An Annotated Bibliography of Affective Natural Language Generation

Paul Piwek

ITRI - University of Brighton, UK
Email: Paul.Piwek@itri.bton.ac.uk

Version 1.3
April, 2003
(Version 1.0 appeared as ITRI Technical Report ITRI-02-02, University of Brighton March 2002)

Contents
1 Introduction 1
2 Affective NLG: Systems and Computational Theories 3
3 Affect in Language: Linguistic, Descriptive and Empirical Work 7

1 Introduction

This annotated bibliography is an attempt to present an overview of papers which are directly dealing with affective natural language generation (affective NLG) and, additionally, studies into affect in language from empirical, descriptive, philosophical and linguistic perspectives. There is a variety of views on how affective NLG and more generally affect in language should be defined. We have collected a number of these definitions. We included all studies which we encountered and which matched at least one of the definitions. In other words, we cast our net fairly widely when collecting papers for this bibliography. Nevertheless, it is impossible to claim completeness: whereas we believe that a substantial part of the studies on affective NLG is included, this cannot be said for studies into the wider field of affect in language. Here, there is such a wealth of papers by psychologists, linguists, sociolinguists and others that we had to content ourselves with a limited survey of the field. It also has to be stressed that we have excluded studies which focus on how affect can be expressed in speech by means of, for instance, pitch, intensity and duration. Roughly speaking, our concern is with how affect influences content selection, grammatical realization and lexical choice.

When one is looking for the definition of a word, an obvious place to start is the dictionary. In the Electronic New Shorter Oxford English Dictionary the, for our purposes, relevant sense of the adjective ‘affective’ is defined in terms of the noun ‘affection’:

*This research is supported by the EC Project NECA IST-2000-28580.
The core notion here is that of emotion. This is reflected in the definition of de Rosis and Grasso (2000) of Affective Natural Language Generation:

\[ \text{NLG that relates to, arises from or deliberately influences emotions or other non-strictly rational aspects of the Hearer.} \]

De Rosis and Grasso consider ‘personality traits, emotions and highly-place values’ to be non-strictly rational aspects of the Hearer. According to them ‘Attitudes’ is the appropriate generic term.

Whereas in the definition of de Rosis and Grasso the hearer is central, Hovy (1988) takes the speaker and her or his opinions as his starting point. According to him affect is concerned with techniques which allow the speaker to slant a text in her or his favour, i.e., to introduce a certain bias into the text. To do this properly the speaker will have to ‘[…] distinguish between what the hearer is likely to find sympathetic, what he or she is likely to dislike, and what he or she is likely not to care about much’ (Hovy, 1988:58). According to Hovy, only three values are needed for this purpose: good, bad and neutral.

In linguistics, a notion closely related to that of affect has gained some currency. The notion in question is that of stance. The term ‘stance’ is defined from the speaker’s perspective:

\[ \text{In addition to communicating propositional content, speakers and writers commonly express personal feelings, attitudes, value judgements, or assessments; that is, they express a ‘stance’. (Biber et al., 1999:966)} \]

This characterization of stance in terms of personal feelings, attitudes, etc. is strongly related to the notion of affect.

The strict separation between stance and propositional content which is suggested by this definition is, however, misleading. Some of the examples which Biber et al. themselves put forward illustrate that this separation cannot be maintained. Sometimes the propositional content and the stance which is expressed by an utterance are the same, as in:

\[ \text{I’m not happy. (Biber et al., 1999:969)} \]

Here the proposition is expressed that the speaker is not happy. This proposition coincides with the communicated stance of the speaker.

In Hunston and Thompson (2000), the term evaluation is used. This edited volume consists of a number of papers on the topic. It starts with a very illuminating introduction by the editors. In this introduction, three functions of evaluation are discerned: (1) to express an opinion, (2) to construct and maintain relations and (3) to organize discourse. Examples of evaluation which illustrate these three functions are provided. All chapters of the book are preceded by a short introduction by the editors.

We have seen that various terms have been used instead of the word ‘affective’. In particular, the following terms are closely related to the term ‘affect’: appraisal, attitude/attitudinal meaning, bias, connotation, evaluation, opinion, slanting and stance.

The few definitions of affect and related terms which we have discussed give a rough impression of the scope of this bibliography. The bibliography is, however, somewhat more focussed than these definitions might suggest. In particular, the primary concern of this bibliography are studies which can be used to inform the construction of NLG algorithms which take affect into account. As background reading, Reiter & Dale (2000) provides a good overview of NLG from an engineering perspective. Readers who are interested in psycholinguistic investigations into NLG are referred to Levelt (1995).

A less strong bias of this bibliography is towards affective NLG for communicating agents, in particular, agents which communicate in a dialogue setting. Bradshaw (1997) contains a collection
of papers on software agents. More recently a volume has appeared on Embodied Conversational Agents edited by Cassell et al. (2000). A special issue of Applied Artificial Intelligence on ‘Animated Interface Agents’ appeared in 1999 (André, 1999). Gratch et al. (2002) reports on a workshop on Virtual Humans at the University of Southern California. Picard (1997) addresses the wider issue of ‘Affective Computing’. She discusses ‘[…] what it might mean for computers to recognize, express, and ‘have’ emotions, as part of efforts to make them more intelligent, friendly, and capable’. (Picard 1997:85)

References


2 Affective NLG: Systems and Computational Theories


Describes two systems which can generate dialogues between presentational agents. The first system uses a top-down planner to generate a complete script for the dialogue, which is then executed by means of the Microsoft Agents Player technology. The second system allows for two agents to provide real-time soccer commentaries (for the RoboCup league).


Describes the architecture of a multimodal believable agent. The agent has a personality and a social role. The paper contains some examples of the Affective Presentation Markup Language (APML).


Describes the Affective Presentation Markup Language (APML).


Includes description of a controlled experiment to study user ability to interpret faces of avatars pre-prepared to express specific emotions.


Describes a generation algorithm which takes as its input a case frame and the speaker’s emotional attitudes (an integer in [-5,5]) towards the events and objects which are to be described. The underlying emotion model is based on appraisal theories (e.g., Ortony et al, 1988; Lazarus, 1991). Firstly, object descriptions are generated. The algorithm selects the description whose default emotional shade (these shades are stored in the lexicon) is closest to that of the speaker’s emotional attitude towards the object. Secondly, the main verb is selected. Verbs are associated with an overall emotional connotation of the verb itself, and with the emotional shades that the verb conveys about each of its arguments. Given the input, semantically adequate verbs are assigned an emotional score by summing the distances between on the one hand the speaker’s emotions about the event and the objects involved, and on the other hand the corresponding emotional connotations associated with the verb and its arguments. The total score associated with a verb (or more precisely, verb group, since also passives are included) is based on the emotional score and the informational score. The informational score is the number of slots from the input frame that are realized by the verb. The precise formula is: Total Score(x) = α Info(x) − (1 − α) EmotScore(x). The verb with the highest score is selected (retaining information from the input is rewarded, whereas a high emotional score, that is a big distance between input and realization emotions, is penalized. α determines the weight of these two factors; it is hypothesized to correspond with the speaker’s personality). The paper contains some evaluation which yielded a statistically significant correlation between human judgements of the emotion of generated expressions and the speaker attitudes used for generation. The research is being carried out in the context of the Mission Rehearsal Exercise (MRE) virtual training environment (see, e.g., Rickel et al., 2002).


Provides a formalization of a theory of informal argumentation (Perelman & Olbrechts-Tyteca, 1969). In a plan-based language a number of argumentative patterns are encoded. The dialogues which are studied involve advice giving in order to change behaviour (in particular, on healthy
nutrition). Conflict situations are considered which arises from differences in opinion between the interlocutors.


Discusses four studies into affective interfaces. In particular, a number of problems which one encounters when carrying out such studies are highlighted.


Contains a detailed description of the PAULINE (Planning And Uttering Language In Natural Environments) system. The general question which Hovy addresses is how to generate a text from a knowledge base given a certain set of situational factors, such as the interlocutors’ factual knowledge, opinions, emotional states, etc. Hovy argues that an intermediate level of rhetorical goals is required. Situational factors cannot be related directly and in a transparent manner to generator production decisions. Hovy discerns rhetorical goals of opinion and rhetorical goals of style. He also stresses that some of these goals are different in nature from the goals that are common in the AI planning literature. The rhetorical goals are never fully satisfied and removed. Rather, during the generation process the extent to which they are being satisfied is continuously monitored (see pp. 136–139).


Describes a model of how speakers go about influencing the image that the listener has of their dialog motivation and their model of the listener. An implementation of the model which simulates everyday dialog in a restricted domain was used to carry out the study.


Paper on the PRACMA dialogue system.


Reports on work which has been carried out in the context of the Oz project at CMU. See also Loyall & Bates (1997) and Reilly (1996). Situational information is propagated during natural language generation by means of features. Features can license various types of rules (information structure, syntax, lexicon, etc). These rules can change features, block them and introduce new ones.


Describes the Hap behaviour based architecture which was developped in the Oz project at CMU. The Hap architecture is intended for agents that “perform actions and use natural language in interactive, animated real-time worlds”.


Includes description of an experiment to study the interaction between human-human relations and agent-human relations.


This paper discusses an experimental multimodal disputation system called Mr. Bengo. The system presents three agents to the user: a judge, a prosecutor and a defense attorney. The defense attorney can be controlled by the user (The user can determine which disputation moves it should make). Each agent’s face is displayed and facial expressions are used to convey the agent’s perception of the dispute (i.e., whether the dispute is going in a direction which yields a favourable or unfavourable conclusion from the agent’s point of view).


Scripted dialogue is defined as a dialogue which is performed by two or more agents on the basis of a description of that dialogue. This script specifies the actions which are performed in the course of the dialogue and their temporal ordering. On the basis of example dialogues, three categories of strategies for influencing the audience of a scripted dialogue are discerned: strategies of information distribution, association and emphasis. The research is carried out in the context of the neca project (see http://www.ai.univie.ac.at/NECA/).


Proposes a representation language for various types and levels of information, including emotion, for specifying agent behaviour.


Describes the multimodal natural language generation (MNLG) platform of the NECA project. The generator allows for variation of the generated language and gestures based on a number of situational parameters, including emotion.


Build on top of the Microsoft agents player technology. Involves a scenario in a virtual coffee shop. The user interacts with an animated agent portraying a waiter. They use Moulin and Rousseau’s (2000) approach to modelling conversation. Furthermore, for their model of emotions they draw on Ortony, Clore and Collins (1988). They use Moffat’s (1997) model of personality to bias an agent’s emotion expression. The personality model takes into account extraversion and agreeableness. Additionally they take into account the social relations between the speaker and the hearer. Some of the ideas are taken from work by Walker et al. (1997). To combine all this information they use so-called filter rules. The paper does not go into the details of natural language realization of the aforementioned information.


Includes models for negotiation behaviour in dialogues. Reports on work which has been carried out in the context of the Oz project at CMU. See also Kantrowitz (1990) and Loyall & Bates (1997).


Describes the Mission Rehearsal Exercise project. The aim is to build a training platform for the US Army. It involves a human user interacting with three autonomous virtual humans in a virtual environment. Further virtual humans are scripted characters. It incorporates the work reported in Fleischman & Hovy (2002) on emotional natural language generation.


Highlights issues in Affective NLG by discussing explanation texts in the domain of drug prescription. They warn for uncritical use of Grice’s maxims and, for instance, aggregation. They raise the issue that sometimes a speaker might not want to aggregate, in order to emphasize some piece of information.


Presents a detailed computational theory (following Brown & Levinson, 1987) for choosing the linguistic form of a speech act (in particular, the level of indirectness) on the basis of the threat of a speech act (computed on the basis of the social distance, the power that the hearer has over the speaker and the ranking of imposition of the speech act).

3 Affect in Language: Linguistic, Descriptive and Empirical Work

Chapter 12 is about ‘The grammatical marking of stance.’ (See the introduction of this bibliography for Biber et al.’s definition of stance).


Contains a dialogue analysis of a fragment of a play by Harold Pinter called ‘The Dumb Waifer’. It provides a nice illustration of how social relations are established/reflected by dialogue acts.


The paper presents a model for representing fine-grained lexical knowledge.


Contains a study on the influence of extraversion/introversion on language production. The basis is a corpus of email texts from persons for whom a categorization on the basis of Eysenck’s personality test was available.


Chapter 6: ‘Speech styles in conversation as an interactive achievement’ by M. Selting (pp. 106–132). Discusses the dynamics of the social relation between interlocutors. Examples are taken from a conversation at the dept. of social services. In particular, the choice between standard language and dialect/informal language is linked to the distance between speaker and hearer. Chapter 7: ‘Discourse control in confrontational interaction’ by J. Thomas (pp. 133–156). Describes how interlocutors can use specific devices (e.g., discourse markers) to signal the relation between them.


Table of Contents:
1. Evaluation: An Introduction (G. Thompson & S. Hunston)
2. Persuasive Rhetoric in Linguistics: A Stylistic Study of Some Features of the Language of Noam Chomsky (M. Hoey)
3. Corpus-Based Analysis of Evaluative Lexis (J. Channell)
4. Adverbial Marking of Stance in Speech and Writing (S. Conrad & D. Biber)
5. A Local Grammar of Evaluation (S. Hunston & J. Sinclair)
8. Beyond exchange: APPRAISAL Systems in English (J. Martin)


Kuno & Kaburaki aim to show that empathy has important influences on syntax. Empathy ‘is the speaker’s identification, with varying degrees (ranging from 0 to 1), with a person who participates in the event that he describes in a sentence’. (p. 628)


In chapter 5 Leech discusses “The Tact Maxim”. Chapter 6 provides “A survey of Interpersonal Rhetoric”.


A theory of emotions which has inspired various implementations.


They apply machine-learning techniques for determining whether given movie reviews are positive or negative.


‘The study of stance examines the expression of emotion, attitude, certainty and doubt in language. Although there have been many studies on stance in recent years, there is no comprehensive study of individual stance markers across a large, multi-register corpus. This study uses a multi-dimensional approach to identify 1) identifying the main patterns of stance use in English, and 2) interpreting these stance patterns. The corpus for the study is the Longman Corpus of Spoken and Written English [...]’ (Source: Linguist List Dissertation Abstract)

Provides a wealth of examples of bias in various contexts: political, courtroom, scientific, sales, etc. Of particular interest are chapters 4 and 5. Chapter 4 deals with 'Indicators of bias in argumentation'. Chapter 5 is entitled 'Biased language'. Walton argues that the use of slanted language is not necessarily bad. A negative judgement is only warranted if the slanting was inappropriate in the context of a specific type of dialogue.


