Producing Affective Language

Content Selection, Message Formulation, and Computational Modelling

Martijn Goudbeek, Nadine Braun, Charlotte Out, Emiel Krahmer

Center for Cognition and Communication

Tilburg University

Tilburg, the Netherlands

m.b.goudbeek@uvt.nl, n.braun@uvt.nl, c.out@uvt.nl, e.j.krahmer@uvt.nl

Abstract— **We introduce a project investigating how emotional states influence language production using both experimental and corpus based approaches. Here, we illustrate our project by asking whether content selection ("deciding what to say") and linguistic realization ("deciding how to say it") are affected by the emotional state of a speaker. We do this first by assessing whether disgusted speakers are more or less prone to align with their dialogue partners than amused speakers. Second, we develop a corpus of emotionally laden soccer reports that, even though they refer to the same event, will differ depending on whether the report comes from the winning or losing team. In both cases, we focus on the production and analysis of referring expressions. Our findings will be used to build an affective natural language generation system.**

Keywords-Emotion and Cognition; Language production; Referring expressions; Natural Language Generation.

I. INTRODUCTION

Spoken language conveys a lot of information about someone's emotional state. For example, angry speakers speak with a loud and high pitched voice, while sad speakers generally speak with a soft and low voice [1][2]; the words used may also vary; even though only a limited number of words can be classified as emotional [3], word use has been shown to be indicative of speaker's feelings. For instance, suicidal poets used relatively more first person singular pronouns, more words referring to death, and fewer references to other people in their poems than non-suicidal poets [4].

The effects of emotion on speech prosody and word production are well established, but the impact of emotion on other aspects of the speech production process is understudied. Our project aims to the conjecture that emotion influences the early content selection and message formulation stages of language production. In particular, we study how language production models can be interfaced with emotion models, and will test predictions made by such a combined model in a series of studies, where we zoom in on referential communication. Based on our findings, we aim to develop a computational model that is capable of generating different linguistic realizations of the same content, as a function of emotional state.

The rest of this paper is organized as follows. Section II describes the processes involved in language production. Section III describes the relationship between emotion and language. Section IV describes the research questions we aim to address in our project "Producing Affective Language". Section V introduces two current projects, one experimental, one corpus based, that address the relationship between emotion and language production. The acknowledgment and conclusion close the article.

II. LANGUAGE PRODUCTION

Speaking is a complex cognitive activity that starts with the conceptual preparation of a message and that ends in articulation. Levelt has described the emerging consensus that speech production takes place in a number of consecutive stages, each of which produces an output representation that provides the necessary information for the next stage [5]-[7]; see also the work by Griffin and Ferreira [8]. Despite different views regarding the exact division of the processes involved (e.g., compare [5] and [6]) the two main processes preceding articulation are generally assumed to be content selection or conceptualization ("deciding what to say") and message formulation or linguistic realization ("deciding how to say it"), as illustrated in Figure 1.
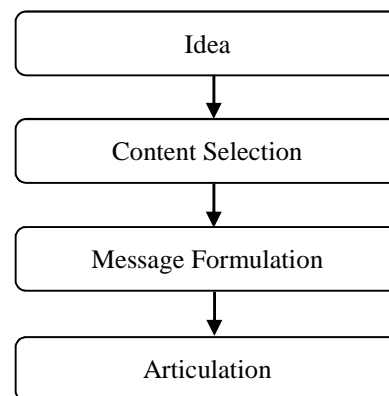

Figure 1. The stages in the speech production process.

Once a speaker has decided what to say and chosen the words to express the message, the relevant linguistic

properties of these words (e.g., their gender, number, whether they are mass or count nouns) are selected and integrated [7]. In contrast to this automatic process of selecting the properties of a word, content selection and message formulation are influenced by perspective taking [6] and dialogue factors [9]. We adjust our words depending on whether we are talking to a colleague ("I study content selection") or to a family member ("I study how we decide what to say").

These psycholinguistic insights have been adopted in the field of natural language generation [10][11], although sometimes using different terminology. Generally, the results from psycholinguistic experiments have proven useful in the development of algorithms for natural language generation. For example, Levelt's *Speaking* [5] and Pechmann's work on overspecification [12] were important influences for Dale and Reiter's [13] incremental algorithm for referring expression generation [11].

Importantly, these phases of the speech production system have traditionally been understood as modules that receive and process input in an automatic and encapsulated way. However, given its basis in affective and social processes together with the accumulating evidence that many cognitive processes are cognitively permeable, it would be surprising if the language production system were entirely isolated from other affective and cognitive processes (see also [14]).

III. EMOTION AND LANGUAGE

Beginning with Bower [15], emotional states have been linked to cognitive effects in a spreading activation approach where an individual's dominant emotional state spreads to semantic nodes related to that state, resulting in the stronger activation of conceptually related nodes [16][17]. Research in this field has provided evidence for the claim that, because they are connected to nodes that indicate a relatively safe situation, positive emotions result in global, heuristic processing [16][18], whereas negative emotions warrant more local, analytical, processing because they activate nodes that indicate a potentially dangerous situation.

In some sense, the influence of emotion on language production is obvious: when we are in a specific affective state, we express that state in words ("I'm angry"), as well as by nonverbal means (we shout and shake our fist). Almost all of this takes place at the levels of Idea and Articulation as depicted in Figure 1. However, the influence of emotion on the early stages of language production proper (content selection and message formulation) is far from trivial, and understudied as well. Our working hypothesis is that the relation between emotion and early language production is affected by specific emotional appraisals [19][20]. These are evaluations of stimuli with respect to a number of dimensions, such as novelty and pleasantness, but also attributions of agency [21], where people are either responsible for their situation and have control over the outcome (high agency) or have no control and thus no

responsibility over the situation (low agency), and the possible violation of moral and social norms [22].

While there is evidence for the influence of emotion on the articulatory stage of speech production, there is surprisingly little work on the earlier stages involving content selection and message formulation. A few notable exceptions exist, of which the work of Kempe et al. [23] is a prime example. They showed that happy speakers were less likely to specify an ambiguous word such as "bat" with a property (such as baseball bat or flying bat, when both are visible) that uniquely identified the intended referent. The authors conclude that a positive mood leads to a less effortful processing style (because a positive state signals a safe situation [16][18]) that causes speakers to spend less mental energy on perspective taking [23].

IV. RESEARCH QUESTIONS

Starting from the idea that emotional appraisals inform other cognitive systems (the "affect as information" theory, see [17]), they are expected to affect the language production process as well. The main research question of this project is if and how the emotional state of language users influences the early processes involved in language production. The emotion part of this question is understood in terms of appraisal theory, and the speech production part is understood in terms of the global version of Levelt's model presented in Figure 1. Here, we will investigate content selection and message formulation by experimentally inducing in discrete emotional states such as amusement and disgust using film fragments and by analyzing the language of emotionally charged sports reporting, comparing reports of teams that won and lost their match.

Future work will be concerned with building a computational model of emotional language production. While a handful of computational models have been developed that address personality-based [24][25] and affective [26][27] text generation, this issue remains largely unexplored. This is unfortunate, given the growing interest in this topic. This kind of application could enable individually tailored reporting with appropriate emotional "shading", which is more likely to be appealing and interesting for readers than straightforward "vanilla" reports [28].

V. CURRENT PROJECTS

In this section, two ongoing projects are described, addressing the research questions introduced above. The first project investigates whether dialogue partners align less when in a positive emotional state (as compared to a negative one). The second project involves a newly developed corpus of soccer reports that describe the same event (a soccer match) from a positive or a negative perspective.

A. Emotional state and Lexical Alignment

Previous research has shown that content selection is affected by alignment processes occurring in interaction (see [9] for a recent model). Speakers tend to select the same properties as their dialogue partners when referring to objects, even if these properties were previously dispreferred [29]. Given that emotions play a central role in our social interactions and in evolution of the speech production system [6], positive emotions will likely result in a less effortful and more egocentric processing style, while negative emotions will result in a more effortful and less egocentric processing style. In addition, emotions that result in different approach or avoidance stances (such as disgust, and amusement) are also expected to influence the extent to which speakers align with their conversation partner.

We will be testing this using a non-scripted version of the interactive reference paradigm [29] in which participants engage in a dialogue and alternatingly describe objects in a director matcher task designed to elicit alignment (as can be seen in Figure 2). First, participant A has to describe the target (1). To uniquely identify the target, she can only use the dispreferred attribute (size), e.g., "the large desk", or, when she uses color redundantly, "the large green desk". By using the property size in her description, she primes participant B to use this property as well in her description (assuming that participants will align in this task). Participant B first identifies the correct object (2) by pointing at it and then describes his object (3) that is finally identified by participant A in (4). Note that in describing his target object (3) participant B can use color ("the red sofa") or size ("the large sofa") or both ("the large red sofa"), in which case one of the properties is redundant, to distinguish the target picture from the distractors. Thus, if participant B uses size more when participant A has done so previously, that would be evidence for alignment.
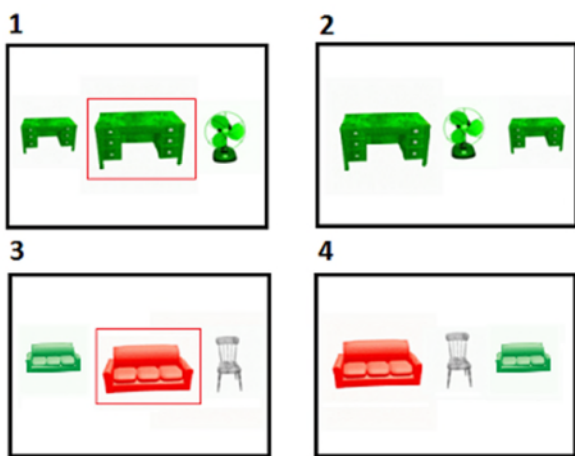

Figure 2. The four tasks that constitute a trial. In frame 1 and 3 the speaker describes the marked object, in frame 2 and 4, the listener identifies the described object.

Methodologically, the role of emotion in cognitive phenomena is often experimentally investigated by first inducing a particular emotion in participants and then asking them to perform a particular task [1]. The most used (and most effective) method has been to use validated film clips, to induce specific emotions [30][31]. In the present study, participants first view an excerpt of an amusing (e.g., the restaurant scene from "When Harry met Sally") or disgusting video (e.g., Devine eating poop in "Pink Flamingo's") and were asked to indicate their level of amusement and disgust on a seven-point scale. A preliminary manipulation check shows that participants report higher levels of amusement (*Mean* = 4.93, *Standard Deviation* = 1.34) than disgust (*Mean* = 2.50, *Standard Deviation* = 1.71) after viewing an amusing video, and, conversely, higher levels of disgust (*Mean* = 6.43, *Standard Deviation* = 1.32) than amusement (*Mean* = 2.13, *Standard Deviation* = 0.94) after viewing a disgusting video (All these differences are statistically significant; $F(1, 56) = 36.70$, $p < .001$ and $F(1, 56) = 206.89$, $p < .001$ respectively).

The proportion of attribute alignment will be used as dependent variable. As indicated, alignment at the level of content selection (i.e., deciding to say "the small chair" versus "the red chair") occurs when participant B uses the dispreferred attribute to describe the target when participant A did use the dispreferred attribute as well. If the amount of alignment indeed depends on whether the speaker is amused or disgusted, that would be evidence for the role of emotion in the conceptualization stage of speech production. Specifically, we predict that disgusted speakers, who should be less egocentric and more willing to engage in effortful processing, will align more with their partner by using the dispreferred attribute size when their partner uses the dispreferred attribute. Conversely, we predict that amused participants, who will be more egocentric and less prone to engage in effortful processing, will align less with their partner and keep using the preferred attribute color, even when their partner uses size.

B. A multilingual corpus of affective soccer reports

Sports reports open up a lot of room for creative language use, starting with the headlines of the match reports [32] and extending to almost every aspect of the report. For many biased ports reports, that is, sport reports that are written from the perspective of one of the competing teams, the point of view of the author of a match report is clearly definable from the beginning. So, it is easy to assume that the different possible outcomes of such a match would also produce different match reports in terms of language and communicated emotion (i.e., different conceptualizations and linguistic realizations). Take for example the following introductory sentences:

"AFC Wimbledon's five-match unbeaten league run came to an end in frustrating fashion tonight as Neal

Ardley's men were beaten by struggling Dagenham & Redbridge." (AFC241115, EO, LOSS, MASC, 2016)

Compared to:

"Daggers recorded a first win in 12 league games with a 1-0 success away to AFC Wimbledon" (DR241115, EO, WIN, MASC, 2016)

Both describe the exact same match and events, with totally different emotional nuances and very different emotions shining through in the texts. For Wimbledon, all the frustration is written out in the long first sentence ("frustrating fashion", "beaten", "struggling"), while the winners limit themselves to a shorter and much more positive -possibly more objective- text ("win", "success away"). These and other differences in biased sports reporting shed light on the language conceptualization and realization process that takes place when writing in an emotional state and would be especially valuable for automatic generation of natural language [33]. Indeed, Hovy [34] describes that taking into account the speaker's emotional state, rhetorical, and communicative goals, is crucial for generating suitable texts for different readers. However, the reality of automatic text generation is that not many NLG systems are able to adapt to the mood of the recipients of the produced text [28] and to convey the mood of the author.

To find out more about the language in texts produced in negative and positive emotional states, we (manually) compiled the Multilingual Affective Soccer Corpus [33]. We collected match reports from 121 different clubs participating in the first and second league in England, Germany, and the Netherlands. A first look at our data shows that reports describing wins are, on average, longer than reports describing losses or ties (777 words versus 715 or 713 words respectively). Of course, length is but one very superficial property of a text. There are likely many other textual elements, such as word choice, grammatical constructions, and pronoun use that potentially differ between biased reports.

We plan to use text analysis tools, such as LIWC [34] to, for example, help to determine the proportions of negative and positive emotion words, such as "frustrating" in example (1) or "success" in example (2). Analyzing this corpus will contribute to the understanding of how different emotional states influence and change (written) language production. We are currently planning a detailed descriptive analysis on surface features, such as already indicated text lengths and emotion words, as well as a more in-depth analysis of, for example, referential expressions and pronouns. Analyzing pronouns possibly sheds light on the focus of the author in the respective outcome of the game. If the match results in a win, does the report focus on the own team's great performance or on the opponent's failure ("us vs. them")? Additionally, we plan to investigate whether there are linguistic features that are related to the affect present in the texts – for example, whether certain grammatical constructions occur more in positive or negative contexts.

VI. CONCLUSION

In this paper, we introduced a project investigating the relationship between emotion and language production, approaching the issue from an experimental as well as a corpus based perspective. We have briefly described the process of language production and argued for the relationship between emotion and language production. To study this relationship, we focus on two aspects of referring expression generation, namely content selection ("deciding what to say") and linguistic realization ("deciding how to say it") and use appraisal theory to generate hypotheses about the influence of specific emotions on these processes. We illustrate our approach by introducing two studies, one experimental and one corpus based, that are currently being conducted in this project. These and similar studies will shed light on the relationship between linguistic and affective processes and will serve as the basis for a computational model of affective language generation.

REFERENCES

[1] K. R. Scherer, "Vocal communication of emotion : A review of research paradigms," Speech. Commun, vol. 40, pp. 227-256, 2003.

[2] J. A. Bachorowski, and M. J. Owren. "Sounds of emotion," Ann NY Acad Sci, vol. 1000, pp. 244-265, 2003.

[3] J. W. Pennebaker, M. R. Mehl, and K. G. Niederhoffer, "Psychological aspects of natural language. use: our words, our selves,". Annu. Rev. Psychol., vol. 54, pp. 547-77, 2003.

[4] S. W. Stirman and J. W. Pennebaker, "Word Use in the Poetry of Suicidal and Non-Suicidal Poets", Psychosom. Med., vol. 63, pp. 517-522, 2001.

[5] W. J. Levelt, Speaking: From Intention to Articulation, Cambridge, MA: MIT Press, 1989.

[6] W. J. Levelt, "Producing spoken language: A blueprint of the speaker", In The neurocognition of language (chapter 4), C. Brown & P. Hagoort, Eds. Oxford, UK: Oxford Press, pp. 83-122, 1999.

[7] W. J. Levelt, "Spoken word production: A theory of lexical access", P. Natl. Acad. Sci. USA, vol. 98, pp. 13464-13471, 2001.

[8] Z. M. Griffin and V. S. Ferreira, "Properties of spoken language production", In The handbook of psycholinguistics, (2nd ed.), M. J. Traxler and M. A. Gernsbacher, Eds. Amsterdam, The Netherlands: Elsevier, pp. 21-59 2006.

[9] M. Pickering and S. Garrod, "An integrated theory of language production and comprehension", Behav. Brain Sci., vol. 36, pp. 329 – 347, 2013.

[10] E. Reiter, "Has a consensus NL generation architecture appeared, and is it psycholinguistically plausible?. In Proc. of the Seventh International Workshop on Natural Language Generation, Kennebunkport, Maine: Association for Computational Linguistics, pp. 163-170, 1994.

[11] E. Krahmer, "What computational linguists can learn from psychologists (and vice versa)", Comput. Linguist., vol. 36, pp. 285-294, 2010.

[12] T. Pechmann, "Incremental speech production and referential overspecification", Linguistics, vol. 27, pp. 89-110, 1989.

[13] R. Dale and E. Reiter, "Computational interpretations of the Gricean maxims in the generation of referring expressions", Cognitive Sci., vol. 19, 233-263, 1995.

[14] G. Vigliocco and R. J. Hartsuiker, "The interplay of meaning, sound, and syntax in sentence production", Psychological Bulletin, vol. 128, pp. 442-472, 2002.

[15] G. H. Bower, "Mood and memory", American psychologist, vol. 36, pp 129-148, 1981.

[16] J. P. Forgas, "Mood and judgment: the affect infusion model (AIM)", Psychol. Bull., vol. 117, pp. 39-66, 1995.

[17] N. Schwarz and G. L. Clore, "Mood as information: 20 years later", Psychol Inq., vol. 14, pp. 296-303, 2003.

[18] B. L. Fredrickson, "The role of positive emotions in positive psychology: The broaden and build theory of positive emotions. Am. Psychol., vol. 56, pp. 218-226, 2001.

[19] N. H. Frijda, "The Emotions", Cambridge, UK: Cambridge University Press, 1986.

[20] K. R. Scherer, "The dynamic architecture of emotion: Evidence for the component process model", Cognition Emotion, vol. 23, pp. 1307-1351, 2009.

[21] C.A. Smith and P. C. Ellsworth, "Patterns of cognitive appraisal in emotion", J. Pers. Soc. Psychol., vol. 48, pp. 813-838, 1985.

[22] C. A. Hutcherson and J. J. Gross, "The moral emotions: A social–functionalist account of anger, disgust, and contempt". J. Pers. Soc. Psychol, vol. 100, pp. 719-737, 2011.

[23] V. Kempe, M. Rookes and L. Swarbrigg, "Speaker emotion can affect ambiguity production", Lang. Cognitive Proc., vol. 28, pp. 1-12, 2012.

[24] F. Mairesse and M. Walker, "PERSONAGE: Personality generation for dialogue", in Ann Meeting of the Association for Computational Linguistics, Association for Computational Linguistics, June 2007, p. 496-503.

[25] M. Walker, J. Sawyer, G. Lin and S. Wing, "Does Personality Matter? Expressive Generation for Dialogue Interaction", in Natural Interaction with Robots, Knowbots and Smartphones, vol 28, J. Mariani, S. Rosset, M. Garnier-Rizet, L Devillers, Eds. New York: Springer, pp 285-301, 2014.

[26] I. van der Sluis and C. Mellish, "Towards empirical evaluation of affective tactical NLG", in Empirical Methods in Natural Language Generation: Data-oriented Methods and Empirical Evaluation, E. Krahmer and M. Theune, Eds. Berlin: Springer, pp. 242-263, 2010.

[27] I. van Der Sluis, C. Mellish and G. Doherty, "Affective text: Generation strategies and emotion measurement issues" Proc. of the Twenty-Fourth International Florida Artificial Intelligence Research Society Conference (FLAIRS 2011), R. Charles Murray, P. M. McCarthy, Eds. The AAAI Press, May 2011, Menlo Park, California. Palm Beach, Florida., pp. 123-128.

[28] S. Mahamood and E. Reiter, E, "Generating affective natural language for parents of neonatal infants" Proc. of the 13th European Workshop on Natural Language Generation (ENLG 2011), C. Gardent and K. Striegnitz, Eds. Association for Computational Linguistics, Sept. 2011, pp. 12-21, 2011.

[29] M. Goudbeek and E. Krahmer, "Alignment in Interactive Reference Production: Content Planning, Modifier Ordering, and Referential Overspecification. Top. Cog. Sci., vol. 4, pp. 269-289, 2012.

[30] J. J. Gross and R. W. Levenson, "Emotion elicitation using films", Cognition Emotion, vol. 9, 87-108, 1995.

[31] A. Schaefer, F. Nils, X. Sanchez and P. Philippot, "Assessing the effectiveness of a large database of emotion-eliciting films: A new tool for emotion researchers", Cognition Emotion, vol. 24, pp. 1153-1172, 2010.

[32] M. K. Smith and M. B. Montgomery, "The semantics of winning and losing", Lang. Soc., vol. 18, pp. 31-57, 1989.

[33] N. Braun, M. Goudbeek and E. Krahmer, "The Multilingual Affective Soccer Corpus (MASC): Compiling a biased parallel corpus on soccer reportage in English, German and Dutch", in Proc. of the Ninth International Natural Language Generation conference (INLG 2016), Association for Computational Linguistics, Sept. 2016.

[34] E. H. Hovy, "Pragmatics and natural language generation", Artificial Intelligence, vol. 43, pp. 153-197, 1990.

[35] J. W. Pennebaker, M. E. Francis and R. J. Booth, "Linguistic inquiry and word count: LIWC 2001", Mahway: Lawrence Erlbaum Associates, 71, 2001.