Who Tunes Accessibility of Referring Expressions in Task-Related Dialogue?

Ellen Gurman Bard
LEL,
University of Edinburgh,
Edinburgh EH8 9LL, UK
ellen@ling.ed.ac.uk

Robin Hill
HCRC,
University of Edinburgh,
Edinburgh EH8 9LW, UK
r.l.hill@ed.ac.uk

Mary Ellen Foster
Informatik VI,
Technische Universität,
München, Germany
foster@in.tum.de

Abstract

Ariel (1988; 1990; 2001) has proposed that the grammatical form of an anaphor can be predicted from the 'deemed' accessibility of its antecedent. The element of judgment in the term 'deemed' is critical: it allows the speaker to reflect an egocentric perspective and frees choice of expression from the actual contingencies of the situation in which it is uttered. Using a screen-based joint tangram construction task (Carletta et al., under revision), we examine the accessibility of 1775 introductory mentions for effects of situation (communication modalities and actions involving the named entity) and of responsibilities assigned to the participants. We find statistically significant effects of three kinds: circumstances readily available to the listener (concurrent movement of the named object); circumstances private to the speaker (hovering the mouse over the object, when the listener cannot see the mouse), and the speaker's role in the joint task. Since egocentrically selected forms may be underspecified, we make a preliminary attempt to discover whether referring expression usage is disabling or irrelevant.

1 Introduction

The question of what a thing shall be called has engaged psychologists and linguists as much as it engages anyone attempting automatic interpretation or generation of referring expressions (Brown, 1958; Dale & Reiter, 1995; Gundel, Hedberg, & Zacharski, 1993; Kranstedt, Lücking, Pfeiffer, Rieser, & Wachsmuth, 2006; Lyons, 1977; Prince, 1981; Van der Sluis & Krahmer, 2007; Walker & Prince, 1996). One very wide-ranging approach, (Ariel, 1988, 1990, 2001), attempts to key elaboration of the form of referring expressions to the 'deemed' accessibility of the referent, that is, to how difficult the producer of the expression estimates it will be to access the referent concept, discourse entity, or extra-linguistic object. Expressions introducing entities deemed completely unfamiliar to the audience should be maximally detailed indefinite NPs including modifiers of various kinds, as in (1). Expressions of intermediate accessibility might be marked by definite articles, deictic expressions, or personal pronouns in that order. Expressions making reference to a single most immediately mentioned entity in focus can be as minimal as so-called clitics (2), unstressed and all but deleted pronouns, or even zero forms (3).

(1) *A Republican governor of a strongly Democratic state.*
(2) A: Where's Arthur?
B: /*z*/ in the garage.
(3) A. And your younger son?
B. *{-}* playing Internet poker.

Accessibility theory provides a unified framework for predicting how forms of referring expressions will respond to givenness, discourse focus, inferrability from local scenarios and the like. As a general notion, accessibility ought to include effects of any available conditions which might draw attention to the correct referent, whatever modality delivers them and whether they are internal or external to the discourse. This paper discusses the accessibility of referring expressions produced during a joint construction task and examines two factors which might draw attention to the correct referent, task related movements and the roles of the participants.

The origins of our questions about these factors lie in the information which human interlocutors might use in determining how to refer. Ariel's notion of accessibility appears to depend on what the speaker supposes is the case, not on what is genuinely easier or more difficult for the

Proceedings of the 12th Workshop on the Semantics and Pragmatics of Dialogue, June 2–4, 2008, London, U.K.

listener. While some approaches to dialogue assume that speakers carefully model their interlocutors, so that initial forms of expression could arise from the interlocutors' needs (Brennan & Clark, 1996; Clark & Krych, 2004; Schober, 1993), there is increasing evidence that we have limited ability to construct, recall, or deploy any such model in a timely fashion (Bard, Anderson et al., 2007; Horton & Gerrig, 2002, 2005a, 2005b; Horton & Keysar, 1996). Interlocutors may behave egocentrically (Bard et al., 2000; Bard & Aylett, 2004; Horton & Keysar, 1996), adopt a global account of affordances of a situation, (Anderson, Bard, Sotillo, Newlands, & Doherty-Sneddon, 1997; Brennan, Chen, Dickinson, Neider, & Zelinsky, In press), or observe information indicative of the listener's knowledge, but fail to act on it (Bard, Anderson et al., 2007; Brennan et al., In press).

The situation for form of referring expressions is mixed. While accessibility of referring expressions is more sensitive to the knowledge of the listener than is clarity of articulation (Bard & Aylett, 2004), other studies show that tendencies to match nomenclature to listener's history or current situation are quite variable (Brennan & Clark, 1996; Horton & Gerrig, 2002, 2005a; Horton & Keysar, 1996; Keysar, Lin, & Barr, 2003). So-called conceptual pacts are actually lexical pacts (Brennan & Clark, 1996), agreements to call objects by certain names, and are the result of negotiation over time, across which accessibility of the referring expression naturally rises. If speakers do track one another's internal states, the accessibility of even introductory mentions will suit the interlocutor's current needs, rather than the speaker's.

The evidence may be inconclusive because the typical paradigms for dialogue studies restrict cooperation to disjointed episodes. Often one participant instructs another to act on or select from an array of potential referents, while the other follows instructions relative to an identical or partially overlapping array. Both responsibilities and activities are clearly distinct. Even when players ultimately exchange roles, the roles are inherently asymmetrical: one has more information and more power to design the communication than the other. Channels for communication are purposely limited; and the knowledge shared between instructor and listener is altered trial by trial in an unpredictable way. To discover whether more robustly cooperative behaviour appears in more cooperative tasks, we have created a corpus of dialogues centred around a shared task which demands joint attention but makes it possible to vary the participants roles.

To study joint action as a model for human-robot cooperation in quasi-industrial settings, the JAST project has developed the Joint Construction Task (Carletta et al., under revision) in which two human players cooperate to construct a two-dimensional tangram on their yoked screens (Figure 1). Each player can manipulate the component parts by mouse actions. Each dyad is assigned to work either with roles (one player managing the task and the other assisting) or without. Mouse actions draw attention not just because they are integral to the construction process, but because, to mimic industrial risks, they are dangerous. If both mice touch the same object, or if two objects overlap, both break. Because each player can act on the tangram parts and sub-constructions, the activity of grasping or moving the named object adds a haptic or praxic modality to spoken forms. Even 'hovering' the mouse over a part without grasping it offers a chance to make a part accessible. The paradigm suggests how accessible initial mentions should be: because tangram parts come in identical pairs, a felicitous first mention should in theory be an indefinite expression like (4) or (5)

(4) Let's get *a red square*

(5) We could try *one pink triangle* first.

To discover how well keyed any change in form of referring expression is to the perceptions of the interlocutor, the design contrasts situations in which each player's mouse cursor is projected onto the other's screen with situations in which each player can see only the resulting movement of the object which the other's mouse 'grasps'. Only in the first case can a player see the mouse 'hovering' over a tangram part which is not actually moving.

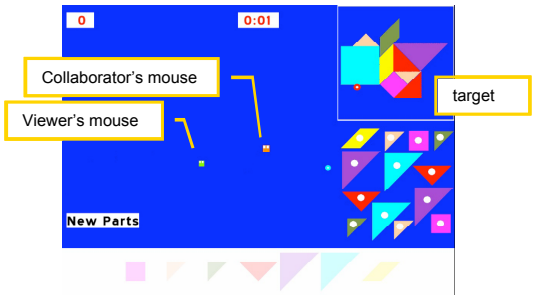

Figure 1. Joint Construction Task shared screen.

If moving a part draws attention, it should also give rise to referring expressions of greater accessibility. Since pointing is associated with shorter, less detailed referring expressions and pointing to closer targets has an even stronger effect (Kranstedt et al., 2006), touching and moving should have a very marked effect on the form of expression. Like Kranstedt et al., and exactly as the definition of deixis would predict (Lyons, 1977), we note the association of the 'hand' location and verbal deixis: in our case a larger proportion of verbal deictics (*this square; these, mine*) than of other forms of expression coincide with mouse-referent overlap (Foster et al., 2008).

To go further, we need to divide overlaps into those where the mouse is moving a part and those where it is merely hovering over it. If the listener's knowledge is of concern, a speaker moving parts and a speaker hovering over them should sometimes make different selections of accessibility level. Since movements of objects will always be visible to the listener in this paradigm, a speaker adjusting to listener knowledge could certainly use deictic forms to refer to parts she is currently moving. In contrast, visibility of the mouse cursor should determine whether a hovering mouse makes an object more accessible: Only when the hovering mouse cursor is visible to the listener can a speaker use it to point to the named object and select a more accessible referring expression. In fact, a speaker might even increase the accessibility of an expression referring to a part which the *listener* is visibly touching or moving. When the hovering mouse is not cross-projected, a listener-sensitive speaker cannot use it to point. If the listener's knowledge is less important than the speaker's, however, the speaker's own hovering movements should attract higher accessibility forms regardless of what the listener can see.

The players' roles suggest further questions. Managers have a primary role in setting the dyad's agenda. They should have more power to designate discourse focus and to change it, for example. If the choice of accessibility level is an overt designation on the speaker's part, then managers should have special powers of designation. Moreover, as we suggested earlier, managers might have less reason to track or adjust to the needs of their partners than role-less players do. Conversely, assistants should have more reason to adjust to the manager's precedents.

In all cases, the answers to our questions should be reflected in distributions of referring expressions across ordered levels of accessibility. Though accessibility bears on the relationships between earlier and later mentions of an entity, it ought to be important to determining the form of introductory mentions, too. By restricting our investigation in this way, and by controlling the objects available for naming, we can test our hypotheses about how a thing shall first be called.

2 Corpus Collection and Coding

2.1 Task

The Joint Construction Task or JCT (Carletta et al., under revision) offers to two collaborating players a target tangram (Figure 1, top right), geometrical shapes for reproducing it (centre right), a work area (centre screen), a counter for breakages (top left), a set of replacement parts (bottom of the screen), and a clock measuring elapsed time (top centre). The players' task is always to construct a replica of the target tangram as quickly, as accurately, and as cheaply in terms of breakages as possible. An accuracy score (top left) appears at the end of each trial.

Participants manipulate objects by left-clicking the mouse and dragging them or by right-clicking and rotating them. Carefully timed collaboration is required. Any part or partially constructed tangram 'held' by both players will break and must be replaced from the spare parts store to complete the trial. Moving an object across another breaks both. Objects can be joined only if each is held by a different player. Objects join permanently wherever they first meet. Inadequate constructions can be purposely broken and rebuilt from spare parts, incurring a cost in both parts and time.

Players' mouse cursors differ in colour and each changes colour when it has grabbed an object, distinguishing grabbing from mere superimposition (hovering).

2.2 Apparatus

Each participant sat approximately 40cm from a separate CRT display in the same sound-attenuated room. Participants faced each other, but direct eye contact was blocked by the monitors between them. Participants were eye-tracked monocularly via two SR-Research EyeLink II head-mounted eye-trackers. Head worn microphones captured speech on individual channels. Continuous audio and video records included a full account of locations and movements of indi-

vidual parts, constructed objects, and cursors. Composite videos recorded all movements and audio.

2.3 Participants, design and materials

Sixty-four Edinburgh University students, paid to participate, were paired into 32 same-sex dyads who had never met before. Four further dyads were discarded because of technical failures. Each dyad participated in 8 experimental conditions produced by the factorial manipulation of three communication modalities: speech, gaze (each player's current eye-track cross-projected onto the other's screen), and mouse cursor (also cross-projected). Participants could always see their own mouse cursor. Without no additional communicative modalities, they saw only the moving parts. Gaze and mouse conditions were pseudo-randomised following a latin square. Speech and non-speech conditions were counterbalanced. Only conditions with speech are analyzed here.

In 16 dyads, one participant was designated manager and the other assistant. The manager was instructed to maintain speed, accuracy, and cost, and to signal the completion of each trial. The assistant was to help. The remaining dyads were assigned no roles but otherwise had the same working instructions. Trials ended when one player declared the construction complete by pressing the spacebar and the other confirmed. An accuracy score reflecting similarity between the built and the target tangrams then appeared across the built exemplar.

Each dyad reproduced 16 different tangrams, 2 per condition. No tangram resembled a nameable object. Each contained 11 parts. All trials used the same set of 13 parts, comprising 2 copies of each of 6 shape-colour combinations (squares or right-angle isosceles triangles differing in size and colour) and a single yellow parallelogram. These initially appeared in 4 different layouts counterbalanced across experimental items. The extra pieces differed from trial to trial.

2.4 Coding referring expressions

Dialogues were transcribed orthographically. Each referring expression was time-stamped for start and end points. Then each expression referring to any on-screen object was coded with a referent identifier linking it to the object. Coders had access to the video and audio track and were allowed to use any material within a dialogue to determine the referent of any expression. All referring expressions were coded for accessibility on the scale given in Table 1. This system represents a modest expansion of a system applied to an earlier corpus of task-related dialogues (Bard & Aylett, 2004) and yielding negligible disagreement between coders.

Table 1 Accessibility Coding Scheme

Level	Definition	Examples
Min	Indefinite NP	*a purple one* *one of the nearest blue pieces*
	Bare nominal	*pink one* *triangles*
	Definite NP	*the red bit* *the other purple one*
	Deictic NP	*those two little kids.*
	Deictic Possessive } Pron	*these* *mine*
Max	Other Pronouns	*it*
	Clitic/inaudible.	*-/z/*

3 Results

3.1 Overall outcome

Figure 2 presents the overall distribution of first mentions across the accessibility scale. Despite the fact that most original parts would be expected to demand an indefinite referring expression to distinguish them from an identical part, only 16% of first mentions were indefinite NPs. The remaining 84% were of higher accessibility. Our question now is whether mouse actions or speaker roles are responsible for this profile.

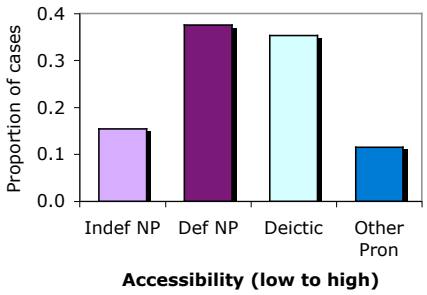

Figure 2. Accessibility of first mentions in the Joint Construction Task

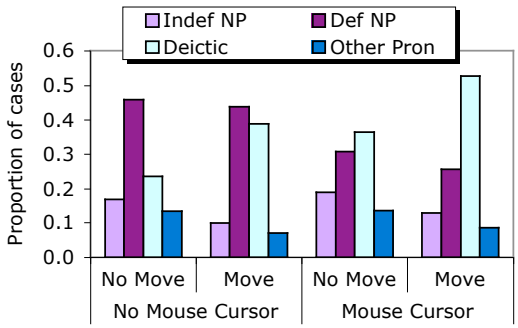

Figure 3. Accessibility of first mentions: Effects of moving the referent object

3.2 Modalities, roles, and accessibility

Method. The conditions critical to our predictions were coded for a multinomial logistic regression which modelled the distribution of first mentions across accessibility categories. This statistic tests the capacity of category variables (like Mouse v No Mouse) to influence ordinal variables (like Accessibility). It constructs regression equations both for the whole ordinal series and for the comparison of each level to some reference level. We use it to ask which actions and modalities change the tendency to produce indefinite referring expressions (the usually expected format) relative to each more accessible category.

The calculations are done on log odds, but for interpretability, we display simple proportions of cases. To reduce the number of empty cells, accessibility categories were collapsed into four levels: Indefinite NPs (including bare nominals), Definite NPs, deictics (including deictic NPs, deictic pronouns and possessive pronouns) and other pronouns (including clitics).

Separate equations were prepared for the Mouse Cursor Cross-Projected (n = 836) and No Mouse Cursor conditions (n = 939). The predictors included the experimental variable Roles Assigned, the participants' mouse actions (the speaker/listener moving part being mentioned, or 'hovering' the mouse over it), and the interactions of Roles Assigned with each movement variable. Gaze cross-projection was not included, as it had proved an ineffective predictor in earlier exploratory regressions. Table 2 shows the significant outcomes. Each effect listed is essentially independent of any effect from any concurrent predictor.

No effect of listener behaviour reached significance. There were effects of the speaker's actions and of Role Assignment.

Table 2. Significant predictors of accessibility. For individual levels of accessibility, $df = 1$. $* = p < .05$; $‡ = p < .01$; $§ = p < .001$.

No Mouse Cursor Cross-Projection						
	-2 Log Likelihood		χ^2	*df*	Cox & Snell	
	268.07		105.00§	27	0.106	
	Speaker Move		Speaker Hover		Speaker Hover x Roles Assigned	
χ^2					276.00	7.42*
	B	Wald	B	Wald	B	Wald
Definites			-1.137	10.85§	1.075	4.99*
Deictics	-0.814	4.78*			1.275	6.17*

Mouse Cursor Cross-Projected						
	-2 Log Likelihood		χ^2	*df*	Cox & Snell	
	258.00		61.34§	27	0.071	
	Speaker Move		Speaker Hover		Speaker Hover x Roles Assigned	
χ^2					266.00	7.77*
	B	Wald	B	Wald	B	Wald
Deictics	-0.722	5.05*				
Pronouns			1.264	6.95‡		

As predicted, actions available to speaker and listener were important: visibly moving the referent (Figure 3) coincided with increased deictic expressions (31% without v 46% with movement overall) at the expense of indefinites (18% v 12% overall) whether or not the speaker's mouse cursor itself was visible. Also, visibly hovering the mouse cursor over the referent (Figure 4) accompanied a significant fall in pronouns (15% v 6%) relative to indefinites (17% v 16%), with deictics the dominant category in both cases (39%, 51%).

Strikingly, actions unavailable to the listener were also important. An *in*visibly hovering mouse accompanied a shift from indefinites (17% v 10%) towards definites (42% v 50%), with the latter as the most common category.

Role assignment influenced the effects of invisible mouse gestures: Only in Manager-Assistant dialogues did introductory mentions shift markedly away from indefinites (22% v 8%) toward deictics (25% v 40%) as well as definites (36% v 48%). Figure 4 shows that in Manager-Assistant dialogues private hovering gestures gave profiles somewhere between their No Roles counterparts (where definite NPs predominate) and dialogues with projected cursors (where deictics rise with public gestures).

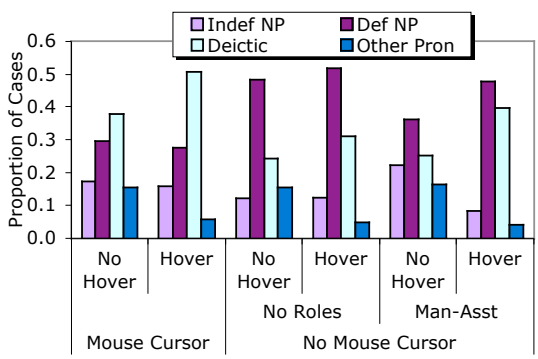

Figure 4. Effects on accessibility of hovering mouse over referent, by cursor visibility and assigned roles.

4 Discussion

This paper asked whether the association between handling a thing and using an accessible format to name it was linked to the speaker's own knowledge or to the knowledge expected to reside with the listener. There are two reasons to believe that the listener is not in charge. First, we found no significant effects of the listener's manipulation of tangram parts on the speaker's form of referring expression, even when the listener's movements were fully visible to the speaker. Second, we did find effects of speakers' mouse gestures which were invisible to the listener.

At the same time, we suggested that if accessibility is an expression of opinion, it should be manipulated by Managers in particular. In the event, Manager-Assistant dialogues showed more egocentric use of accessibility than no-role dialogues: in these dialogues the presence of a gesture invisible to interlocutor all but eliminated indefinite introductory mentions, in favour of definites and deictics.

While the effect of movement is a praxic or haptic form of deixis, the effects of private gestures and of role ought to be counterproductive. Though all the conditions examined here yield tangrams of equal similarity to their models, the costs do follow this prediction. Trials without mouse cross-projection took longer than those with it (205 v 187sec; $F_1 = 11.45$, $df = 1, 30$, $p = .002$) and incurred more breakages (1.8 v 2.3: $F_1 = 4.52$, $df = 1, 30$, $p = .008$) to achieve equal accuracy (92.1 v 91.9). Manager-Assistant trials took longer than No Role trials (216 v 175sec: $F_1 = 10.67$, $df = 1, 30$, $p = .003$) to give similar performance (Accuracy: 93.8 v 91.2; Breakages: 2.0 v 2.1). The latter finding is the stronger argument, because additional breakages require additional time to fix.

Nonetheless, the picture is far from complete. We see three major issues.

First, the results fall some way short of a clear case for managerial insensitivity. The Role Assignment results were based on expressions produced by both participants. Analyses comparing managers with assistants are made difficult by small or empty cells. Both players show the pattern found in Figure 4. Accordingly, we have no particular evidence contrasting managers with their assistants, though we can distinguish manager-assistant dyads from the dyads who had no assigned roles.

As we suggested earlier, however, one result of role differences is to give precedence to one individual. The manager decided what should happen next. To cooperate, the assistant had to conform to the manager's choices. Conforming to the manager's referential habits, for social reasons, or through structural priming, could make the assistant appear to designate with invisible gestures, too. In essence, the assistant can achieve a tendency toward use of definites or deictics where they might not otherwise appear to be unwarranted and then employ private gestures to

accompany these instances. In contrast, No Role dyads might follow a mixture of styles or compete to control the task plan or the naming habits. If so, manager and assistant should have more similar profiles in than No Role players. Quantitatively, this seems to be the case.

Second, though it is clear that speakers' private and public actions associate with particular levels of accessibility, it is not clear that their effects are all increases in accessibility. For example, Figures 3 and 4 show a tendency, significant only with hovering, for speaker actions not to collocate with the highest levels of accessibility in first mentions: pronominal or clitic introductory mentions are used less often when the mouse overlaps the referent part than when it does not. Thus, the haptic or ostensive functions of mouse movements are specific to definite and deictic usage: they literally turn *a triangle* into *this* but they do not turn *this* into *it*. For this reason, the single accessibility continuum might be viewed as the result of a set of different referential phenomena, for example, demonstration or givenness in context, bearing on speakers' choices with different degrees of force.

Finally, there is the issue of the discourse history within which the introductory mentions are set. Clearly, some first mentions do not refer to totally discourse-new or completely unpredictable entities (Prince, 1981). There is no doubt that other forces work on the choice of referring expressions.

We do not yet know how the sequence of external events – construction of the tangram, for example, affects the forms of introductory referring expressions. In theory, it is possible that corpus dialogues went well because speakers were attending to the same objects regardless of the form of referring expressions. Our eye-tracking results from this corpus suggest, however, that alignment between players was far from perfect. If players were already attending to the same objects, whatever either said, we should find very high levels of overlapping gaze. To study the relationship between interlocutors' gaze patterns, we have used a cross-recurrence analysis (Richardson, Dale, & Kirkham, 2007), which shows how behaviours can be entrained even if they are not synchronized to the extremely fine levels that eyetrackers detect. This technique shows what percentage of interlocutors' gaze fixations are on the same objects but separated by various lags in one direction or another. Almost without exception ((Bard, Hill, Nicol, & Carletta, 2007), maximal shared view was simultaneous, but it was far from complete: participants showed a maximum of 36% gaze at the same objects in conditions with speech (as against 40% without). It would seem that referring expressions still have some work to do when joint attention is required.

We began by discussing referring expressions in the light of speakers' ability to maintain models of their listeners' knowledge that update quickly enough to be the basis of initial mentions. Our speakers did not appear to use such models. Ultimately, of course, they could have had ample opportunity to produce adequate reference by subsequent joint adjustment. The additional time taken for dialogues in the more egocentric conditions suggests that this could be the case.

If so, the speakers' behaviour is another example of joint responsibility for dialogue being effectively shared rather than duplicated across interlocutors (Bard, Anderson et al., 2007). In this kind of responsibility structure, rather than a fully articulated model of common ground, a simple and risky egocentric process guides production. Speakers were free to designate invisibly just because their partners were under an obligation to object to inadequate expressions.

Acknowledgments

This work was funded by EU Project JAST (FP6-003747-IP). The authors are grateful to the JCT programmers, Tim Taylor, Craig Nicol, Joe Eddy and Jonathan Kilgour for their contributions to the software, to Jean Carletta who managed the program development, to the reference coders for guessing what the speaker had in mind, and to JP de Ruiter for helpful discussions.

References

Anderson, A. H., Bard, E. G., Sotillo, C., Newlands, A., & Doherty-Sneddon, G. (1997). Limited visual control of the intelligibility of speech in face-to-face dialogue. *Perception and Psychophysics, 59*(4), 580-592.

Ariel, M. (1988). Referring and accessibility. *Journal of Linguistics, 24*, 65-87.

Ariel, M. (1990). *Accessing Noun-Phrase Antecedents*. London.: Routledge/Croom Helm.

Ariel, M. (2001). Accessibility theory: An overview. In T. Sanders, J. Schilperoord & W. Spooren (Eds.), *Text representation:Linguistic and psycholinguistic aspects.* (pp. 29-87). Amsterdam: John Benjamins.

Bard, E. G., Anderson, A. H., Chen, Y., Nicholson, H. B. M., Havard, C., & Dalzel-Job, S. (2007). Let's you do that: Sharing the cognitive

burdens of dialogue. *Journal of Memory and Language, 57*(4), 616-641.

Bard, E. G., Anderson, A. H., Sotillo, C., Aylett, M., Doherty-Sneddon, G., & Newlands, A. (2000). Controlling the intelligibility of referring expressions in dialogue. *Journal of Memory and Language, 42*, 1-22.

Bard, E. G., & Aylett, M. P. (2004). Referential form, word duration, and modelling the listener in spoken dialogue. In J. C. Trueswell & M. K. 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., Nicol, C., & Carletta, J. (2007). *Look here: Does dialogue align gaze in dynamic joint action?* Paper presented at the AMLaP2007, Turku, Finland.

Brennan, S. E., Chen, X., Dickinson, C., Neider, M., & Zelinsky, G. (In press). Coordinating cognition: The costs and benefits of shared gaze during collaborative search. *Cognition*.

Brennan, S. E., & Clark, H. H. (1996). Conceptual pacts and lexical choice in conversation. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 11*, 1482-1493.

Brown, R. (1958). How shall a thing be called. *Psychological Review, 65*, 14-21.

Carletta, J., Nicol, C., Taylor, T., Hill, R., de Ruiter, J. P., & Bard, E. G. (under revision). Eyetracking for two-person tasks with manipulation of a virtual world. *Behavior Research Methods, Instruments, and Computers*.

Clark, H. H., & Krych, M. A. (2004). Speaking while monitoring addressees for understanding. *Journal of Memory and Language, 50*, 62–68.

Dale, R., & Reiter, E. (1995). Computational inerpretations of the Gricean maxims in the generation of referring expressions. *Cognitive Science, 19*(2), 233-263.

Foster, M. E., Bard, E. G., Guhe, M., Hill, R., Oberlander, J., & Knoll, A. (2008). *The roles of haptic-ostensive referring expressions in cooperative task-based human-robot dialogue.* Paper presented at the Human Robot Interaction, Amsterdam.

Gundel, J. K., Hedberg, N., & Zacharski, R. (1993). Cognitive status and the form of referring expressions in discourse. *Language, 69*, 274-230.

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*(4), 589-606.

Horton, W. S., & Gerrig, R. J. (2005a). Conversational common ground and memory processes in language production. *Discourse Processes, 40*(1), 1-35.

Horton, W. S., & Gerrig, R. J. (2005b). The impact of memory demands on audience design during language production. *Cognition, 96*(2), 127-142.

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

Keysar, B., Lin, S., & Barr, D. J. (2003). Limits on theory of mind use in adults. *Cognition, 89*, 25–41.

Kranstedt, A., Lücking, A., Pfeiffer, T., Rieser, H., & Wachsmuth, I. (2006). Deictic Object Reference in Task-oriented Dialogue. In G. Rickheit & I. Wachsmuth (Eds.), *Situated Communication,* (pp. 155-207). Berlin: Mouton de Gruyter.

Lyons, J. (1977). *Semantics* (Vol. 2): Cambridge, UK.

Prince, E. F. (1981). Toward a taxonomy of given-new information. In P. Cole (Ed.), *Radical Pragmatics* (pp. 223-256). New York: Academic Press.

Richardson, D. C., Dale, R., & Kirkham, N. (2007). The art of conversation is coordination: common ground and the coupling of eye movements during dialogue. *Psychological Science, 18*(5), 407-413.

Schober, M. (1993). Spatial perspective-taking in conversation. *Cognition, 47*(1), 1-24.

Van der Sluis, I., & Krahmer, E. (2007). Generating multimodal referring expressions. *Discourse Processes, 44*(3), 145-174.

Walker, M. A., & Prince, E. F. (1996). A bilateral approach to Givenness: A hearer-status algorithm and a centering algorithm. In T. Freitheim & J. Gundel (Eds.), *Reference and referent accessibility*. (pp. 291-306). Amsterdan: John Benjamins.