted for 1034 labels, while leave-taking alone accounted for 786 labels, making a total of 1820 acts of greeting/introduction and leave-taking, or 16.7% of all dialog acts tagged in the corpus. The leave-taking totals include 194 Leave-taking Introductions – utterances which introduce the closure of the dialog. These utterances could be included in the Discourse Structuring dimension, in which case the total for GIL drops to 1626 or 15% of all dialog act labels, which is the most conservative estimate of the proportion of GIL tags in the corpus. The total SOM acts in the corpus including SOM categories outside GIL from the ISO standard amounts to 1824 or 17%. In terms of the prevalence of the new greeting tags, in 187 conversations the hay (How are you?) tag appeared 68 times, the ntmy (Nice to meet you) tag appeared 101 times, and the extra greet tag appeared 66 times (each conversation contained two initialGreets). The response tags repHay and repNtmy appeared less frequently, with 49 instances of repHay and 25 of repNtmy. For the leavetaking tags, there were 139 wntmy (It was nice to meet you) tags and 47 repWntmy tags.

## 5 RESULTS AND CONCLUSIONS

There is a high proportion of SOM acts in the ADELE corpus, and GIL acts contribute greatly to this total. Petukova reports SOM acts in task-based corpora as ranging from 0.5 to 7.8% of total dialog acts [3], compared with 17% in the ADELE corpus. Most SOM in ADELE are greetings/introductions and leavetaking. Increasing interest in friendly interfaces strengthens the need for greater understanding and more accurate modelling of social dialogue. There are large areas of such dialogue which are not well represented in dialog annotation schemes, ranging from simple politeness formulae, such as the greeting and leavetaking acts treated here, to larger concerns of how to represent the relationship building and maintenance functions integral to casual social talk.

## ACKNOWLEDGMENTS

This work is supported by Science Foundation Ireland through the ADAPT Centre, Trinity College Dublin, and CHISTERA JOKER.

## REFERENCES

- [1] Mark G. Core and James Allen. 1997. Coding dialogs with the DAMSL annotation scheme. In *AAAI fall symposium on communicative action in humans and machines*. Boston, MA, 28–35.
- [2] D. Jurafsky, C. Van Ess-dykema, et al. 1997. Switchboard Discourse Language Modeling Project (Final Report). (1997).
- [3] Volha Viktoraua Petukhova. 2011. Multidimensional dialogue modelling. *Pro-motor: H. Bunt. Tilburg* 1 (2011). <http://arno.uvt.nl/show.cgi?fid=115903>
- [4] Elizabeth Shriberg, Raj Dhillon, Sonali Bhagat, Jeremy Ang, and Hannah Carvey. 2004. *The ICSI meeting recorder dialog act (MRDA) corpus*. Technical Report. ICSI, Berkeley.
- [5] D. Traum. 1999. Speech acts for dialogue agents. *Foundations of rational agency* 14 (1999), 169–202.