Crosslinguistic priming in interactive reference: Evidence for conceptual alignment in speech production

Anne Vallinghs, Martijn Goudbeek, Emiel Krahmer
Tilburg center for Cognition and Communication (TiCC), School of Humanities, Tilburg University, The Netherlands
a.j.p.m.vullinghs@uvt.nl, m.b.goudbeek@uvt.nl, e.j.krahmer@uvt.nl

Abstract
Bilingual participants offer a unique opportunity to study how concepts and their relations are stored in the brain. We focus on this in the context of the spoken production of referring expressions. Previous work on reference in interaction showed that speakers adapt their descriptions to those produced earlier in the interaction. For example, when participants hear a description that contains a dispreferred attribute (such as orientation in “the chair seen from the side”), they were more likely to use that attribute in future references. The adaptation found here is not claimed to take place at the conceptual level (i.e., participants adapt to the attribute “orientation” and not to the value “seen from the side”). However, so far, convincing evidence of this claim has been lacking, because it has proven difficult to rule out adaptation at the lexical or syntactic level. A crosslinguistic study was set up to provide evidence for conceptual adaptation. In our study, Spanish/Dutch bilinguals listened to Spanish descriptions that all used postmodifiers and referred in Dutch using premodifiers. The results showed that, even without syntactic or lexical cues, speakers adapted their Dutch descriptions with the (Spanish) attributes they had listened to, providing evidence for adaptation at the conceptual level.

Index Terms: speech production, referring expressions, crosslinguistic alignment

1. Introduction
During conversations, speakers are known to adapt their speech to their addressee [1, 2, 3]. It has been shown, for instance, that dialogue participants are likely to converge on a global pitch range (e.g., [4, 5]), and the same has been argued for speech intensity (e.g., [6]) and speech rate (e.g., [7]). Of course, the speaking process involves more than auditory speech production alone. In Levelt’s Blueprint of the Speaker [8, 9], auditory speech is the result of a complex process also involving Conceptualization (in which the speaker decides what to say) and Formulation (in which the speaker decides on the grammatical and lexical encoding of the conceptualized message). One may wonder whether adaptation also occurs on these levels, and it has indeed repeatedly been shown that adaptation during dialogue may also occur at the level of the Formulator.

For example, in one seminal study by Levelt and Kelter ([10], Experiment 3), the researchers called shops in Nijmegen and asked them (in Dutch) either “At what time does your shop close?” or “What time does your shop close?”. They found that whether shopkeepers included a proposition (e.g., “At five o’clock”) or not (“Five o’clock”) was influenced by the syntactic structure of the question. This “correspondence effect”, as the authors call it, offers clear evidence for syntactic adaptation. A similar phenomenon has been observed at the lexical level. Branigan, Pickering, Pearson, and McLean [11] for example, found that if one dialogue participant refers to a couch as a “sofa”, the next speaker is more likely to use the word “sofa” as well (even though couch is generally more preferred), offering evidence for lexical alignment. Brennan and Clark [12], to give a last example, showed that conversation partners in a referential task may form a “conceptual pact” on how to refer to an object. For example, they may agree to refer to a tangram figure that resembles a skater as “the ice skater”, with both conversation partners eventually using the same lexical description when repeatedly referring to this object.

While the phenomenon of lexical and syntactic adaptation is thus firmly established, convincing evidence for adaptation on the conceptual level has mostly been lacking. One important reason for this is that it is difficult to show that adaptation occurs at the conceptual level, in such a way that adaptation at the syntactic and lexical level can be ruled out. In this paper, we make a case for conceptual adaptation during reference production. To rule out adaptation at other levels of speech production, we study Spanish/Dutch bilinguals in an interactive reference understanding and production experiment, where participants hear Spanish descriptions and produce Dutch ones. Since the descriptions participants hear and produce differ both in terms of their syntactic structures and lexical items, any effects of adaptation can only be attributed to the conceptualizer.

2. Background
When people communicate, they frequently refer to objects or persons, thus allowing them to link their speech to their surroundings. They may achieve this, for instance, by producing definite descriptions such as “the brown desk” or “the man with the glasses”. Given their ubiquity in human communication, it is hardly surprising that the production of referring expressions has received considerable scholarly attention, both from an experimental, psycholinguistic perspective and a computational linguistics one. A central question in this kind of research is how human speakers decide upon the contents of their referring expressions: how do speakers determine which attributes of an object or person to include in a description? [13].

Computational studies of reference often approach the production of descriptions as a choice problem, where a domain-dependent preference order determines which attributes (e.g., color, size or orientation) are preferably chosen for inclusion in a description. Arguably, the best known of these models is the Incremental Algorithm [14], which indeed assumes the existence of a complete preference order. This algorithm first tries
out preferred attributes (such as color) before less preferred attributes (such as size or orientation) are considered. Whenever a target object (for which a description is generated) has a different value for an attribute than some of the other objects in a visual scene (the distractors), e.g., the target is blue, and at least one distractor has a different color, it is included in the description under construction. In this way, the Incremental Algorithm models the earlier findings from, for instance, Pechmann [15], that show that speakers prefer absolute attributes such as color in their object descriptions (“the blue chair”) over relative attributes such as size (“the large chair”) and even include them when a less preferred property would have been sufficient.

Clearly, the Incremental Algorithm does not take descriptions that were produced earlier in an interaction into account, and in that sense one could argue that the algorithm is “addresser-blind” [16]. Note that the same can be said for most other referring expression generation algorithms discussed in the recent survey of Krahmer and van Deemter [17]. However, Goudbeek and Krahmer [18, 19] have shown that when speakers produce descriptions in an interactive setting, they indeed do not solely rely on preferences. In particular, they found that when prior descriptions in an interaction used dispreferred attributes to refer to an object, participants were later on more likely to rely on that dispreferred attribute as well. For example, in a furniture domain, where participants could always uniquely identify an object by using its color (e.g., “the blue chair”) or its orientation (“the chair seen from the front”), Goudbeek and Krahmer found that when participants had heard descriptions containing the attribute orientation, they were more likely to use that attribute themselves later on, even though we know from earlier studies that without a preceding interaction speakers strongly prefer color over orientation [20, 21].

Goudbeek and Krahmer suggest that their findings are evidence for conceptual adaptation: participants were confronted (or primed) with descriptions such as “the chair seen from the front”, while the target could be, for instance, a couch facing left. Even though in their experiments, the prime and the target did thus not have the same values for the relevant attributes (and prime and target were always separated by a pair of unrelated fillers), it is conceivable that attribute values occurring elsewhere in the experiment in one way or another managed to lexically prime values over longer distances. In addition, given that primes and targets were always produced in the same language (Dutch), it cannot be ruled out that some amount of syntactic adaptation might partly account for their findings as well. This means that, as it stands, the suggestion from Goudbeek and Krahmer concerning conceptual adaptation should be considered as nothing more than a conjecture.

In this paper, however, we present more direct evidence for this conjecture, using a crosslinguistic version of the interactive reference production paradigm, inspired by earlier crosslinguistic priming experiments [22, 23, 24]. In this experiment, Spanish/Dutch bilingual participants are exposed to primes in Spanish (e.g., “la silla roja; English (literally): “the chair red”), and subsequently have to describe (after two filler items) a target in Dutch (e.g., “de grote bank”; English: “the large couch”). Given that Spanish and Dutch use different words, lexical adaptation is unlikely to occur (and in any case no attribute-value pairs are literally repeated). In addition, since Spanish and Dutch use different syntactic structures, with Spanish descriptions containing post-modifiers (adjectives after the head noun) and Dutch descriptions containing pre-modifiers, syntactic adaptation is similarly unlike to arise. Nevertheless, in this paper we show that the same kind of adaptation effects arise in this crosslinguistic priming experiment as did in the monolingual priming experiments of Goudbeek and Krahmer [18, 19] suggesting that conceptual adaptation indeed occurs.

3. Method

Spanish/Dutch bilinguals alternatively interpret and produce expressions referring to furniture items and people using the paradigm proposed by Goudbeek & Krahmer [18, 19]; the references that they hear are produced in Spanish and contain either preferred (color) or less preferred (size) attributes, and we measure the frequency with which participants use these attributes themselves in the Dutch descriptions that they produce.

3.1. Participants

26 people (of which 23 were female) between 19 and 58 years old participated in this experiment, part of them in exchange for course credit, others for a remuneration of 5 Euro. This distinction is made because of the multiple locations the participants were recruited. The bulk of the participants was recruited at an institute for secondary teacher training (Spanish) in Tilburg, the Netherlands, others were Spanish speaking students from the participant pool of Tilburg University and the remainder were Spanish students minoring in Dutch recruited at the Universidad Complutense de Madrid in Spain. All were Spanish/Dutch bilinguals, including native Spanish speakers with Dutch as second language and native Dutch speakers with formal instruction in Spanish for 4 years or more. All had normal hearing and normal (or corrected to normal) vision. Two participants were not included in the analysis due to their insufficient understanding of the task.

3.2. Materials

The stimuli for this experiment consisted of a set of furniture items (the furniture domain) and a set of black and white images of mathematicians (the people domain), which have frequently been used in previous research on the production of referring expressions (e.g., [20]) and which were also used in Goudbeek and Krahmer [18, 19]. The primes and targets were always one of five furniture items (chair, desk, fan, sofa, television), varying in both color and size. An overview of the attributes and possible values is provided in Table 1. Each target could at all times be distinguished by either color, size, or both. The resulting description, therefore, depended solely on the participants’ choice. The fillers were included to blur the link between prime and target and to distract attention from the goal of the experiment. The experiment contained 80 trials, of which 56 were critical ones and 24 fillers. Of the critical trials, 28 had the color attribute as prime (14 red, 14 blue) and 28 had the size attribute as prime (14 large, 14 small).

<table>
<thead>
<tr>
<th>Attribute</th>
<th>Possible value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Type</td>
<td>chair, desk, fan, sofa, television</td>
</tr>
<tr>
<td>Color</td>
<td>blue, red</td>
</tr>
<tr>
<td>Size</td>
<td>large, small</td>
</tr>
</tbody>
</table>

3.3. Procedure

The participants took part in an interactive understanding and referring-task, where they were shown items from the furniture
Figure 1: The experimental paradigm, from upper left to lower right. Participants first identify a description (that uses size or color), then describe a filler picture, then identify a filler picture, and finally describe the target picture, using size or color (or both).

Figure 2 displays the proportion of use of the attributes color and size for the two priming conditions (color and size). When participants heard a speaker use size in their Spanish descriptions, their Dutch descriptions contained more size attributes ($M = .54, SD = .39$) than when they were primed with color ($M = .40, SD = .36$). Similarly, when our participants were primed with color in Spanish, they used color more ($M = .85, SD = .26$) than when they were primed with size ($M = .72, SD = .37$) in their referring expressions. The figure also shows that the attribute color is preferred over size, which is in line with previous research [15, 21, 20].

To investigate these effects, we conducted a mixed analysis of variance with prime (color, size) and response (color, size) as categorical within variables and the proportion of attribute use as dependent variable. To control for possible order effects, we added the order in which the participants were primed (size first, color first) as (categorical) between participants variable. The results confirm the observation that color is in general preferred over size with a significant main effect of response ($F_{1,22} = 14.63, p < .001, \eta^2 = .40$), indicating that our speakers used color more than size. Importantly, while there was no significant main effect of prime per se: $F_{1,22} = 0.46, p = .83, \eta^2 = .002$, there was a significant interaction between prime and
response \( (F[1,22] = 4.98, p = .036, \eta^2 = 0.18) \) showing that prime indeed affected the extent to which an attribute was used.

We also found a significant interaction between order and response \( (F[1,22] = 9.39, p = .006, \eta^2 = 0.30) \), showing that when participants first heard descriptions containing color, they tended to stick with color in the second part of the experiment as well (see figure 3a), thus not adapting to the attributes used in the descriptions in the primes. In contrast, when they were first confronted with descriptions containing size they not only adapted to these descriptions but they also did change their descriptions in the second part of the experiment (Figure 3b).

5. Discussion and Conclusion

This experiment investigated whether speakers adapt to their dialogue partners at the conceptual level. Spanish/Dutch bilinguals listened to Spanish descriptions of furniture items using either color or size and described these items in Dutch. Crucially, the Spanish descriptions always used a postmodifier while the Dutch descriptions always used a premodifier so that there could be no adaptation at the syntactic level. Because Spanish and Dutch use different lexical realization for the types and attributes used, lexical alignment is excluded as well. In addition, the Dutch targets that followed the Spanish prime always contained another type and another attribute - value pair (e.g., “la silla grande” (the big chair) was followed by “de kleine televisie” (the small television)). Nevertheless, the results showed that participants adapted to the attributes used in the prime descriptions. When participants heard descriptions containing size, they were more likely to use size attributes in their referring expressions. Similarly, when participants were confronted with color descriptions, they were more likely to use color in their referring expressions.

Our participants listened to color and size descriptions in two blocks; they either first heard descriptions containing size or first heard descriptions containing color. There was a strong effect of the order in which participants were confronted with the attributes color or size. When they first identified objects based on descriptions that used size, they were more likely to use size in their own descriptions when primed to do so, than when they first identified objects based on the preferred property color. It appears that participants that start with a highly preferred property such as color stick to their referential strategy more than participants that start with a less preferred property.

In sum, speakers appear to be sensitive to the way an object is conceptualized (in terms of size rather than color, for instance) rather than to a specific property (such as being small or being red). These findings extend the results of studies that show adaptation at the lexical [11, 12] and syntactic level [10] and show that adaptation can also take place at the conceptual level of Levelt’s [8, 9] model of the speaker. Dialogue partners not only adapt to each other in terms of the words and constructions they use, but also in terms of how they look at things.

Our results can be also be understood in terms of Pickering and Garrod’s Interactive Alignment model [2, 1]. In addition to alignment at the phonological, lexical, and syntactic level, this model allows for adaptation at the level of semantic representations (i.e., “what dialogue partners want to talk about”, our conceptual level). This model explicitly allows for interactions between the various levels of adaptation, which is why our previous results could also be explained by adaptation at the syntactic or lexical level. However, because the design of our current experiment explicitly prevents alignment at other levels, our results, to the best of our knowledge, provide the first experimental support for adaptation at the conceptual level in reference production.

6. Acknowledgements

The research reported in this paper forms part of the VICI project “Bridging the gap between psycholinguistics and Computational linguistics: the case of referring expressions”, funded by the Netherlands Organization for Scientific Research (NWO grant 277-70-007). We thank Ingrid Masson for lending her voice for the stimulus material and Lieke van Maastricht for her assistance in running the experiments in Madrid.
7. References


