

Accessibility, Duration, and Modeling the Listener in Spoken Dialogue

E. G. Bard* and M. P. Aylett*[†]

* Human Communication Research Centre and Department of Theoretical and Applied Linguistics,
Adam Ferguson Building, University of Edinburgh, EH8 9LL, U.K.
ellen@ling.ed.ac.uk

[†]Centre for Speech Technology Research
2 Buccleuch Place, University of Edinburgh, Edinburgh EH8 9LW, U.K.
matthewa@cogsci.ed.ac.uk

Abstract

Referring expressions are thought to be tailored to the needs of the listener, even when those needs might be very costly to assess, but tests of this claim seldom manipulate listener's and speaker's knowledge independently. The Map Task enables us to do so. We examine two 'tailoring' changes in repeated mentions of landmark names: falling clarity of word articulation and rising accessibility of referring expression. Clarity results replicate Bard et al. (2000). Standardized word duration fell for speaker-Given listener-New items (Expt 1). Hence it was unimportant whether the listener heard an earlier mention. Reduction between mentions was no greater when it could be inferred that the listener could see the named item (Expt 2), and no less when the listener explicitly declared that they could not (Expt 3). Hence it was unimportant whether the listener could see the landmark. Reduction was unaffected by whether the repeater could see the mentioned landmark (Expt 4). Articulation thus depends only on what the speaker has heard previously. In contrast, accessibility was more sensitive both to listener (Expt 1) and speaker knowledge (Expt 4). The results conform most closely to a Dual Process model: fast, automatic, word-by-word processes let the speaker's own experience prime articulation, while computationally costly assessments of listener knowledge control influence referring expression design where competing tasks permit.

1. Introduction

Speakers are said to design their utterances to suit the needs of their listeners, insofar as those needs can be known (Ariel, 1990; Clark and Marshall, 1981; Gundel, Hedberg, and Zacharski, 1993; Lindblom, 1990). Certainly there is variation in form. Clarity of pronunciation varies with predictability from local context (Hunnicut, 1985; Lieberman, 1963) and with repeated mention (Fowler and Housum, 1987). Forms of referring expression differ in elaboration with the more readily interpreted, those having more accessible antecedents, syntactically simpler (*a blacksmith's cottage v it*) (Ariel, 1990, Fowler, Levy, and Brown, 1997; Gundel, J.K., Hedberg, N., and Zacharski, R., 1993; Vonk, Hustinx, and Simmons, 1992). Yet maintaining an incrementally updated model of what the listener knows, including the established common ground, and of what the listener needs to know is a considerable cognitive task. Because speaker's and listener's knowledge overlap and because it may be impossible to assess the listener's knowledge accurately, it is suggested that speakers often default to an account of their own knowledge as a proxy for the listener's (Clark and Marshall, 1981). In fact, many discussions of this topic simply assume that the two are the same: they describe or manipulate the speaker's knowledge without independently manipulating the listener's (see Keysar, 1997).

This paper presents two attempts to examine the hypothesis that referring expressions are genuinely tailored to the speaker's model of the addressee. One deals with the articulation of individual words, the other with the form of referring expressions. They have different implications for psychological models of dialogue. Current models of the production of language indicate that noun phrase structure and articulation are generated within units of different sizes, phonological phrases or tone

groups on the one hand and phonological words, lexical words, or syllables on the other (Wheeldon and Lahiri, 1997; Levelt and Wheeldon, 1994). Speech appears to be produced in a cascade, with smaller units being prepared for articulation as the succeeding larger unit is being designed. Thus, incrementally updating a listener model in order to articulate each phonological word would impose a much heavier computational burden on a speaker, than updating it phrase by phrase.

We will first present hypotheses which the literature offers us for the way in which speakers manage the task of modeling listeners while planning and producing speech. Then we will report two studies which test these hypotheses on materials from a single corpus. Both make the same comparisons. Bard et al. (2000) excised a balanced sample of spontaneously uttered words, and measured their intelligibility to naïve listeners as well as their duration. The new results in the present paper report duration and accessibility of all suitable referring expressions in the corpus. Finally, we will discuss the implications of the comparison for the nature of listener modeling in on-line utterance generation.

2. Modeling listener knowledge while speaking

The literature offers several versions of the hypothesis that what we say is tailored to the needs of our listeners. They can be arranged in order of the computational demands they would impose on speakers.

Lindblom's *H-and-H Hypothesis* (1990) makes the heaviest demands. It posits that speakers adjust the articulation of spoken words to the knowledge which the listener can currently recruit to decoding the speech signal. Thus, speakers hyper-articulate when listeners lack such auxiliary information and hypo-articulate when redundancy is high. There is ample evidence that linguistic environments which provide more redundancy

contain word tokens articulated with greater speed and less precision. (Bard and Anderson, 1983, 1994; Fowler and Housum, 1987; Lieberman, 1963; Samuel and Troicki, 1998). The question is whether this relationship depends on the speakers updating and consulting a model of the listener's current knowledge each time they prepare the prosodic character of a phonological word or the articulation of its syllables. Though the H-and-H view does not preclude defaulting to the speaker's own knowledge, it is framed under in terms of genuine listener knowledge. To adjust articulation on line to a non-default account of listener knowledge, speakers should observe listeners continuously for signs of misunderstanding or disagreement. Wherever speaker's knowledge and listener's knowledge differ, listener's knowledge should take precedence. In effect the H-and-H Hypothesis corresponds to a *Negligible Defaulting Hypothesis*.

The second alternative arose from a consideration of how speakers might manage the many tasks involved in generating appropriate utterances in dialogue. Brown and Dell (1987) propose a modular division between the initial formulation of utterances, a process based on speaker knowledge, and the monitoring and revision of output, processes based on a model of listener knowledge, or more precisely, of common ground. Called the *Monitoring and Adjustment Hypothesis* by Horton and Keysar (1996), this model defaults first and pays later – if necessary. Because responsibility for tailoring utterances to the listener's needs is shared by the interlocutors (Carletta and Mellish, 1996), the speaker's attention can initially be devoted to utterance planning rather than to listener modeling. Faultless utterances, those for which speaker- and listener-knowledge are alike, are produced quickly and accurately. Poorly designed utterances can be revised in response to explicit requests from the listener, which are received well after the initial planning of the faulty utterance is complete. If the Monitoring and Adjustment Hypothesis holds, post-feedback utterances should reflect any aspects of listener-knowledge which the feedback has conveyed. Otherwise, listener-knowledge should be irrelevant to production.

The third proposal deals with co-presence, middle- or long-term characteristics of listeners which affect likely overlap with speakers' own knowledge. Various kinds of 'co-presence' in social or regional background (Isaacs & Clark, 1987; Fussell & Krauss, 1992), physical location during the interaction (Schober, 1993), or recent experiences (Schober & Clark, 1989; Wilkes-Gibbs & Clark, 1992; Brennan & Clark, 1996) are taken into account. Although this work is usually interpreted as showing that the 'initial design' (Horton & Keysar, 1996) of conversational speech is sensitive to listeners' needs, it does not directly address on-line processes. Since most discussions of this notion focus on lasting characteristics of listeners, we assume that it is also intended to reduce the number of occasions in a dialogue when a speaker must update a model of the listener. If so, speakers should attend to evidence for and against co-presence, and defaults could hold for some undefined time after positive evidence. We will call this the *Co-presence Default Hypothesis*.

Finally, Bard et al. (2000) develop a suggestion of Brown and Dell (1987) which we will call the *Dual Process Model*. It proposes a division between fast and automatic processes, which have no computational cost,

and slower, more costly processes requiring inference or attention. The former include priming, an unconscious process that allows the performance of an activity or the recognition of a stimulus to reduce the reaction time for or the duration of a behaviour. (Balota, Boland, and Shields, 1989; McKoon and Ratcliff, 1980; Mitchell and Brown, 1988; O'Seaghdha, 1997.). Only the speaker's own experience is effective in priming. The latter include all those complex forms of reasoning usually implicated in the ability to construct a model of the listener. In competition with this set are the computations which underlie the ability to plan a dialogue or keep track of a shared task. When there is competition for time and attention, the second set of processes may suffer (Horton and Keysar, 1996), leaving the speaker with only cost-free defaults in the form of his own knowledge.

Of these four hypotheses, the second and third make roughly the same predictions for speakers' ability to tailor form of referring expression and word articulation to the listener's needs. Where speaker and listener have different pertinent knowledge which the speaker might access, Monitoring and Adjustment would predict that both form of expression and articulation will reflect the speaker's own knowledge until some corrective feedback points out the discrepancy. Co-presence tells us that long- or mid-term information is available for the whole language production process.

The other two hypotheses might distinguish between the two measures. H-and-H, the Negligible Defaulting Hypothesis, makes no comment on units larger than words. What is essential is that an account of listener needs is available for each lexical or phonological word. This could be provided in two ways. In the more complex, speakers must conduct two parallel series of updates on the listener model: word-by-word while uttering one phrase and simultaneously, as if that phrase were complete, while constructing the next. Thus different states of listener knowledge would have to be modeled at the same time. This alternative seems so demanding that to preserve the essential predictions of H-and-H, word-by-word operations would have to take precedence, leaving phrase-by-phrase operations either impoverished or non-existent. Thus word intelligibility and duration should be the more sensitive to listener needs. In the simpler arrangement, the redundancy of each word would be assessed as part of the design process preceding the construction of their phrase. Thus clarity and accessibility should be equally sensitive to the listener's knowledge, because they are designed around the same reasoning about that knowledge.

The Dual Process Hypothesis makes a straightforward prediction. Here, the critical issue is the scale or duration of the process and not the stage when it occurs. Under the time pressures imposed by real conversations, smaller scale processes involved in articulatory design of phonological words should seldom allow scope for costly interaction with the listener model, and would have to be controlled by speaker knowledge. Larger-scale processes, like planning an NP, could cycle slowly enough to permit updating the listener model, drawing inferences from it, and the like. Accordingly, form of referring expression could be more sensitive than duration to any records of listener knowledge which speakers maintain. This hypothesis does not predict uniformly good tailoring of referring expressions to listener knowledge, however,

because the task of updating the listener model may have an inherently low priority.

3. Studies of intelligibility and accessibility

3.1. Given-ness and referring expressions

To test the foregoing hypotheses, we made use of the effects of Given status. Word tokens in expressions introducing New items are longer and clearer than those referring to Given items (Fowler and Housum, 1987). Forms of referring expression are known to differ in elaboration so that changes with repeated mention are usually abbreviations (*a blacksmith's cottage...it*) which can be assigned a place in a scale of referential accessibility (Ariel, 1990, Gundel., Hedberg, and Zacharski, 1993) To compare the two systems, we used a corpus of spontaneous speech designed to vary what each interlocutor could see, coded to what each had mentioned or heard mentioned, and to what feedback each had given the other. Thus it was possible to select items which were Given to one or both interlocutors on the basis of what each saw, said or heard in the dialogue. Table X summarizes the comparisons which formed the basis of 4 experiments.

3.2. Method

3.2.1. Materials

All materials were drawn from the HCRC Map Task Corpus (Anderson et al., 1991), 128 unscripted dialogues from 64 pairs of Glasgow University undergraduates communicating routes defined by labeled cartoon landmarks on schematic maps of imaginary locations. Instruction Giver's and Follower's maps for any dialogue matched only in alternate landmarks. Participants knew that their maps differed but not where or how. In no case could either player see the other's map. The corpus was balanced for familiarity of participants and for ability to see the interlocutor's face. Each participant served as Instruction Giver for the same route to two different Followers and as Instruction Follower for two different routes.

Digital stereo recordings with one channel per speaker were segmented at word boundaries. All the words of any expression referring to a landmark were coded for the appropriate landmark, tagged for part-of-speech, and parsed.

Bard et al. (2000) excerpted individual words from references to the labeled schematic landmarks around which the route is defined in cases where both mentions used at least some of the same open class words (*the rift valley...the rift valley; the rift valley...the valley*). Except where the design of the experiment dictated otherwise, items were restricted to the Instruction Giver's initial encounter with a map and were balanced for familiarity of interlocutors and for the availability of a visual channel. Items forming part of disfluencies or interrupted by overlapping speech were excluded.

The present study examined all expressions which refer to landmarks that were mentioned more than once within a dialogue, with the exception of those which were ambiguous as to accessibility. Note that the items for which duration measures are suitable, like those assessed for intelligibility, must include the same words in both

Score	Definition	Examples
0	numeral + noun indef art + sequence	<i>one mountain</i> <i>a mountain</i>
1	def article + nominal possessive +	<i>the mountain</i> <i>my one</i>
2	possess pron deictic pron deictic adj + nominal	<i>mine</i> <i>that</i> <i>this mountain</i>
3	other pron	<i>it</i>

Table 1. Accessibility scale for referring expressions

mentions. Repetitions using different words in different mentions (*the rift valley...it*), may only be assessed via accessibility.

3.2.2. Dependent variables

Intelligibility loss. Individual open class items from matching repeated mentions were excerpted from context, as were control tokens of the same landmark names read in lists by the original speakers. A standard set of phonetic conventions was used to determine the positions of word boundaries (Laver et al., 1989). All words were overlaid with noise and presented to panels of 9 to 15 naïve informants for identification. The tokens of a word were distributed among informants by Latin square. Intelligibility is the percentage of listeners identifying a word. Intelligibility loss is the difference between the intelligibility of the clearer read token and of the more reduced running speech token of the same word. (See Bard et al. for further details.)

Duration loss. Both studies used normalized duration (Campbell and Isard, 1991). The normalization makes use of the distributions of lengths typical of each phoneme and assigns to each word token a value k representing its position in the expected log length distribution for words of its dictionary phoneme composition and stress pattern. The k -score makes it possible to compare length-relative-to-expected-length for words of quite different composition. All comparisons were based on the difference between the k -durations of a read control form and the corresponding item in running speech.

Accessibility. The 27 items with relative clauses in their first mentions were excluded because of a conflict in coding schemes. All other first and second mentions of landmarks (N = 1136) were classed by accessibility on the scale displayed in Table 1.

3.3. Experiment 1: Listener identity

3.3.1. Design

Experiment 1 examines introductory mentions of the same shared landmarks in Givers' two trials with the same map. In the first trial, the landmark was New to the discourse for both players. In the second, it was Given for the speaker, an Instruction Giver who had mentioned it before, but New to each successive listener (hence the value 'no' in each of the 'how Given status is achieved – by listener' cells in Table 6). The identity of the listener and the state of progress through the map route were both route pre-printed on their maps. Thus, if the Negligible

Measure	Introduction	
	1	2
Word articulation:		
Intelligibility loss	0.072	0.182
<i>k</i> -duration loss	0.498	0.558
Referring expression form:		
Accessibility	0.466	0.552

Table 2. Effects of re-introductions to new listeners on articulatory clarity (intelligibility or length loss relative to citation form) and on form of referring expression (accessibility).

expressions are tailored to the listener's needs even when these differ from the speaker's, then introductory mentions of the same landmark should not differ in clarity: no Given-ness effect should be warranted because the named entities are not Given for the listener. Similarly, accessibility of referring expressions should not change. The Monitoring and Adjustment prediction is for changes in both, because the listener in the second trial has had no opportunity for feedback. The Dual Process prediction is that intelligibility will be insensitive to the listener's knowledge and fall, because it depends on the speaker's previous mention. Only accessibility ought to reflect the listener's knowledge and remain unchanged.

3.3.2. Results

Table 2 begins with the intelligibility results of Bard et al.: second introductions show significant loss of intelligibility relative to first introductions (i.e., a greater difference between the carefully pronounced form read in a list and the token produced in running speech). The present study also finds changes in articulation: second introductions are significantly shorter than first (i.e., increasingly different from citation forms) for 239 pairs of words on repeated introductions, ($F_2(1, 238) = 12.48$; $p < 0.0005$). In contrast, accessibility does not increase on average over 116 pairs of introductory mentions ($F_2(1,115) < 1$). Thus, duration appears to reflect the Given status of the item for the speaker, while form of referring expression reflects the fact that the freshly introduced landmark is New for each listener. Greater sensitivity in form of referring expression is predicted only by the Dual Process Hypothesis.

3.4. Repeater identity: inferred listener knowledge

3.4.1. Design

Experiment 2 compared repeated mentions of shared landmarks within and between speakers. As Table 6 shows, in self-repetitions the second token refers to a landmark which is Given to the repeating speaker because he or she has seen the item, and both uttered and heard the original mention. The landmark's status vis-à-vis the listener is less certain. In other-repetitions, the second token is Given to the current speaker only by virtue of having been heard before, but Given to current listener who has mentioned the item, heard it mentioned, and must have been able to see the landmark to introduce it at all.

Measure	Repetition	
	Self	Other
Word articulation:		
Intelligibility loss	0.081	0.081
<i>k</i> -duration loss	0.127	0.192
Referring expression form:		
Accessibility	0.878	0.745

Table 3. Effects of self- v other-repetition on change in articulatory clarity (intelligibility and duration loss relative to citation form) and in form of referring expression (accessibility) with repeated mention

Negligible Defaulting should promise that either articulation or both articulation and form of referring expression will be show greater change in such cross-speaker repetitions, because an updated account of the listener's knowledge should include the inference that the item was Given to that player in those three ways. In contrast, Monitoring and Adjustment predicts no effect of original speaker on any measure, because no corrective feedback is involved. Co-presence will make the same prediction if we assume that it is satisfied by common experience of the discourse (a 'yes' in both 'Heard' columns in Table 6) without inferring what the listener can currently see. Dual Process predicts that any effect will be found in accessibility, which is designed over intervals long enough to permit inferences to be made.

3.4.2. Results

Table 3 shows that changes in articulatory clarity were the same in self- and other-repetition. Like the intelligibility results of earlier experiments, *k*-duration fell relative to citation form controls with repeated mention (mention: $F_2(1,691) = 63.75$, $p < 0.0001$) but showed no difference between the 263 other-repetitions and the 430 self-repetitions. (mention x prior speaker: *n.s.*). Accessibility for 90 other-repetitions and the 430 self-repetitions. behaved in the same way ($F_2(1,269) = 177.12$, $p < 0.0001$; mention x prior speaker: *n.s.*). Once more the listener's experience was not the critical factor, and repetitions of any mention which the speaker has heard are treated alike..

3.5. Feedback: signalled listener knowledge

3.5.1. Design

Experiment 3 provides a more direct test of the effects of listener knowledge. When one speaker introduces an unshared landmark, the listener, who lacks it, may provide corrective feedback indicating the discrepancy between the players' maps. Sometimes, however, that listener fails to find or signal the discrepancy. To test for the effects of feedback on second mentions, we use repeated mentions by the same speaker with and without accurate intervening feedback from the listener.

It is difficult to see how a cooperative speaker, in the usual sense, could ignore such overt evidence. We assume that Negligible Defaulting and Co-presence joint Monitoring and Adjustment predicting that feedback will make a difference to the nature of subsequent mentions. In

Measure	Visibility to listener	
	Not denied	Denied
Word articulation:		
Intelligibility loss	-0.080	0.080
<i>k</i> -duration loss	0.070	0.140
Referring expression form:		
Accessibility	0.470	0.410

Table 4. Effects of feedback about listener's ability to see an entity on changes in articulatory clarity (intelligibility and duration loss relative to a citation form) and in form of referring expression (accessibility) with repeated mention.

fact, the repetitions with feedback are the only ones where Monitoring and Adjustment does predict an effect of listener knowledge. In all these cases, cooperative behaviour would yield a more restricted effect of repetition where the listener has denied ability to find the object, -- that is, less change in intelligibility or accessibility across repetitions. Only Dual Process, which holds that feedback may be unimportant, could account for failure to mitigate of the effects of repetition on form and articulation.

3.5.2. Results

Table 4 begins with results from Bard et al. Which require further comment. Intelligibility loss vis-à-vis a clear control form should have increased more where listeners offered no negative feedback and less where they denied having the named object on their maps. In fact, the reverse was true, with a significant interaction of mention and feedback because of increased intelligibility with no-denial repetitions and decreased with denial. However, the root of the difference lay in the first mentions, not the second, whose absolute intelligibility scores were indistinguishable. In the present study, no such complication is found. For the 73 repeated words with intervening denial and the 122 without, *k*-duration loss increased with repetition significantly and equally (mention: $F_2(1, 193) = 9.45, p = .0024$; mention x denial *n.s.*). Form of referring expression showed the same pattern: the change toward more accessible referring expressions on second mention was no more limited for the 44 cases with intervening denials than for the 86 without ($F_2(1,128) = 18.49, p < .0001$; mention x denial: *n.s.*). Feedback that should block defaulting does not do so. Only what the repeater has seen, heard, and said seems to play a role.

3.6. Repeater knowledge

3.6.1. Design

What the repeater knows is the subject of Experiment 4, in particular what the speaker can see. Here only cross-

speaker repetitions were used, but now the landmark in question might be shared by both speakers or absent from the repeater's map. As Table 6 shows, the original introducer, the listener at the point of second mention, can see the item, has mentioned it, and has heard it mentioned. The repeater has also heard it mentioned. Experiments 2

Measure	Visibility to speaker	
	Seen	Unseen
Word articulation:		
Intelligibility loss	0.151	0.181
<i>k</i> -duration loss	0.114	0.183
Referring expression form:		
Accessibility	0.745	0.240

Table 5. Effects of speaker's ability to see named entity on change in articulatory clarity (intelligibility and duration loss relative to a citation form) and in form of referring expression (accessibility) with repeated mention.

and 3 have already shown that the original introducer's ability to see the named item does not bear on the manner of repeated mention. What we ask here is whether the speaker's visual knowledge of the named entity is also unimportant or whether articulation and form or referring expression are influenced by this kind of knowledge. If what the repeater can see is an important addition to speaker-Given status, then intelligibility loss across repeated mentions will be greater for shared landmarks, where the speaker has more knowledge of the entity than for unshared.

The Negligible Default Hypothesis predicts no effect of what the speaker can see, because the more important listener knowledge is constant across conditions. Co-presence would seem to make the same prediction. Monitoring and Adjustment allows for speaker knowledge having direct effects on articulation or referring expression design. Dual Process makes the assumption that articulation is keyed to speaker knowledge by fast priming processes. It is not clear whether visual stimuli prime word duration. Thus far, illustrations have all been via perceiving or producing the repeated word. Dual process does allow for slower, costly access to additional information, and so would allow for effects of speaker knowledge on accessibility.

3.6.2. Results

Table 5 shows the effects of repetition. Bard et al. found a robust effect of repetition on intelligibility loss vis-à-vis citation forms, but no tendency toward greater change where the repeater could see the landmark. The present results have the same interpretation: *k*-duration falls with repeated mention (mention: $F_2(1, 224) = 12.37, p < .0005$) but there is no significant difference between the outcome for the 144 shared repetitions and the 82 unshared (mention x introducer: *n.s.*).

In contrast, form of referring expression shows the speaker-centric result. Second mentions are made in more accessible forms in both cases (mention: $F_2(1,138) = 24.67, p < .0001$), but the increase is greater for the 90 cases where the repeater can see the landmark than for the 50 where he or she cannot. (mention x sharing: $F_2(1,138) = 6.48, p < .02$). This outcome is certainly not indicative of careful adjustment to listeners alone. Nor does it indicate overall attention to speaker knowledge as a proxy for listener knowledge. It conforms best to the notion that different processes design the form and articulation of referring expressions with the former sensitive to a wider range of information.

4. Discussion

Table 6 summarizes the results reported here and in Bard et al. (2000). Each of the experiments tests for an effect on repeated mentions of some aspect of speaker or listener knowledge. Experiment 1 pitted the speaker's experience in having seen the mentioned landmark, mentioned it, and heard it mentioned against the new listener's ignorance of the item as the landmark was introduced in a second trial with a map. Experiment 2 pitted the speaker's own experience in seeing and hearing against the listener's under two conditions, when those listeners to the repetition had produced the original mention so that it might be inferred that they could see the landmark, and when they had not. Experiment 3 pitted the speaker's experience of seeing, saying, and hearing against the listener's declared inability to see the item in question. Experiment 4 kept the listener's knowledge constant as well as the speaker's experience in hearing a prior mention, but manipulated the speaker's ability to see the landmark.

In all these cases, as the shaded cells of Table 6 show, the repeating speaker had heard the original mention. In all cases the measures of word articulation were sensitive only to what the speaker had heard. These are exactly the results found by Bard et al. (2000) for a balanced but restricted sample of materials and with intelligibility as dependent variable. Thus, reductions in articulatory detail with repeated mention are conditioned by what the repeaters have heard mentioned. There is no indication that models of the listener are consulted except insofar as

they conform exactly to the speaker's memory for what he or she has heard.

Form of referring expression showed a different pattern. It behaved like articulation in being insensitive to information which should have been of use in updating a model of the listener: either an indication that the listener could see the landmark under discussion or a direct statement to the effect that he or she did not (Experiments 2 and 3). Yet it did show two effects which articulation did not. In Experiment 1 accessibility of referring expression did not increase with re-introductions to new listeners. In this case, form of referring expression was tailored to the listener's needs. In Experiment 4, accessibility was enhanced more for repeated mentions of landmarks which the speaker could see than for repeated mentions of items which the speaker had only heard mentioned. Thus, accessibility is more sensitive than articulation but not in a way which support claims for the tailoring of referring expressions to listeners' needs.

Why should accessibility have these characteristics? The current results indicate that form or referring expression does not respond on-line to changes in co-presence, whether via feedback or inference. Nor does accessibility, which seems to be designed before articulation, show the characteristics that Monitoring and Adjustment would predict for initial design. In Dell and Brown's account, early processes like design of referring expressions should, if anything, be less sensitive to listener knowledge than later processes like articulation. This certainly is not the case here: referring expressions patterned like duration when the two should have differed.

Experiment	Effects on repeated mention (by dependent variable)		How Given status achieved					
			By speaker			By listener		
	Word articulation	Referring expression form	Said	Sees	Heard	Said	Sees	Heard
1: different listeners	Speaker	Listener	yes	yes	yes	no	no	no
2: same/ different speakers	Speaker	Speaker	no / yes	yes	yes	yes Û / no Û	yes (inferred) / ?	yes
3: +/- negative feedback	Speaker	Speaker	yes	yes	yes	no	no (declared) / yes (inferred)	yes
4. speaker +/- sees	Speaker	Speaker (additional)	no	no / yes	yes	yes	yes (inferred)	yes

Table 2. Speaker-knowledge and listener-knowledge effects on repeated mentions of landmark names. Word articulation results in terms of intelligibility (Bard et al, 2000) agree with current results of standardized word duration (*k*). Shaded cells indicate conditions in common across all experiments where repeated mentions lost clarity. Form of referring expression in terms of accessibility shows additional sensitivity to conditions in the doubly boxed cells. (yes = condition holds; no = condition does not hold; / = contrast manipulated in experiment).

Furthermore, referring expressions patterned differently from duration where the two should have been alike in reflecting the speaker's knowledge.

We would argue that Map Task speakers demonstrated the effects of competing demands on their attention, as the Dual Process Hypothesis predicts. Unlike the fast automatic processes which affect articulation and are keyed to speaker memory, slower processes compete for attention with the task in hand. Consequently only the factors grossly related to that task -- who is participating and what is on the speaker's own map -- have a noticeable effect.

We have argued elsewhere (Bard et al., 2000) that the difficulty of the communicative task may well influence the degree to which speakers appear to be modeling their listeners. We noted the Map Task is more difficult than other tasks where more cooperative behaviour is reported. For example, the tangram task involves a fixed set of shapes and players usually know that the match between their shapes will be complete and that none will have to be re-used. Hence the problem becomes easier with every trial. In contrast, the Map Task does not make it clear at the outset how many landmarks will determine each route, how many are on the map but irrelevant, how many match between players' maps, how many are duplicated on a single map, and how many have to be revisited as the task advances. If listener modeling competes for attention with task management, we might well expect the Map Task and the more complex of everyday communicative tasks to show little tendency toward tailoring for the listener. It remains to be seen whether direct manipulation of extended communicative tasks will change speakers' priorities (see Horton and Keysar, 1996, for a simple example). It also remains to be seen whether speakers will be more sensitive to fine differences in listener knowledge in any task if some kind of external record-keeping eases the computational burden. The Dual Process Hypothesis predicts that both task and memory load should have effects on the design of referring expressions, but that neither should affect the articulation of individual words.

5. Acknowledgment

This work was supported by ESRC (UK) main grant to the Human Communication Research Centre.

6. References

- Anderson, A., Bader, M., Bard, E.G., Boyle, E., Doherty, G., et al., 1991. The H.C.R.C. Map Task Corpus. *Language and Speech*, 34:351-366.
- Ariel, M., 1990. *Accessing Noun-Phrase Antecedents*. London: Routledge/Croom Helm.
- Balota, D. A., Boland, J. E., and L. W. Shields, 1989. Priming in pronunciation: Beyond pattern-recognition and onset latency. *Journal of Memory and Language*, 28:14-36.
- Bard, E. G., and A. Anderson, 1983. The unintelligibility of speech to children. *Journal of Child Language*, 10:265-292.
- Bard, E. G., and A. Anderson, 1994. The unintelligibility of speech to children: Effects of referent availability. *Journal of Child Language*, 21: 623-648.
- Bard, E. G., Anderson, A., Sotillo, C., Aylett, M., Doherty-Sneddon, G., and A. Newlands, 2000. Controlling the intelligibility of referring expressions in dialogue. *Journal of Memory and Language*, 42:1-22.
- Brennan, S., and H. H. Clark, 1996. Conceptual pacts and lexical choice in conversation. *Journal of Experimental Psychology: Learning Memory and Cognition*, 22:1482-1493.
- Brown, P., and Dell, G., 1987. Adapting production to comprehension -- the explicit mention of instruments. *Cognitive Psychology*, 19:441-472.
- Campbell, W. N., and S.D. Isard, 1991. Segment durations in a syllable frame. *Journal of Phonetics*, 19:37-47.
- Carletta, J., and C. Mellish, 1996. Risk taking and recovery in task-oriented dialogue. *Journal of Pragmatics*, 26:71-107.
- Clark, H. H., and C. R. Marshall, 1981. Definite reference and mutual knowledge. In A.K. Joshi, B. Webber, and I. Sag (eds.), *Elements Of Discourse Understanding*. Cambridge: Cambridge University Press.
- Dell, G., and P. Brown, 1991. Mechanisms for listener-adaptation in language production: Limiting the role of the "model of the listener". In D. J. Napoli and J. A. Kegl (eds.), *Bridges Between Psychology And Linguistics*. (pp. 111-222). Hillsdale: Erlbaum.
- Fowler, C., and J. Housum, 1987. Talkers' signalling of 'new' and 'old' words in speech and listeners' perception and use of the distinction. *Journal of Memory and Language*, 26:489-504.
- Fowler, C., Levy, E. and J. Brown, 1997. Reductions of spoken words in certain discourse contexts. *Journal of Memory and Language*, 37:24-40.
- Fussell, S., and R. Krauss, 1992. Coordination of knowledge in communication: Effects of speakers' assumptions about what others know. *Journal of Personality and Social Psychology*, 62: 378-391.
- Gundel, J.K., Hedberg, N., and R. Zacharski, 1993. Cognitive status and the form of referring expressions in discourse. *Language*, 69: 274-307.
- Horton, W., and B. Keysar, 1996. When do speakers take into account common ground? *Cognition*, 59:91-117.
- Hunnicuttt, S., 1985. Intelligibility vs. redundancy—conditions of dependency. *Language and Speech*, 28:47-56.
- Isaacs, E., and H. H. Clark, 1987. References in conversation between experts and novices. *Journal of Experimental Psychology: General*, 116:26-37.
- Laver, J., Hieronymus, J., Bennett, C., Cohin, I., Dalby, J., Davies, D., Hiller, S., McAllister, M., and E. Purves, 1989. *The ATR/CSTR Speech Database Project (Report 3)*. Edinburgh: Centre for Speech Technology Research, Edinburgh University.
- Levelt, W., and L. Wheeldon, 1994. Do speakers have access to a mental syllabary? *Cognition*, 50:239-269.
- Lieberman, P., 1963. Some effects of the semantic and grammatical context on the production and perception of speech. *Language and Speech*, 6:172-175.
- Lindblom, B., 1990. Explaining variation: a sketch of the H and H theory. In W. Hardcastle and A. Marchal (eds.), *Speech Production And Speech Modelling* (pp. 403-439). Dordrecht, Netherlands: Kluwer Academic Publishers.
- Keysar, B., 1997. Unconfounding common ground. *Discourse Processes*, 24:253-270.
- O'Seaghdha, P. G., 1997. Conjoint and dissociable effects of syntactic and semantic context. *Journal of*

- Experimental Psychology: Learning, Memory, and Cognition*, 23:807-828.
- McKoon, G., and R. Ratcliff, 1980. The comprehension processes and memory structures involved in anaphoric reference. *Journal of Verbal Learning and Verbal Behavior*, 19:668-682.
- Mitchell, D. B., and A. S. Brown, 1988. Persistent repetition priming in picture naming and its dissociation from recognition memory. *Journal of Experimental Psychology: Learning Memory and Cognition*, 14:213-222
- Schober, M. 1993. Spatial perspective taking in conversation. *Cognition*, 47: 1-24.
- Schober, M., and H. H. Clark, 1989. Understanding by addressees and overhearers. *Cognitive Psychology*, 21:211-232.
- Samuel, A., and M. Troicki, 1998. Articulation quality is inversely related to redundancy when children or adults have verbal control. *Journal of Memory and Language*, 39:175-194.
- Vonk, W., Hustinx, L., and W. Simmons, 1992. The use of referential expressions in structuring discourse. *Language and Cognitive Processes*, 7:301-33.
- Wheeldon, L., and A. Lahiri, 1997. Prosodic units in speech production. *Journal of Memory and Language*, 37:356-81
- Wilkes-Gibbs, D., and H.H. Clark, 1992. Coordinating beliefs in conversation.. *Journal of Memory and Language*, 31:183-194