University of Groningen

Accessibility and Reference Production

Vogels, Jorrig; Krahmer, Emiel; Maes, Alfons

Published in:
The Oxford Handbook of Reference

DOI:
[10.1093/oxfordhb/9780199687305.013.16](https://doi.org/10.1093/oxfordhb/9780199687305.013.16)

IMPORTANT NOTE: You are advised to consult the publisher's version (publisher's PDF) if you wish to cite from it. Please check the document version below.

Document Version
Publisher's PDF, also known as Version of record

Publication date:
2019

[Link to publication in University of Groningen/UMCG research database](#)

Citation for published version (APA):

Vogels, J., Krahmer, E., & Maes, A. (2019). Accessibility and Reference Production: The Interplay Between Linguistic and Non-Linguistic Factors. In J. Gundel, & B. Abbott (Eds.), *The Oxford Handbook of Reference* Oxford University Press. <https://doi.org/10.1093/oxfordhb/9780199687305.013.16>

Copyright

Other than for strictly personal use, it is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license (like Creative Commons).

The publication may also be distributed here under the terms of Article 25fa of the Dutch Copyright Act, indicated by the "Taverne" license. More information can be found on the University of Groningen website: <https://www.rug.nl/library/open-access/self-archiving-pure/taverne-amendment>.

Take-down policy

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

Downloaded from the University of Groningen/UMCG research database (Pure): <http://www.rug.nl/research/portal>. For technical reasons the number of authors shown on this cover page is limited to 10 maximum.

Oxford Handbooks Online

Accessibility and Reference Production: The Interplay Between Linguistic and Non-Linguistic Factors

Jorrig Vogels, Emiel Krahmer, and Alfons Maes

The Oxford Handbook of Reference

Edited by Jeanette Gundel and Barbara Abbott

Print Publication Date: Feb 2019 Subject: Linguistics, Semantics

Online Publication Date: Mar 2019 DOI: 10.1093/oxfordhb/9780199687305.013.16

Abstract and Keywords

This chapter reviews recent research on speakers' referential choices in discourse. It focuses on the choice to mention a certain referent first on the one hand, and to produce a pronoun or more elaborate noun phrase on the other. Whereas traditional theories have relied mainly on influence of linguistic context to explain these choices, recent psycholinguistic studies have started investigating the effect of non-linguistic factors as well. Results of these studies suggest that the two referential choices (choice of referent and choice of referring expression) are largely driven by different sets of factors, contrary to the claim that both are related to accessibility of mental representations in memory. The chapter concludes that these findings are best explained by assuming referents can be accessible on different levels: a local and a more global discourse level. Referential choices may be influenced by both levels of accessibility, but with different probabilities for different choices.

Keywords: reference, referring expressions, accessibility, referential choices, non-linguistic factors, psycholinguistic studies

Jorrig Vogels, Emiel Krahmer, and Alfons Maes

16.1 Introduction

REFERENCE is an essential part of language. When we speak, we talk about something. The act of referring can be seen as forming a link between the speaker's mind and the outside world. For example, a speaker asking 'could you hand me that stapler?' is expressing her¹ intention to get hold of a physical object in the world by referring to that object with a linguistic expression (in this case, the definite noun phrase 'the stapler'). The things we refer to are, however, not always physical objects (including people), nor do they need to be part of the outside world. For example, we can refer to objects that are

Accessibility and Reference Production: The Interplay Between Linguistic and Non-Linguistic Factors

not present in the direct physical environment ('I left the stapler in the office'), or objects that only exist in our imagination ('the stapler I dreamt about last night'). We can refer to objects that existed in the past ('the cake that I ate yesterday'), or will exist in the future ('the cake that I will bake tomorrow'). We can also refer to events ('last night's dinner party'), locations ('the picturesque town of Tilburg'), and abstract concepts ('the financial crisis'), to name a few. In none of these situations is the thing that is being referred to (the referent) an object in the directly perceivable world. It would therefore be better to say that we refer to conceptualizations in our minds, rather than to objects in the outside world (e.g., Jackendoff 2002; Johnson-Laird 1983). Even in those cases where the referent is present in the world, reference is still mediated (p. 338) by a conceptualization of the referent (which may be wrong, as in 'Could you hand me that stapler?' 'That's not a stapler, that's a hole punch.').

One of the main questions in linguistic research is how speakers put such conceptualizations into language. In Levelt's model of language production (Levelt 1989), a speaker who wants to communicate about a certain entity has to make a number of important decisions. First, she has to decide which information to include in the utterance, that is, she needs to select the content of the message to be expressed. Once a relevant concept has been selected, it has to be put into a grammatical structure. Given that speech proceeds serially, this structure ultimately has to map on a linear order of words. That is, one thing has to be mentioned before another. Hence, a speaker needs to *choose a concept that will be referred to first*. Second, the speaker has to decide which linguistic form she is going to use to refer to a certain concept. That is, she has to *choose a referring expression*.

This chapter reviews research that is concerned with the factors that influence a speaker's referential choices in language production, that is, how they choose which referent they mention first, and how they choose what type of referring expression to use for a certain referent. In particular, it explores the small amount of data that exists on how linguistic factors (e.g., grammatical function; topicality) interact with non-linguistic factors (e.g., perceptual salience; animacy) in the choice of a referring expression in discourse. Research on the choice of referring expressions has been primarily concerned with the role of the linguistic context. By contrast, the influence of non-linguistic factors on the choice of referent for first mention (or, more generally, choice of word order) has been investigated quite extensively. Therefore, this chapter focuses on the question to what degree such non-linguistic factors also influence the choice of referring expression. The area of interest is the choice of a particular type of referential form in a discourse (e.g., a pronoun or a full noun phrase), rather than the selection of semantic content to include in an expression to describe something (e.g., how speakers choose between 'the large stapler', 'the red stapler', and 'the large red stapler').

This chapter is organized as follows: §16.2 briefly discusses some terminological issues associated with the notion of *accessibility*, which is assumed to underlie many referential choices. §16.3 reviews the most important research on these referential choices. §16.4 discusses recent studies that investigated interactions between linguistic and non-

linguistic factors in the choice of referring expression in discourse. In §16.5, different viewpoints on the nature of accessibility effects described in the literature are discussed, and a tentative proposal of how seemingly contradictory effects might be unified is presented. Finally, §16.6 concludes and provides suggestions for further research.

16.2 Accessibility and related terms

Both the choice of referent for first mention and the choice of referring expression have been related to the degree of activation of the conceptualizations—or mental representations—of referents in memory. The higher the activation of a certain representation, the more likely it is to appear early in the linguistic structure, and the higher the likelihood that the expression referring to it is more attenuated (e.g., Levelt 1989). This activation status has been described with a variety of terms, such as *accessibility* (Ariel 1990; Bock and Warren 1985), *salience* (Osgood 1971; Sridhar 1988), *cognitive status* (Gundel, Hedberg, and Zacharski 1993), *givenness* (Chafe 1976; Gundel et al. 1993; Prince 1981b), *topicality* (Givón 1983) and *focus of attention* (Grosz, Joshi, and Weinstein 1995), each with slightly different assumptions and viewpoints.

Some of these terms, such as givenness and topicality, emphasize the importance of information structure in the discourse. For example, when a referent was the topic of the preceding sentence, with topic being defined as what the sentence is about (Reinhart 1982), its representation in memory is likely to be highly activated. Other terms—such as focus of attention and cognitive status—emphasize the importance of cognitive capacities. For example, it seems likely that those referents that are attended to are more activated, since they may be actively maintained in memory (Foraker and McElree 2007). To remain implicit as to the source of the activation, we will use the more general term accessibility to refer to the activation of mental representations in the memories of speakers and hearers, whatever the cause. For the sake of brevity, we will often use ‘the accessibility of a referent’ as shorthand for ‘the accessibility of the mental representation of a referent’.

Crucially, the notion of accessibility concerns activation of non-linguistic representations, rather than activation of lexical items in the mental lexicon (cf. Arnold 2010). To distinguish the two types of activation, Bock and Warren (1985) speak of *conceptual accessibility*, which they define as “the ease with which the mental representation of some potential referent can be activated in or retrieved from memory” (1985: 50), as opposed to *lexical accessibility*, which refers to “the ease with which the representations of word forms can be recovered from memory” (1985: 52). In this chapter, the term *accessibility* is used to refer to conceptual accessibility, unless explicitly specified otherwise. Furthermore, the term *salience* is reserved for properties of the referent itself rather than of its representation in memory. These properties can be linguistic, as when the referent is mentioned in a prominent or non-prominent syntactic position,² or non-linguistic (e.g., perceptual), as in the size or color of the physical object that is referred to. They can also be determined by the context, such as the preceding discourse or the physical environment, or they can be intrinsic to the referent, such as animacy. Finally,

the terms *topicality* and *givenness* are taken to denote factors that contribute to an entity's (linguistic) salience, while *focus of attention* is used as a speaker- or hearer-internal factor that might influence accessibility directly.

Thus, accessibility is thought to be a determining factor both in the choice of referent for first mention and in the choice of referring expression. However, research has revealed differences in how exactly accessibility affects these choices. Notably, there may be differences in the degree to which the two types of referential choice are affected (p. 340) by certain factors, linguistic as well as non-linguistic (e.g., Fukumura and van Gompel 2010; Rohde and Kehler 2014; Stevenson, Crawley, and Kleinman 1994). The next section reviews relevant literature on the role of accessibility in the choice of referent for first mention (§16.3.1) and in the choice of referring expression (§16.3.2).

16.3 Factors influencing referential choices

16.3.1 Effects of accessibility on the choice of referent for first mention

When people speak (or write), seemingly unordered thoughts and concepts have to be put into the linear order that characterizes language. As a first decision, speakers have to choose a *starting point* from which their utterance is going to unfold (MacWhinney 1977). The question is how they do this. Research suggests that the salience of the concepts people talk about plays an important role. It has frequently been observed that when people describe simple visual scenes, salient entities are mentioned earlier than less salient entities. For example, entities that are perceptually more prominent are typically found to have a higher likelihood to be produced early in the sentence than perceptually less prominent entities (e.g., Flores d'Arcais 1975; Gleitman, January, Nappa, and Trueswell 2007; Osgood 1971; Sridhar 1988; Tomlin 1997). Gleitman et al. (2007), for instance, presented participants with simple scenes, for example, of a dog chasing a man, and very briefly showed a not consciously noticeable attentional cue (a black square), either on the dog or on the man. They found that the scenes were more likely to be described with active sentences ('the dog chases the man') when the cue was on the dog, and more likely to be described with passive sentences ('the man is chased by the dog'), when the cue was on the man.

Another important factor that has been found to influence order of mention is animacy. Human or animate entities tend to be produced before inanimate entities in a sentence (e.g., Branigan, Pickering, and Tanaka 2008; McDonald, Bock, and Kelly 1993; Prat-Sala and Branigan 2000; Van Nice and Dietrich 2003a). For example, in a series of sentence recall experiments by McDonald et al. (1993), participants were more likely to recall originally passive sentences with an inanimate subject as active sentences with an animate subject.

Accessibility and Reference Production: The Interplay Between Linguistic and Non-Linguistic Factors

Other studies have concentrated on the production of sentences within a discourse. Here, it has been found that given information is more likely to be mentioned first in a sentence than new information (Bock and Irwin 1980; Ferreira and Yoshita 2003). For example, Ferreira and Yoshita (2003) found that participants were more likely to recall Japanese dative constructions such that given information preceded new information, even when the original sentences had the reverse order. In addition, information that is (p. 341) predictable from the discourse context may be produced early in the sentence (Arnold 2001; Givón 1983), although predictable entities may also be postponed to a less prominent position, due to a preference to start an utterance with the most important (i.e., most newsworthy) information (Givón 1983, 1988; Gundel 1988).

These saliency effects on the choice of referent for first mention have generally been attributed to the conceptual accessibility of referents' mental representations in the speaker's memory (e.g., Bock and Warren 1985; Prat-Sala and Branigan 2000). If we assume that language production proceeds incrementally (e.g., Kempen and Hoenkamp 1987; Levelt 1989), speakers start producing an utterance before the planning of that utterance is completed. Because highly accessible referents are more easily retrieved from memory, they are subsequently mentioned earlier in the sentence. However, the exact relation between degree of accessibility and order of mention is not yet completely clear. With respect to perceptual salience, for example, Gleitman et al. (2007) argued that when a speaker's visual attention is focused on a referent, this speeds up the retrieval of lemma representations, and hence directly affects order of mention. Other researchers argue instead that in picture descriptions, visual attention does not influence the mentioning of entities directly. Rather, the order in which entities are produced is determined by the global apprehension of the scene, for example, who does what to whom (Bock, Irwin, and Davidson 2004; Griffin and Bock 2000). It is in this apprehension stage that visual attention may have an influence (see also Vogels, Krahmer, and Maes 2013a).

The relation between accessibility and the positioning of concepts in the sentence may also be mediated by grammatical function or topichood. In English, for example, it has been found that the most accessible concept is typically made the subject (e.g., Bock and Warren 1985; McDonald et al. 1993). This suggests that accessibility determines which entity becomes the subject of the sentence, which in turn is preferably produced in the sentence-initial position, but that it does not determine sentence position directly. However, subject and sentence-initial position are highly confounded in English, which makes the exact relation between accessibility and sentence position unclear. In languages in which word order is more free than in English, such as Greek (Branigan and Feleki 1999), German (Kempen and Harbusch 2004), Hungarian (É. Kiss 2002), Italian, and Spanish (Brunetti 2009), accessibility has been found to affect sentence position independently of grammatical function. However, in such languages, accessibility may instead affect the likelihood of a referent becoming a topic. This referent will in turn be more likely to occupy the topic position, which is often the first position in the sentence (Lambrecht 1994). In a study of the Algonquian language Odawa, Christianson and Ferreira (2005) were able to disentangle effects on both grammatical function and

topichood from those on linear order. They found that accessible entities in Odawa were not directly promoted to the sentence-initial position, but were given prominent syntactic functions via the priming of a particular syntactic structure.

Different saliency factors may also interact. Combining picture descriptions with the presence of a discourse context, Prat-Sala and Branigan (2000) found that a referent's (p. 342) intrinsic salience (in this case, animacy) and its givenness in the discourse interact in the decision to mention this referent first in an utterance: for given entities, the influence of intrinsic salience was greater than for new entities. This led Prat-Sala and Branigan to distinguish two types of accessibility: a referent's *inherent accessibility* refers to activation in memory caused by its intrinsic properties—such as animacy or concreteness—which are assumed to be stable across contexts. Within a discourse, this inherent activation can be supplemented by the referent's *derived accessibility*, a temporary level of activation caused by the salience of the referent in the discourse, such as whether it is given or topical. Thus, a referent's derived accessibility adds to its inherent accessibility. However, if the two types of accessibility run counter to each other, such as when the referent is inanimate but given, derived accessibility may override inherent accessibility in strong enough contexts (Prat-Sala and Branigan 2000).

Central to this chapter is the question whether the non-linguistic factors that have been found to affect the choice of referent for first mention, such as visual salience and animacy, also affect the choice of referring expression. This is the topic of the next subsection.

16.3.2 Effects of accessibility on the choice of referring expression

Language provides many possible ways to refer to something, ranging, for example in English, from very elaborate expressions such as full definite descriptions with modifiers (e.g., 'the large old-fashioned red stapler with the little scratch on the top') to very short ones such as pronouns (e.g., 'it'). For references to objects in simple scene descriptions (so-called 'one-shot' references), it has been found that the perceptual context of the referent affects the choice of a referring expression (e.g., Brown-Schmidt and Tanenhaus 2006; Ferreira, Slevc, and Rogers 2005; Osgood 1971; Sedivy 2003; Sridhar 1988). For example, speakers use more specific referring expressions, such as 'the empty martini glass', when the visual context contains multiple objects of the same type, as a less specific expression, such as 'the martini glass', may be ambiguous. In addition, which objects in the scene are taken into account as possible competitors depends on what part of the scene people are focusing on (e.g., Beun and Cremers 1998; Brown-Schmidt and Tanenhaus 2008). Objects outside of the focused area are less likely to be considered (see also Koolen, Krahmer, and Swerts 2015). Speakers also often overspecify their referring expressions, adding information that reflects salient properties such as color (e.g., Arts, Maes, Noordman, and Jansen 2011; Engelhardt, Bailey, and Ferreira 2006; Koolen, Goudbeek, and Krahmer 2013). For example, in a referential description task described in Koolen et al. (2013), speakers often used color information when providing definite

Accessibility and Reference Production: The Interplay Between Linguistic and Non-Linguistic Factors

descriptions for objects in a scene, sometimes even when all objects in the scene had the same color.

(p. 343) A different line of research has focused on the production of referring expressions in discourse. In contrast to research on the choice of referent for first mention, research on the choice of referring expressions in discourse has mainly concentrated on the influence of linguistic factors. For example, in both experimental and corpus studies of discourse, it has been found that referring expressions tend to become shorter when the same object is referred to multiple times. This includes the use of fewer words (e.g., Clark and Wilkes-Gibbs 1986), acoustic reduction (e.g., Aylett and Turk 2004; Bard et al. 2000), and the choice for less specific types of referring expressions, such as pronouns (Ariel 1990). Discourse factors that have been identified as influencing the choice of referring expression include, among others, recency (e.g., Clark and Sengul 1979), topicality (e.g., Givón 1983), first mention (e.g., Gernsbacher and Hargreaves 1988), grammatical function (e.g., Brennan 1995; Gordon, Grosz, and Gilliom 1993), syntactic parallelism (e.g., Arnold 1998), competition (e.g., Ariel 2001; Arnold and Griffin 2007), protagonist hood (e.g., Karmiloff-Smith 1981; Morrow 1985), episode shifts (e.g., Anderson, Garrod, and Sanford 1983; Vonk, Hustinx, and Simons 1992), and thematic role (e.g., Arnold 2001; Stevenson et al. 1994). For example, a referent that has recently been mentioned is likely to be referred to with a pronoun. Hence, in the second sentence of (1a) a pronoun will generally be preferred to refer to Alison, who is mentioned in the directly preceding sentence.³

(1)

- a. Alison_i was working in the garden. Suddenly,
 {she_i/#Alison_i} was hit by a swishing tree branch.
- b. Emily was helping Alison_i in the garden. Suddenly,
 {#she_i/Alison_i} was hit by a swishing tree branch.

Here, repeating the name could give rise to the impression that the discourse contains two people named Alison. However, in (1b) a name would be preferred over a pronoun to refer to Alison, despite the fact that Alison is still the most recently mentioned entity. This is because there is a competing entity, Emily, which is mentioned in subject position and which is more topical (i.e., the sentence is more about Emily than about Alison).

Even though research on one-shot references in scene descriptions has revealed the importance of the perceptual context, interactions between linguistic and non-linguistic factors have not received much attention in research on the choice of referring expressions in discourse. The selection of a certain type of expression in discourse has traditionally been seen as a highly addressee-oriented process (e.g., Clark, Schreuder, and Buttrick 1983). Thus, in theories of reference (e.g., Ariel 1990; Chafe 1994; Givón 1983; Gundel et al. 1993) speakers are commonly believed to choose referring expressions such that these aid the addressee in retrieving the correct antecedent. Since

Accessibility and Reference Production: The Interplay Between Linguistic and Non-Linguistic Factors

the linguistic context can generally be assumed to be available for both speaker and (p. 344) addressee, speakers are assumed to choose referring expressions in accordance with the information from the discourse that is part of the common ground of speaker and addressee. The different theoretical frameworks all share the idea that, in a discourse, some entities are focused on more than others (both by the speaker and the addressee), and that this has an impact on the choice of referring expression. The degree of accessibility (or topicality/givenness/focus of attention) is presented as a property of mental representations, which is influenced by—but by no means identical to—the salience of referents in the preceding discourse. Indeed, all frameworks acknowledge that referents that have not been mentioned previously can still be accessible, for instance from world knowledge (e.g., ‘*the king* will visit my hometown tomorrow’) or from the physical context (e.g., ‘*that man* over there’). For example, expressions such as unheralded pronouns (pronouns without a linguistic antecedent) and deictics (e.g., ‘that one’, often accompanied by a pointing gesture) are dependent on the configuration of objects in the physical environment of the interlocutors (e.g., Clark and Marshall 1981; Clark et al. 1983; Greene, Gerrig, McKoon, and Ratcliff 1994; Jarvella and Klein 1982; Piwek, Beun, and Cremers 2008).

However, the discourse context has been considered the most important factor driving the activation of mental representations. For example, Ariel (2001) claims that:

[...] it is the discoursal rather than the physical salience of the entities involved which determines the degree of accessibility assigned to particular mental representations [...]. Although the physical context does affect the discourse model of the speakers, mental representations are a direct product of our discourse model only.

Ariel (2001: 31)

A similar assumption is reflected in the account of Brennan and Clark (1996). They argue that while factors such as perceptual salience may influence the choice of a referring expression, what is most important, at least in dialogue, is whether the referent has been mentioned recently or frequently in the discourse.

In research on accessibility, referring expressions have therefore been investigated mainly as anaphors, that is, expressions that have an antecedent in the preceding discourse (e.g., Brennan 1995; Chafe 1976; Clark and Marshall 1981; Gundel et al. 1993; Prince 1981b). For example, when the linguistic context contains two possible referents of the same gender, using a pronoun (e.g., ‘he’ or ‘she’) to point out one of these referents would, at least in English, result in ambiguity for the addressee. Therefore, speakers are likely to avoid this type of ambiguity by choosing a more specific referring expression, something that has also been found experimentally (e.g., Arnold, Eisenband, Brown-Schmidt, and Trueswell 2000; Arnold and Griffin 2007; Fukumura, van Gompel, and Pickering 2010). The influence of non-linguistic factors, such as the visual context and intrinsic properties of referents, is likely to be more subtle, and it is only recently that these factors have been systematically investigated in studies of reference production in

discourse. The next section discusses a number of relatively recent psycholinguistic studies that have investigated interactions between linguistic and non-linguistic factors in the choice of referring expression.

(p. 345) 16.4 The influence of non-linguistic factors on the choice of referring expression

While research on the choice of referring expression in discourse has focused on linguistic factors, perceptually and conceptually salient entities are likely to attract attention (e.g., Coco and Keller 2010; Henderson and Ferreira 2004; New, Cosmides, and Tooby 2007; Pratt, Radulescu, Guo, and Abrams 2010), and may therefore influence referent accessibility (Arnold and Griffin 2007). The few existing studies that have investigated how non-linguistic factors interact with linguistic factors in the choice of referring expression in discourse have yielded different results. The most important of these are summarized below.

Arnold and Griffin (2007) conducted a story completion experiment in which they varied the number of possible referents in the discourse. Participants were presented with a scene accompanied by a context sentence, after which they had to describe a second scene, referring to a target referent. There were three types of stimuli: (1) a competitor character was present both visually and linguistically; (2) no competitor was present; (3) a competitor was present both visually and linguistically in the first scene but not visually present in the second. Target referent and competitor were referentially unambiguous (i.e., of different genders). Arnold and Griffin found that participants used fewer pronouns to refer to the target character in condition (1) than in condition (2), suggesting that the presence of a competitor decreased the target referent's accessibility. However, no difference in pronoun use was found between conditions (1) and (3), suggesting that it was the linguistic presence of the competitor in the context sentence that affected accessibility rather than its visual presence during sentence production.

Fukumura et al. (2010) argued that the lack of an effect of visual salience in Arnold and Griffin's (2007) experiment could be due to a too weak manipulation of visual salience or a too strong manipulation of linguistic salience. In addition, because speakers in this experiment were not talking to an addressee, there may have been no incentive to take the visual context into account. Fukumura et al. conducted a new experiment in which they remedied these potential problems. Participants were presented with two consecutive pictures showing either two toy characters (competitor present condition) or one toy character (competitor absent condition), as shown in Figure 16.1. The first picture was combined with a written context sentence in which either only the target character was mentioned (competitor not mentioned condition) or both characters were mentioned (competitor mentioned condition). To reduce the dominance of the linguistic context, the target character was mentioned as the possessor in a genitive phrase and the competitor (if present) in a *by*-phrase (e.g., 'The pirate's carpet had been cleaned (by a prince)'). In the second picture, the target character **(p. 346)** performed a simple action.

Accessibility and Reference Production: The Interplay Between Linguistic and Non-Linguistic Factors

Participants described this picture to a confederate, who then acted out the description using the real toys. The results showed that participants used fewer pronouns when the competitor had been mentioned, replicating the linguistic salience effect found by Arnold and Griffin (2007). More importantly, participants also used fewer pronouns to refer to the target referent when the competitor was visually present than when it was not visually present. In addition, the effect of the visual context was larger in the condition where the competitor was linguistically present than in the condition where the competitor was not mentioned at all, indicating that the visual context had a greater effect when the linguistic context was less compelling (i.e., when the competitor was also linguistically present).

Figure 16.1 Example of a stimulus item from Fukumura et al. (2010) in two conditions (a) a competitor character is visually present; (b) no competitor character is visually present).

Reprinted from Fukumura, Kumiko, van Gompel, Roger P. G., and Pickering, Martin J. (2010). 'The use of visual context during the production of referring expressions', *The Quarterly Journal of Experimental Psychology* 63(9): 1700–15. doi: 10.1080/17470210903490969. © The Experimental Psychology Society, reprinted by permission of Taylor & Francis Ltd, www.tandfonline.com.

These results suggest that speakers use both linguistic and visual information simultaneously when choosing a referring expression: the presence of a visually presented competitor referent influences pronoun use independently of the presence of a linguistically presented competitor referent. One explanation for this effect may be (p. 347) that speakers use more specific expressions because they believe that their addressee has more difficulty identifying the correct referent when multiple possible referents are present (Fukumura et

al. 2010). Alternatively, the effect may be explained by interference from the competitor referent in the retrieval of the target referent's representation. When multiple referents are present, representations of these have to be kept in memory. Since working memory has a restricted capacity (Baddeley 1986), the presence of an additional possible referent may decrease the speaker's memory activation of the target referent, resulting in more specific expressions (cf. Arnold and Griffin 2007). In addition, such an interference effect might be stronger the more similar the target and competitor are. Indeed, Fukumura, van Gompel, Harley, and Pickering (2011), found that speakers used more explicit referring expressions (and fewer pronouns) for the target referent when a different-gender competitor character was in a similar situation in the visual context (e.g., both characters were sitting on a horse), than when the competitor was in a different situation (e.g.,

Accessibility and Reference Production: The Interplay Between Linguistic and Non-Linguistic Factors

standing on the ground). A similar effect was found when target and competitor had the same gender (Fukumura, Hyönä, and Scholfield 2013), which was shown to be independent of a general avoidance of gender-ambiguous pronouns. Fukumura et al.'s (2011) explanation for this finding was that a shared semantic feature that is activated during reference (e.g., sitting on a horse; being male) causes the accessibility of the entities that have this feature to be lower than when only one entity has this feature. This decreased accessibility in turn results in the use of more explicit referring expressions (see also Fukumura et al. 2013).

The reduced salience of a referent due to referential competition (i.e., the presence of another possible referent) may however be a different type of salience than salience caused by a referent's perceptual properties, such as its size (Flores d'Arcais 1975; Osgood 1971; Sridhar 1988). To investigate the influence of the referent's visual salience within a scene on the choice of referring expression in a discourse, Vogels, Krahmer, and Maes (2013a) conducted two story completion experiments in which the relative size of the characters in a scene was manipulated by presenting the target referent either on the foreground or in the background, as shown in Figure 16.2. The scenes always contained two characters, keeping the amount of referential competition constant. Discourse salience was manipulated by either mentioning the referent as the subject of the directly preceding context sentence or in a prepositional phrase in the penultimate context sentence. The results showed that although speakers were more likely to mention visually foregrounded characters first in their utterances, they were not more likely to pronominalize these characters. Discourse salience, on the other hand, had a strong effect on pronoun use: for discourse-salient referents, pronouns were highly preferred, which left little room for visual salience to take effect. A second experiment in which the relative linguistic salience of the two characters was reduced (by mentioning both characters as subjects in both context sentences) again showed no increased likelihood of pronominalization for visually foregrounded characters. Thus, Vogels et al. (2013a) found no evidence that visual foregrounding of a referent leads to more attenuated referential forms within a linguistic context.

Accessibility and Reference Production: The Interplay Between Linguistic and Non-Linguistic Factors

Figure 16.2 Example of a stimulus item in four conditions in Vogels et al. (2013a), Experiment 1.

Reprinted from Vogels, Jorrig, Krahmer, Emiel, and Maes, Alfons (2013a). 'Who is where referred to how, and why? The influence of visual saliency on referent accessibility in spoken language production', *Language and Cognitive Processes* 28(9): 1323–49. doi: 10.1080/01690965.2012.682072. Reprinted by permission of Taylor & Francis Ltd, www.tandfonline.com.

(p. 348) Finally, a corpus study by Duan, Elsner, and de Marneffe (2013) did reveal some effects of several visual saliency variables on the choice of referring expression. However, these effects were mainly found on the choice between longer and shorter definite descriptions, and not on the choice for a pronoun versus a definite description. This suggests that the use of pronouns may indeed be more dependent on linguistic salience. In sum, it remains unclear how exactly the visual context influences referent

accessibility in discourse. However, one consistent finding seems to be that the linguistic context often remains dominant in the choice to produce a pronoun.

Another source of accessibility consists of higher-level conceptual properties of referents, such as animacy, individuation, and concreteness (e.g., Brown-Schmidt, Byron, and Tanenhaus 2005; Dahl and Fraurud 1996; Fukumura and van Gompel 2011; Maes 1997; Yamamoto 1999). A few studies have investigated the interaction between animacy and linguistic factors in reference production. For example, in a story (p. 349) completion experiment, Fukumura and van Gompel (2011) found that speakers were more likely to pronominalize animate entities than inanimate entities. Thus, when producing a continuation to the sentences given in (2a), participants were more likely to use a pronoun ('they') when they referred to 'the hikers' than when they referred to 'the canoes'. The same held when grammatical roles were reversed, such as in (2b), suggesting that the effect of animacy on pronominalization was independent of the grammatical function of the antecedent. Fukumura and van Gompel also found that animacy affected the choice of referent: participants were more likely to refer to the animate NP than to the inanimate NP in their continuations.

(2)

- a. The hikers carried the canoes downstream.
Sometimes . . .
- b. The canoes carried the hikers downstream.
Sometimes . . .

Vogels, Krahmer, and Maes (2013b) additionally showed that the animacy effect on the choice of referring expression is likely to be driven by the conceptualization of the referent in the (perceptual) context, rather than by the lexical semantics of the antecedent: In an animation retelling experiment, simple geometrical figures were made perceptually animate or perceptually inanimate. For example, Figure 16.3 shows a series of stills from an animation in which a green circle moves in such a way that it appears to be climbing up a slope. Entities that were made to look animate in this way were more likely to be pronominalized than perceptually inanimate entities (e.g., a circle that rolled down a slope). This effect was independent of the entity's lexical animacy (e.g., whether it was called a 'stone' or a 'hiker'). Interestingly, whether speakers mentioned the entity as the subject of their utterances seemed to be influenced more by its lexical than by its perceptual animacy. This suggests that animacy may have differential effects on the choice of referent and the choice of referring expression. At the very least, these results indicate that the speaker's choice of referring expression in discourse is not only affected by purely linguistic factors: the referent's saliency based on perceptual or conceptual factors also has an influence. Still, the perceptual animacy effect in Vogels et al. (2013b) was only present when the referent was not salient in the discourse, again suggesting that linguistic salience is a stronger factor than non-linguistic salience.

Figure 16.3 Four stills taken from an animation used in Vogels et al. (2013b), Experiment 1. Letters indicate order; arrows indicate movement. Both were not shown during the experiment.

From Vogels, Jorrig, Krahmer, Emiel, and Maes, Alfons (2013b). 'When a stone tries to climb up a slope: the interplay between lexical and perceptual animacy in referential choices', *Frontiers in Psychology* 4: 154. doi: 10.3389/fpsyg.2013.00154.

In sum, while it seems clear that the choice of referring expression in discourse is strongly affected by the referent's salience in the linguistic context, as measured by factors such as grammatical function and topichood (e.g., Fukumura and van Gompel 2010; Rohde and Kehler 2014), there is some evidence that it can also be influenced by non-linguistic factors, although the findings are not always straightforward. Why do some studies find clear effects of non-linguistic factors on referential

choices, independently of linguistic factors, while others find no effects? One reason might be the influence of the nature of the task. For example, with respect to the influence of the visual context, the way scenes are presented may have an effect on accessibility. In a series of picture description studies by Van Nice and Dietrich (2003b), participants (p. 350) either produced descriptions with the scene in view or were only allowed to start describing after the picture had disappeared. The results showed an interaction between the animacy of the agent character and the animacy of the patient character on word order when speakers had to speak from memory (i.e., when the picture disappeared): there were more passive sentences when the patient character was animate than when it was inanimate, but only when the agent character was inanimate. Crucially, this interaction was absent when participants described the pictures in view. Van Nice and Dietrich argued that when people speak from memory instead of describing things that are present in the immediate context, they process information from multiple referents simultaneously, allowing different types of information to interact. Hence, it might be the case that the effects of more subtle, non-linguistic factors on referential choices become more manifest in cases of more compacted processing.

Montag and MacDonald (2013) investigated interactions between task context, visual salience, and animacy on the production of relative clauses in picture descriptions. Participants watched scenes containing multiple objects—both animate and (p. 351) inanimate—in various degrees of visual salience (e.g., size), and answered questions inviting them to describe a target object, such as 'What is red?' Montag and MacDonald found that whereas animate target objects were almost always described with a passive

Accessibility and Reference Production: The Interplay Between Linguistic and Non-Linguistic Factors

relative clause (e.g., ‘the man that is being thrown’), inanimate target objects were described with both active (e.g., ‘the ball that the man is throwing’) and passive (e.g., ‘the ball that is being thrown’) relative clauses. This choice was influenced by the object’s visual salience. Interestingly, inanimate objects that were more visually salient were more likely to be described with active relative clauses. According to Montag and MacDonald, this is surprising, because the standard hypothesis is that more salient referents are assigned to higher grammatical functions, such as subject. In active relative clauses, the target referent is the object of the relative clause, so these clause types should actually be less preferred for salient objects according to this hypothesis. Montag and MacDonald argued that this effect could be explained by assuming that differences in visual salience changed the nature of the task: less salient objects take longer to locate, and this longer search time increases the chance that other objects will be focused on as well. This may lead speakers to more explicitly contrast the target object from other objects, for which a passive relative clause—with the target object as the subject and the possibility to leave out the agent character from the description—may be best suited.

Applied to the choice of referring expression, this argument would predict that speakers also use more specific referring expressions when search times to locate the referent are longer, since such expressions may more explicitly contrast possible referents from each other. This reasoning would be in line with Arnold and Griffin’s (2007) proposal that more specific expressions are used when the speaker has to distribute her attention over multiple possible referents. Alternatively, speakers may also use visual information to determine whether to disambiguate referring expressions for an addressee. For example, Brown-Schmidt and Tanenhaus (2008) found in an unscripted referential communication task speakers were more likely to use ambiguous referring expressions (e.g., ‘the green rectangle’ when there are two green rectangles) when the target referent was more salient in the visual context than its possible competitor referents. Salience was mainly determined by the proximity of the competitors to the target and whether the competitors were relevant to the task. Eye-tracking analyses revealed that these were the same factors that addressees used to interpret ambiguous referring expressions. For example, they were less likely to take into account competitors that were far away or irrelevant. This suggests that speakers use less specific expressions when they believe that their addressee is focusing on a restricted referential domain that excludes possible referents that are not relevant to the task.

Indeed, the role of the addressee is likely to be an important factor in reference production. In another experiment by Brown-Schmidt and Tanenhaus (2008), the experimenter occasionally interrupted the referential communication task to recalibrate the addressee’s eye-tracker. During this calibration, the addressee was asked to look at different objects that sometimes started with the same sounds and were therefore temporarily ambiguous (e.g. ‘clown’/‘cloud’). As predicted, the results showed that in (p. 352) these calibration sessions, addressees were considering a much larger visual domain for the identification of the referent than in the referential communication task, where the domain was restricted by the previous interactions with the speaker.

Accessibility and Reference Production: The Interplay Between Linguistic and Non-Linguistic Factors

If speakers are sensitive to the specific role the addressee is taking up in a task, this should influence their production of referring expressions. Studies of reference production have indeed provided evidence that the addressee's degree of engagement makes a difference in the use of referring expressions (e.g., Arts et al. 2011; Maes, Arts, and Noordman 2004; Rosa, Finch, Bergeson, and Arnold 2013). Especially when it is clear to the speaker what the addressee's task is (e.g., whether he has to operate an alarm clock or assist a surgeon in a life-saving operation; Arts et al. 2011), speakers may be more motivated to take this information into account in their referring expressions. After all, in most everyday communicative situations speakers have the intention to convey a message with a particular goal in mind (see also Montag and MacDonald 2013).

With respect to the choice of referring expressions in anaphoric contexts, it has been found that just the presence or absence of an addressee affects the use of pronouns (Kantola and van Gompel 2011). This may explain some of the divergent findings in research on the effect of non-linguistic factors on the choice of referring expression. For example, no addressee was present in the studies by Arnold and Griffin (2007) and Vogels et al. (2013a), in which no effect of the visual context was found. By contrast, speakers in the study by Fukumura et al. (2010) produced their story continuations for an addressee, who had to recreate the stories based on these continuations. Here, the visual context did have an effect on the type of referring expression. A straightforward interpretation of these findings is that speakers are more likely to take the visual context into account when there is an addressee for whom this visual information might be helpful. Such an explanation would suggest that visual salience effects on the choice of referring expression are related to audience design rather than to speaker-internal cognitive constraints (cf. Arnold and Griffin 2007; Fukumura et al. 2010).

Still, there is no conclusive evidence yet as to the role of addressee-oriented processes in the effects of non-linguistic factors on reference production. For example, even when no addressee is present, speakers sometimes show linguistic behavior that would otherwise be considered a form of audience design (e.g., Koolen, Gatt, Goudbeek, and Krahmer 2011; Van der Wege 2009). This makes it difficult to determine whether certain referential choices are really adaptations towards an addressee. In addition, given that the studies discussed above manipulated visual salience in different ways (i.e., varying the number of characters in the scene versus varying the perceptual prominence of the characters in the scene), different effects of visual salience on reference production may be difficult to compare. Future research should investigate the degree to which the nature of the task and the role of the addressee influence the role of non-linguistic accessibility factors in the choice of referring expressions.

To sum up, when speakers refer to an entity they appeal to a mental representation of the referent. This representation is assumed to have a certain degree of activation, which can be determined by many different factors, either linguistic (e.g., structural (p. 353) properties of a previous mention of the referent in the discourse) or non-linguistic (e.g., the referent's visual salience or animacy). The precise influence of these factors is probably highly dependent on the nature of the task. Still, it generally holds that the

higher the activation of a representation, the more accessible the referent, that is, the easier it can be retrieved from memory. Assuming that spoken language production proceeds incrementally, the element that is retrieved quickest is produced first (within grammatical constraints). Thus, more accessible referents are more likely to be chosen for first mention in an utterance. Because highly accessible referents do not need a lot of linguistic encoding to be retrieved by an addressee, speakers are more likely to use more attenuated expressions, such as pronouns, to refer to such referents.

16.5 Dissociating the choice of referent and the choice of referring expression

So far, accessibility has been defined as the degree of activation of a referent's mental representation, which may be influenced by many different factors. However, as has become clear, research on the choice of referent for first mention and research on the choice of referring expressions have generally focused on a different set of factors. This raises the question whether the two referential choices are actually driven by the same notion of accessibility.

The choice of what to mention first has generally been considered a speaker-internal process of activating and selecting (lexical) concepts (Levelt 1989). As discussed above, a concept is assumed to have a high activation, for example, when it has been mentioned recently (Ferreira and Yoshita 2003), when it is perceptually salient (Gleitman et al. 2007), or when it is inherently salient (e.g., animate; Branigan et al. 2008). As soon as this activation reaches a certain threshold, the concept is passed down to the level of linguistic formulation, where it is assigned a prominent syntactic position (either directly, or through the assignment of grammatical functions or information structure). Thus, salient entities are preferred to be mentioned first in the sentence because they are highly accessible for the speaker. Givón (1976) proposes that the different saliency factors—such as animacy and givenness—combine to form a hierarchy of topicality. Since people tend to talk about animate agents, for example, such entities are likely to be the topic of the sentence, and hence to occur in a prominent (e.g., sentence-initial) position. Alternatively, what these saliency factors may have in common is that they attract the speaker's attention (e.g., Gleitman et al. 2007; Myachykov, Garrod, and Scheepers 2009; Tomlin 1997). Perceptual attention may be captured by, for example, large, foregrounded, animate, or moving objects (p. 354) (e.g., Flores d'Arcais 1975; Mazza, Turatto, and Umiltà 2005; New et al. 2007; Pratt et al. 2010). In a discourse, elements in a prominent syntactic function (e.g., subject) may be in the focus of attention (e.g., Grosz et al. 1995). Because what is attended to is easier to retrieve, it is more likely to be talked about first.

By contrast, the choice for a certain type of expression has traditionally been seen as a highly addressee-oriented process (e.g., Ariel 1990; Clark et al. 1983; Gundel et al. 1993). A speaker chooses a referring expression such that her addressee is able to readily resolve what is being referred to. Therefore, the speaker has to make assumptions about what information is in common ground with her addressee. Speakers then attenuate their

Accessibility and Reference Production: The Interplay Between Linguistic and Non-Linguistic Factors

expressions when they believe that the referent is highly accessible for their addressee. Accessibility in this view is primarily determined by properties of the referent in the discourse, which are most likely to be available to both speaker and addressee.

In sum, how speakers choose a referent for first mention has often been explained in terms of the speaker's focus of attention, or the topic the speaker wants to talk about; conversely, the choice of referring expression has been attributed to the speaker's assumption about the addressee's focus of attention, or what the addressee expects the speaker to be talking about. However, if both referential choices (what to refer to first and how to refer to it) are determined by the same concept of accessibility, one would predict that those factors that make a referent accessible will increase the probability that it will be mentioned first as well as the probability that it will be referred to with an attenuated expression. This implies that what is accessible is what is most likely to be talked about next, that is, most predictable.

Indeed, Givón (1983) argues that more accessible topics are more likely to be talked about again and hence are more predictable. Because predictable entities need less explicit encoding, they can be referred to with more attenuated expressions (see also Jaeger 2010). This idea is also reflected in *centering theory* (Grosz et al. 1995). One of the assumptions in centering theory is that speakers seek to produce a maximally coherent discourse to minimize inferences addressees have to make about the relations between consecutive utterances. To this end, speakers try to avoid too many shifts to a different backward-looking center (the entity that is currently in the focus of attention) across utterances, that is, they tend to keep talking about the same thing. Speakers are also assumed to choose certain referring expressions to signal whether they continue to talk about the same thing. If any entity in the current utterance is pronominalized, this should at least be the backward-looking center, that is, the entity that is most likely to be talked about again.

Whereas in centering theory it is primarily the subject of the preceding sentence that is most likely to be talked about again, and hence to be pronominalized, this does not always have to be the case. For example, in (3), 'he' most likely refers to Stan, despite John being the subject of the preceding sentence, because it is likely that the second sentence is providing the reason why Stan was admired.

(3)

John admired Stan_i. He_i could devour an entire cake in twenty seconds.

(p. 355) Thus, in a predictability account, Stan is the most accessible entity in the first sentence of (3), because he is most likely to be mentioned next. Therefore, a subsequent reference to Stan is most likely to be in the form of a pronoun.

That accessibility is related to predictability has also been proposed by Arnold (1998, 2001) in her *expectancy hypothesis*: accessibility correlates with the probability that a referent will be mentioned again (Arnold 2008). Hearers make predictions about what

Accessibility and Reference Production: The Interplay Between Linguistic and Non-Linguistic Factors

speakers are going to say based on the preceding discourse and on their experience with what types of entities (e.g., subjects; animates) speakers are likely to keep on talking about. Hearers will be more likely to perceive these entities as being accessible for the speaker, who in turn will be more likely to use attenuated expressions to refer to them (for speaker-internal reasons). In this way, the speaker's choice of referring expression aids the hearer's comprehension (whether so intended or not), since it signals whether the information was predictable or not. Evidence for this view was found by Arnold (2001), who used a story continuation task to elicit references to entities with either a source or a goal role, counterbalanced for grammatical function (e.g., 'Lisa_{SOURCE} gave the leftover pie to Brendan_{GOAL}' or 'Marguerite_{GOAL} caught a cold from Eduardo_{SOURCE} two days before Christmas'). In their continuations, participants were more likely to refer to the goal entity than to the source entity. While grammatical subjects were most likely to be pronominalized, participants were also more likely to use pronouns to refer to the goal entity than to the source entity, at least for non-subject referents. Arnold (2001) concluded that goal entities are more accessible than source entities because they are more likely to be mentioned next. Hence, speakers will (inadvertently) signal this accessibility to their addressees by using attenuated expressions.

Other researchers question the view that what the speaker refers to first, and how she refers to it, are both determined by the predictability of a referential event. For example, like Arnold (2001), Stevenson et al. (1994) found in a story completion study that participants were most likely to continue the story with the character that was in a goal role (e.g., 'John_{SOURCE} passed the comic to Bill_{GOAL}'), a patient role (e.g., 'Joseph_{AGENT} hit Patrick_{PATIENT}'), or either a stimulus or experiencer role (e.g., 'Ken_{STIMULUS} impressed Geoff_{EXPERIENCER}') depending on the connective that followed the introductory sentence (e.g., 'because' or 'so'). However, the choice of whether to use a pronoun or not was only affected by the referent's syntactic position, with participants being more likely to use a pronoun when referring to the first mentioned character than to the second mentioned character.

More recently, additional evidence that the choice of referring expression is independent from the likelihood that the referent will be mentioned next was obtained by several recent studies (e.g., Chiriacescu 2011; Fukumura and van Gompel 2010; Kaiser, Li, and Holsinger 2011; Karimi, Fukumura, Ferreira, and Pickering 2014; Rohde and Kehler 2014; Vogels et al. 2013a). For example, Fukumura and van Gompel (2010) found a strong bias for continuing sentences containing either Stimulus-Experiencer verbs such as 'scare' or Experiencer-Stimulus verbs such as 'fear' with the character that was the stimulus rather than the experiencer. However, as in Stevenson et al. (1994), pronoun use was only affected by the antecedent's syntactic position. These findings (p. 356) suggest that accessibility-related factors affect the choice of referent and the choice of referential form differently (cf. Rohde and Kehler 2014). This does not seem to support an account in which the accessibility of referents equals their predictability in context (e.g., Arnold 2001, 2008; Givón 1983). The reason is that if effects of accessibility could be explained by predictability alone, it should have been found that what is most predictable, that is,

what is most likely to be mentioned next, is also most likely to be referred to with more attenuated expressions.

Kehler et al. (2008) offer a Bayesian approach to pronoun production and interpretation that captures the dissociation between the choice to pronominalize and the likelihood of next mention, presented in (4).

(4)

$$P(\text{referent}|\text{pronoun}) = \frac{P(\text{pronoun}|\text{referent})P(\text{referent})}{P(\text{pronoun})}$$

In this account, $P(\text{referent}|\text{pronoun})$ represents the probability that a pronoun refers to a certain referent. This probability is dependent on both the probability that given a certain referent, a pronoun is used to refer to it ($P(\text{pronoun}|\text{referent})$) and the probability that that referent is mentioned ($P(\text{referent})$). Crucially, $P(\text{pronoun}|\text{referent})$ is taken to be independent of $P(\text{referent})$, enabling both referential choices to be determined by different factors.

Based on the opposition between what speakers refer to first and how they refer to it, some researchers have concluded that accessibility is only related to structural linguistic properties of the referent's previous mention in the discourse, and not affected by predictability based on semantic factors (e.g., Fukumura and van Gompel 2010; Rohde and Kehler 2014). However, given that non-linguistic factors may affect the choice of referring expression, as discussed above, it is not clear how these effects would fit in such an account, since they would not be related to accessibility in this view.

Other researchers propose instead that accessibility is a multifaceted notion, capturing different constraints that may interact and sometimes work against each other (e.g., Arnold 1998). For example, Kaiser and Trueswell (2008) propose that referents can be accessible on different levels of representation. On the one hand, a referent can be accessible in a syntactico-semantic representation of the preceding sentence, which includes information about linguistic factors such as grammatical function and thematic role. On the other hand, it can be accessible in a mental model of the discourse, which involves a global representation about the event being described (Johnson-Laird 1983; Kintsch and Van Dijk 1978). Kaiser and Trueswell argue for a form-specific multiple constraints approach to reference, in which different referential choices may be either more or less sensitive to different accessibility-related factors. This means that there can be differences between specific referential forms in the degree to which they are sensitive to certain factors (e.g., Brown-Schmidt et al. 2005; Kaiser and Trueswell 2008). For example, while the choice between pronouns and more elaborate referring expressions may be affected by discourse salience, the choice between full and reduced pronouns in languages such as Dutch and Estonian may be more related to contrastiveness or importance (Kaiser 2010b, 2011b; Kaiser and Trueswell 2004; Vogels,

Accessibility and Reference Production: The Interplay Between Linguistic and Non-Linguistic Factors

Maes, and Krahmer 2014). Here, salient (e.g., given; animate) referents may actually be more likely to be referred to with the full forms than with the reduced forms.

Such findings are not in accordance with the claim that the accessibility scales proposed by Ariel (1990), Givón (1983), and Gundel et al. (1993) map on a hierarchy of referential forms that goes from less to more attenuated expressions. Rather, they suggest that the choice of referring expression may not be a unified phenomenon, as a linear accessibility scale makes it appear. Which type of referring expression is chosen is not just a function of its association with a particular accessibility status; multiple factors may play a role, which may each have different effects on different expressions.

In the same vein, some factors may be more important for the choice of referent for first mention, while others are more important for the choice of referring expression. Indeed, even though both linguistic and non-linguistic saliency factors are assumed to underlie the accessibility of mental representations, and hence referential choices, recent research suggests that these factors may affect these referential choices differently. For example, while Vogels et al. (2013a) found no convincing evidence for visually salient entities being pronominalized more often, visual salience did affect which referent speakers were more likely to mention first, as the subject of their utterance. With respect to discourse salience, whether referents were explicitly introduced as the topic of the discourse influenced the choice of referent for first mention, whereas pronominalization was mainly affected by whether referents were local topics.

These findings support a division of accessibility effects along the line of a global mental model versus a local syntactico-semantic representation, as proposed by Kaiser and Trueswell (2008). In this view, the selection of a referent for first mention may be largely driven by the speaker's *global conceptualization of the discourse*, perhaps in the form of a mental model (Johnson-Laird 1983; Van Nice and Dietrich 2003a). In a narrative discourse, this model may contain information about, for example, who or what the main or most important character is and about who does what to whom and why (Griffin and Bock 2000). It may also contain contextual information about the event being described and the communicative situation (e.g., 'who is my addressee?'; Galati and Brennan 2010). Who or what is conceptualized as the main character may in turn be influenced by whether the character was linguistically introduced as such (e.g., 'Once upon a time there was a little dragon...'), by whether it is perceptually salient (e.g., visually foregrounded), or by whether it is conceptually salient (e.g., human agents are more likely to be main characters than stones or snowflakes). These factors make a referent accessible in a global representation of the discourse, although they can be in conflict, for example when the discourse topic is not a human agent.

The choice of a particular referring expression, on the other hand, may depend more on a *local model of the discourse*, involving primarily linguistic factors (cf. Grosz 1977; Grosz et al. 1995). For example, pronouns are likely to be used when the antecedent was the subject or the topic of the directly preceding sentence. This is supported by the fact that even studies that took into account non-linguistic factors consistently found (p. 358) that

Accessibility and Reference Production: The Interplay Between Linguistic and Non-Linguistic Factors

the linguistic context had a stronger effect. However, this does not mean that there is no room for more global factors to play a role. For example, speakers may still be more likely to pronominalize those entities that they have conceptualized as animate or human based on non-linguistic information. This might be part of speakers' mental model of the event, in which entities higher in animacy are more accessible, for example because they are more likely to be main characters and more important for the event to be described.

The distinction between local and global accessibility may partly overlap with the distinction between derived and inherent accessibility proposed by Prat-Sala and Branigan (2000): inherent properties of referents such as animacy are likely to influence a global model of the discourse, while properties derived from context such as topicality may be more likely to affect the local discourse model. However, Vogels et al. (2013b) showed that this distinction is not always easy to make, because inherent properties such as animacy may also be influenced by the context. As mentioned above, they found that entities that had animate names (but were not necessarily perceived as animate) had no increased probability of pronoun use when the entity's discourse salience was controlled for, but that they were more likely candidates for the subject role than entities with inanimate names. The opposite pattern was observed for perceptual animacy, where entities that were perceived as animate were more likely to be pronominalized than entities that were perceived as inanimate, although they were not more likely to be mentioned as the subject of the sentence. These findings suggest that the differences in accessibility effects between different referential choices cannot easily be explained by a distinction between inherent and derived accessibility. After all, it holds for animacy as well as for other factors that what determines accessibility is not the properties of entities per se, but how the entities are *conceptualized* (Jackendoff 2002).

Another possible division of accessibility is to distinguish a referent's accessibility for the speaker from the referent's accessibility for the addressee (e.g., Arnold 2008; Bard et al. 2000; Galati and Brennan 2010). Indeed, research suggests a difference between the choice of referent and the choice of referring expression in the degree to which they are influenced by speaker- and addressee-oriented factors. For example, studies on the linearization of constituents have suggested that the choice of referent for first mention is determined by the degree to which the speaker attends to the referent (e.g., Gleitman et al. 2007; Tomlin 1997). However, other findings (e.g., Vogels, Krahmer, and Maes 2015) suggest that entities receiving less attention from the speaker are not necessarily referred to with less attenuated expressions. Rather, the choice of referring expression may depend on a model of the discourse the speaker maintains to be able to select referring expressions that a (hypothetical) addressee can interpret correctly (e.g., Hendriks, Koster, and Hoeks 2014). In other words, the choice of a referent for first mention may be mainly driven by the referent's accessibility for the speaker, while the choice of referring expression is more strongly affected by the speaker's assumptions (p. 359) about the referent's accessibility for the addressee. This would be in line with how these referential choices are traditionally viewed in the literature.

Accessibility and Reference Production: The Interplay Between Linguistic and Non-Linguistic Factors

Still, such a division between speaker- and addressee-oriented accessibility would be highly confounded with a division between local and global accessibility. After all, the local discourse context constitutes highly concrete evidence for the speaker about the knowledge of the addressee: in most cases, a speaker may safely assume that her addressee has internalized a record of the immediately past discourse with this speaker that is similar to her own. The addressee's global representation of the discourse is probably more abstract and subjective, which makes it harder for the speaker to build assumptions on it. Therefore, it might be the case that effects of non-linguistic factors such as animacy and visual salience are more related to accessibility in the speaker's own mind.

The distinction between a local and a global model of the discourse in referential choices might be visualized as in Figure 16.4 (Vogels 2014). As shown by the arrows, both models influence the choice of referent as well as the choice of referring expression. However, the amount of influence differs, in accordance with the idea that accessibility-related factors affect both referential choices differently. This could be represented by giving weight values to the connections. For example, if referent (p. 360) accessibility in the global discourse model is more influential in the choice of referent, the weight of the corresponding connection would be higher than that of the arrow departing from the local discourse model. Conversely, if local discourse factors strongly affect the choice of certain types of referring expression, the link from the local discourse model to particular referring expression types should have a higher weight than the link from the global discourse model.

Figure 16.4 Visualization of a possible unified model of effects of accessibility from local and global discourse models on the choice of referent and the choice of referring expression type. Solid lines mean 'provides input for'; dashed lines mean 'influences'.

Different types of referring expressions—such as full and reduced pronouns—may also differ in their sensitivity to both local and global discourse salience, in line with Kaiser and Trueswell's (2008) form-specific multiple-constraints approach. In addition, the same type of referring expression may be sensitive to both global and local sources of

accessibility to differing degrees. Again, global factors affecting the choice of referring expression, such as the referent's animacy, may run counter to a classical accessibility account (e.g., Ariel 1990), such that referents that should count as more accessible are actually preferred to be referred to with the less reduced forms. The model in Figure 16.4 extends this account to the choice of referent for first mention: what is most likely to be mentioned next may be affected by both levels of accessibility, but the degree to which

these levels are involved differs, and is not the same as for the choice of a particular referring expression.

How this tentative model might be integrated into a general model of language production (e.g., Levelt 1989) is shown by the grey colored connections in Figure 16.4. As soon as a speaker has chosen a referent for next mention, it is probably assigned a grammatical function in the grammatical encoding stage (Levelt 1989). In this stage, the referent will also be assigned the selected type of referring expression. The lexical item required to express the selected referent is retrieved from the mental lexicon. The ease of retrieval of lexical items (lexical accessibility) may also influence grammatical encoding, for example, whether the item is produced as the subject of the sentence. The assignment of grammatical functions in the grammatical encoding stage may in turn result in an update of the local discourse model, with more prominent grammatical functions increasing the referent's accessibility in this model.

The influence of factors affecting the global mental model will often be in line with local factors. For example, the global discourse topic is likely to also be the topic of individual utterances (e.g., Givón 1983). Global factors may also affect the local discourse model indirectly: a globally accessible entity is more likely to be mentioned first in a particular utterance, which in turn increases its local accessibility (via the connection between grammatical encoding and the local discourse model in Figure 16.4). This will affect the referential form with which this entity is referred to in the next utterance. In other cases, the two types of factors may be in conflict, such as when the local topic is inanimate (see Vogels et al. 2013b), or when the local topic is not the same as the global discourse topic or the protagonist (e.g., Poesio, Stevenson, Di Eugenio, and Hitzeman 2004; Van Vliet 2008). In sum, although local discourse factors might dominate in the choice to pronominalize, more global conceptual factors can still play a role.

(p. 361) 16.6 Open issues

The model presented in the previous section is still somewhat rough and imprecise, and in its current state does not make specific predictions about how a certain referential choice will be influenced by different factors. Clearly, further research has to be conducted to investigate the exact role of various global and local factors in referential choices. In this section, some questions that still await an answer are discussed briefly.

One question is what the locus of accessibility effects in the human mind is. Accessibility is believed to be a property of representations in memory (e.g., Ariel 1990; Bock and Warren 1985), but the exact relation between reference and memory often remains implicit (cf. Chafe 1994; Van Nice and Dietrich 2003b). One idea is that the degree of accessibility relates to whether a referent is represented in long-term or in short-term memory (e.g., Gundel et al. 1993): referents are more accessible when they are in short-term memory. Alternatively, accessibility may only be a property of working memory. Van Rij, van Rijn, and Hendriks (2013) propose a computational model of the production and comprehension of pronouns, in which accessibility is modeled as a combination of base

Accessibility and Reference Production: The Interplay Between Linguistic and Non-Linguistic Factors

and spreading activation of entities in working memory. The entity with the highest activation is pronominalized.

However, if referents can be accessible on multiple levels (e.g., global versus local), as suggested above, it is not clear how these different levels of accessibility would connect to a single memory representation. It is therefore an open question whether it is possible for these different levels to be subsumed under a single notion of accessibility, or whether they actually represent different cognitive processes. Since the role of accessibility in reference production has been studied mostly by looking at the output (i.e., the referring expressions produced), few direct measures of accessibility are available. Future research in this field may therefore benefit from studies that tap into the cognitive processes underlying reference production themselves, for example by using eye-tracking or manipulating memory load.

Second, as discussed above, it is becoming clearer that the influence of different factors on reference production is highly dependent on the specific communicative situation in which an utterance is produced, on the type of task, the speaker's goals and the role of the addressee. Therefore, these environmental variables cannot be ignored when investigating referential choices. Given that researchers in reference production use a variety of production tasks, this poses a challenge when one wants to compare different studies. For example, many studies use written completion tasks (e.g., Anderson et al. 1983; Fukumura and van Gompel 2011; Kaiser and Trueswell 2004; Stevenson et al. 1994; Vonk et al. 1992, while other studies focused on spoken language production (cf. Arnold 2001; Arnold and Griffin 2007; Fukumura et al. 2010; Kaiser et al. 2011; Vogels et al. 2013a, b). Written tasks might not elicit the same kind of results as would spoken language production experiments. For instance, when people speak, they have less opportunity to reflect on what they are saying than (p. 362) when they write. Hence, spoken tasks might reveal more about people's initial, automatic linguistic choices than written tasks (Arnold 2001). In addition, written texts generally have a less clear addressee than spoken communication, which may also influence referential choices.

Third, studies differ in the degree to which they resemble naturalistic communicative situations. Many story completion tasks are not embedded within a visual context, while everyday communicative situations often are. If visual context is used in an experiment, it is often highly artificial or cartoonish. On the other hand, naturalistic scenes quickly become cluttered, which may influence reference production (e.g., Coco and Keller 2009; Koolen et al. 2015). For instance, it increases the number of factors that may affect referential choices, which may obscure effects of other factors. Moreover, people often talk about things that are not immediately present in the perceptual context. Therefore, speaking from memory rather than describing scenes might actually reflect the more frequently occurring type of language production (Christianson and Ferreira 2005).

Completion tasks also suffer from shifts in modality or breaks in the speech flow. It would be more naturalistic to have participants tell their own stories than to first have them listen to or read (aloud) part of the story and then have them switch to speaking from

Accessibility and Reference Production: The Interplay Between Linguistic and Non-Linguistic Factors

their own imagination, which also involves a switch from comprehension to production. However, this latter setup may be necessary to systematically investigate effects of the linguistic context. Therefore, we would not argue for or against one or the other method. However, we would argue that the nature of the task is something that should be taken into account when discussing the outcomes of any experiment. It would be interesting to conduct experiments in highly naturalistic, unconstrained contexts for comparison, since it is conceivable that in such contexts non-linguistic cues play a more important role. These contexts may also include other types of discourse than narratives, as well as longer discourses than the ones typically used in story completion experiments (cf. Rohde and Kehler 2014). In addition, experimental research might be complemented with data from corpus studies or other naturally collected data to see whether the results would generalize to language use outside the lab (Brown-Schmidt and Tanenhaus 2008). Reference production should also be studied in other groups of participants than university students. Work that has already been done includes research on young children (e.g., Hendriks et al. 2014; Matthews, Lieven, Theakston, and Tomasello 2006), elderly people (e.g., Hendriks et al. 2014), Alzheimer patients (Almor, Kempler, MacDonald, Andersen, and Tyler 1999), and people with autism (e.g., Arnold, Bennetto, and Diehl 2009).

Finally, most research on referring expression production has focused on the use of third person singular personal pronouns versus more specific types of referring expression (e.g., full NPs) in references to persons and objects. The range of possible referents and referring expressions is, however, much larger than that. To name a few, one could investigate references to spatial locations or points in time (e.g., 'here', 'there', 'tomorrow'), to events and other abstract entities (e.g., 'yesterday's thunderstorm', 'my dream'), to substances (e.g., 'the mud') to parts of objects (e.g., 'the sheep's nose') or to sounds (e.g., 'that noise'); one could also investigate the use of different types of definite (p. 363) and indefinite descriptions (e.g., with specific or generic referents), deictics (e.g., 'this', 'that'), second person pronouns (e.g., 'you'), plural expressions or zero anaphora. For all these types of referents and referring expressions, the influence of linguistic and non-linguistic variables on accessibility may be different. In addition, accessibility effects can be different in other dimensions than the choice of referential form, such as acoustic reduction, disfluency, and gesture. Recent work in psycholinguistics has already started to explore some of these directions (e.g., Arnold and Tanenhaus 2011; Hoetjes, Koolen, Goudbeek, Krahmer, and Swerts 2015; Kaiser et al. 2011; Watson, Arnold, and Tanenhaus 2008).

Insights from a broader view on referential choices, for example, one that takes into account non-linguistic factors as well as all kinds of environmental factors, may also be beneficial for computational models of reference production. Most models of referring expression generation (REG) focus on content selection for initial definite descriptions used for object identification outside of a discourse context (e.g., Dale and Reiter 1995). Existing models of generating referring expressions in discourse generally implement only a basic account of linguistic salience, for example based on centering theory (e.g., Kibble 1999; Krahmer and Theune 2002; Passonneau 1996), although algorithms for

Accessibility and Reference Production: The Interplay Between Linguistic and Non-Linguistic Factors

generating pronouns in texts have been developed that also take into account factors such as the antecedent's ontological type (Strube and Wolters 2000) and global discourse structure (Callaway and Lester 2002; McCoy and Strube 1999; Salmon-Alt and Romary 2000). Recently, the GREC challenges program (Generating Referring Expressions in Context; Belz, Kow, Viethen, and Gatt 2010) has started to evaluate natural language generation systems that produce referring expressions in discourse, including pronominal expressions. One of the aims of these systems is to generate human-like references within a context, making use of psycholinguistic data (see also Krahmer and Van Deemter, Chapter 19 of this volume). For human-like generation of referring expressions, future models should also take into account influences of factors that go beyond the local discourse context, incorporating the perceptual context, speaker as well as addressee goals, and the nature of the task.

In conclusion, this chapter concerned the question of how speakers choose what they mention first in an utterance and how they refer to it, as well as the role the accessibility of mental representations plays therein. Traditionally, effects of accessibility on the choice of a referent for first mention have been ascribed to the amount of attention the speaker has allocated to a referent, which may be influenced by linguistic, perceptual, or conceptual factors. By contrast, effects of accessibility on the choice of referential form have been explained as the speaker's effort to choose expressions in accordance with the information that is in common ground between speaker and addressee, which is largely determined by linguistic properties of the directly preceding discourse. The research reviewed here suggests that the choice of referring expression may also be influenced by non-linguistic factors, interacting with linguistic factors, although the effects may be dependent on the task. At the same time, it is clear that there are indeed differences in the effects of accessibility on both referential choices. This has consequences for our understanding of the notion of accessibility, because it (p. 364) suggests that there are multiple types or levels of accessibility, to which different language production processes can be sensitive in different ways. These findings must be incorporated into both theoretical and computational models of reference.

Acknowledgments

This chapter is based on parts of the first author's PhD thesis, which was written under supervision of the second and third author. The first author is the main contributor to the text of this chapter. The authors are grateful to Josefin Lindgren for her comments on an earlier version of this chapter.

Notes:

⁽¹⁾ Following common practice, feminine forms will be used to refer to speakers, and masculine forms to refer to hearers.

Accessibility and Reference Production: The Interplay Between Linguistic and Non-Linguistic Factors

(²) Depending on the language, prominent syntactic positions include the subject, topic, or preverbal position, for example, while object, focus, and post-verbal positions may be considered non-prominent.

(³) Indexes mark coreferentiality. Hashes (#) mark low acceptability given the context.

Jorrig Vogels

Jorrig Vogels is a postdoctoral researcher at the Center for Language and Cognition, University of Groningen (CLCG), where he works on a personal NOW VENI project on the role of perspective taking in the production of referring expressions. He is interested in both experimental and theoretical approaches to reference production and comprehension, especially in the influence of factors that exceed the local linguistic context. He received his PhD from Tilburg University in 2014.

Emiel Krahmer

Emiel Krahmer is a full professor of Communication, Cognition, and Computation in the Tilburg School of Humanities and Digital Sciences. In his work he studies how people communicate with each other, both in text and in speech, with the aim of subsequently improving the way computers communicate with human users. He was PI on various externally funded research projects, including an NWO VICI roject on the production of referring expressions, which ran from 2007 to 2013.

Alfons Maes

Alfons Maes is a full Professor and Head of the Department of Communication and Digital Sciences, Tilburg School of Humanities and Digital Sciences. His work includes research on multimodal aspects of human communication with studies on visual metaphor, visual health communication, multimodal navigation communication, and information visualization. He received his PhD in 1991.