# Studying the Lexicon of Dialogue Acts

**Nicole Novielli**[1], **Carlo Strapparava**[2]

[1] Università degli Studi di Bari
Dipartimento di Informatica
via Orabona 4 - 70125 Bari, Italy
novielli@di.uniba.it

[2] FBK- irst, Trento
Istituto per la Ricerca Scientifica e Tecnologica
via Sommarive 18 - I-38050 Povo Trento, Italy
strappa@fbk.eu

**Abstract**

Dialogue Acts have been well studied in linguistics and attracted computational linguistics research for a long time: they constitute the basis of everyday conversations and can be identified with the communicative goal of a given utterance (e.g. asking for information, stating facts, expressing opinions, agreeing or disagreeing). Even if not constituting any deep understanding of the dialogue, automatic dialogue act labeling is a task that can be relevant for a wide range of applications in both human-computer and human-human interaction. We present a qualitative analysis of the lexicon of Dialogue Acts: we explore the relationship between the communicative goal of an utterance and its affective content as well as the salience of specific word classes for each speech act. The experiments described in this paper fit in the scope of a research study whose long-term goal is to build an unsupervised classifier that simply exploits the lexical semantics of utterances for automatically annotate dialogues with the proper speech acts.

## 1. Introduction

Dialogue Acts (DA) (Core and Allen, 1997) constitute the basis of everyday conversations and can be identified with the communicative goal of a given utterance (Austin, 1962): asking for information, stating facts, expressing opinions, agreeing or disagreeing with the interlocutor.

There is a large number of applications that could benefit from automatic DA annotation: dialogue systems, blog analysis, automatic meeting summarization, user profiling by mean of dialogue pattern analysis, and so on. In this kind of applications, the system should be able to understand the communication dynamics, that is understanding who is telling what to whom.

The task of automatic DA recognition has been addressed with promising results by studies developed in supervised frameworks (Stolcke et al., 2000; Samuel et al., 1998; Reithinger et al., 1996). Rather than improving the performance of supervised approaches, the long term goal of our research is to define DA lexical profiles that can be used in an unsupervised framework for automatic labelling of natural dialogues with the proper speech acts.

In the present paper, we exploit the Switchboard corpus of telephone conversations (Godfrey et al., 1992) in order to better understand what are the most salient lexical features for each DA. Even if prosody and intonation surely play a role (see, for example (Stolcke et al., 2000; Warnke et al., 1997)), we decided to focus on text analysis because language and words are what people use to convey their communicative intentions. Moreover, in recent years a large amount of material about natural language interactions on the Web has become available, raising the attractiveness of empirical methods of analyses on this field and text is just what we have at disposal in such a scenario. In particular, we describe a qualitative study of the lexicon aimed at investigating the relationship between the DA and the affective load of a given utterance, as well as the role played by lexical categories and their salience with respect to each DA.

## 2. Dataset

To run our experiments, we exploited the Switchboard corpus of English task-free telephone conversations (Godfrey et al., 1992), which involve couples of randomly selected strangers talking informally about general interest topics. Complete transcripts are distributed by the Linguistic Data Consortium. A part of them is annotated using DA labels (overall 1155 dialogues, for a total of 205,000 utterances and 1.4 million words).

**Labelling** A Dialogue Act can be identified with the communicative goal of a given utterance i.e. it represents its meaning at the level of illocutionary force (Austin, 1962). Researchers use different labels and definitions to address the communicative goal of a sentence: Searle (1969) talks about *speech act*; Schegloff (1968) and Sacks (1974) refer to the concept of *adjacency pair part*; Power (1979) adopts the definition of *game move*; Cohen and Levesque (1995) focus more on the role speech acts play in interagent communication.

Traditionally, the NLP community has employed DA annotation approaches with the drawback of being domain oriented. Only recently, some efforts have been made towards unification of DA annotation (Traum, 2000). In this study we refer to DAMSL (Dialogue Act Markup in Several Layers) a domain-independent annotation framework (Core

| Label | Description | Example | % |
|---|---|---|---|
| INFO-REQUEST | Utterances that are pragmatically, semantically, and syntactically questions | *'What did you do when your kids were growing up?'* | 7% |
| STATEMENT | Descriptive, narrative, personal statements | *'I usually eat a lot of fruit'* | 57% |
| S-OPINION | Directed opinion statements | *'I think he deserves it.'* | 20% |
| AGREE-ACCEPT | Acceptance of a proposal, plan or opinion | *'That's right'* | 9% |
| REJECT | Disagreement with a proposal, plan, or opinion | *'I'm sorry no'* | .3% |
| OPENING | Dialogue opening or self-introduction | *'Hello, my name is Imma'* | .2% |
| CLOSING | Dialogue closing (e.g. farewell and wishes) | *'It's been nice talking to you.'* | 2% |
| KIND-ATT | Kind attitude (e.g. thanking and apology) | *'Thank you very much.'* | .1% |
| GEN-ANS | Generic answers to an Info-Request | *'Yes', 'No', 'I don't know'* | 4% |
| total cases | | | 131,265 |

Table 1: The set of labels employed for Dialogue Acts and their distribution in the corpus.

and Allen, 1997). DA annotation is out of the scope of the present study hence we used already annotated data. In particular, the Switchboard employs the SWBD-DAMSL revision of the DAMSL scheme (Jurafsky et al., 1997).

Table 1 shows the set of labels we employ: it maintains the DAMSL main peculiarity of being domain-independent and the semantics of the SWBD-DAMSL labels used for the original Switchboard annotation. Thus, the original Switchboard annotation has been automatically converted in our set of tags as shown in Table 2.

| Label | SWBD-DAMSL |
|---|---|
| INFO-REQ | *Yes-No question* (*qy*), *Wh-Question* (qw), *Declarative Yes-No-Question* (qy^d), *Declarative Wh-Question* (qw^d), *Alternative ('or') question* (qr) *and OR-clause* (qrr) , *Open-Question* (qo), *Declarative* (^d) *and Tag questions* (^g) |
| STATEMENT | *Statement-non-opinion* (sd) |
| S-OPINION | *Statement-opinion* (sv) |
| AGREE-ACC | *Agreement /accept* (aa) |
| REJECT | *Agreeement /reject* (ar) |
| OPENING | *Conventional-opening* (fp) |
| CLOSING | *Conventional-closing* (fc) |
| KIND-ATT | *Thanking* (ft) *and Apology* (fa) |
| GEN-ANS | *Yes answers* (ny), *No answers* (nn), *Affirmative non-yes answers* (na) *Negative non-no answers* (ng) |

Table 2: The Dialogue Act set of labels with their mapping with the SWBD-DAMSL correspondent categories

## 3. Dialogue Act recognition: experimental setup and results

Is it possible to automatically annotate natural dialogues with the proper dialogue acts? What is the role played by lexical semantics in conveying the communicative goal of an utterance? To answer these questions we conducted some experiments in both a supervised and an unsupervised frameworks (see Novielli and Strapparava (2009) for details).

In summary, for the supervised framework, we used the Support Vector Machine (SVM) (Vapnik, 1995), a state-of-the art technique that has been successfully employed in several problems, including text classification. We randomly split the two corpora in 80/20 train/test partitions.

A first version of our unsupervised framework was set up using the same partitions. Schematically, our unsupervised methodology is: (i) building a semantic similarity space in which words, set of words, text fragments can be represented homogeneously, (ii) finding seeds (words) that properly represent dialogue acts and considering their representations in the similarity space, and (iii) checking the similarity of the utterances. To get a similarity space with the required characteristics, we used Latent Semantic Analysis (LSA). LSA is a corpus-based measure of semantic similarity proposed by Landauer (Landauer et al., 1998). In LSA, term co-occurrences in a corpus are captured by means of a dimensionality reduction operated by a singular value decomposition (SVD) on the term-by-document matrix $\mathbf{T}$ representing the corpus. For representing a word set or a sentence in the LSA space we use the *pseudo-document* representation technique, as described by Berry (1992), using also a *tf.idf* weighting scheme (Gliozzo and Strapparava, 2005). Starting from the sets of seeds representing the dialogue acts, we build the corresponding vectors in the LSA space and then we compare the utterances to find the communicative act with the highest similarity.

The seeds are general and language-independent: they are defined by considering only the communicative goal and the specific semantics of each dialogue act, just avoiding the overlapping between seed groups as much as possible. Since our aim is to design an approach that is as general as possible, we do not consider domain words that could make easier the classification. Table 3 shows some examples of sets of seeds with the corresponding DAs. To allow comparison with SVM, the performance is measured on the same test set partition used in the supervised experiment.

To reduce data sparseness, we used a POS-tagger and a morphological analyzer (Pianta et al., 2008) and we used

| Label | Seeds |
|---|---|
| INFO-REQ | Question_mark |
| S-OPINION | Verbs which directly express opinion or evaluation (guess, think, suppose, affect) |
| AGREE-ACC | yep, yeah, absolutely, correct |
| OPENING | Expressions of greetings (hi, hello), words and markers related to self-introduction formula |
| KIND-ATT | Lexicon which directly expresses wishes (wish), apologies (apologize), thanking (thank) and sorry-for (sorry, excuse) |

Table 3: Some examples of sets of seeds

lemmata instead of tokens in the format lemma#POS, with no further feature selection, in both experimental settings.
We evaluated the performance in terms of precision, recall and F1-measure (Novielli and Strapparava, 2009) according to the DA labels given by annotators. Consistently with our goal of defining a general method for DA annotation, we compared the performance on the Switchboard corpus with the results on an Italian corpus of human-computer interactions (Clarizio et al., 2006). The seeds are the same for both languages, which is coherent with our goal of defining a language-independent method. As a baseline we consider the most frequent label assignment (respectively 37% for Italian, 57% for English) for the supervised experiment and random DA selection (11%) for the unsupervised one.
We got .71 and .77 of F1 respectively for the Italian and the English corpus in the supervised condition, and .66 and .68 for the unsupervised one. Both results are significantly above the baselines and are comparable to the state of the art (Stolcke et al., 2000; Samuel et al., 1998; Reithinger et al., 1996; Poesio and Mikheev, 1998). This is particularly encouraging, especially considering that we focus only on written text.
The error analysis highlights that the main cause of error is the misclassification of many utterances as STATEMENT: statements are usually quite long and it is highly likely that they contain lexical features that characterize other DAs. This is particularly true for the S-OPINIONs, which are mostly misclassified as statements: the only significative difference between the two labels seems to be the wider usage of 'slanted' and affectively loaded lexicon when conveying an opinion.
Recognition of such cases could be improved by enriching the data preprocessing, e.g. by exploiting information about lexicon polarity and subjectivity parameters or information about word class use. In the following section we present a qualitative study of the lexicon employed in formulating dialogue acts.

## 4. Studying the lexicon of Dialogue Acts

To better understand what are the distinctive lexical features of each DA so as to improve the performance of our unsupervised approach, we performed a qualitative analysis to investigate:

(a) the relationship between the affective load of a given utterance and the communicative intention it conveys (i.e. the DA);

(b) the salience of word categories for each DA.

### 4.1. Affective load of Dialogue Acts

Sensing emotions from text is an appealing task for computational linguistics (Strapparava and Mihalcea, 2007): it is becoming a fundamental issue in several domains such as human-computer interaction (see, for example, (Conati, 2002; Picard and Klein, 2001; Clarizio et al., 2006)) or sentiment analysis for opinion mining (e.g. (Pang and Lee, 2008)). A first attempt to exploit affective information in dialogue act disambiguation has been made by Bosma and André (2004), with promising results. In their study, the recognition of emotions is based on sensory inputs that evaluate physiological user input.
In this section, we present the results of a qualitative study aimed at investigating the affective load of DAs. To the best of our knowledge, this is the first attempt to study the relationship between the communicative goal of an utterance and its affective load by applying lexical similarity techniques to textual input.
We calculated the affective load of each DA label using the methodology described in (Strapparava and Mihalcea, 2008). The idea underlying the method is the distinction between direct and indirect affective words. For direct affective words, authors refer to the WordNet Affect (Strapparava and Valitutti, 2004) lexicon, which is exploited to represent emotions in an LSA space acquired from the British National Corpus [1]. This LSA space is then used to check the affective load of indirect affective words.
Results (see Table 4) are quite encouraging and show that a relationship exists between the communicative goal of an utterance and its affective load: S-OPINION is the DA with the highest affective load, immediately followed by KINDATT due to the high frequency of politeness expressions in such utterances (see Table 5 for examples).

| Label | Affective Load |
|---|---|
| S-OPINION | .1439 |
| KIND-ATT | .1411 |
| STATEMENT | .1300 |
| INFO-REQ | .1142 |
| CLOSING | .0671 |
| REJECT | .0644 |
| OPENING | .0439 |
| AGREE-ACC | .0408 |
| GEN-ANS | .0331 |

Table 4: Affective load of DA labels

[1] http://www.hcu.ox.ac.uk/bnc/

| |
|---|
| S-OPINION |
| Gosh uh, it's getting pathetic now, absolutely pathetic. |
| They're just horrid, you'll have nightmares, you know. |
| That's no way to make a decision on some terrible problem. |
| They are just gems of shows. Really, fabulous in every way . |
| And, oh, that is so good. Delicious. |
| KIND-ATTITUDE |
| I'm sorry, I really feel strongly about this. |
| Sorry, now I'm probably going to upset you. |
| I hate to do it on this call. |

Table 5: Examples of slanted lexicon in S-OPINION and KIND-ATT (b)

### 4.2. Identifying dominant lexical categories in Dialogue Acts

We conducted a qualitative investigation of the lexicon of each DA to better understand what are the most distinctive lexical features (i.e. word classes) for classification. We followed the methodology described in (Mihalcea and Pulman, 2009) to calculate a score associated with a given class of words, in order to evaluate the relevance of each class with respect to a specific DA.

Let $C$ be a class of words $C = W_1, W_2, ..., W_n$ and $da$ the generic dialogue act, belonging to the Dialogue Act set employed for this study (see Table 1). We can build the corpus $DA$ including all utterances in our data set that have been labeled as $da$ (e.g. the complete set of all INFO-REQUEST), as well as the complementary corpus $\neg DA$, which includes all the utterances annotated differently. We compute the *dominance score* for the class $C$ in the generic dialogue act *DA* as

$$Dominance_{DA}(C) = \frac{Coverage_{DA}(C)}{Coverage_{\neg DA}(C)} \quad (1)$$

The class coverage for the *DA* is calculated as

$$Coverage_{DA}(C) = \frac{\sum_{W_i \in C} Frequency_{DA}(W_i)}{Size_{DA}}$$

where $Frequency_{DA}(W_i)$ is the total number of occurrences of all words in $C$ in $DA$ and $Size_{DA}$ is the dimension of $DA$ in words. Analogously, the class coverage for the rest of the corpus $\neg DA$ is calculated as

$$Coverage_{\neg DA}(C) = \frac{\sum_{W_i \in C} Frequency_{\neg DA}(W_i)}{Size_{\neg DA}}$$

A dominance score close to 1 indicates that $C$ has a similar distribution for both $DA$ and the rest of the corpus (that is, $C$ is not salient for $da$). On the contrary, a score significantly higher than 1 indicates a high salience of a class of words for a given DA.

In our study, we refer to the word classes defined in the Linguistic Inquiry and Word Count (LIWC) taxonomy, developed in the scope of psycholinguistic research (Pennebaker and Francis, 2001). We do not consider domain specific categories of words (e.g. School, Money, Leisure etc.) in order to make the analysis consistent with our goal of defining a domain-independent approach for DA annotation.

Table 6 shows the ranking for the most salient word classes for each DA with their dominance score. Sample words for each class are provided in Table 7. Results are particularly interesting and confirm our findings about the higher affective load for S-OPINION and KIND-ATTITUDE labels. In particular, negative emotions seem to prevail in the expression of opinions while words referring to both, positive and negative affective states, are used for kind-attitude expressions. Also, the class FEEL is relevant to both labels. Of course, and according to Austin's definition of 'Behabitives' (Austin, 1962), the fact that affective loaded lexicon is used in the formulation of politeness expression of KIND-ATTITUDE doesn't necessary mean that the speaker is reporting about an emotion actually felt while speaking (as in 'I'm sorry' or in 'I'm pleased to announce you...'). Still, we believe that such an information about affective lexicon use in both opinions and kind attitude expressions should be exploited to improve the DA classification performance. This is one of the direction we intend to follow in our future research.

Moreover, it is interesting to see a clear distinction in the lexicon used for STATEMENTs and S-OPINIONs, because the confounding between these two labels is the main cause of error of our DA classifier. In particular, statements are mainly expressed using the past tense, the first person pronouns and expressions of inclusion (e.g. 'also', 'altogether', 'plus') while opinions are mainly expressed using the future tense. Also, when formulating statements people talk about facts, using lexicon related to physical actions (MOTION), the five senses and the perception of the world (SENSES). On the contrary, when expressing opinions people mainly refer to their feelings (FEEL) and beliefs (COGMECH). This result confirms the descriptive/narrative nature of statements (Austin, 1962; Searle, 1969) in contrast with the subjective connotation of opinions, which are rather connected to appraisal and evaluation.

There is also a clear distinction in the lexicon used for expressing agreement and disagreement: ASSENT, CERTAIN and OPTIM categories are highly salient for the AGREE-ACCEPT label while negation (NEGATE) and exclamations (METAPH) are salient for REJECT.

OPENING and CLOSING share the common characteristic of being used for meta-communication goals (respectively, for beginning and ending the interaction). Hence, they both show linguistic features related to their role, like the lexicon included in the COMM and HEAR category (e.g. verbs like 'call', 'chat', 'discuss', 'talk'). For example, the category HEAR is particularly salient for CLOSING because the most common way of closing the dialogue, in the Switchboard corpus, is to use sentences like 'Its been nice talking to you'.

Finally, the YOU and OTHREF categories seem to be relevant for the INFO-REQUEST, which clearly indicates that

| Opinion | | Statement | | Kind-Att | |
|---|---|---|---|---|---|
| FUTURE | 2.00 | PAST | 2.17 | NEGEMO | 19.14 |
| NEGEMO | 1.85 | I,SELF,WE | ∼2 | AFFECT | 7.95 |
| SAD | 1.69 | INCL | 1.41 | POSEMO | 5.43 |
| INSIGHT | 1.56 | SEE | 1.30 | COMM | 4.51 |
| ANGER | 1.54 | MOTION | 1.25 | INHIB | 2.68 |
| DISCREP | 1.47 | HEAR | 1.18 | ANGER | 2.61 |
| OPTIM | 1.49 | SENSES | 1.17 | SELF, FEEL | ∼2.3 |
| FEEL | 1.44 | | | ANX | 1.87 |
| SWEAR | 1.40 | | | | |
| COGMECH | 1.37 | | | | |
| **Reject** | | **Agree-acc** | | **Opening** | |
| NEGATE | 14.54 | ASSENT | 75.32 | COMM | 27.65 |
| METAPH | 1.91 | CERTAIN | 4.64 | ASSENT | 3.22 |
| NEGEMO | 1.60 | POSEMO | 2.67 | SOCIAL | 3.10 |
| INHIB | 1.22 | AFFECT | 2.22 | CAUSE | 3.02 |
| | | OPTIM | 2.12 | HEAR | 2.10 |
| **Closing** | | **Info-Req** | | **Gen-Ans** | |
| HEAR | 8.10 | YOU | 3.73 | ASSENT | 38.21 |
| ASSENT | 6.75 | CAUSE | 1.88 | NEGATE | 7.15 |
| COMM | 6.42 | OTHREF | 1.73 | | |

Table 6: Dominant word classes for each DA with their scores

the attentional focus (Pennebaker and Francis, 2001) of questions is on the interlocutor rather than on the speaker.

| Class | Sample words |
|---|---|
| PAST | had, ago, became, called, did, disliked |
| FUTURE | be, I'll, may, might, will, won't, you'll |
| ASSENT | accept, alright, fine, yep, yeah |
| NEGATE | aren't, don't, neither, no, never, zero |
| AFFECT | wrong, warm, sorrow, romantic, unpleasant |
| NEGEMO | abandon, anger, boring, cry, danger, depressed |
| POSEMO | won, wealth, triumph, treasure, wisdom, sweet |
| INSIGHT | believe, think, know, see, understand, feels |
| COGMECH | acknowledge, admit, become, believe, discern |
| FEEL | tries, senses, pain, hold, grab, feel |
| I | I, myself, mine |
| SELF | our, myself, mine, ours |
| WE | us, we, our, ourselves |
| YOU | you, thou |
| INCL | also, altogether, and, here, plus |
| MOTION | go, approach, bring, carry, cross, drive |
| SENSES | witness, touch, tell, talk, look, listen, perceive |
| HEAR | talk, ask, call, discuss, ear, listen, say, tell |
| METAPH | god, die, sacred, mercy, sin, dead, hell |
| CERTAIN | always, all, very, truly, completely, totally |
| OPTIM | best, ready, hope, accepts, proud, won, super, |
| COMM | admit, blame, call, chat, describe, discuss |
| SOCIAL | ya, ye, you, you'd, you'll, your |

Table 7: LIWC word classes with sample words

## 5. Conclusion

The long-term goal of our research is to define an unsupervised approach for DA labelling. The method has to be independent from the language, domain, size, interaction scenario of the referred corpus, focusing only on lexical analysis. In our previous work (Novielli and Strapparava, 2009) some preliminary steps have been done toward the achievement of this goal. In this paper we proposed a qualitative study of the lexicon of dialogue acts in order to better understand what are the most salient and distinctive lexical features for DA profiling. In particular we investigated the relationship between the affective load of utterances and their communicative goal. Finally the analysis of word classes dominance highlighted interesting lexical patterns for DAs. As a direction for future work, we plan to exploit the findings of the present study to improve the performance of our unsupervised method (Novielli and Strapparava, 2009) (e.g. by enriching the preprocessing with information about the affective load of sentences or by exploiting the salience of word classes).

## 6. References


J. Austin. 1962. *How to do Things with Words*. Oxford University Press, New York.

M. Berry. 1992. Large-scale sparse singular value computations. *International Journal of Supercomputer Applications*, 6(1).

W. Bosma and E. André. 2004. Exploiting emotions to disambiguate dialogue acts. In *IUI '04: Proceedings of the 9th international conference on Intelligent user interfaces*, pages 85–92, New York, NY, USA. ACM.

G. Clarizio, I. Mazzotta, N. Novielli, and F. deRosis. 2006. Social attitude towards a conversational character. In *Proceedings of the 15th IEEE International Symposium on Robot and Human Interactive Communication*, pages 2–7, Hatfield, UK, September.

P. R. Cohen and H. J. Levesque. 1995. Communicative actions for artificial agents. In *in Proceedings of the First International Conference on Multi-Agent Systems*, pages 65–72. AAAI Press.

C. Conati. 2002. Probabilistic assessment of user's emotions in educational games. *Applied Artificial Intelligence*, 16:555–575.

M. Core and J. Allen. 1997. Coding dialogs with the DAMSL annotation scheme. In *Working Notes of the AAAI Fall Symposium on Communicative Action in Humans and Machines*, pages 28–35, Cambridge, MA, November.

A. Gliozzo and C. Strapparava. 2005. Domains kernels for text categorization. In *Proc. of the Ninth Conference on Computational Natural Language Learning (CoNLL-2005)*, pages 56–63, University of Michigan, Ann Arbor, June.

J. Godfrey, E. Holliman, and J. McDaniel. 1992. SWITCHBOARD: Telephone speech corpus for research and development. In *Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pages 517–520, San Francisco, CA. IEEE.

D. Jurafsky, E. Shriberg, and D. Biasca. 1997. Switchboard SWBD-DAMSL shallow-discourse-function annotation coders manual, draft 13. Technical Report 97-01, University of Colorado Institute of Cognitive Science.

T. K. Landauer, P. Foltz, and D. Laham. 1998. Introduction to latent semantic analysis. *Discourse Processes*, 25.

R. Mihalcea and S. Pulman. 2009. Linguistic ethnography: Identifying dominant word classes in text. In *Proceeding of Computational Linguistics and Intelligent Text Processing (CICLing-09)*.

N. Novielli and C. Strapparava. 2009. Towards unsupervised recognition of dialogue acts. In *NAACL HLT 2009, Student Research Workshop*.

B. Pang and L. Lee. 2008. Opinion mining and sentiment analysis. *Foundations and Trends in Information Retrieval*, 2(1-2):1–135.

J. Pennebaker and M. Francis. 2001. Linguistic inquiry and word count: LIWC. Erlbaum Publishers.

E. Pianta, C. Girardi, and R. Zanoli. 2008. The TextPro tool suite. In *Proceedings of LREC-08*, Marrakech, Morocco, May.

R. W. Picard and J. Klein. 2001. Computers that recognise and respond to user emotion: Theoretical and practical implications. Technical report, MIT Media Lab.

M. Poesio and A. Mikheev. 1998. The predictive power of game structure in dialogue act recognition: Experimental results using maximum entropy estimation. In *Proceedings of ICSLP-98*, Sydney, December.

R. Power. 1979. The organisation of purposeful dialogues. *Linguistics*, 17:107–152.

N. Reithinger, M. Kipp, R. Engel, and M. Klesen. 1996. Predicting dialogue acts for a speech-to-speech translation system. In *Proceedings of the International Conference on Spoken Language Processing*, pages 654–657.

H. Sacks, E. Schegloff, and G. Jefferson. 1974. A simplest systematics for the organization of turn-taking for conversation. *Language*, 50(4):696–735.

K. Samuel, S. Carberry, and K. Vijay-Shanker. 1998. Dialogue act tagging with transformation-based learning. In *Proceedings of the 17th international conference on Computational linguistics*, pages 1150–1156, Morristown, NJ, USA. Association for Computational Linguistics.

E. Schegloff. 1968. Sequencing in conversational openings. *American Anthropologist*, 70:1075–1095.

J. Searle. 1969. *Speech Acts: An Essay in the Philosophy of Language*. Cambridge University Press, Cambridge, London.

A. Stolcke, N. Coccaro, R. Bates, P. Taylor, C. Van Ess-Dykema, K. Ries, E. Shriberg, D. Jurafsky, R. Martin, and M. Meteer. 2000. Dialogue act modeling for automatic tagging and recognition of conversational speech. *Computational Linguistics*, 26(3):339–373.

C. Strapparava and R. Mihalcea. 2007. SemEval-2007 task 14: Affective Text. In *Proceedings of the $4^{th}$ International Workshop on Semantic Evaluations (SemEval 2007)*, pages 70–74, Prague, June.

C. Strapparava and R. Mihalcea. 2008. Learning to identify emotions in text. In *SAC '08: Proceedings of the 2008 ACM symposium on Applied computing*, pages 1556–1560, New York, NY, USA. ACM.

C. Strapparava and A. Valitutti. 2004. WordNet-Affect: an affective extension of WordNet. In *Proceedings of LREC*, volume 4, pages 1083–1086.

D. Traum. 2000. 20 questions for dialogue act taxonomies. *Journal of Semantics*, 17(1):7–30.

V. Vapnik. 1995. *The Nature of Statistical Learning Theory*. Springer-Verlag.

V. Warnke, R. Kompe, H. Niemann, and E. Nöth. 1997. Integrated dialog act segmentation and classification using prosodic features and language models. In *Proceedings of 5th European Conference on Speech Communication and Technology*, volume 1, pages 207–210, Rhodes, Greece.