



This article is part of the topic “Memory and Common Ground Processes in Language Use,” Sarah Brown-Schmidt, William S. Horton and Melissa C. Duff (Topic Editors). For a full listing of topic papers, see: [http://onlinelibrary.wiley.com/journal/10.1111/\(ISSN\)1756-8765/earlyview](http://onlinelibrary.wiley.com/journal/10.1111/(ISSN)1756-8765/earlyview)

## Revisiting the Memory-Based Processing Approach to Common Ground

William S. Horton,<sup>a</sup> Richard J. Gerrig<sup>b</sup>

<sup>a</sup>*Department of Psychology, Northwestern University*

<sup>b</sup>*Department of Psychology, Stony Brook University*

Received 17 November 2014; received in revised form 27 May 2015; accepted 30 June 2015

---

### Abstract

Horton and Gerrig (2005a) outlined a memory-based processing model of conversational common ground that provided a description of how speakers could both strategically and automatically gain access to information about others through domain-general memory processes acting over ordinary memory traces. In this article, we revisit this account, reviewing empirical findings that address aspects of this memory-based model. In doing so, we also take the opportunity to clarify what we believe this approach implies about the cognitive psychology of common ground, and just as important, what it does not imply. We also highlight related areas of research demonstrating how general cognitive processes can constrain access to relevant knowledge in ways that shape both language production and comprehension.

**Keywords:** Common ground; Audience design; Conversation; Memory-based processing; Resonance; Episodic memory

---

### 1. Introduction

In classic pragmatic models, speaking felicitously requires that individuals design utterances in ways that reflect the *common ground* they share with addressees. Similarly, addressees interpret utterances in light of this same common ground. As described by

---

Correspondence should be sent to William S. Horton, Northwestern University, 2029 Sheridan Road, Evanston, IL 60208-2710, USA. E-mail: [whorton@northwestern.edu](mailto:whorton@northwestern.edu)

Schiffer (1972), Grice (1975), and others, interlocutors can only successfully communicate intended meanings by taking such common ground into account. From a psycholinguistic perspective, however, understanding *how* individuals consider common ground as part of the “working out” of intended meanings has been immensely challenging, to say the least (for an overview, see Brennan & Hanna, 2009). In Horton and Gerrig (2005a) (hereafter H&G), we articulated a particular model of conversational common ground that focused on the role of ordinary memory processes in shaping the information available to speakers during language production. Our goal was to demonstrate how theorists could gain traction on the issue of common ground by viewing it through the lens of well-understood cognitive psychological accounts of memory.

In this article, we revisit this account, reviewing empirical findings that address aspects of this memory-based model. In doing so, we take the opportunity to clarify what we believe this approach implies about the cognitive psychology of common ground, and just as important, what it does not imply. We also highlight other lines of research demonstrating how general cognitive processes can constrain access to relevant knowledge during both language production and comprehension.

## 2. Memory-based processing and common ground

As part of a more general account of definite reference, Clark and Marshall (1981) articulated what almost certainly remains the most influential attempt to provide a psychologically motivated model of common ground. Clark and Marshall rejected the idea that successful reference required complex meta-representations of mutual knowledge (“I know that you know that I know. . .”). Instead, they proposed that language users rely upon a set of simpler “co-presence heuristics” that direct attention toward types of evidence that support inferences about commonality. In particular, inferences about personal common ground—that is, information shared between individuals—could be derived from evidence for past physical and linguistic co-presence, whereas inferences about communal common ground—that is, information shared within communities—could be derived from evidence for joint membership in particular social or cultural groups (Clark, 1996).

A central component of Clark and Marshall’s (1978, 1981) account was the notion that information about co-presence is encoded via special-purpose memory representations, either in the form of “reference diaries,” a type of episodic memory relevant for personal common ground, or “reference encyclopedias,” a type of semantic memory relevant for communal common ground. In particular, these memory representations were seen as encoding information about *triple co-presence*, or moments in which the speaker, addressee, and potential referent were all “openly present together” (Clark & Marshall, 1981, p. 32). For example, Nadia’s mental diary might encode the fact that she and her brother were co-present when her new neighbor stopped by to introduce himself as “Bill.” With this in mind, when later talking to her brother, Nadia could refer to “Bill” without having to ensure that the neighbor’s name was indeed part of their mutual knowledge; support for the appropriateness of the bare name would be directly available in her diary entry

for the previous encounter. By maintaining information about triple co-presence, such memory representations provide a basis upon which language users may infer common ground for purposes of reference resolution and production.

While Clark and Marshall (1981) contained a variety of important insights, part of the motivation behind H&G was to take seriously the particular idea that inferences about common ground necessarily involve access to information stored in memory. Drawing upon examples from a corpus of spoken conversations, we framed our arguments in the context of *audience design*—or how speakers tailor utterances for addressees. Within this context, our account aimed to provide a description of both the memory processes that could support felicitous utterances as well as the types of memory representations upon which these processes are likely to act.

On the representation side, our interest in articulating a domain-general approach to common ground led us to reject the notion of dedicated, special-purpose memory representations. In particular, we argued that the representational demands of Clark and Marshall's (1978, 1981) reference diaries—despite being less than that required by full-blown representations of mutual knowledge—were still too computationally burdensome to form the basis for a psychologically plausible account of common ground, given that they would seem to require that individuals encode triple co-presence with respect to each and every combination of interlocutor and referent. In their stead, we proposed that language use involves access to representations functionally identical to the kinds of episodic traces people routinely encode. These traces are “ordinary” in the sense that they are not focused on anything as specific as triple co-presence—that is, unlike reference diaries, they are not dedicated to the task of directly tagging information to support inferences about common ground. Instead, other people, events, and objects become representationally bound together merely through co-activation or other forms of semantic or episodic association related to the structure and content of these experiences.

On the processing side, we highlighted the theoretical and methodological importance of distinguishing between two types of processes relevant for audience design. We used the term *commonality assessment* to refer to the means by which language users gain access to information relevant to common ground, and *message formation* to refer to how speakers produce utterances that reflect this information. Importantly, we described how both automatic and strategic processes could contribute to commonality assessment and message formation. In fact, our most important claim was that automatic memory processes (which function in a domain-general fashion) can explain many apparent instances of audience design.

In H&G, the description of automatic commonality assessment was inspired by cue-driven retrieval processes found in global matching models of recognition memory (Gillund & Shiffrin, 1984; Hintzman, 1986; Ratcliff, 1978; Ratcliff & McKoon, 1988). In such models, particular configurations of information function as probes to memory, triggering a parallel search for any stored information that shares overlapping features with, or is in some way similar to, the probe cue. This automatic search process has been termed *resonance* (Ratcliff, 1978), based on the metaphor of a tuning fork (i.e., stored memories) vibrating at particular frequencies in response to being “struck” by a configuration of cues, with resonance strength being a function of the number of overlapping features

between the stimulus cues and target memory traces. Importantly, when this resonance reaches some activation threshold, which itself is a function of the recency and frequency with which a memory has been previously retrieved, that knowledge can become accessible in a way that influences other processes. In recognition models, this generally takes the form of a particular stimulus being recognized as “old” or familiar from previous encounters.

On this account, commonality assessment often emerges from a speaker’s automatic recognition that certain information can be treated as familiar or not within a particular context. This emergent familiarity can influence message formation as well, with speakers being more likely to use particular forms of reference if the relevant linguistic representations are sufficiently accessible in the moment. Consider again the situation in which Nadia’s brother was present when her new neighbor stopped by to introduce himself. Once an episodic trace of this event is part of Nadia’s memory, her brother will be a salient cue for the automatic retrieval of the contents of that entire interaction as represented in memory, including the neighbor’s name. The emergence of this information will facilitate audience design without Nadia having to engage in special considerations of common ground. Conversely, this information would likely not reach the same threshold of accessibility in the context of someone (such as Nadia’s sister) who was not present for the original event.

The notion of accessibility “in the moment” is important for understanding how resonance shapes the form and content of utterances with respect to common ground. Given enough time, capacity, and effort, language users should nearly always be able to engage in strategic assessments of commonality by constructing or elaborating upon appropriate representations. However, a critical motivation for the memory-based approach was the assumption that utterances in fluent dialog are frequently generated and understood under conversational pressure. For that reason, the product of automatic processes might often represent the limits of possible real-time common ground assessment. Thus, speakers’ utterance plans may incorporate addressee-specific knowledge only if resonance strength is sufficiently robust or if the activation threshold for that information is low.

As documented in H&G, the just-in-time nature of this automatic retrieval process helps explain why particular types of errors with respect to common ground arise in normal conversation. Indeed, the same mechanisms can account both for instances in which people appear to assume too little common ground (due to the insufficient accessibility of partner-relevant information arising from low cue overlap and/or high activation thresholds) and instances in which people appear to assume too much common ground (due to high cue overlap and/or especially low activation thresholds). The latter case is especially interesting because it illustrates how information could become (inappropriately) accessible due to the influence of entire configurations of cues in ways that go beyond direct encoding of triple co-presence. For example, one may mistakenly assume that a coworker is aware of a colleague’s illness given the availability of a constellation of associated cues (the office setting; other mutually known colleagues with whom you might have discussed the illness) that conspire, in a way, to make the fact of the illness highly accessible in the context of this coworker, despite the lack of direct co-presence for this information.

Because these memory processes are automatic, they should operate in all instances of language production. However, H&G also suggested that speakers may sometimes choose to undertake strategic consideration of representations of shared experiences. Indeed, our corpus analysis turned up several instances in which what appeared to be effortful evaluation of common ground could be observed on the surface of the conversation.

We expect that ordinary memory processes will help explain many circumstances in which speakers opt to engage strategic processes. Recall Nadia, her brother, and her neighbor Bill. Suppose that Nadia knows someone else named Bill who is also associated in memory with her brother. As she begins to plan an utterance, her memory processes may provide her with evidence for both Bills, creating a context for her to produce a referring phrase beyond the simple proper name (e.g., “my neighbor Bill”). More generally, the overall constellation of memories encoded with respect to a particular addressee—as well as the strength of those memory representations—will influence the likelihood that speakers may be prompted to undertake strategic control over both commonality assessment and message formation. To be sure, other forces also affect these decisions. For example, speakers might be more likely to engage strategic effort when speaking to young children (e.g., Snow, 1972) or addressees who hail from different cities or cultures (e.g., Isaacs & Clark, 1990). However, a memory-based account should help explain circumstances in which speakers engage strategic effort with addressees who are familiar to them.

In sum, the memory-based model of common ground outlined in H&G provided a description of how individuals could strategically as well as automatically gain access to information about others through domain-general memory processes acting over ordinary memory traces. Our goal was to show how this approach could provide a useful starting point for understanding the full range of behaviors that speakers display with respect to successful and unsuccessful instances of audience design. With this overview in mind, we now briefly consider evidence supporting aspects of this approach.

### 3. Empirical assessment of the memory-based view

A variety of findings are consistent with the idea that processes of memory encoding and retrieval constrain how people manage the demands of audience design. For example, Horton and Gerrig (2005b) demonstrated that speakers are more likely to show evidence for audience design to the extent that they have access to memory representations that provide more distinct cues concerning “who knows what.” In our referential communication task, participants, acting as Directors, helped each of two partners, acting as Matchers, to arrange sets of picture cards. In an *orthogonal* condition, each Matcher arranged cards from distinct card categories (e.g., Matcher A only saw dogs while Matcher B only saw fish), whereas, in an *overlapping* condition, each Matcher arranged unique cards from the same categories (e.g., Matcher A and Matcher B each saw different fish). In a subsequent round with each Matcher that involved the full set of cards, Directors in the orthogonal condition showed stronger evidence for audience design, often by elaborating more

when describing cards that were new for a given Matcher. These results provided support for the claim that evidence for audience design will depend in many circumstances on the extent to which speakers have access to suitable memory representations.

A similar conclusion emerges from a set of experiments reported by Gorman, Gegg-Harrison, Marsh, and Tanenhaus (2013). Pairs of participants learned names for a set of novel images, some of which were learned by both individuals (shared names), whereas others were learned only by the participant playing the role of Director (privileged names). On subsequent test trials, the Director instructed the partner which image to select from a series of arrays. In an initial experiment, Directors either learned the shared names together with their partner (*shared experience* condition), or they were simply told that their partner learned the same subset of shared names separately (*alone* condition). As one would expect, Directors' descriptions on test trials were more likely to mention shared names than privileged names. However, when the shared names had been learned alone, Directors used names more frequently for privileged items, and also more frequently elaborated upon names with additional descriptive information, suggesting uncertainty about the status of name knowledge. In a follow-up experiment, when Directors had the shared experience of learning some names with one partner but in the test phase interacted with a third party (who learned the shared names independently), the likelihood of naming privileged items or supplementing names with descriptions was intermediate between the alone and shared experience conditions. Consistent with the memory-based account, the direct experience of learning the names together seems to have resulted in memories of that experience that more clearly delineated between shared and privileged information, but such representations were not as available in the context of a third party with whom they did not have this experience.

More generally, H&G's emphasis on timely access to information from memory is also consistent with a variety of findings highlighting the impact of cognitive accessibility on language production—often in ways that are independent of purely communicative considerations (for a review, see Arnold, 2010). Accessibility effects have been found in domains such as word articulation and duration (Bard & Aylett, 2004; Kahn & Arnold, 2015), syntax (Ferreira & Dell, 2000; Slevc, 2011), reference production (Bard, Hill, Foster, & Arai, 2014; Fukumura & van Gompel, 2012; Knutsen & Le Bigot, 2012), and lexical choices in expert–novice interactions (Jucks, Becker, & Bromme, 2008). Not all of these findings, of course, emerge from situations involving access to information stored in memory. For example, simple visual accessibility in the discourse context can influence speakers' decisions when and how to refer to entities independent of addressees' access to the same information (Fukumura, van Gompel, & Pickering, 2010; Jucks et al., 2008; Knutsen & Le Bigot, 2012). Taken together, though, such findings highlight the importance of understanding how ordinary cognitive processes make information available to individuals as they produce utterances.

Beyond these existing findings, the memory-based approach has some broader implications for research on common ground. First, for both real-life conversations and experimental tasks, it is important to ask whether speakers will have had appropriate opportunities to encode information into memory. Particularly for experimental tasks,



researchers should be able to make data-based assertions that participants were able to encode with sufficient rigor the information that comprises “common ground.” Furthermore, not all participants will be equally able to encode information into memory. Researchers should thus be mindful of individual differences in memory or domain-general control functions when they draw conclusions about the use of common ground in particular task situations (Brown-Schmidt, 2009; Wardlow, 2013). Second, we must look at the time course with which various memory representations become accessible as a function of the exigencies of speech production. Theories of memory, for example, often focus on the time course with which information becomes available (e.g., Ratcliff & Starns, 2013; Tibon & Levy, 2014). We suggest that research on common ground should be mindful of the implications of those theories.

The H&G approach also focuses attention on the roles of automatic and strategic processes. Extant research leaves open the question of how researchers might properly assert that a particular utterance is the product of just automatic processes or both automatic and strategic processes. In experimental tasks, researchers have often made claims for automaticity based on the details of particular paradigms, most commonly eye tracking (Brown-Schmidt, Gunlogson, & Tanenhaus, 2008; Ferguson & Breheny, 2012; Hanna, Tanenhaus, & Trueswell, 2003; Keysar, Barr, Balin, & Brauner, 2000). In these studies, early eye movements to perspective-appropriate information, for example, is taken as evidence that common ground has an immediate and potentially automatic influence upon language processing. Still, our memory-based approach should lead to more general predictions that rely on configurations of memory cues. For example, with respect to Horton and Gerrig’s (2005b) paradigm with orthogonal versus overlapping card displays, we could assert that speakers’ utterances likely reflected strategic processes (in addition to automatic processes) when the sets overlapped.

Finally, we might ask what evidence could count *against* the claim that automatic uses of common ground emerge only from ordinary memory processes. To disprove that claim, researchers would have to document examples of audience design that could not be attributed to ordinary memory processes. In sum, researchers could provide evidence for types of representations not used in other domains, or they could provide empirical evidence that specialized memory processes devoted to encoding and retrieving person-specific information only function in the domain of audience design.

#### **4. Clarifications and extensions of the memory-based view**

H&G outlined the dual processes of commonality assessment and message formation, and for each we illustrated possibilities for automatic and strategic assessment of common ground to influence speakers’ utterance planning. Thus, the memory-based view is not a prescriptive theory: Rather than suggesting what speakers *must* do, we outlined a range of possibilities. In fact, our strongest claims centered on circumstances in which information from long-term memory might not be sufficiently accessible to influence language production in the moment. We noted, for example, that “when associations between

individuals and other information are weak (or missing altogether), the processes of commonality assessment will not be able to reveal, within any reasonable time course, that certain information can be treated as co-present” (p. 18). Thus, speakers’ ability to design their utterances felicitously will depend substantially on conversation- or task-specific variations in the accessibility of memory representations.

We re-emphasize this analysis because we wish to consider some instances in which other researchers have attached prescriptive claims to H&G. As we shall see, in some cases researchers have attributed weaker claims than we intended for our memory-based approach. In other cases, the claims have been stronger (i.e., more restrictive). We hope that, by reviewing these cases in which the memory-based theory has been represented in ways that depart from H&G, we can help to clarify the general theoretical territory relevant to our claims about memory and audience design.

In several instances, researchers have weakened H&G’s claims by glossing the account as focusing on “mere association” (Barr, Jackson, & Phillips, 2014, p. 405). In fact, the critical contrast between H&G and earlier theories was on the types of representations, rather than their content. As we have noted, H&G rejected Clark and Marshall’s (1978, 1981) assertion that common ground requires reference diaries. We argued, instead, that both commonality assessment and message formation could function by drawing upon “the ordinary episodic memory traces that people encode as experiences unfold” (p. 8). Indeed, we worked through several examples of speakers’ use of information encoded in rich episodic representations. On our view, these “ordinary episodic memory traces” encode the same types of information that Clark and Marshall suggested reference diaries would encode, with the notable exception of information specifically about triple co-presence. To be sure, our account highlighted the importance of associations in memory. However, we did not intend to suggest that such associations were sufficient to support common ground assessment in all cases.

Nonetheless, the claim of “mere association” has made several appearances. For example, Brown-Schmidt (2012) introduced the memory-based account as asserting “that simple associations between partners and referents can facilitate language use in ways that may relieve interlocutors of the need to query explicit representations of joint knowledge” (pp. 64–65). In that context, she suggested, “What this research does not address, however, is whether representations of common ground itself might be of the rich, diary-like, episodic representations originally suggested (Clark & Marshall, 1978, 1981), and if so, whether these rich representations might ever play a role in real-time processing” (p. 65). In fact, the contrast between H&G and Clark and Marshall is only apt if “diarylike” has special meaning, aside from the content that people ordinarily encode in episodic representations. Although we do not presume purposeful, dedicated encoding of triple co-presence for evaluating common ground, we do believe in “rich” representations. In fact, as we noted earlier, resonance may enhance the accessibility of the entire contents of relevant memory traces.

A similar misunderstanding has appeared in research bringing neuroscience evidence to bear on issues of common ground. Rubin, Brown-Schmidt, Duff, Tranel, and Cohen (2011) conducted research with participants who had hippocampal amnesia, obtaining



evidence to support the conclusion that common ground implicates declarative memory. However, in their theoretical development, Rubin et al. described H&G, thusly: “Horton and Gerrig (2005a,b) suggested that low-level, automatic, cue-based associations between partners may serve as the basis for much common ground. Although they do not specify the kind of memory that would support such associations, they contrast it with the ‘reference diary’ kind of explicit representation advocated by Clark and his colleagues (e.g., Clark & Marshall, 1978)” (p. 1575). As we have noted, H&G clearly proposed ordinary episodic memory traces as “the kind of memory that would support such associations.” In other neuroscience contexts, this aspect of our theoretical position has been recognized. For example, Gupta, Tranel, and Duff (2012) acknowledged that the memory-based processing account of audience design relies on declarative memory.

Researchers have also attributed claims to H&G that are more restrictive than what we intended. For example, there has been disagreement with respect to whether our memory-based theory counts as an *egocentric* theory. A variety of studies have obtained findings suggesting that language processing is strongly influenced by an individual’s own knowledge, at least during the earliest stages of processing (e.g., Keysar et al., 2000). Based on such findings, Keysar and colleagues (Barr & Keysar, 2006; Horton & Keysar, 1996) proposed an *anchoring and adjustment* account of audience design in which utterance plans are initially egocentric, with speakers only making adjustments based on (often effortful) assessments of common ground in a subsequent stage of processing.

H&G should not be seen as entailing a commitment to a distinctly egocentric component to production (e.g., Köymen, Schmerse, Lieven, & Tomasello, 2014). On our view, utterance plans, from the earliest moments of speech planning, will indeed generally be influenced by information readily accessible from memory, and sometimes this will consist primarily of an individual’s private knowledge. However, other types of context-relevant knowledge can also become accessible at the same time, through the same means. In this respect, the memory-based account eliminates the need to postulate separate stages. Note that, as with anchoring and adjustment, the memory-based approach emphasizes the importance of considering the time course with which language users gain access to information relevant to common ground. However, we conclude that although speakers may sometimes appear to be egocentric as a function of information that is most accessible from memory, that appearance does not reflect a design feature of language production.

In fact, one of our key claims was that memory cues will often enable speakers to produce appropriate utterances without any strategic attempts to assess common ground. This aspect of H&G has also sometimes been misconstrued. For example, Shintel and Keysar (2009) offered their own memory-based account of audience design. To individuate their claims, they offered this footnote: “Horton and Gerrig (2005b) suggest that individuals are trying to assess common ground, but are constrained by memory (e.g., p. 141). In that sense, their account differs from ours” (p. 271). In fact, an important claim of H&G was that felicitous utterances often emerge without speakers “trying to assess common ground.” Moreover, in our theory, the constraint from memory refers to the structure and content of speakers’ representations. As we noted earlier, Horton and Gerrig (2005b) reported data in which speakers showed more evidence for audience design when appropriate memory representations

were relatively more accessible. The theoretical analysis focused on information that became available without speakers' engaging in any effortful assessments of common ground. In fact, we argued that speakers could quite possibly have produced utterances that were more felicitous had they engaged strategic effort.

The fact that participants in Horton and Gerrig (2005b) could have expended more effort to assess common ground, but sometimes failed to do so, illustrates why it is important for researchers to exercise caution when drawing general conclusions about audience design. We might have used our results as an occasion to argue that speakers are insufficiently attentive to common ground. Rather, it seems more appropriate to wonder exactly what properties of the communicative situation led participants to behave as they did (cf. Horton & Gerrig, 2002).

## 5. The role of ordinary cognitive processes

A major goal of H&G was to argue that speakers' assessments of common ground did not require special cognitive processes. In a similar fashion, a variety of findings have emerged emphasizing the role of general cognitive constraints on common ground. In considering these findings, we highlight other ideas that support the general claim that common ground emerges within the limits of ordinary cognition.

One set of findings comes from work in language comprehension. Although H&G largely focused on audience design in the context of language production, the effects of memory encoding and retrieval, as domain-general processes, influence language interpretation as well (Metzing & Brennan, 2003). For example, Horton and Slaten (2012) carried out an eye tracking study in which two pre-recorded speakers alternated giving instructions to participants to click on a series of Tangram shapes presented via computer. An initial association phase established different patterns of speaker-item mappings in memory. Then in a test phase, when the current speaker was associated with only one of the two Tangrams presented on experimental trials, participants were more likely to look to that shape before the point of linguistic disambiguation. Speaker identity functioned as a cue for the retrieval of relevant mappings from memory, resulting in cue-driven anticipation of reference.

A similar result was obtained by Barr et al. (2014), who asked pairs of friends to play a communication game in which one friend, as the addressee, had to select a target person from sets of photos of individuals, some of whom were known to both friends (*shared*), others were known only to the addressee (*privileged*), whereas others were completely unfamiliar. On some trials the addressee's friend identified the target individual, whereas on other trials an experimental assistant identified the target. Importantly, the experimental procedure separated out the role of the "designer" of this message from the actual sender of that message. Examining addressees' eye fixations to the photo sets, Barr et al. (2014) found that addressees were more likely to look toward shared targets when hearing their friend's voice than the assistant's voice, regardless of whether the friend also designed the message. They interpreted their data as indicating that reference resolution relies on ordinary episodic memory processes.

Barr et al. (2014) did, however, offer a brief criticism of H&G in the context of the opening anecdote to their paper, in which they described a situation in which one partner asks another, of their son, “Is he still having problems with Kevin?” (p. 404). They suggest that the addressee may require “mental gymnastics” to determine that the referent of *Kevin* is a toddler in the son’s preschool rather than a colleague with the same name. Barr et al. (2014) assert that memory-based processing “cannot explain what led you to decide toddler Kevin was a ‘better’ candidate than adult Kevin, given that memory processes favored the latter” (p. 405). Despite their overall focus on episodic priming, this criticism suggests a misreading of the memory-based approach. H&G made a careful distinction between automatic and strategic processes, and also noted instances in which automatic processes can apparently yield errors (i.e., infelicitous utterances). Just so, we would expect that listener’s ordinary memory processes would sometimes yield incorrect hypotheses that might require strategic assessment and correction, like momentarily misunderstanding of the referent of “Kevin.”

To explain their results, Barr et al. (2014) argued that people will have encoded episodic representations that specifically associate voices and names in memory. Given evidence that people do, in fact, encode person-specific information for spoken words (Creel & Tumlin, 2011; Goldinger, 1998), this is a compelling analysis. However, we suggest that a more general memory-based account for comprehension and common ground can be found in *compound cue* approaches to memory retrieval (see also Horton & Slaten, 2012). On these accounts, multiple cues combine to provide access to information in long-term memory (Ratcliff & McKoon, 1988). Thus, we would expect that the speaker’s identity and a particular referring expression would regularly provide a compound cue to make appropriate information in memory more accessible. (Of course, as with other applications of memory processes, compound cuing would not inevitably provide correct comprehension.) Compound cue approaches have proved valuable in a range of circumstances involving memory retrieval (e.g., Lohnas & Kahana, 2014; McKoon & Ratcliff, 2012) and are consistent with other findings showing that the joint consideration of multiple types of cues can shape the memory representations available to individuals for purposes of perspective taking (Galati & Avraamides, 2015).

This focus on compound cues coheres strongly with more general “constraint-based” accounts of perspective integration in language comprehension (for a review, see Brown-Schmidt & Hanna, 2011) in asserting that the likelihood of observing the influence of particular cues on interpretation will depend greatly upon the salience or relevance of those cues within particular contexts. For example, Brown-Schmidt (2012) introduced the notion of “gradient” representations of common ground, suggesting that particular context-dependent discourse factors, such as listeners being asked a specific question about object identity, could result in more robust representations of partner–referent associations. Some communicative situations will provide individuals with stronger evidence for what other people know, allowing them to act accordingly (Brown-Schmidt et al., 2008; Hanna et al., 2003). Gann and Barr (2014) suggested a similar idea in the context of production by asserting that speakers might “use the strength of the memory signal as a cue to determine how much effort they allocate to planning” (p. 757). Although this phrasing makes the planning process seem overly

strategic, we agree with the general point that the strength and immediacy with which relevant memory representations become accessible is likely to be crucial for facilitating attention to partner-relevant information.

This conclusion is reinforced by other work illustrating the limits on the information participants can acquire as particular experiences unfold. For example, Meagher and Fowler (2014) had participants engage in a joint map navigation task wherein, halfway through the experiment, the participants changed rooms, changed partners, changed both, or changed neither. Measurements of word durations for critical map landmarks produced both before and after these changes revealed that participants who remained in the same room tended to shorten words, whereas participants who switched rooms showed evidence for lengthening after the change. The new room appeared to render previous utterances less accessible, resulting in longer word durations. Changing partners, however, did not perturb the overall tendency to attenuate word length across repetitions. The fact that word duration showed stronger context sensitivity to room changes than to partner changes would seem, at first blush, to go against the predictions of audience design based on partner-specific associations in memory. However, Meagher and Fowler suggested that partner-based memory traces may not have been strong enough in the circumstances instantiated in their study (in which the partner was behind a barrier and partner role mattered more than partner identity) to influence retrieval in ways that could have a reliable impact upon word production.

Similar evidence comes from work exploring the development of episodic memories in young children. Newcombe, Balcomb, Ferrara, Hansen, and Koski (2014) found that young children could bind particular contextual information (such as features of the room or the experimenter name) to their episodic memories of a toy being hidden in a particular location. However, as memory cues for recall of the hidden toy, these types of contextual associations were relatively fragile compared to features of the toy itself. More generally, the importance of understanding how strongly particular circumstances support the retrieval of context-relevant information for purposes of audience design is reinforced by the findings from Brown-Schmidt and Horton (2014), which failed to replicate Horton's (2007) facilitative effect of partner-specific associations on picture naming. The relatively arbitrary partner-item associations instantiated in that paradigm may have been too tenuous to allow partner identity to provide a reliable cue to appropriate memory representations.

Researchers have also examined other types of cognitive processes to argue against special-process accounts of audience design. For example, Wardlow Lane and Ferreira (2008) demonstrated the importance of general mechanisms of attention allocation within a privileged knowledge paradigm (see also Brennan & Hanna, 2009). In their studies, speakers sometimes described visually co-present objects in ways that were influenced by the presence of objects that the addressee could not see. In their Experiment 3, Wardlow Lane and Ferreira created circumstances that changed the salience of those privileged objects and found that speakers' utterances were more affected by privileged information when it had been made highly salient. They glossed their results as consistent with the philosophy of the memory-based account of commonality assessment: "Horton and Gerrig focused on the effects of memory retrieval, showing that when a cue to memory retrieval (namely, a particular conversational participant) was highly effective for retrieving

partner-specific information, speakers produced utterances in accordance with addressee knowledge. The present results instead focus on a different general cognitive mechanism, namely attention allocation” (p. 1479).

Finally, a strong emphasis on domain-general mechanisms motivates the influential *interactive alignment* model proposed by Pickering and Garrod (2004, 2013). In this model, felicitous utterances are facilitated by low-level priming mechanisms, which cause interlocutors’ discourse models to become more similar to one another over the course of an interaction. As described by Pickering and Garrod (2004), this activation of similar discourse representations across individuals can be seen as establishing the basis for an “implicit” common ground, enabling speakers in dialog to produce and understand context-appropriate utterances without having to construct meta-representations to capture how their knowledge overlaps. Pickering and Garrod’s description of cross-interlocutor priming of relevant linguistic representations is similar to claims made for resonance as part of the memory-based processing approach to common ground. The difference is mostly one of focus: Whereas interactive alignment describes the priming of representations currently active in working memory, the memory-based approach emphasizes the cue-based retrieval of information from long-term memory.

## 6. Conclusions

In general, speakers have the goal of producing utterances that are felicitously designed for particular addressees. H&G represented an attempt to explain speakers’ successes and failures at achieving this goal with reference only to ordinary memory representations and processes. In this article, we have reviewed evidence consistent with our proposal and have also noted instances in which our proposal has been misconstrued, with the hope of clarifying the general theoretical landscape with respect to audience design. Meanwhile, several areas of research support the contention that audience design relies on ordinary cognitive processes—and importantly, these ordinary processes are likely to go beyond the memory processes featured in our original analysis. In addition, we described extensions of memory-based processing to the realm of message interpretation. Overall, we are encouraged by the range of accumulated evidence that highlights how memory access in particular, and cognitive access more generally, can influence fundamental aspects of language processing. We remain impressed by the ways in which subtle adjustments by both speakers and addressees can emerge from these ordinary processes.

## References

- Arnold, J. E. (2010). How speakers refer: The role of accessibility. *Language and Linguistics Compass*, 4, 187–203.
- Bard, E. G., & Aylett, M. (2004). Referential form, duration, and modeling 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, MA: MIT Press.



- Bard, E. G., Hill, R. L., Foster, M. E., & Arai, M. (2014). Tuning accessibility of referring expressions in situated dialogue. *Language, Cognition and Neuroscience*, 29, 928–949.
- Barr, D. J., Jackson, L., & Phillips, I. (2014). Using a voice to put a name to a face: The psycholinguistics of proper name comprehension. *Journal of Experimental Psychology: General*, 143, 404–413.
- Barr, D. J., & Keysar, B. (2006). Perspective taking and the coordination of meaning in language use. In M. J. Traxler & M. A. Gernsbacher (Eds.), *Handbook of psycholinguistics* (2nd ed., pp. 901–938). Amsterdam: Academic Press.
- Brennan, S. E., & Hanna, J. E. (2009). Partner-specific adaptation in dialogue. *Topics in Cognitive Science*, 1, 274–291.
- Brown-Schmidt, S. (2009). The role of executive function in perspective taking during online language comprehension. *Psychonomic Bulletin & Review*, 16, 893–900.
- Brown-Schmidt, S. (2012). Beyond common and privileged: Gradient representations of common ground in real-time language use. *Language and Cognitive Processes*, 27, 62–89.
- Brown-Schmidt, S., Gunlogson, C., & Tanenhaus, M. K. (2008). Addressees distinguish shared from private information when interpreting questions during interactive conversation. *Cognition*, 107, 1122–1134.
- Brown-Schmidt, S., & Hanna, J. E. (2011). Talking in another person's shoes: Incremental perspective-taking in language processing. *Dialogue and Discourse*, 2, 11–33.
- Brown-Schmidt, S., & Horton, W. S. (2014). The influence of partner-specific memory associations on picture naming: A failure to replicate Horton (2007). *PLoS ONE*, 9(10), e109035. doi:10.1371/journal.pone.0109035
- Clark, H. H. (1996). *Using language*. Cambridge, UK: Cambridge University Press.
- Clark, H. H., & Marshall, C. R. (1978). Reference diaries. In D. L. Waltz (Ed.), *Theoretical issues in natural language processing* (Vol. II, pp. 57–63). New York: Association for Computing Machinery.
- Clark, H. H., & Marshall, C. R. (1981). Definite reference and mutual knowledge. In A. K. Joshi, B. L. Webber, & I. A. Sag (Eds.), *Elements of discourse understanding* (pp. 10–63). Cambridge, UK: Cambridge University Press.
- Creel, S. C., & Tumlin, M. A. (2011). On-line acoustic and semantic interpretation of talker information. *Journal of Memory and Language*, 65, 264–285.
- Ferguson, H. J., & Breheny, R. (2012). Listeners' eyes reveal spontaneous sensitivity to others' perspectives. *Journal of Experimental Social Psychology*, 48, 257–263.
- Ferreira, V. S., & Dell, G. S. (2000). The effect of ambiguity and lexical availability on syntactic and lexical production. *Cognitive Psychology*, 40, 296–340.
- 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., & Pickering, M. J. (2010). The use of visual context during the production of referring expressions. *Quarterly Journal of Experimental Psychology*, 63, 1700–1715.
- Galati, A., & Avraamides, M. N. (2015). Social and representational cues jointly influence spatial perspective-taking. *Cognitive Science*, 39, 739–765.
- Gann, T. M., & Barr, D. J. (2014). Speaking from experience: Audience design as expert performance. *Language, Cognition and Neuroscience*, 29, 744–760.
- Gillund, G., & Shiffrin, R. M. (1984). A retrieval model for both recognition and recall. *Psychological Review*, 19, 1–65.
- Goldinger, S. D. (1998). Echoes of echoes? An episodic theory of lexical access. *Psychological Review*, 105, 251–279.
- Gorman, K. S., Gegg-Harrison, W., Marsh, C. R., & Tanenhaus, M. K. (2013). 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. L. Morgan (Eds.), *Syntax and semantics: Speech acts* (Vol. 3, pp. 41–58). New York: Academic Press.



- Gupta, R., Tranel, D., & Duff, M. C. (2012). Ventromedial prefrontal cortex damage does not impair the development and use of common ground in social interaction: Implications for cognitive theory of mind. *Neuropsychologia*, 50, 145–152.
- Hanna, J. E., Tanenhaus, M. K., & Trueswell, J. C. (2003). The effects of common ground and perspective on domains of referential interpretation. *Journal of Memory and Language*, 49, 43–61.
- Hintzman, D. (1986). "Schema abstraction" in a multiple-trace memory model. *Psychological Review*, 93, 411–428.
- Horton, W. S. (2007). The influence of partner-specific memory associations on language production: Evidence from picture naming. *Language and Cognitive Processes*, 22, 1114–1139.
- 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. (2005a). Conversational common ground and memory processes in language production. *Discourse Processes*, 40, 1–35.
- Horton, W. S., & Gerrig, R. J. (2005b). The impact of memory demands on audience design during language production. *Cognition*, 96, 127–142.
- Horton, W. S., & Keysar, B. (1996). When do speakers take into account common ground? *Cognition*, 59, 91–117.
- Horton, W. S., & Slaten, D. G. (2012). Anticipating who will say what: The influence of speaker-specific memory associations on reference resolution. *Memory and Cognition*, 40, 113–126.
- Isaacs, E., & Clark, H. H. (1990). References in conversation between experts and novices. *Journal of Experimental Psychology: General*, 116, 26–37.
- 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.
- Kahn, J. M., & Arnold, J. E. (2015). Articulatory and lexical repetition effects on durational reduction: Speaker experience vs. common ground. *Language, Cognition and Neuroscience*, 30, 103–119.
- Keysar, B., Barr, D. J., Balin, J. A., & Brauner, J. S. (2000). Taking perspective in conversation: The role of mutual knowledge in comprehension. *Psychological Science*, 11, 32–39.
- Knutsen, D., & Le Bigot, L. (2012). Managing dialogue: How information availability affects collaborative reference production. *Journal of Memory and Language*, 67, 326–341.
- Köymen, B., Schmerse, D., Lieven, E., & Tomasello, M. (2014). Young children create partner-specific referential pacts with peers. *Developmental Psychology*, 50, 2334–2342.
- Lohnas, L. J., & Kahana, M. J. (2014). Compound cuing in free recall. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 40, 12–24.
- McKoon, G., & Ratcliff, R. (2012). Aging and IQ effects on associative recognition and priming in item recognition. *Journal of Memory and Language*, 66, 416–437.
- Meagher, B. R., & Fowler, C. A. (2014). Embedded articulation: Shifts in location influence speech production. *Language, Cognition and Neuroscience*, 29, 561–567.
- Metzing, C., & Brennan, S. E. (2003). When conceptual pacts are broken: Partner-specific effects on the comprehension of referring expressions. *Journal of Memory and Language*, 49, 201–213.
- Newcombe, N. S., Balcomb, F., Ferrara, K., Hansen, M., & Koski, J. (2014). Two rooms, two representations? Episodic-like memory in toddlers and preschoolers. *Developmental Science*, 17, 743–756.
- Pickering, M. J., & Garrod, S. (2004). Toward a mechanistic psychology of dialogue. *Behavioral and Brain Sciences*, 27, 167–226.
- Pickering, M. J., & Garrod, S. (2013). An integrated theory of production and comprehension. *Behavioral and Brain Sciences*, 36, 329–392.
- Ratcliff, R. (1978). A theory of memory retrieval. *Psychological Review*, 85, 59–108.
- Ratcliff, R., & McKoon, G. (1988). A retrieval theory of priming in memory. *Psychological Review*, 95, 385–408.
- Ratcliff, R., & Starns, J. J. (2013). Modeling confidence judgments, response times, and multiple choices in decision making: Recognition memory and motion discrimination. *Psychological Review*, 120, 697–719.

- Rubin, R. D., Brown-Schmidt, S., Duff, M. C., Tranel, D., & Cohen, N. J. (2011). How do I remember that I know you know that I know? *Psychological Science*, 22, 1574–1582.
- Schiffer, S. (1972). *Meaning*. Oxford, UK: Clarendon Press.
- Shintel, H., & Keysar, B. (2009). Less is more: A minimalist account of joint action in communication. *Topics in Cognitive Science*, 1, 260–273.
- Slevc, L. R. (2011). Saying what's on your mind: Working memory effects on sentence production. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 37, 1503–1514.
- Snow, C. E. (1972). Mothers' speech to children learning language. *Child Development*, 43, 549–565.
- Tibon, R., & Levy, D. A. (2014). The time course of episodic associative retrieval: Electrophysiological correlates of cued recall of unimodal and crossmodal pair-associate learning. *Cognitive, Affective, & Behavioral Neuroscience*, 14, 220–235.
- Wardlow, L. (2013). Individual differences in speakers' perspective taking: The roles of executive control and working memory. *Psychonomic Bulletin & Review*, 20, 766–772.
- Wardlow Lane, L., & Ferreira, V. S. (2008). Speaker-external versus speaker-internal forces on utterance form: Do cognitive demands override threats to referential success? *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 34, 1466–1481.