

# Classifying Speech Acts using Verbal Response Modes

Andrew Lampert †‡

Robert Dale ‡

Cécile Paris †

†CSIRO ICT Centre  
Locked Bag 17, North Ryde  
NSW 1670 Australia

firstname.lastname@csiro.au

‡Centre for Language Technology  
Macquarie University  
NSW 2109 Australia

[alampert,rdale]@ics.mq.edu.au

## Abstract

The driving vision for our work is to provide intelligent, automated assistance to users in understanding the status of their email conversations. Our approach is to create tools that enable the detection and connection of speech acts across email messages. We thus require a mechanism for tagging email utterances with some indication of their dialogic function. However, existing dialog act taxonomies as used in computational linguistics tend to be too task- or application-specific for the wide range of acts we find represented in email conversation. The Verbal Response Modes (VRM) taxonomy of speech acts, widely applied for discourse analysis in linguistics and psychology, is distinguished from other speech act taxonomies by its construction from cross-cutting principles of classification, which ensure universal applicability across any domain of discourse. The taxonomy categorises on two dimensions, characterised as literal meaning and pragmatic meaning. In this paper, we describe a statistical classifier that automatically identifies the literal meaning category of utterances using the VRM classification. We achieve an accuracy of 60.8% using linguistic features derived from VRM's human annotation guidelines. Accuracy is improved to 79.8% using additional features.

## 1 Introduction

It is well documented in the literature that users are increasingly using email for managing requests

and commitments in the workplace (Bellotti et al., 2003). It has also been widely reported that users commonly feel overloaded when managing multiple ongoing tasks through email communication e.g. (Whittaker and Sidner, 1996).

Given significant task-centred email usage, one approach to alleviating email overload in the workplace is to draw on Speech Act Theory (Searle, 1969) to analyse the intention behind email messages and use this information to help users process and prioritise their email. The basic tenet of Speech Act Theory is that when we utter something, we also act. Examples of such acts can include stating, questioning or advising.

The idea of identifying and exploiting patterns of communicative acts in conversations is not new. Two decades ago, Flores and Winograd (1986) proposed that workplace workflow could be seen as a process of creating and maintaining networks of conversations in which requests and commitments lead to successful completion of work.

Recently, these ideas have begun to be applied to email messages. Existing work analysing speech acts in email messages differs as to whether speech acts should be annotated at the message level, e.g., (Cohen et al., 2004; Leuski, 2004), or at the utterance or sentence level, e.g., (Corston-Oliver et al., 2004). Our thesis is that a single email message may contain multiple commitments on a range of tasks, and so our work focuses on utterance-level classification, with the aim of being able to connect together the rich tapestry of threads that connect individual email messages.

Verbal Response Modes (VRM) (Stiles, 1992) is a principled taxonomy of speech acts for classifying the literal and pragmatic meaning of utterances. The hypothesis we pose in this work is that VRM annotation can be learned to create a classi-

fier of literal utterance meaning.

The driving vision for our work is to eventually provide intelligent, automated assistance to email users in understanding the status of their current email conversations and tasks. We wish to assist users to identify outstanding tasks easily (both for themselves and their correspondents) through automatically flagging incomplete conversations, such as requests or commitments that remain unfulfilled. This capability should lead to novel forms of conversation-based search, summarisation and navigation for collections of email messages and for other textual, computer-mediated conversations. The work described here represents our first steps towards this vision.

This paper is structured as follows. First, in Section 2, we describe related work on automatically classifying speech and dialogue acts. In Section 3 we introduce the VRM taxonomy, comparing and contrasting it with other speech act taxonomies in Section 4. Then, in Section 5, we describe our statistical VRM classifier, and in Section 6 we present what we believe are the first results in the field for automatic VRM classification of the literal meaning of utterances. Finally, in Section 7 we discuss our results. Section 8 presents some concluding remarks and pointers to future work.

## 2 Related Work

There is much existing work that explores automated processing of speech and dialogue acts. This collection of work has predominantly focused around two related problems: *dialogue act prediction* and *dialogue act recognition*. Our work focuses on the second problem, and more specifically on speech act recognition.

Examples of dialogue act recognition include work by Core (1998) which uses previous and current utterance information to predict possible annotations from the DAMSL scheme (Core and Allen, 1997). Similar work by Chu-Carroll (1998) on statistical “discourse act” recognition also uses features from the current utterance and discourse history to achieve accuracy of around 51% for a set of 15 discourse acts. In particular, Chu-Carroll’s results were significantly improved by taking into account the syntactic form of each utterance.

The use of n-gram language models is also a popular approach. Reithinger and Klesen (1997) apply n-gram language models to the VERBMobil corpus (Alexandersson et al.,

1998) and report tagging accuracy of 74.7% for a set of 18 dialogue acts. In common with our own work, Webb et al. (2005) approach dialogue act classification using only intra-utterance features. They found that using only features derived from n-gram cue phrases performed moderately well on the SWITCHBOARD corpus of spoken dialogue (Godfrey et al., 1992).

To our knowledge, however, there has been no previous work that attempts to identify VRM categories for utterances automatically.

## 3 Verbal Response Modes

Verbal Response Modes (VRM) is a principled taxonomy of speech acts that can be used to classify literal and pragmatic meaning within utterances. Each utterance is coded twice: once for its *literal meaning*, and once for its communicative intent or *pragmatic meaning*. The same VRM categories are used in each case.

Under the VRM system, every utterance from a speaker can be considered to concern either the *speaker’s* or the *other’s* experience. For example, in the utterance “I like pragmatics.”, the *source of experience* is the speaker. In contrast, the source of experience for the utterance “Do you like pragmatics?” is the other interlocutor.

Further, in making an utterance, the speaker may need to make presumptions about experience. For example, in saying “Do you like pragmatics?”, the speaker does not need to presume to know what the other person is, was, will be, or should be thinking, feeling, perceiving or intending. Such utterances require a *presumption of experience* of the speaker only. In contrast, the utterance “Like pragmatics!” attempts to impose an experience (a liking for pragmatics) on the other interlocutor, and has a presumption of experience for the other.

Finally a speaker may represent the experience either from their own personal point of view, or from a viewpoint that is shared or held in common with the other interlocutor. The three example utterances above all use the speaker’s *frame of reference* because the experience is understood from the speaker’s point of view. In contrast, the utterance “You like pragmatics.” takes the other’s frame of reference, representing the experience as the other interlocutor views it.

These three principles — *source of experience*, *presumption about experience* and *frame of reference* — form the basis of the VRM taxonomy.

The principles are dichotomous — each can take the value of *speaker* or *other* (other interlocutor) — and thus define eight mutually exclusive VRM categories, as shown in Table 1.

With 8 VRM modes and 2 separate dimensions of coding (literal and pragmatic meaning), there are 64 possible form-intent combinations. The 8 combinations of codings in which the literal and pragmatic meanings coincide are referred to as *pure modes*. The other 56 modes are labelled *mixed modes*. An example of a mixed-mode utterance is “Can you pass the sugar?” which is coded QA. This is read *Question in service of Advice*, meaning that the utterance has a Question form (literal meaning) but Advice intent (pragmatic meaning). In this way, the VRM taxonomy is designed to simply and consistently classify and distinguish direct and indirect speech acts.

#### 4 Comparison with Other Speech Act Taxonomies

There are, of course, alternate speech and dialogue act taxonomies, some of which have been applied within natural language processing applications. Unfortunately, many of these taxonomies tend to offer competing, rather than complementary approaches to classifying speech acts, making it difficult to compare experimental results and analyses that are based on different taxonomies. It would clearly be desirable to unambiguously relate categories between different taxonomies.

One specific drawback of many speech and dialogue act taxonomies, including taxonomies such as those developed in the VERBMOBIL project (Alexandersson et al., 1998), is that they are domain or application specific in their definition and coverage of speech act categories. This often stems from the taxonomy being developed and used in a rather *ad hoc*, empirical manner for analysing discourse and utterances from a single or small set of application domains.

While the VRM research grew from studying therapist interventions in psychotherapy (Stiles, 1992; Wiser and Goldfried, 1996), the VRM system has been applied to a variety of discourse genres. These include: American Presidential speeches (Stiles et al., 1983), doctor-patient interactions (Meeuswesen et al., 1991), courtroom interrogations (McGaughey and Stiles, 1983), business negotiations (Ulijn and Verweij, 2000), persuasive discourse (Kline et al., 1990) and tele-

vision commercials (Rak and McMullen, 1987). VRM coding assumes only that there is a *speaker* and an intended audience (*other*), and thus can be applied to any domain of discourse.

The wide applicability of VRM is also due to its basis of clearly defined, domain-independent, systematic principles of classification. This ensures that the VRM categories are both extensive and exhaustive, meaning that all utterances can be meaningfully classified with exactly one VRM category<sup>1</sup>. In contrast, even widely-applied, complex taxonomies such as DAMSL (Core and Allen, 1997) resort to the inclusion of an *other* category within the speech act component, to be able to classify utterances across domains.

In addition, the VRM principles facilitate more rigorous and comparable coding of utterances from which higher-level discourse properties can be reliably calculated, including characterisation of the roles played by discourse participants. If required, the eight VRM modes can also be further divided to identify additional features of interest (for example, the Question category could be split to distinguish open and closed questions). Importantly, this can be done within the existing framework of categories, without losing the principled basis of classification, or the ability to compare directly with other VRM analyses.

Table 2 compares the VRM categories with Searle’s five major speech act categories (1969; 1979). Searle’s categories are largely subsumed under the subset of VRM categories that offer the *speaker’s* source of experience and/or frame of reference (Disclosure, Edifications, Advice and Questions). The coverage of Searle’s speech acts seems more limited, given that the other VRMs (Reflection, Interpretation, Confirmation and Acknowledgement), all *other* on at least two principles, have no direct equivalents in Searle’s system, except for some Interpretations which might map to specific subcategories of Declaration.

#### 5 Building a VRM Classifier

As discussed earlier, the VRM system codes both the literal and pragmatic meaning of utterances. The pragmatic meaning conveys the speaker’s actual intention, and such meaning is often hidden or

<sup>1</sup>The only exceptions are utterances that are inaudible or incomprehensible in spoken dialogue, which are coded Uncodable (U).

Source of Experience	Presumption about Experience	Frame of Reference	VRM Mode	Description
Speaker	Speaker	Speaker	Disclosure (D)	Reveals thoughts, feelings, perceptions or intentions. E.g., <i>I like pragmatics.</i>
		Other	Edification (E)	States objective information. E.g., <i>He hates pragmatics.</i>
	Other	Speaker	Advisement (A)	Attempts to guide behaviour; suggestions, commands, permission, prohibition. E.g., <i>Study pragmatics!</i>
		Other	Confirmation (C)	Compares speaker's experience with other's; agreement, disagreement, shared experience or belief. E.g., <i>We both like pragmatics.</i>
Other	Speaker	Speaker	Question (Q)	Requests information or guidance. E.g., <i>Do you like pragmatics?</i>
		Other	Acknowledgement (K)	Conveys receipt of or receptiveness to other's communication; simple acceptance, salutations. E.g., <i>Yes.</i>
	Other	Speaker	Interpretation (I)	Explains or labels the other; judgements or evaluations of the other's experience or behaviour. E.g., <i>You're a good student.</i>
		Other	Reflection (R)	Puts other's experience into words; repetitions, re-statements, clarifications. E.g., <i>You dislike pragmatics.</i>

Table 1: The Taxonomy of Verbal Response Modes from (Stiles, 1992)

Searle's Classification	Corresponding VRM
Commissive	Disclosure
Expressive	Disclosure
Representative	Edification
Directive	Advisement; Question
Declaration	Interpretation; Disclosure; Edification

Table 2: A comparison of VRM categories with Searle's speech acts

highly dependent on discourse context and background knowledge. Because we classify utterances using only intra-utterance features, we cannot currently encode any information about the discourse context, so could not yet plausibly tackle the prediction of pragmatic meaning. Discerning literal meaning, while somewhat simpler, is

akin to classifying direct speech acts and is widely recognised as a challenging computational task.

### 5.1 Corpus of VRM Annotated Utterances

Included with the VRM coding manual (Stiles, 1992) is a VRM coder training application for training human annotators. This software, which is freely available online<sup>2</sup>, includes transcripts of spoken dialogues from various domains segmented into utterances, with each utterance annotated with two VRM categories that classify both its literal and pragmatic meaning.

These transcripts were pre-processed to remove instructional text and parenthetical text that was not actually part of a spoken and coded utter-

<sup>2</sup>The VRM coder training application and its data files are available to download from <http://www.users.muohio.edu/stileswb/archive.htmlx>

ance. Several additional example utterances were extracted from the coding manual to increase the number of instances of under-represented VRM categories (notably *Confirmations* and *Interpretations*).

The final corpus contained 1368 annotated utterances from 14 dialogues and several sets of isolated utterances. Table 3 shows the frequency of each VRM mode in the corpus.

VRM	Instances	Percentage
Disclosure	395	28.9%
Edification	391	28.6%
Advisement	73	5.3%
Confirmation	21	1.5%
Question	218	15.9%
Acknowledgement	97	7.1%
Interpretation	64	4.7%
Reflection	109	8.0%

Table 3: The distribution of VRMs in the corpus

## 5.2 Features for Classification

The VRM annotation guide provides detailed instructions to guide humans in correctly classifying the literal meaning of utterances. These suggested features are shown in Table 4.

We have attempted to map these features to computable features for training our statistical VRM classifier. Our resulting set of features is shown in Table 5 and includes several additional features not identified by Stiles that we use to further characterise utterances. These additional features include:

- **Utterance Length:** The number of words in the utterance.
- **First Word:** The first word in each utterance, represented as a series of independent boolean features (one for each unique first word present in the corpus).
- **Last Token:** The last token in each utterance – either the final punctuation (if present) or the final word in the utterance. As for the First Word features, these are represented as a series of independent boolean features.
- **Bigrams:** Bigrams extracted from each utterance, with a variable threshold for including only frequent bigrams (above a specified threshold) in the final feature set.

VRM Category	Form Criteria
Disclosure	Declarative; 1st person singular or plural where <i>other</i> is not a referent.
Edification	Declarative; 3rd person.
Advisement	Imperative or 2nd person with verb of permission, prohibition or obligation.
Confirmation	1st person plural where referent includes the other (i.e., “we” refers to both <i>speaker</i> and <i>other</i> ).
Question	Interrogative, with inverted subject-verb order or interrogative words.
Acknowledgement	Non-lexical or contentless utterances; terms of address or salutation.
Interpretation	2nd person; verb implies an attribute or ability of the other; terms of evaluation.
Reflection	2nd person; verb implies internal experience or volitional action.

Table 4: VRM form criteria from (Stiles, 1992)

The intuition for including the utterance length as a feature is that different VRMs are often associated with longer or shorter utterances - e.g., Acknowledgement utterances are often short, while Edifications are often longer.

To compute our utterance features, we made use of the Connexor Functional Dependency Grammar (FDG) parser (Tapanainen and Jarvinen, 1997) for grammatical analysis and to extract syntactic dependency information for the words in each utterance. We also used the morphological tags assigned by Connexor. This information was used to calculate utterance features as follows:

- **Functional Dependencies:** Dependency functions were used to identify main subjects and main verbs within utterances, as required for features including the 1st/2nd/3rd person subject, inverted subject-verb order and imperative verbs.
- **Syntactic Functions:** Syntactic function in-

formation was determined using the Connexor parser. This information was used to identify the main utterance subject where dependency information was not available.

- **Morphology:** Morphological tags, also generated by Connexor, were used to distinguish between first and third person pronouns, as well as between singular and plural forms of first person pronouns. Additionally, we used morphological tags from Connexor to identify imperative verbs.
- **Hand-constructed word lists:** Several of the features used relate to closed sets of common lexical items (e.g., verbs of permission, interrogative words, variations of “yes” and “no”). For these features, we employ hand-constructed simple lists, using online thesauri to expand our lists from an initial set of seed words. While some of the lists are not exhaustive, they seem to help our results and involved only a small amount of effort; none took more than an hour to construct.

<i>Feature</i>	<i>Likely VRM</i>
1st person singular subject	D,Q
1st person plural singular subject	D,C
3rd person subject	E,Q
2nd person subject	A,Q,I,R
Inverted subject-verb order	Q
Imperative verb	A
Verbs of permission, prohibition, obligation	A
Interrogative words	Q
Non-lexical content	K
Yes/No variants	K
Terms of evaluation	I
Utterance length	all
First word	all
Last token	all
Bi-grams	all

Table 5: Features used in VRM Classifier

## 6 Results

Our classification results using several different learning algorithms and variations in feature sets are summarised in Table 6. We experimented with using only the linguistic features suggested

by Stiles, using only the additional features he identified, and using a combination of all features shown in Table 5. All our results were validated using stratified 10-fold cross validation.

We used supervised learning methods implemented in Weka (Witten and Frank, 2005) to train our classifier. Through experimentation, we found that Weka’s Support Vector Machine implementation (SMO) provided the best classification performance. Encouragingly, other relatively simple approaches, such as a Bayesian Network classifier using the K2 hill-climbing search algorithm, also performed reasonably well.

The baseline against which we compare our classifier’s performance is a OneR (one rule) classifier using an identical feature set. This baseline system is a one-level decision tree, (i.e., based on a set of rules that test only the single most discriminative feature). As shown in Table 6, the accuracy of this baseline varies from 42.76% to 49.27%, depending on the exact features used. Regardless of features or algorithms, our classifier performs significantly better than the baseline system.

<i>Algorithm</i>	<i>Feature Set</i>	<i>Accuracy</i>	<i>Mean Abs Error</i>
SVM	All	79.75%	0.19
SVM	Only Stiles’	60.82%	0.20
SVM	No Stiles’	74.49%	0.19
Bayes Net	All	78.51%	0.06
Bayes Net	Only Stiles’	60.16%	0.12
Bayes Net	No Stiles’	75.68%	0.07
Baseline	All	49.27%	0.36
Baseline	Only Stiles’	49.27%	0.36
Baseline	No Stiles’	42.76%	0.38

Table 6: VRM classifier results

Another tunable parameter was the level of pruning of n-grams from our feature set according to their frequency of occurrence. Heuristically, we determined that a cut-off of 5 (i.e., only n-grams that occur five or more times in our corpus of utterances were included as features) gave us the highest accuracy for the learning algorithms tested.

## 7 Discussion

This work appears to be the first attempt to automatically classify utterances according to their literal meaning with VRM categories. There are thus no direct comparisons to be easily drawn for

our results. In classifying only the literal meaning of utterances, we have focused on a simpler task than classifying in-context meaning of utterances which some systems attempt.

Our results do, however, compare favourably with previous dialogue act classification work, and clearly validate our hypothesis that VRM annotation can be learned. Previous dialogue act classification results include Webb et al. (2005) who reported peak accuracy of around 71% with a variety of n-gram, word position and utterance length information on the SWITCHBOARD corpus using the 42-act DAMSL-SWBD taxonomy. Earlier work by Stolcke et al. (2000) obtained similar results using a more sophisticated combination of hidden markov models and n-gram language models with the same taxonomy on the same corpus. Reithinger and Klesen (1997) report a tagging accuracy of 74.7% for a set of 18 dialogue acts over the much larger VERBMOBIL corpus (more than 223,000 utterances, compared with only 1368 utterances in our own corpus).

<i>VRM</i>	<i>Instances</i>	<i>Precision</i>	<i>Recall</i>
D	395	0.905	0.848
E	391	0.808	0.872
A	73	0.701	0.644
C	21	0.533	0.762
Q	218	0.839	0.885
K	97	0.740	0.763
I	64	0.537	0.453
R	109	0.589	0.514

Table 7: Precision and recall for each VRM

In performing an error analysis of our results, we see that classification accuracy for Interpretations and Reflections is lower than for other classes, as shown in Table 7. In particular, our confusion matrix shows a substantial number of transposed classifications between these two VRMs. Interestingly, Stiles makes note that these two VRMs are very similar, differing only on one principle (frame of reference), and that they are often difficult to distinguish in practice. Additionally, some Reflections repeat all or part of the *other's* utterance, or finish the *other's* previous sentence. It is impossible for our current classifier to detect such phenomena, since it looks at utterances in isolation, not in the context of a larger discourse. We plan to address this in future work.

Our results also provide support for using both

linguistic and statistical features in classifying VRMs. In the cases where our feature set consists of only the linguistic features identified by Stiles, our results are substantially worse. Similarly, when only n-gram, word position and utterance length features are used, classifier performance also suffers. Table 6 shows that our best results are obtained when both types of features are included.

Finally, another clear trend in the performance of our classifier is that the VRMs for which we have more utterance data are classified substantially more accurately.

## 8 Conclusion

Supporting the hypothesis posed, our results suggest that classifying utterances using Verbal Response Modes is a plausible approach to computationally identifying literal meaning. This is a promising result that supports our intention to apply VRM classification as part of our longer-term aim to construct an application that exploits speech act connections across email messages.

While difficult to compare directly, our classification accuracy of 79.75% is clearly competitive with previous speech and dialogue act classification work. This is particularly encouraging considering that utterances are currently being classified in isolation, without any regard for the discourse context in which they occur.

In future work we plan to apply our classifier to email, exploiting features of email messages such as header information in the process. We also plan to incorporate discourse context features into our classification and to explore the classification of pragmatic utterance meanings.

## References

- Jan Alexandersson, Bianka Buschbeck-Wolf, Tsutomu Fujinami, Michael Kipp, Stephan Koch, Elisabeth Maier, Norbert Reithinger, Birte Schmitz, and Melanie Siegel. 1998. Dialogue acts in VERBMOBIL-2. Technical Report Verbmobil-Report 226, DFKI Saarbruecken, Universitt Stuttgart, Technische Universitt Berlin, Universitt des Saarlandes. 2nd Edition.
- Victoria Bellotti, Nicolas Ducheneaut, Mark Howard, and Ian Smith. 2003. Taking email to task: The design and evaluation of a task management centred email tool. In *Computer Human Interaction Conference*, CHI, Ft Lauderdale, Florida, USA, April 5-10.

- Jennifer Chu-Carroll, 1998. *Applying Machine Learning to Discourse Processing. Papers from the 1998 AAAI Spring Symposium.*, chapter A statistical model for discourse act recognition in dialogue interactions, pages 12–17. The AAAI Press, Menlo Park, California.
- William W Cohen, Vitor R Carvalho, and Tom M Mitchell. 2004. Learning to classify email into "speech acts". In Dekang Lin and Dekai Wu, editors, *Conference on Empirical Methods in Natural Language Processing*, pages 309–316, Barcelona, Spain. Association for Computational Linguistics.
- Mark Core and James Allen. 1997. Coding dialogs with the DAMSL annotation scheme. In *AAAI Fall Symposium on Communicative Action in Humans and Machines*, pages 28–35, Cambridge, MA, November.
- Mark Core. 1998. Predicting DAMSL utterance tags. In *Proceedings of the AAAI-98 Spring Symposium on Applying Machine Learning to Discourse Processing*.
- Simon H Corston-Oliver, Eric Ringger, Michael Gamon, and Richard Campbell. 2004. Task-focused summarization of email. In *ACL-04 Workshop: Text Summarization Branches Out*, July.
- John J Godfrey, Edward C Holliman, and Jane McDaniel. 1992. SWITCHBOARD: Telephone speech corpus for research and development. In *Proceedings of the IEEE Conference on Acoustics, Speech and Signal Processing*, volume 1, pages 517–520, San Francisco, CA, March.
- Susan L Kline, C L Hennen, and K M Farrell. 1990. Cognitive complexity and verbal response mode use in discussion. *Communication Quarterly*, 38:350–360.
- Anton Leuski. 2004. Email is a stage: discovering people roles from email archives. In *Proceedings of Annual ACM Conference on Research and Development in Information Retrieval*, Sheffield, UK.
- Karen J McGaughey and William B Stiles. 1983. Courtroom interrogation of rape victims: Verbal response mode use by attorneys and witnesses during direct examination vs. cross-examination. *Journal of Applied Social Psychology*, 13:78–87.
- Ludwein Meeuswesen, Cas Schaap, and Cees van der Staak. 1991. Verbal analysis of doctor-patient communication. *Social Science and Medicine*, 32(10):1143–50.
- Diana S Rak and Linda M McMullen. 1987. Sex-role stereotyping in television commercials: A verbal response mode and content analysis. *Canadian Journal of Behavioural Science*, 19:25–39.
- Norbert Reithinger and Martin Klesen. 1997. Dialogue act classification using language models. In *Proceedings of Eurospeech '97*, pages 2235–2238, Rhodes, Greece.
- John R Searle. 1969. *Speech Acts : An Essay in the Philosophy of Language*. Cambridge University Press.
- John R Searle. 1979. *Expression and Meaning*. Cambridge University Press.
- William B Stiles, Melinda L Au, Mary Ann Martello, and Julia A Perlmutter. 1983. American campaign oratory: Verbal response mode use by candidates in the 1980 american presidential primaries. *Social Behaviour and Personality*, 11:39–43.
- William B Stiles. 1992. *Describing Talk: a taxonomy of verbal response modes*. SAGE Series in Interpersonal Communication. SAGE Publications. ISBN: 0803944659.
- Andreas Stolcke, Klaus Ries, Noah Coccaro, Elizabeth Shriberg, Rebecca Bates, Daniel Jurafsky, Paul Taylor, Rachel Martin, Marie Meteer, and Carol Van Ess-Dykema. 2000. Dialogue act modeling for automatic tagging and recognition of conversational speech. *Computational Linguistics*, 26(3):339–371.
- Pasi Tapanainen and Timo Jarvinen. 1997. A non-projective dependency parser. In *Proceedings of the 5th Conference on Applied Natural Language Processing*, pages 64–71, Washington D.C. Association for Computational Linguistics.
- Jan M Ulijn and Maurits J Verweij. 2000. Question behaviour in monocultural and intercultural business negotiations: the Dutch-Spanish connection. *Discourse Studies*, 2(1):141–172.
- Nick Webb, Mark Hepple, and Yorick Wilks. 2005. Dialogue act classification based on intra-utterance features. In *Proceedings of the AAAI Workshop on Spoken Language Understanding*.
- Steve Whittaker and Candace Sidner. 1996. Email overload: exploring personal information management of email. In *ACM Computer Human Interaction conference*, pages 276–283. ACM Press.
- Terry Winograd and Fernando Flores. 1986. *Understanding Computers and Cognition*. Ablex Publishing Corporation, Norwood, New Jersey, USA, 1st edition. ISBN: 0-89391-050-3.
- Susan Wiser and Marvin R Goldfried. 1996. Verbal interventions in significant psychodynamic-interpersonal and cognitive-behavioral therapy sessions. *Psychotherapy Research*, 6(4):309–319.
- Ian Witten and Eibe Frank. 2005. *Data Mining: Practical machine learning tools and techniques*. Morgan Kaufmann, San Francisco, 2nd edition.