BRIEF REPORT

Interface of linguistic and visual information during audience design

Kumiko Fukumura

School of Psychological Sciences and Health

University of Strathclyde

Graham Hills Building

40 George Street

Glasgow G1 1QE

kumiko.fukumura@strath.ac.uk

Tel: +44 141 548 2696

Abstract

Evidence suggests that speakers can take account of the addressee's needs when referring. However, what representations drive the speaker's *audience design* has been less clear. The current study thus aims to go beyond previous studies by investigating the interplay between the visual and linguistic context during audience design. Speakers repeated subordinate descriptions (e.g. *firefighter*) given in the prior linguistic context less and used basic-level descriptions (e.g., *man*) more when the addressee did not hear the linguistic context than s/he did. But crucially, this effect happened only when the referent lacked the visual attributes associated with the expressions (e.g., the referent was in plain clothes rather than in a firefighter uniform), so there was no other contextual cue available for the identification of the referent. This suggested that speakers flexibly use different contextual cues to help their addressee map the referring expression onto the intended referent. Additionally, speakers used fewer pronouns when the addressee did not hear the linguistic antecedent than s/he did. This suggests that although speakers may be egocentric when taking the referent's accessibility into account during anaphoric reference (Fukumura & Van Gompel, 2012), they can avoid pronouns when the linguistic antecedents were not shared with their addressee during initial reference.

Key words: audience design, referential communication, language production, pronouns

1. Introduction

When referring, the same entity can be described with different degrees of specificity. A horse can be described with a basic-level description like *horse,* a more specific, subordinate-level description like *stallion* (Brennan & Clark, 1996; Levelt, 1989) or semantically reduced pronouns such as *he* or *it* (Ariel, 1990; Gundel, Hedberg, & Zacharski, 1993). How do speakers decide which referring expression to use? For successful reference, the intended referent must be easily identifiable to the addressee (Grice, 1975). Therefore, it makes sense to assume that the choice of referring expressions is driven by the speaker's communicative effort to help the addressee's comprehension or *audience design* (Clark & Murphy, 1982).

Indeed, research suggests that speakers can avoid ambiguous descriptions by adopting the addressee's visual perspective (e.g., Horton & Keysar, 1996; Nadig & Sedivy, 2002). Also, speakers can take account of the addressee's expert knowledge (e.g., Isaacs & Clark, 1987; but see Jucks, Becker, & Bromme, 2008) or their prior referential exchange with their addressee (e.g., Brennan & Clark, 1998; Galati & Brennan, 2010; Gorman, Gegg-Harrison, Marsh, & Tanenhaus, 2013; Heller, Gorman, & Tanenhaus, 2012; Horton & Gerrig, 2002; 2005; Wu & Keysar, 2007) when choosing different descriptions. The current study aims to go beyond these studies by focusing on the interplay between different representations during the speaker's audience design.

According to theories of language production, production processes commence with the activation of the meaning of a to-be-produced word (e.g., Dell & O'Seaghdha, 1992; Levelt, Roelofs, & Meyer, 1999). But given that an entity can be expressed by different words, how do speakers select a particular word meaning? Although different production models make different assumptions about semantic representations (see Griffin & Ferreira, 2006, for a review), in Fukumura, Hyönä, and Scholfield (2013), we discussed how a particular word meaning (or *lexical concept*, e.g., Levelt, et al.) can be identified on the basis of attributes or properties of the referent's non-linguistic representation. For instance, for an entity to be identified as a *firefighter*, speakers must identify the referent's attributes that can be associated with the lexical concept FIREFIGHTER (e.g., *being a firefighter, being in a firefighter uniform*). To understand how speakers choose a particular lexical concept, we thus need to understand how speakers identify the relevant attributes

of the referent. To this end, consider the diagram in Fig.1, which illustrates how the linguistic as well as visual context can influence the speaker's conceptual representation.

In this model, speakers may refer to a person as FIREFIGHTER if they believe the person is a firefighter, because he was identified as FIREFIGHTER in the preceding linguistic context. But speakers may also call someone a FIREFIGHTER even if they do not know whether that person really *is* a firefighter, that is, if the person has visible attributes associated with the role, for instance, wearing a firefighter uniform. A critical question is how speakers take account of these different contextual cues to assist their addressee.

------------------------------Insert Figure 1 about here ------------------------------

According to some models of audience design (e.g., Brennan & Clark, 1996; Clark & Wilkes-Gibbs, 1986; Wilkes-Gibbs & Clark, 1992), speakers choose descriptions by conforming to conceptualizations established with their partner in previous conversations. Crucially, this effect, known as *lexical entrainment* (Garrod & Anderson, 1987), appears to be *person-specific* – when speaking to a new addressee, speakers tend to modify or abandon those descriptions used with the previous addressee, possibly to seek a new *conceptual pact* (Brennan & Clark, 1996). Although these studies typically examined references to abstract figures (e.g. Clark & Wilkes-Gibbs, 1986; Gorman et al., 2013; Wu & Keysar, 2007), evidence for conceptual pacts has been found even for common objects (Brennan & Clark; 1996; Van der Wege, 2009). In Brennan and Clark, speakers avoided subordinate descriptions (*loafer* rather than *shoe*) for a new addressee, even though the descriptions would have identified the referent without prior mention. We may therefore wonder whether such effect arises because speakers follow simple heuristics: they repeat a particular expression (e.g., FIREFIGHTER) more when the addressee also heard its prior mention than when she or he did not, irrespective of whether the visual context provides cues for identification (or the referent is in a firefighter uniform or in plain clothes).

However, speakers may not simply rely upon shared linguistic experience during audience design; they can also take account of other information that they share with their addressee (Clark & Marshall, 1981). In our example, speakers can take account of the referent's attributes in the *linguistic as well as visual* context to facilitate the word-referent mapping for their addressee. If the referent is in a firefighter

uniform, speakers may frequently say FIREFIGHTER, regardless of whether the addressee has also heard prior mention of the word, because the person's visual feature, *being in a firefighter-uniform*, makes him the plausible referent (Perner, Mauer, & Hildenbrand, 2011). However, when the referent has no such visible attributes, the availability of the linguistic antecedent to the addressee may strongly influence the speaker's repetition of the antecedent. If the referent lacks the relevant visual feature *and* the addressee does not know the referent's identity, there will be no salient feature that links the referent's representation to lexical concept FIREFIGHTER, making it difficult for the addressee to identify the referent. Thus, the effect of the shared status of the linguistic context should be strongest when the referent lacks the visual attribute associated with the description.

Note that this assumes that speakers' descriptions are affected by audience design. But the extent to which audience design constrains their referential choice has remained controversial, because many studies have shown that speakers can be *egocentric* during referential communication (e.g., Bard & Aylett, 2005; Fukumura & Van Gompel, 2012; Horton & Keysar, 1996; Jucks et al., 2008; Rossnagel, 2004). Specifically, Fukumura and Van Gompel recently showed that speakers produce pronouns (e.g., *he* or *she*) more frequently when the referent is accessible to them. Having read aloud a description of a picture of toy characters to their addressee (1), the speaker heard a sentence, which referred to the subject (2a) or prepositional object (2b) in the first sentence. Importantly, this second sentence was also heard by the addressee or it was only heard by the speaker. The speaker then described another picture showing an action carried out by the character introduced in the prepositional object position in the first sentence to their addressee (3). The addressee then acted out the speaker's description using the toy characters.

1. The mermaid is waiting for a taxi with the admiral.

2a. She is sitting on a bench.

2b. He is sitting in a wheelchair.

3. e.g., He / the admiral stands up.

Fukumura and Van Gompel found that speakers produced more pronouns (e.g., *he*) and thus fewer repeated descriptions (*the admiral*) in (3) when the second sentence mentioned the referent (2b) than the other

referential candidate (2a), *regardless of* whether the addressee also heard this sentence. That is, speakers choose pronouns when the referent is accessible to them, without considering how accessible it is to their addressee. Such a finding is, however, inconsistent with the widely-held assumption that speakers take the addressee's discourse model into account when choosing referring expressions (Ariel, 1990; Chafe, 1976; Givón, 1983; Grosz, Joshi, & Weinstein, 1995; Gundel et al., 1993); speakers should use reduced referring expressions like pronouns more when the referent is highly accessible to their addressee, whereas more explicit expressions such as definite descriptions should be favoured when the referent is less accessible to their addressee.

The second goal of the current study is thus to investigate if pronoun use is affected by the addressee's needs under different circumstances. Although speakers may be insensitive to the accessibility of the referent for their addressee, they may be sensitive to how *identifiable* the referent is to their addressee. Filik, Sanford, and Leuthold (2008) found that pronouns without immediate linguistic antecedents quickly disrupt comprehension processes, which led them to argue that the semantic interpretation of third person singular pronouns, like *he* or *she*, immediately requires a linguistic antecedent. Although corpus analyses have identified specific circumstances under which pronouns could be felicitous without linguistic antecedents (Gerrig, Horton, & Stent, 2011; Gundel, Hedberg, & Zachariski, 2005), it remains unclear whether speakers reduce pronoun use when there is no linguistic antecedent for their addressee. Do speakers avoid such pronouns for their addressee?

2. Method

2.1. Participants

Twenty four pairs of naïve participants were recruited from the University of Strathclyde student community in exchange of course credit or money. They were all reported to be native speakers of British English, to be aged less than 30, and to have no reading difficulty.

2. 2. Materials and procedure

A pair of participants drew lots to decide who would be the speaker or addressee. They then sat side-by-side at a table, each facing a computer screen. At the beginning of each trial, both the speaker and addressee saw a photograph of two human toy characters (the top panel of Fig.2) on the computer screen. The addressee then received the toys from the experimenter and recreated the scene depicted in the photo on the table. The speaker then pressed a key, which triggered the auditory presentation of a context sentence (4).

(4) The firefighter was visiting the school.

The sentence was pre-recorded by a female native speaker of British English, and referred to one of the characters in the scene, or the *target,* using a subordinate-level expression (e.g., *firefighter*) (see Appendix). Unlike in Fukumura and Van Gompel (2012), the non-target character was never introduced in the sentence.

------------------------------Insert Figure 2 about here ------------------------------

In the *attribute-present* condition, the target had visible attributes associated with the expression given in the context sentence (e.g., wearing a firefighter uniform for *firefighter*, Fig.2A), whereas in the *attribute-absent* condition, it did not (e.g., wearing plain clothes, Fig.2B). The other character in the scene always had a different gender from the target (so a basic term like *the man* or a pronoun like *he* was unambiguous) and its physical characteristics made it an implausible candidate for the role (e.g., small child for *firefighter*).

Crucially, in the *shared* condition, the sentence was presented via loudspeakers, so that both the speaker and addressee heard the sentence, whereas in the *privileged* condition, the sentence was presented via headphones that the speaker was wearing, so only the speaker heard the sentence. The presentation of the context sentence was manipulated in blocks so that speakers were clear when the sentence was shared with their addressee (the order of blocks was counterbalanced) (Fukumura & Van Gompel, 2012).

Having listened to the sentence, the speaker pressed a button, which led to the presentation of a second photograph (e.g., the bottom panel of Fig.2), which depicted an action carried out by the target. The addressee could not see the second photograph, and the speaker was asked to describe the action, such that

the addressee could act out the descriptions using the toys. For instance, speakers typically described the example stimuli in Fig.2 as:

(5) The firefighter/the man/he raised his arm.

To ensure that participants paid attention to the auditory sentence, a multiple-choice written comprehension question (e.g., "Where was the firefighter visiting? SCHOOL   HOSPITAL") was presented after the speaker's target description. In the shared condition, both the speaker and addressee answered the question by pressing a key, whereas in the privileged condition, only the speaker responded to the question. There were 24 experimental trials as well as 34 filler trials that had varied numbers of characters in the linguistic and visual context. Additionally, there were six practice trials before the start of each block, where the speaker and the addressee swapped their tasks. The experiment lasted around 45 minutes. Each participant's speech was recorded for transcriptions and analyses carried out later.

2.3. Design

A 2 (Visual Attributes: present vs. absent) × 2 (Sharedness: shared vs. privileged) repeated measures design was used. Together with 34 filler items, 24 experimental items were randomly distributed across four lists, each containing six items from each condition, and one version of each item. Twenty-four pairs of participants were randomly assigned to one of the lists.

2.4. Scoring

Speakers' descriptions of the target character were scored as a (1) *repeated description* when the description in the context sentence was repeated ($N = 385$); (2) *basic description* when speakers used basic-level descriptions like *the guy*, *the man, the woman* ($N = 118$), including cases where participants modified basic noun phrases as in *the big guy* or *the man with the tie* ($N = 18$); and (3) *pronoun* when participants used a third-person singular pronoun (*he, she*) ($N = 46$). Responses were excluded if participants did not refer to the target as the agent/subject in their description ($N = 12$); repaired the initial choice of expressions (e.g., *The guy, the firefighter*; *The woman, the witch*) ($N = 4$); used *they* or *both* ($N = 5$), produced imperative instructions (e.g., *Raise the man's left arm*) ($N = 5$) or used a proper name (*Johnny Ramone* for a rock-star) ($N = 1$). In total, 27 trials (4.7%) were excluded.

3. Results

Fig. 3 reports the mean percentage of different descriptions. We used logit mixed effects modelling (Baayen, Davidson, & Bates, 2008; Jarger, 2008) to analyze the data as functions of Sharedness (shared vs. privileged) and Attributes (present vs. absent). Because the number of each description was dependent on alternative descriptions, we analyzed the log-odds of each description relative to others. Sharedness and Attributes were both centred so that the results could be interpreted in the same way as in traditional ANOVAs.

-----------------------------Insert Figure 3 about here ------------------------------

First, we analyzed the log-odds of producing a repeated description relative to a pronoun and a basic description. Because the maximal random effect structure, or the model containing by-participants and by-items random intercepts and slopes for Sharedness and Attributes as well as the interaction between the two, did not converge, we followed what Barr, Levy, Scheepers, and Tily (2013) called "best path" algorithm (p.265): we determined which random slope showed strongest evidence for improvement against the intercepts-only model, and then tested for inclusion of other random slopes against the model with the best random effect. We did this by adopting a liberal alpha level as suggested by Barr et al. ($p = .20$). Thus, the main analyses contained by-participants and by-items random intercepts as well as by-participants random slopes for Sharedness and Attributes and a by-item random slope for Attributes. Table 1 reports the results. Participants produced more repeated descriptions in the shared (76%) than in the privileged condition (63%). They also produced more repeated descriptions in the attribute-present (85%) than in the attribute-absent condition (53%). Crucially, there was a significant Sharedness × Attributes interaction. Simple effects (including by-participants and by-items intercepts and slopes for Sharedness) further revealed that in the attribute-present condition, there was no effect of Sharedness, whereas in the attribute-absent condition, repeated descriptions were more frequent in the shared (67%) than in the privileged condition (40%).

-----------------------------Insert Table 1 about here ------------------------------

Next, we analyzed the log-odds of basic-level descriptions relative to repeated descriptions and pronouns with the maximal random effect structure. Table 2 summarises the fixed effects. There were more

basic descriptions in the privileged (34%) than in the shared condition (10%). There were more basic descriptions in the attribute-absent (39%) than in the attribute-present condition (6%). There was no significant Sharedness × Attributes interaction.

-----------------------------Insert Table 2 about here -----------------------------

Finally, we analyzed the log-odds of pronouns relative to repeated and basic descriptions. The maximal random model did not converge, so as before, we followed the best-path algorism and included by-participants and by-items random intercepts and a by-items random slope for Sharedness. Table 3 reports the results. Participants produced more pronouns in the shared (14%) than in the privileged condition (3%). Neither Attributes nor the Sharedness × Attributes interaction significantly affected the use of pronouns, however.

-----------------------------Insert Table 3 about here -----------------------------

4. Discussion

First, speakers repeated subordinate antecedents (e.g., *firefighter*) less when the antecedents were not shared with their addressee than when they were. Also, speakers used fewer subordinate descriptions for referents without relevant visible attributes (being in plain clothes for *firefighter*) than for referents with visible attributes (being in a firefighter uniform). But novel evidence for audience design comes from the interaction between the two: the effect of shared linguistic antecedents was found only for referents without visible attributes, that is, when there were no visible attributes that could link the referent to the antecedents. When the visual context provided relevant cues for identification, speakers frequently repeated subordinate antecedents, regardless of whether the addressee also heard the antecedent. This interaction indicates that speakers use linguistic as well as visual attributes to facilitate the word-referent mapping for the addressee.

Second, speakers produced fewer pronouns and used more explicit expressions when addressees did not hear the antecedent than when they did. When the addressee did not hear the antecedent, that is, when speakers introduced the referent for the first time, they favoured more explicit expressions, presumably because they provided more identifying information about the referent than pronouns. Speakers may have also been aware that pronouns without linguistic antecedents are hard to process (Filik et al., 2008). This appears to contrast with Fukumura and Van Gompel (2012), who showed that speakers produced more

pronouns when the immediately preceding sentence re-mentioned the referent (2b), than when the sentence re-mentioned the competitor (2a), regardless of whether the addressee also heard that sentence. In that study, the antecedent was always introduced to both the speaker and the addressee (1). One possibility is that speakers are egocentric during *anaphoric reference,* because the referent's accessibility is harder to model for the addressee than its identifiability. Researchers have shown that the use of anaphoric pronouns is affected by many different discourse factors (e.g., Anderson, Garrod, & Sanford, 1983; Arnold, 2001; Ariel, 1990; Brennan, Friedman, & Pollard, 1987; Fukumura & Van Gompel, 2011, 2014; Fukumura, Van Gompel, & Pickering, 2010; Fukumura, Van Gompel, Harley, & Pickering, 2011; Givón, 1983; Karimi, Fukumura, Ferreira, & Pickering, 2014). So to choose anaphoric pronouns from the addressee's distinct discourse model, speakers need to keep track of the detailed representations of the shared discourse, whilst ignoring the representations of the non-shared discourse, which could impose high processing demands. To use the addressee's discourse model in Fukumura and Van Gompel (2012), speakers had to consider not only whether the second sentence that they heard was shared with their addressee but also which of the two characters introduced in the first sentence was re-mentioned in the second sentence, and if the addressee did not hear the second sentence, speakers had to ignore the mention of the character in the second sentence (if speakers really wanted to be accurate, they then had to remember how the referent was mentioned in the shared first sentence). In contrast, in our study, only the target was linguistically introduced, and all speakers needed to know was whether the addressee also heard the sentence that introduced the target. If the task demand in the current study had been higher – e.g., some linguistic antecedents were shared but others were not –, then speakers might not have been able to avoid pronouns when the linguistic antecedents were not shared with their addressee.

Although speakers repeated linguistic antecedents least when the addressee did not hear the preceding sentence and the referent lacked the associated visual attributes, they nevertheless frequently repeated antecedents that did not straightforwardly identify the referent (40%). The availability of the linguistic context was manipulated in blocks, so it must have been clear to speakers when the addressee did not hear the linguistic antecedent. Speakers might have simply forgotten to take the addressee's discourse model into account in some trials, but they might have also experienced egocentric interference: because

speakers themselves heard the linguistic antecedents, the antecedents were highly available to speakers, so the speaker's use of subordinate descriptions was also affected by lexical priming. Although this could be further examined by having an additional condition in which speakers do not hear any linguistic antecedent or hear a basic-level term instead, it seems highly likely that the choice of different descriptions was also modulated by speaker-internal factors in the light of previous research. For instance, Horton and Keysar (1996) showed that speakers often produce contrastive descriptions (e.g., *large circle*) even when the addressee cannot see the contrasting referential alternative (e.g., a small circle), suggesting that speakers are also affected by their own privileged contextual information. Moreover, Wu and Keysar (2007) showed that the speakers tend to *overestimate* the addressee's knowledge when choosing names and descriptions for abstract objects (see also Heller et al., 2012).

The finding that the shared status of subordinate descriptions mattered only when the referent lacked the relevant visual attributes highlights the role of visual context in audience design. But the shared status of visual attributes was not independently manipulated, so we may thus wonder about the extent to which the effect of visual context was due to audience design. Did speakers repeat subordinate antecedents more for the referent with relevant visual attributes than the referent without visual attributes because the presence of the visual attributes made the production of subordinate descriptions easier or because speakers believed subordinate descriptions identified the referent with visible attributes better for their addressee? Although both possibilities are plausible and not mutually exclusive, future research may manipulate the availability of the referent's visual properties to the addressee. If speakers can take account of the referent's attributes for their addressee, the shared visual attributes should affect the speakers' descriptions more than non-shared visual attributes, in keeping with previous research that showed that speakers adapt their descriptions depending on their addressee's visual perspective (e.g., Horton & Keysar, 1996; Lockridge & Brennan, 2002).

When the referent lacked visible attributes for the subordinate antecedents, speakers repeated the antecedents less and produced basic descriptions more. However, their pronoun use was unaffected by the referent's visual attributes, indicating that the choice of pronouns is affected by different representations from those that determine the production of subordinate antecedents. According to our model discussed

earlier, when referring to a man with a repeated description like *firefighter*, speakers take account of the conceptual fit between the referent's representation and lexical concept FIREFIGHTER. When referring to the same character with pronoun *he*, however, speakers are primarily concerned with the referent's attributes associated with lexical concept HE (e.g., *being single*, *being human*, *being male*), but not with those associated with the lexical concept for *firefighter* that the pronoun is replacing.

Many previous studies have examined speakers' audience design in situations where interlocutors take turns to refer to abstract or complex category exemplars over multiple trials. The key findings from these studies were that after repeated reference, interlocutors' descriptions are gradually attenuated, but when their conversational partner is replaced with a new addressee, speakers avoid or elaborate those brief descriptions developed with their old addressee. Although much research has examined how shared referential experience affects speakers' estimation of their partner's knowledge (e.g., Horton & Gerrig, 2002, 2005; Gorman et al., 2013; Wu & Keysar, 2007), the exact representations that support audience design have been less clear. The current study goes beyond these studies by highlighting the interface between the shared linguistic context and the referent's visual properties. Speakers in the previous studies may have avoided brief descriptions developed with their previous addressee after a partner change, because they knew that the old expressions would not easily identify the referent to their new addressee *on the basis of the visual context alone*. That is, audience design is supported by speakers' sensitivity to the mapping between the referent's visual representation and a particular lexical concept, and their awareness of how linguistic and visual context can reinforce that.

In conclusion, the current study makes two important contributions. First, we shed new light on the interface between the shared linguistic context and the referent's visual properties, demonstrating how flexibly speakers use different contextual cues to map the referring expression onto the referent for their addressee. Second, we showed that speakers avoid pronouns when their addressee did not hear the linguistic antecedents, indicating speakers can use pronouns cooperatively by taking the addressee's needs into account during initial mention.

References

Anderson, A., Garrod, S. C., & Sanford, A. J. (1983). The accessibility of pronominal antecedents as a

function of episode shifts in narrative text. *Quarterly Journal of Experimental Psychology Section a-

Human Experimental Psychology, 35*, 427-440.

Arnold, J. E. (2001). The effect of thematic roles on pronoun use and frequency of reference continuation.

Discourse Processes, 31, 137-162.

Ariel, M. (1990). *Accessing noun-phrase antecedents*. London: Routledge.

Bard, E. G., & Aylett, M. (2005). Referential form, duration, and modelling the listener in spoken

dialogue. In J. Trueswell & M. Tanenhaus (Eds.), *Approaches to studying world-situated

language use: Bridging the language-as-product and language-as-action traditions* (pp. 173-

191). Cambridge: MIT Press.

Baayen, R. H., Davidson, D. J., & Bates, D. M. (2008). Mixed-effects modelling with crossed random

effects for subjects and items. *Journal of Memory and Language, 59*, 390-412.

Barr, D. J., Levy, R., Scheepers, C., & Tily, H.J. (2013). Random effects structure for confirmatory

hypothesis testing: keep it maximal. *Journal of Memory and Language, 68*, 255-278.

Brennan, S. E., & Clark, H. H. (1996). Conceptual pacts and lexical choice in conversation. *Journal of

Experimental Psychology: Learning, Memory and Cognition, 22*, 1482-1493.

Brennan, S. E., Friedman, M. W., & Pollard, C. J. (1987). A centering approach to pronouns. In *Proceedings

of the 25th annual meeting of the association for computational linguistics* (pp. 155-162).

Cambridge, MA: Association for Computational Linguistics.

Clark, H.H., & Murphy, G.L. (1982). Audience design in meaning and reference. In J.F. LeNy & W.

Kintsch (Eds.), *Language and Comprehension* (pp. 287-299). Amsterdam: North-Holland.

Clark, H. H., & Marshall, C. R. (1981). Definite reference and mutual knowledge. In A. K. Joshi, B. Webber

& I. Sag (Eds.), *Elements of discourse understanding* (pp. 10-63). Cambridge: Cambridge University

Press.

Clark, H. H., & Wilkes-Gibbs, D. (1986). Referring as a collaborative process. *Cognition, 22*, 1-39.

Chafe, W. L. (1976). Givenness, contrastiveness, definiteness, subjects, topics, and point of view. In C. N.

Li (Ed.), *Subject and topic* (Vol. 25-56). New York: Academic Press.

Dell, G. S., & O'Seaghdha, P. G. (1992). Stages of lexical access in language production. *Cognition, 42*, 287-314.

Filik, R., Sanford, A., & Leuthold, H. (2008). Processing pronouns without antecedents: evidence from event-related brain potentials. *Journal of Cognitive Neuroscience. 20*, 1315-26.

Fukumura, K., & Van Gompel, R.P.G. (2011). The effects of animacy in the choice of referring expressions. *Language and Cognitive Processes, 26*, 1472-1504.

Fukumura, K. & Van Gompel, R.P.G. (2012). Producing pronouns and definite noun phrases: Do speakers use the addressee's discourse model? *Cognitive Science, 36,* 1289-1311.

Fukumura, K., & Van Gompel, R.P.G. (in press). Effects of order of mention and grammatical role on anaphor resolution. *Journal of Experimental Psychology: Learning, Memory and Cognition.*

Fukumura, K., Van Gompel, R.P.G., & Pickering, M. J. (2010). The use of visual context during the production of referring expressions. *Quarterly Journal of Experimental Psychology, 63, 1700–1715.*

Fukumura, K., Hyönä, J., & Scholfield, M. (2013). Gender affects semantic competition: the effect of gender in a non-gender marking language. *Journal of Experimental Psychology: Learning, Memory and Cognition, 39,* 1012-1021.

Galati, A., & Brennan, S. E. (2010). Attenuating information in spoken communication: For the speaker, or for the addressee? *Journal of Memory and Language, 62*, 35-51.

Garrod, S., & Anderson, A. (1987). Saying what you mean in dialog: a study in conceptual and semantic coordination. *Cognition*, 27, 181-218.

Gerrig, R. J., Horton, W. S., & Stent, A. (2011). Production and comprehension of unheralded pronouns: A corpus analysis. *Discourse Processes*, *48*, 161-182.

Givón, T. (1983). Topic continuity in discourse: An introduction. In T. Givón (Ed.), *Topic Continuity in Discourse: A Quantitative Cross-Language Study* (pp. 1-41). Amsterdam: John Benjamins.

Gorman, K. S., Gegg-Harrison, W., Marsh, C.R. & Tanenhaus, M.K. (2012). What's learned together stays together: Speakers' choice of referring expression reflects shared experience. *Journal of Experimental Psychology: Learning, Memory and Cognition, 39*, 843-853.

Grice, H. P. (1975). Logic and conversation. In P. Cole & J. Morgan (Eds.), *Syntax and Semantics Vol. III: Speech acts* (pp. 41-58). New York: Academic Press.

Griffin, Z. M., & Ferreira, V. S. (2006). Properties of Spoken Language Production. In M. J. Traxler & M. A. Gernsbacher (Eds.), *Handbook of Psycholinguistics* (2nd ed., pp. 21-59). London: Elsevier.

Grosz, B. J., Joshi, A. K., & Weinstein, S. (1995). Centering: A framework for modelling the local coherence of discourse. *Computational Linguistics, 21*, 203-225.

Gundel, J. K., Hedberg, N., & Zacharski, R. (1993). Cognitive status and the form of anaphoric expressions in discourse. *Language, 69*, 274-307.

Gundel, J. K., Hedberg, N., & Zacharski, R. (2005). Pronouns without explicit antecedents: How do we know when a pronoun is referential? In A. Branco, T. McEnery, & R. Mitkov (Eds.) Anaphora processing: Linguistic, cognitive and computational modelling (pp. 351–364). Amsterdam: Benjamins.

Heller, D., Gorman, K. S., & Tanenhaus, M. K. (2012). To name or to describe: Shared knowledge affects referential form. *Topics in Cognitive Science*, *4*, 290–305.

Horton, W. S., & Keysar, B. (1996). When do speakers take into account common ground? *Cognition, 59*, 91-117.

Horton, W. S., & Gerrig, R. J. (2002). Speakers' experiences and audience design: knowing when and knowing how to adjust utterances to addressees. *Journal of Memory and Language, 47*, 589-606.

Horton, W. S., & Gerrig, R. J. (2005). The impact of memory demands on audience design during language production. *Cognition, 96*, 127-142.

Isaacs, E. A., & Clark, H. H. (1987). References in conversation between experts and novices. *Journal of Experimental Psychology: General, 116*, 26-37.

Jaeger, T. F. (2008). Categorical data analysis: Away from ANOVAs (transformation or not) and towards logit mixed models. *Journal of Memory and Language*, *59*, 434-446.

Jucks, R., Becker, B. M., & Bromme, R. (2008). Lexical entrainment in written discourse: Is experts' word use adapted to the addressee? *Discourse Processes*, *45*, 497-518.

Karimi, H., Fukumura, K., Ferreira, F., & Pickering, M.J. (in press). The effect of noun phrase length on the form of referring expressions. *Memory & Cognition.*

Levelt, W. J. M. (1989). *Speaking: From intention to articulation*. Cambridge, CA: MIT Press.

Levelt, W. J. M., Roelofs, A., & Meyer, A. S. (1999). A theory of lexical access in speech production. *Behavioral and Brain Sciences, 22*, 1-38.

Lockridge, C. B., & Brennan, S. E. (2002). Addressees' needs influence speakers' early syntactic choices. *Psychonomic Bulletin & Review, 9*, 550-557.

Perner, J., Mauer, M., & Hildenbrand, M. (2011). Identity: key to children's understanding of belief. *Science*, *22*, 474-7

Rossnagel, C. S. (2004). Lost in thought: Cognitive load and the processing of addressees' feedback in verbal communication. *Experimental Psychology*, *51*, 191-200.

Van der Wege, M. M. (2009). Lexical entrainment and lexical differentiation in reference phrase choice. *Journal of Memory and Language*, *60*, 448-463.

Wilkes-Gibbs, D., & Clark, H. H. (1992). Coordinating beliefs in conversation. *Journal of Memory and Language, 31*, 183–194.

Wu, S., & Keysar, B. (2007). The effect of information overlap on communication effectiveness. *Cognitive Science, 31,* 169–181.

Appendix

The firefighter was visiting the school. IN FIREFIGHTER UNIFORM / PLAIN CLOTHES [RAISE ARM]

The doctor took a walk in the garden. IN WHITE COAT WITH STETHOSCOPE / PLAIN CLOTHES [DROP ICE CREAM FROM HAND]

The policewoman fought in the battle. IN POLICE UNIFORM / PLAIN CLOTHES [RAISE GUN]

The prince had a good look at the golf course. WEARING CROWN AND GOLD FROCK/ NO CROWN, PLAIN CLOTHES [JUMP OFF FOUNTAIN]

The bride arrived at the reception. IN WHITE WEDDING DRESS / NORMAL DRESS [PICK UP A DRINK]

The photographer was on the beach. WITH CAMERA / WITHOUT CAMERA [TAKE OFF HAT]

The astronaut trained for landing. IN ASTRONAULT SUIT / PLAIN CLOTHES [PUT DOWN FLAG]

The nurse came to the farm. IN NURSE UNIFORM / PLAIN CLOTHES [PICK UP A BASKET OF APPLES]

The pianist was checking the concert hall. SITTING AT PIANO / WITHOUT PIANO [STANDS UP]

The princess attended a party. WEARING CROWN AND DRESS WITH WIRE HOOP / NO CROWN, WEARING SIMPLE DRESS [LOWER UMBRELLA]

The cameraman walked into the pub. HAVING VIDEO CAMERA / WITHOUT VIDEO CAMERA [SIT DOWN]

The rock-star was getting drunk. HAVING ELECTRIC GUITAR / WITHOUT ELECTRIC GUITAR [GET OFF TABLE]

The queen was defending the castle. WEARING CROWN AND DRESS WITH WIRE HOOP / WEARING NO CROWN AND SIMPLE DRESS [RAISE SWORD]

The boxer returned from the match. IN BOXER COSTUME / PLAIN CLOTHES [OPEN WASHING MACHINE]

The racing driver was running away from the fire. IN RACING DRIVER COSTUME / PLAIN CLOTHES [FALL DOWN ON FACE]

The singer was invited to the football match. HAVING MICROPHOHE / WITHOUT MICROPHONE [STAND UP ON BENCH]

The captain went to the theatre. IN CAPTAIN COSTUME / PLAIN CLOTHES WITH HAT [TAKE OFF HAT]

The maid was in the courtyard. IN MAID COSTUME / PLAIN CLOTHES [HANG UP TOWEL]

The nun was cleaning the floor. IN NUN COSTUME / PLAIN CLOTHES [TURN AROUND]

The footballer enjoyed a good cup of tea. IN FOOTBALLER UNIFORM / PLAIN CLOTHES [STAND UP]

The policeman was in the library. IN POLICEMAN UNIFORM / PLAIN CLOTHES [READ BOOK]

The witch was sailing to an island. WITCH FROCK AND HAT / WITHOUT WITCH HAT AND FROCK [LOOK AT MAP]

The chef was looking for some flowers. IN CHEF UNIFORM / PLAIN CLOTHES [PICK UP FLOWERS]

The king took part in a fire drill. CROWN AND FANCY FROCK / PLAIN CLOTHES [EMPTY EXTINGUISHER]

Representation of a person

Referent

Visual Context

Linguistic Context

In a firefighter uniform

FIRE-
FIGHTER
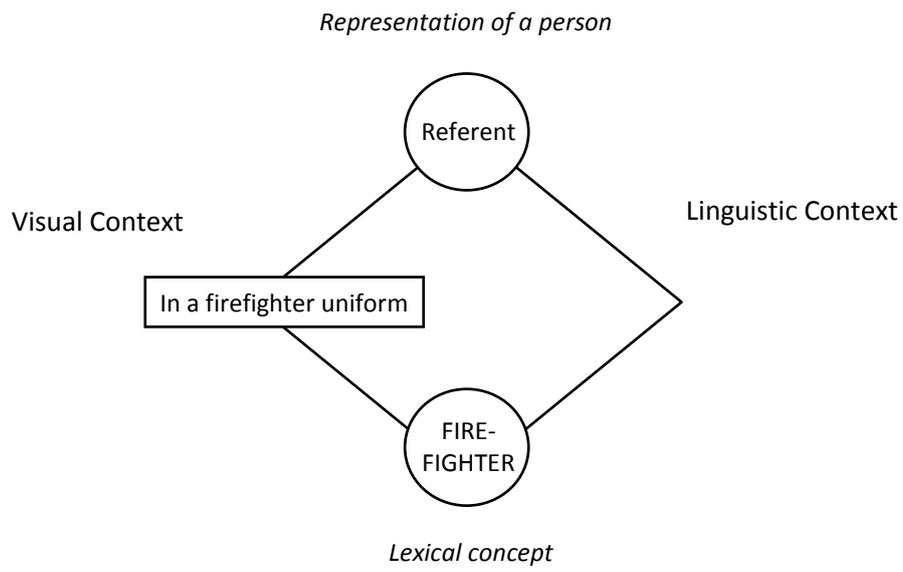
Lexical concept

Fig.1. Diagram of the linkage between the referent's representation and the lexical concept of the to-be-produced word.

(A) Attribute present   (B) Attribute absent


Fig.2. Example stimuli.
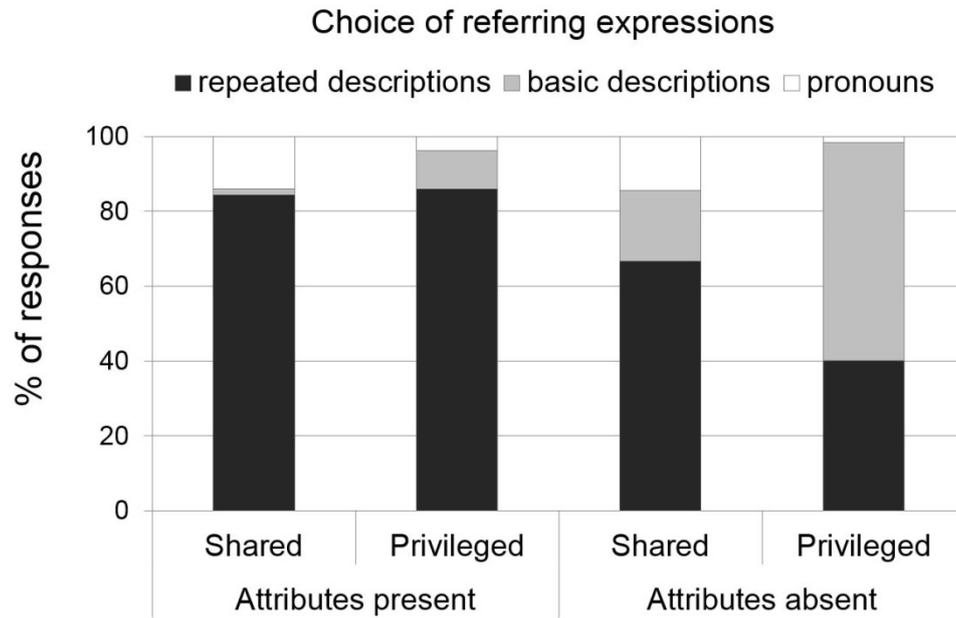
Fig.3. Mean percentages of repeated descriptions, basic descriptions and pronouns by condition.

Table 1. Summary of the fixed effects for the production of repeated descriptions

	β	*SE*	z	*p*
(Intercept)	1.55	0.40	3.83	<.001
Sharedness	-0.75	0.30	-2.50	.013
Attributes	-1.02	0.23	-4.51	<.001
Shareness × Attributes	-0.50	0.14	-3.47	<.001

Effect of Sharedness in the Attribute-Present condition

	β	*SE*	z	*p*
(Intercept)	2.77	0.55	5.07	<.001
Sharedness	-0.01	0.75	-0.02	.985

Effect of Sharedness in the Attribute-Absent condition

	β	*SE*	z	*p*
(Intercept)	1.55	0.59	2.65	.008
Sharedness	-2.46	0.72	-3.44	<.001

Table 2. Summary of the fixed effects for the production of basic descriptions

	β	*SE*	z	*p*
(Intercept)	-3.64	0.62	-5.87	<.001
Sharedness	2.19	0.46	4.80	<.001
Attributes	1.84	0.40	4.56	<.001
Sharedness × Attributes	0.12	0.35	0.34	.736

Table 3. Summary of the fixed effects for pronoun responses

	β	*SE*	z	*p*
(Intercept)	-3.38	0.47	-7.14	<.001
Sharedness	-0.90	0.25	-3.62	<.001
Attributes	-0.09	0.18	-0.49	.623
Sharedness × Attributes	0.01	0.18	0.04	.967