Do Speaker’s Emotions influence their Language Production? Studying the Influence of Disgust and Amusement on Alignment in Interactive Reference

Charlotte Out (c.out@uvt.nl)
Tilburg center for Cognition and Communication, Tilburg University, PO Box 90153, 5000 LE, The Netherlands

Martijn Goudbeek (m.b.goudbeek@uvt.nl)
Tilburg center for Cognition and Communication, Tilburg University, PO Box 90153, 5000 LE, The Netherlands

Emiel Krahmer (e.j.krahmer@uvt.nl)
Tilburg center for Cognition and Communication, Tilburg University, PO Box 90153, 5000 LE, The Netherlands

Abstract
The influence of emotion on (the early stages of) speech production processes, notably content selection has received little scholarly attention. Goudbeek & Krahmer (2012) found evidence for alignment at the conceptual level: speakers may start using a dispreferred attribute over a preferred attribute in their referring expressions when they are primed by a pre-recorded female voice in a preceding interaction. The current study aimed to assess the role of emotion (using amusement and disgust) in alignment, while simultaneously replicating this finding in a more naturalistic setting involving two human participants in naturalistic dialogue. Our results replicate the findings by Goudbeek & Krahmer (2012), generalizing their findings to a much more naturalistic setting. In addition, we found that amused, but not disgusted speakers tend to use the preferred attribute more to describe objects to their conversational partner.

Keywords: alignment; egocentricity bias, attentional bias, emotion; amusement; disgust; speech production; referential expressions, psycholinguistics.

Introduction
Several effects of emotion on various processes in speech have been studied extensively, including effects on articulation and pronoun use. For instance, speakers often signal their emotional state in their prosody, by sobbing, crying or shouting (Bachorowski, 1999; Goudbeek & Scherer, 2010) and depressed writers have been shown to use more first person singular pronouns (Pajak & Trzebiński, 2014; Stirmar & Pennebaker, 2001). However, the impact of emotion on other aspects of the speech production processes has received little attention. In this study, we aim to investigate the role of emotion in the earlier processes in speech production, in particular on content selection stage (“deciding what to say”) of language production, focusing on referential expressions.

Emotion and speech production
As far as we know, only a few studies have looked at the relationship between emotional state and content selection. For example, Kempe, Rookes and Swarbrigg (2012) looked at the effect of speaker emotion (positive or negative emotion) on ambiguity avoidance in the production of referring expressions. In their experiment, emotion was induced by a positive or negative video, accompanied by emotion congruent classical music. After the emotion manipulation, participants were asked to uniquely describe four pictures on the sheet. In the critical trials, two of the four pictures could be described in a linguistically ambiguous way, e.g. as a “bat” which could either be a flying bat or a baseball bat. They found that speakers in a positive state were less likely to disambiguate the second linguistically ambiguous picture, that is, they were more likely to use the word ‘bat’ for both the flying bat and the baseball bat. These findings suggest that positive emotions might increase ambiguity in referring expressions, which could be the result of an attentional shift in the speaker (Beukeboom and Semin, 2006). Attentional bias It has been generally accepted that positive emotions (e.g., amusement) broaden attention, whereas negative emotions (e.g., sadness) narrows attention (see Frederickson, 2001). However, Harmon-Jones, Gable, and Price (2013) state that not valence, but the motivational intensity of emotions influence attention: emotions of low motivational intensity (e.g., sadness) broaden cognitive scope and emotions of high motivational intensity (e.g., disgust) narrow cognitive scope. They found that individuals exposed to disgusting pictures (compared to neutral pictures) who did a global-local letter task (Navon, 1977) responded slower to global than to local targets (Gable and Harmon-Jones, 2010), supporting the hypothesis that emotions of high motivational intensity narrow attention and make people focus more on details.

Egocentricity bias Egocentricity of speakers has been known to influence content selection. The egocentricity bias is the tendency of individuals to use their own perspective as reference point to the world (Ross & Sicy, 1979). Many authors claim that although individuals are often able to adjust to the perspective of the listener, they initially act egoistically (Epley, Morewedge, and Keysar, 2004; Horton and Keysar, 1996), although some beg to differ (see for example Bezuidenhout, 2013). According to Converse, Lin, Keysar and Epley (2008) and Clore and Hutsinger (2007), individuals in a positive state are less likely to adopt to the perspective of another person than individuals in a negative state, the shift of perspective to the listener might be impaired because positive emotions promote automatic responses.
Although these studies suggest an interesting link between emotion and (the early stages of) speech production, many questions remain. For example, what is the effect of emotion on the language production of speakers and listeners in an interaction?

Alignment in interaction

In the current study, we focus on the effect of emotion on the amount of alignment between conversational partners in referential expressions. As argued by Garrod and Pickering (2004), one of the ways conversational partners can align to each other is by using the same attributes to refer to an object as their conversational partner. For example, if the other person just referred to an object in terms of its size (the large table), the speaker would be more likely to use size as well in a subsequent reference, because the previous use of size would prime this attribute. However, this is at odds with another tendency that has been reported in the literature, namely that speakers prefer to use certain attributes that are more “absolute” in their meaning over attributes that are less so (e.g. color over size, Martin, 1969; Pechmann, 1989). Inspired by observations such as these, Dale and Reiter (1995) developed the Incremental Algorithm which assumes a fixed preference order of attributes to determine in what order certain attributes are used in the generation of referential expressions. This Incremental Algorithm states that when individuals describe an object, they will first use the most preferred attribute and matching value, e.g., color and then “red”, leading to the red chair. When this is not sufficient to single out the target object (e.g., there are multiple red chairs), the speaker will proceed by adding a less preferred attribute, e.g., size, leading to the large red chair. The speaker will continue adding attributes until the listener is able to identify which chair she is talking about. Dale and Reiter’s (1995) Incremental Algorithm thus predicts that speakers will never use a dispreferred attribute when a preferred attribute is sufficient for identification. However, Goudbeek and Krahmer (2012) primed speakers with dispreferred attributes (attributes that were used earlier in an interaction) – and investigated whether they would stick to their preferences or align by incorporating the dispreferred attribute that was used by their conversational partner. In their study, participants listened to a pre-recorded female voice referring to one of three furniture objects, using either a preferred (color; “the red chair”) or dispreferred (orientation, “the chair seen from the side”) attribute. They subsequently indicated which image (the target) matched this description. When they were asked to describe a new target object, they tended to use the same type of attribute that they were primed with before, even when they could also use the preferred attribute to distinguish the target. These results show that speakers may thus use dispreferred attributes over preferred ones when they are primed to do so. In this paper, we study whether the emotions amusement and disgust might inhibit or promote this tendency. In addition, we aim to replicate the findings by Goudbeek and Krahmer (2012) using a more naturalistic elicitation paradigm.

The present study

Based on the results from Goudbeek and Krahmer (2012) we predict that speakers will indeed align with their dialogue partners and start using the dispreferred attribute in their referring expressions when primed to do so. With respect to the effect of emotion, previous research (e.g., Kempe et al., 2013, Beukeboom & Semin, 2006) indicates that the emotional state of a speaker influences the content selection process of language production, and thus potentially the degree to which speakers align with respect to the attributes they use in interaction.

However, mainly the influence of emotions differing in valence (positive vs. negative) on speech production has been studied (see Kempe et al., 2012; Converse et al., 2008) which severely limits our understanding into the role of emotion in speech production. After all, emotions can be differentiated in other ways, which might influence speech production as well.

We induced amusement and disgust, two emotions that differ on multiple appraisals, among which is valence, but also approach/avoidance, potency/control and, possibly, intensity (see Scherer, 2013). Amusement is a positive emotion that occurs when a person experiences something entertaining (e.g., a joke) and feels pleasant (Tong, 2015). Disgust is an emotion that is elicited when a person is confronted with something they deem repulsive, for instance bodily fluids (vomit, pus, urine). We will study the effect of disgust and amusement on alignment in an interactive referential task. Will amused speakers or disgusted speakers align more with their conversational partners, even when they use a dispreferred way to refer to a target?

Methods

Participants

A total of 140 Dutch-speaking university students (36 male), participated in the experiment in pairs (n = 70).

Materials

Stimuli Following Goudbeek and Krahmer (2012), we used pictures taken from the TUNA corpus (van Deemter, Gatt, van der Sluis, & Power, 2012), depicting front-facing furniture items (a fan, a chair, a couch, and a desk) in four different colors (blue, green, red, and grey) and two different sizes (large or small). Participants were asked to uniquely identify the target picture (accompanied by two distractors) to their conversational partner. Previous studies (Gatt et al., 2007; Goudbeek & Krahmer, 2012) indicate the well-known preference of participants for color in their description of the target picture.

There were three types of trials: color trials, size trials and filler trials. Each participant pair was presented with 60 trials, divided into two blocks, consisting of 20 color trials
and 20 size trials. Additionally, each block included 10 filler trials, all containing large pictures of furniture in greyscale. Four versions were created, containing different orders of trials.

**Mood questions** To control for participants’ mood before the experiment, we asked the participants to rate their mood before they watched the video. They indicated on a 1 to 7 scale how much they experienced each of the following moods: happy/sad, pleasant/unpleasant, satisfied/unsatisfied, content/discontent, cheerful/sullen, in high spirits/low-spirited (Krahmer, van Dorst, & Ummelen, 2014, based on Mackie and Worth, 1991 and Bohner et al. 1992; English translations of Dutch originals).

**Manipulation check** To check whether emotion induction was successful, we asked participants after viewing the video to report how much amusement and disgust (and pride, anger, sadness, disgust, surprise and fear) they experienced on a 1 (“not at all”) to 7 (“extremely”) point Likert scale.

**Other-participant questions** After the director-matcher task, participants rated on a 7-point Likert scale (ranging from 1: “not at all” till 7: “very”) how much they liked the other participant, how empathic they felt towards them, and how much they thought they got along. Finally, they indicated if they knew the participant, choosing either “no”, “yes, a little”, or “yes, very well”.

**Videos** To counter the possibility of film specific effects, two different disgust-inducing and two different amusement videos were shown. Based on existing literature, we used four videos that were moderate to highly successful in inducing the corresponding emotions amusement and disgust, respectively. The amusement videos were “When Harry met Sally” (1989) and “There’s Something About Mary” (1998). The disgust videos were “Trainspotting” (1996) and “Pink Flamingos” (1972). We selected these videos because they have been used effectively in recent work (e.g., Hewig, Hagemann, Seifert, Gollwitzer, Naumann, and Bartussek, 2005; Fajula, Bonin-Guilaume, Jouve, and Blin, 2013; Harlé and Sanfey, 2010; Schaeffer, Nils, Sanchez, and Philippot, 2010; and Rottenberg, Ray, and Gross, 2007).

**Procedure**

After the participants had read and signed the consent form, they were sent to separate cubicles and filled in the demographics and answer the mood questions. The participants were informed that they were going to view an (emotional) video and were instructed to pay attention to the video and keep their eyes on the screen, because they would need this information in the video later in the experiment. After viewing the video (the emotion induction), they answered the questions of the manipulation check with respect to their current emotional state. Subsequently, they went into a new room together with the other participant. To enhance the emotion manipulation, participants discussed the video they watched with each other for approximately 2 minutes. They were instructed to focus on describing what they saw in the video, and telling the other participant what they thought and felt while viewing the video. They then filled in the mood questionnaire again and go on to perform the director-matcher task together.

Each trial consisted of four turns. First, participant A, the director, described the target picture (framed by a red border on the screen) to participant B, the matcher. Depending on the trial, participant A used (was forced to use) either a preferred or dispreferred attribute to describe the target picture to participant B. In the color trials, participant A could only use the preferred attribute color to distinguish the target picture from the distractors. For example, the target picture was a large blue fan, and the distractors were a large red couch and a large red fan. Therefore, participant A had to use color to describe the target picture. In the size trials, participant A could only use the dispreferred attribute size to distinguish the target picture from the distractors. For example, the target picture was a large green desk, and the distractors were a small green desk and a small green fan. Therefore, participant A had to use size to describe the target picture (see Figure 1, square 1).

Second, participant B, the matcher, saw the same pictures on their screen in a different order than participant A. After listening to the description of participant A, they indicated the matching picture by pressing the key of the corresponding number on their keyboard, e.g., “2” (see Figure 1, square 2). When participant A knew that participant B had selected an answer, she pressed “Enter” and the participants viewed a new screen.
Third, the participants switched roles: now participant B was the director and participant A the matcher. In contrast to the previous turn, the combination of pictures on this screen gave participant B the choice to use either the preferred or dispreferred attribute to distinguish the target picture from the distractors. For example, the target picture was a large red couch and the distractors were a small grey chair and a small blue desk. Participant B could either use the preferred attribute (“the red couch”) or use the dispreferred attribute (“the large couch”) to distinguish the target picture from the distractors (see Figure 1, square 3). In case participant B aligned with participant A, they used color when participant A (i.e., in color trials) used the preferred attribute, and size when participant A used the dispreferred attribute (i.e., in size trials).

Fourth, participant A, now the matcher, selected the picture that matched participant B’s description by pressing the key of the corresponding number on the keyboard, e.g., “2” (see Figure 1, square 4). When participant B knew that participant A had selected an answer, participant B pressed “Enter”, marking the end of the trial. After participant B had pressed “Enter”, a new trial appeared and the procedure was repeated. Following the director-matcher task, both participants filled in the questions about the other participant. They got debriefed and received compensation (credits or money).

**Results**

**Manipulation check**

First, we tested whether the emotion manipulation was effective. We performed a one-way analysis of variance with Emotion Video (Amusement vs. Disgust) as independent variable and Emotion Scale (amusement vs. disgust) as dependent variable. As expected, we found a significant effect of Emotion on amusement, $F(1, 138) = 88.89, p < .0001$, and disgust, $F(1, 138) = 255.47, p < .0001$. The mean scores of the combined videos per emotion indicate that participants who viewed an amusing video reported higher levels of amusement ($M = 4.89, SD = 1.38$) than disgust ($M = 2.60, SD = 1.49$). Participants who viewed a disgusting video reported a higher level of disgust ($M = 6.26, SD = 1.38$) than amusement ($M = 2.51, SD = 1.39$). This indicates that the emotion manipulation had the desired effect.

**Analyses**

To statistically evaluate the effects of emotion, prime, and attribute, we conducted an analysis of variance with the proportion of attribute use as dependent variable and Emotion (Amusement vs. Disgust) as between subject factor, and Prime (Color vs. Size) and Attribute (preferred vs. dispreferred) as within-subject factors. The results of this analysis can be found in Table 1.

A significant main effect was found for Prime, $F(1, 68) = 47.36, p < .0001, \eta^2 = .41$, indicating that the prime indeed influences the selection of attribute. Mean scores (with standard deviations) of the proportion of attribute use as influenced by prime can be found in Table 2.

A significant main effect was found for Attribute, $F(1, 68) = 33.67, p < .001, \eta^2 = .33$, confirming that the preferred property color ($M = .80, SE = .03$) is indeed preferred over size ($M = .52, SE = .03$).

**The effect of emotion** The interaction between Emotion and Attribute is significant, $F(1, 68) = 5.01, p = .028, \eta^2 = .07$. A one-way analysis of variance with Emotion and Attribute shows a significant effect of Emotion for the use of preferred attribute, $F(1,68) = 5.54, p = .022$, regardless of prime. Amused individuals showed a preference for the preferred attribute color ($M = .86, SD = .19$) over the dispreferred attribute size ($M = .48, SD = .22$). Disgusted participants did not show a preference for color ($M = .73, SD = .28$) or size ($M = .56, SD = .25$), $F(1,68) = 1.91, p = .172$. The three-way interaction between Emotion, Prime and Attribute was not significant ($F(1,68) = 3.34, p = .072$), but there was a significant interaction between Emotion and Attribute ($F(1,68) = 5.01, p = .022$). The proportions of preferred and dispreferred attributes as a function of Prime and Emotion are shown in Figure 2a (for Amusement) and Figure 2b (for Disgust).
Table 1. Summary of statistical analysis

<table>
<thead>
<tr>
<th></th>
<th>$F$</th>
<th>$p \leq$</th>
<th>$\eta^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Emotion</td>
<td>0.05</td>
<td>.376</td>
<td>.01</td>
</tr>
<tr>
<td>Prime</td>
<td>47.36</td>
<td>.001</td>
<td>.41</td>
</tr>
<tr>
<td>Attribute</td>
<td>33.67</td>
<td>.001</td>
<td>.33</td>
</tr>
<tr>
<td>Emotion x Prime</td>
<td>3.34</td>
<td>.072</td>
<td>.05</td>
</tr>
<tr>
<td>Emotion x Attribute</td>
<td>5.01</td>
<td>.028</td>
<td>.07</td>
</tr>
<tr>
<td>Prime x Attribute</td>
<td>119.89</td>
<td>.001</td>
<td>.64</td>
</tr>
<tr>
<td>Emotion x Prime x Attribute</td>
<td>0.31</td>
<td>.580</td>
<td>.01</td>
</tr>
</tbody>
</table>

Table 2. Proportions of preferred and dispreferred attributes per Prime (color vs. size)

<table>
<thead>
<tr>
<th>Prime</th>
<th>Attribute</th>
<th>$M$</th>
<th>$SD$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Color</td>
<td>Preferred</td>
<td>0.85</td>
<td>0.21</td>
</tr>
<tr>
<td></td>
<td>Dispreferred</td>
<td>0.31</td>
<td>0.28</td>
</tr>
<tr>
<td>Size</td>
<td>Preferred</td>
<td>0.75</td>
<td>0.31</td>
</tr>
<tr>
<td></td>
<td>Dispreferred</td>
<td>0.72</td>
<td>0.29</td>
</tr>
</tbody>
</table>

Discussion

The aim of this study was twofold. One, investigating the effect of emotion on alignment in interactive reference production. Two, providing a more naturalistic replication of the results by Goudbeek and Krahmer (2012) by investigating alignment on dispreferred properties in a truly naturalistic version of the interactive alignment paradigm.

Regarding the effect of emotion on attribute use, we found that amused speakers have a stronger preference for the preferred attribute (color) over the dispreferred attribute (size) than disgusted speakers. This finding can be explained by the theory that individuals in a positive state tend to process information more shallow and global than individuals in a negative state (e.g., Beukeboom & Semin, 2006) therefore using the preferred attribute because it is the default.

Regarding the effect of emotion on alignment, we did not find a statistically significant interaction between emotion, prime and attribute. However, upon inspecting our data, we did observe some interesting trends. We found that amused speakers aligned with their conversational partner regardless of prime. In other words, amused speakers aligned when they were primed with color and when they were primed with size (Figure 2a). This is in line with Harmon-Jones et al. (2013): amusement, an emotion of low emotional intensity, broadens the cognitive scope and therefore, speakers align with their conversational partners, regardless of prime. However, our amused speakers still used the preferred attribute color more (Figure 2a), supporting Clore and Hutsinger (2007), who state that speakers in a positive state find it harder to shift to the perspective of their conversational partner.

Disgusted speakers aligned when they were primed with color and when they were primed with size as well (Figure 2b). However, disgusted speakers primed with size (opposed to the amused speakers primed with size) use the dispreferred attribute more than the preferred attribute, indicating that disgusted speakers have an even stronger tendency to align than amused individuals. This might be explained by the theory that individuals in a negative state have a narrower scope of attention (Beukeboom & Semin, 2006). A narrow scope of attention might cause disgusted speakers to focus more on the words of their conversational partner than amused speakers who have a broad focus. The increased attention for the conversational partner in turn results in more alignment, regardless of prime (color or size). However, the result that disgusted speakers align more when they are primed with size than their amused peers can also be explained by the egocentricity bias (see Kempe et al., 2012). If amused speakers are more egocentric than disgusted speakers, they will rely more on their own perspective, using the preferred attribute color more, regardless of prime. This might not be the case for disgusted speakers, who are less self-focused and therefore align with their conversational partner, even when the prime was a dispreferred attribute.

The results of this study are perfectly in line with those of Goudbeek and Krahmer (2012). First, we found that participants generally used the preferred attribute color over the dispreferred attribute size. Second, participants primed with size used the dispreferred attribute size more than when they were primed with color. This is an interesting result, because the paradigm used in the study by Goudbeek & Krahmer (2012) was much more artificial than the one in our current study. In their experiment, speakers interacted with a computer and were primed by a pre-recorded computerized female voice. In our study, two human participants interacted in pairs in a relatively natural setting: they were asked to interact normally, without restrictions. The participants who primed did this naturally and unconsciously, by being only able to use the preferred or dispreferred attribute to describe the target picture.

The preliminary evidence for the differential effects of amusement and disgust on attribute choice in referential expressions should lead to further explorations of the effect of (various) emotions on language production in human interactions. Future studies could, for example, focus on finding the (inter- and intrapersonal) mechanisms that might underlie the preference of emotional speakers to use either preferred over dispreferred attributes. In our study, amused speakers preferred color much more than disgusted speakers did, which implies that the emotional state of a speaker influences her attribute preferences. These and similar studies should result in a more detailed picture of the underlying mechanisms of the language production of emotional speakers.

Acknowledgments

The research reported in this article is part of the project “Producing Affective Language: Content Selection, Message Formulation and Computational Modelling” funded by the Netherlands Organization for Scientific
Research (NWO via Grant PR-14-87). We thank Nadine Braun for her support and insights.

References


