

Dialogue Systems, Spatial Tasks and Elderly Users: A Review of Research into Elderspeak

Kavita E. Thomas



SFB/TR 8 Report No. 021-02/2010

Report Series of the Transregional Collaborative Research Center SFB/TR 8 Spatial Cognition
Universität Bremen / Universität Freiburg

Contact Address:

Dr. Thomas Barkowsky
SFB/TR 8
Universität Bremen
P.O.Box 330 440
28334 Bremen, Germany

Tel +49-421-218-64233
Fax +49-421-218-64239
barkowsky@sfbtr8.uni-bremen.de
www.sfbtr8.uni-bremen.de

Dialogue Systems, Spatial Tasks and Elderly Users:
A Review of Research Into Elderspeak
Kavita E. Thomas

Introduction

Elderspeak is a register which is frequently used by non-elderly speakers addressing elderly listeners. It is frequently used in nursing homes and hospitals where frail elders are found. Elderspeak consists of some or all of the following characteristics: (1) using a singsong voice which changes pitch and tone and exaggerates words, (2) simplifying the length and complexity of sentences, (3) speaking more slowly, (4) using limited vocabulary, (5) repeating or paraphrasing what has just been said, (6) using terms of endearment like “dear”, and (7) using rising pitch at the end of statements so that they sound like questions. It is similar to the “baby talk” register used with small children and is based on the stereotype that the elderly person is cognitively reduced and can occur even in cases when there is no evidence that the elderly person in question cannot understand or respond to the standard register (Kemper et al, 1995).

Dialogue systems which interact with elderly users need to be aware of the myriad issues involved with elderly addressees in order to produce language which best maximises their comprehension while also encouraging independent behaviour. Furthermore, dialogue systems need to avoid using features of elderspeak which undermine elderly users' perception of their own capabilities. In order to focus on these issues, research into how younger speakers address older hearers, as well as how elderly speakers address elderly hearers will be presented. In the review which follows, different issues to do with elderspeak are discussed, leading to a set of guidelines for dialogue systems interacting with elderly users. The specific focus of this review is to isolate salient research questions which emerge for dialogue systems assisting elderly users with spatial tasks like route instruction dialogues so that systems can best tailor language to older users.

Issues involved with elderspeak

Negative aspects of elderspeak

Researchers in gerontology have shown that the use of elderspeak regardless of the addressee's cognitive skills can diminish elderly people's confidence in their own abilities and competence, which in turn erodes their self-esteem (O'Conner & Rigby, 1996; Ryan, Bourhis & Knops, 1991; Kemper et al, 1998). Many features of elderspeak cause elderly people to lose confidence in their own communication abilities even when they don't have any comprehension problems because the subliminal priming of negative age stereotypes have the detrimental effects of self-stereotyping for memory, handwriting, walking, etc (Levy, 2003; Rodin & Langer,

1980). Additionally, Kemper & Harden (1999) have shown that most aspects of elderspeak actually achieve the reverse goal and decrease comprehension, as unimportant words are often stressed without need and rising end intonations make interpreting statements harder. Similarly talking too slowly affects elderly people's ability to focus on the main point and retain information.

Kemper & Harden (1999) also indicate that older and younger addressees differ in the feedback they produce in response to route instructions; older listeners repeat instructions and request clarification more often than younger listeners and also express confusion more than younger listeners. Although they do this regardless of whether the speaker was young or old, this feedback occurs much more when they listen to younger speakers than with older speakers. Furthermore, elderly listeners indicate less self-confidence when listening to instructions from younger speakers.

Given that Kemper et al (1998) show that young people use elderspeak when addressing elderly listeners, and that some aspects of elderspeak erode elderly listeners' confidence in their own abilities (O'Conner & Rigby, 1996; Ryan, Bourhis & Knops, 1991; Kemper et al, 1998), this explains why elderly listeners differ in feedback to younger and older speakers. Elderly route instructors, unlike younger ones, did not vary prosody, fluency, grammatical complexity, semantic content, style, or use elderspeak markers when addressing young vs. older listeners.

However both young and old listeners performed less accurately with older speakers, despite the fact that older speakers tend to use a simplified speech style regardless of addressee when they themselves are speakers. The poor performance by both old and young in response to older speakers can be explained by a tendency of older speakers to use vague or ambiguous references and also ignore clarification requests.

Positive aspects of elderspeak

However some aspects of elderspeak can actually improve comprehension, e.g., repeating and paraphrasing what has been said, and adopting simple grammatical structure with explicit content and low propositional density while not shortening sentences (Kemper 2001). Additionally using a varied vocabulary with a high type to token ratio makes the conversation more interesting and improves elderly listeners' comprehension (Kemper et al, 1995). Young listeners were shown to be more accurate than older listeners in response to older speakers, while older listeners were nearly as accurate as young listeners in response to young speakers (Kemper et al, 1995), showing that some aspects of elderspeak can actually improve comprehension for elderly listeners. These findings provide a guideline for how best to facilitate comprehension with elderly addressees. However self-evaluation questionnaires following the task showed that older people report more receptive and expressive problems when paired with young speakers despite the fact that they performed more accurately in response to younger speakers. This shows that while some aspects of

elderspeak are beneficial for comprehension, others are detrimental to older people's confidence in their own abilities.

Elderspeak and spatial tasks

The experiments run by Kemper et al (1995, 1999, 2001) are of particular interest to researchers of spatial dialogue who have elderly end-users, as they investigate the effectiveness of particular characteristics of elderspeak and its ability to hinder or facilitate comprehension in a route instruction dialogue involving maps. In these studies, elderly and young listeners are evaluated with respect to their ability to correctly trace the instructed route on their maps following a route instruction dialogue in which both elderly and young participants act as speakers. Their findings have led to the above-mentioned guidelines for facilitating comprehension with elderly listeners.

However Kemper et al don't focus on spatial strategy differences adopted by route instructors based on the age group of their addressees, or likewise, whether these various strategies were successful or not. Instead, their work focuses on general dialogue strategies like semantic elaboration via varying lexical choice (e.g., “street” and “road”), repetition and paraphrasing, and comprehension checks, all of which they found improved older addressees' comprehension. They propose that the ideal form of elderspeak would provide semantic elaborations and providing several location checks per map (e.g., “So you're at the intersection of Pine and Main?”) to improve accuracy in the route-tracing task (Kemper, 2001). Semantic elaborations involve repeating or expanding 50% of map directions, e.g., “turn right onto Main St.” might be expanded as “that is, continue for two blocks, and then take your second right onto Main Street”. While these are useful general dialogic strategies to account for from the perspective of dialogue system building where the participants are elderly, this still leaves a gap where spatial language strategies need to be evaluated for their usefulness in improving elderly addressees' comprehension in spatial tasks. Tenbrink et al (2010) and Goschler et al (2008) show that people adapt spatial strategies like choosing and switching between egocentric and exocentric spatial perspective and also varying features of spatial granularity to their addressee in route-finding dialogues, where their addressee groups are humans vs. a dialogue system. What is not known is whether people particularly adapt spatial strategies to older addressees, and then whether these adaptations facilitate comprehension for older listeners engaged in spatial task-oriented dialogues.

The role of individual differences

However, regarding dialogue strategies like semantic elaboration, attention to addressees' individual differences and abilities is crucial (Kemper, 1998). Torrey et al (2006) found that semantic elaboration as a dialogic device should be tailored to listener expertise in experiments where a robotic chef instructs both expert and novice cooks. They advocate adaptivity based on Clark & Wilkes-Gibbs' (1986) least

collaborative effort concept and use addressee-oriented design in order to maintain positive affect between speakers and listeners, where providing too little information can be interpreted as a sign that the speaker has no concern for the addressees' needs while too much information is irritating and patronizing. They use expertise in the task at hand as the gauge for determining the extent of semantic elaboration of instructions in task-oriented dialogues. They found that providing detailed explanations for novices improved information-exchange (measured as the number of questions addressees asked) and performance of novices on the object-identification task, but did not influence the information-exchange or performance of experts. In the same object-identification task with added incentives for speed, experts ranked robots as more authoritative and less patronizing and also more effective at communicating in dialogues without elaborations, in direct contrast to novices' rankings.

In spatial instruction dialogues involving unknown environments and goal locations, background knowledge and expertise don't seem as central a role as in task instruction dialogues where the task can involve identifying known objects. I.e., if someone knows what a cleaver is, a description doesn't help, but in a route description task for an unknown map/environment, expertise is less of a factor. However the effects of expertise may crop up in cases where elderly users have been using a dialogue system for assistance over a longer period of time, provided there is some repetition in tasks. Then the system would need to account for what it has already told the user in order to keep from presenting known or redundant information.

The effects of cognitive factors

Furthermore, commonalities across elderly people in terms of their communication needs might be a more driving concern in optimizing comprehension for elderly addressees than individual differences. Research into language comprehension with elderly addressees has shown that memory decline is fairly common and can cause language comprehension problems (Wingfield and Stine-Morrow, 2000). Kemper et al (2008) report that older adults also have more difficulty comprehending target information in the presence of distractors. However Stine-Morrow et al (2004) show that when reading expository and narrative texts, older adults matched younger readers for comprehension, allocating more attention to situation model features on their first reading of the passage, but showing less facilitation compared to younger readers at word-level processing, indicating that representation at this level is not as firmly established during reading or decays more rapidly.

Beyond social perceptions of elderly people which result in elderspeak, research in neurobiology has shown that there are mental differences which occur with age in spatial tasks (Sanders et al, 2008; Moffat et al, 2001). Moffat et al use virtual environments to assess spatial navigation in the elderly and found that elderly participants both made more errors and took longer to complete the route-finding

task. Sanders et al examine egocentric navigation and a floor maze task to assess exocentric navigation in healthy elderly people, looking at the relationships between navigational tests and cognitive factors representing executive function and attention, verbal ability and memory. Egocentric skills on the local route recall test were associated with executive function and attention and memory factors while exocentric skills were related to executive function and attention. That is, they found that cognitive processes may be differently activated based on spatial perspective. In terms of speakers' adaptivity to individual needs, although young speakers tend to use elderspeak regardless of their elderly addressee's individual abilities or needs, speakers (both old and young) adjust their language to compensate for listeners with serious communicative deficiencies (Kemper et al 1998).

This research into both the social perceptions of elderly people and findings from neurobiology and psychology shows that elderly people have a range of difficulties with spatial tasks and have specific communication needs which need to be accounted for in order for them to best comprehend instructions.

Considerations for dialogue systems

Considering cognitive assistance systems for elderly people in their homes, Torrey et al (2005) argue that speech technologies should accommodate to their users' needs, taking into account the findings on positive and negative aspects of elderspeak. Given that the mechanics of speaking and hearing decline with age, some examples of these kinds of accommodation are adjusting volume, intonation and sentence structure, and modelling elderly speech in speech recognition software. Mueller and Wasinger (2002) present a multimodal interface for a mobile pedestrian navigation system which adapts to different user groups' needs and takes into account the cognitive load of the user as well as age. They argue that current speech recognition systems have trouble recognising elderly voices, and spatial acoustic models which also model elderly users will enable speaker clustering as well as better speech recognition for elderly speakers. Wolters et al (2009) however argue that individual differences play a role in deciding whether or not to adapt acoustic models to elderly voices. Wolters et al (2009b) also argue that in terms of interaction behaviour, cognitive abilities and gender did not predict group behaviour, and they advocate adapting dialogue systems to observed behaviour rather than age. In terms of dialogue strategies, Wolters et al (2009c) found in an appointment booking domain that reducing the number of options presented and using explicit confirmation strategies did not improve task success, and conclude that accurately assessing the cognitive demands of the given task should be performed before considering whether to adopt user-group based strategies.

Ideally dialogue systems should be able to accommodate based on an estimate of their particular user's needs and Torrey et al (2005) argue that individual differences should be accounted for to avoid patronizing, demoralising or confusing their addressees. Influenced by work on politeness theory, which describes the

balance between maintaining positive face (protecting individual worth) and maintaining negative face (protecting individual autonomy) when communicating with others (Brown & Levinson, 1987), Torrey et al advocate avoiding categories of impolite talk. Hummert and Ryan (1996) proposed a model of patronizing talk which varies orthogonally on the dimensions of control (low vs. high) and care (low vs. high), distinguishing between overly nurturing talk, which involves high control and high care, directive talk, which involves high control but low care, overly personal talk, which involves high care but low control, and superficial talk, which involves both low care and control. However older adults may require an additional security dimension (Parmelee & Lawton, 1990), thus complicating the picture, as caregivers may favour the security goal over being polite (Wahl, 1991). Feedback from older listeners as a result are often ignored, as demonstrations of independence through behaviour or speech are often ignored by caregivers or receive dependence-supportive feedback by caregivers (e.g., “I told you not to do that; you always get it wrong”), while dependent behaviours elicit positive responses (Baltes & Wahl, 1996; 1992). This results in the nonuse of competence by older people and undermines their autonomy, since they often receive help whether they need it or not.

Given these findings, Torrey et al (2005) emphasize the need for optimizing for autonomy as well as accommodating to individual addressees' needs. They argue that the three functions of elderspeak—nurturance, comprehension and control-- vary depending on how appropriate the accommodation is to the addressee, and the typical use of elderspeak focuses on nurturing, aiding comprehension and controlling the addressee, resulting in overaccommodation based on stereotypes rather than individual needs.

Likewise, while some older people might be cognitively or physically compromised and need comprehension-facilitating aspects of elderspeak, the control and politeness aspects need to be carefully monitored to maintain autonomy and support independent behaviour and avoid enforcing negative self-stereotyping. Torrey et al (2005) propose that speech technology should avoid aiming for a nurturing tone and instead employ prosodic convergence, i.e., a gradual adjustment in nonverbal features that mirrors features of the addressee's speech; this type of convergence can be seen as socially cohesive since it communicates a desire for social approval (Giles et al, 1991). Other prosodic features which can be adapted are stress patterns and presentation rates. Cohen and Faulkner (1981) show that contrastive stress can facilitate recall of stressed names, while Wingfield et al (1989) show that slow rates of presentation facilitate prose processing. Torrey et al recommend that speech synthesis should maintain natural distinctions in prosody, including word stress while using simple sentence structure without embedded clauses and providing semantic elaborations like repetition, expansion and comprehension checks. However as Torrey et al (2006) show, semantic elaborations need to be tailored to the addressee.

In terms of managing control, Torrey et al (2005) suggest an adjustable autonomy strategy, where assistance software lets the user do as much of the task as

they are capable of, providing assistance only when the user can't perform some action. They note that as a heuristic in optimizing system behaviour, adjustable autonomy is a very different metric from task efficiency, as autonomy keeps the control in the hands of the user as long as the user is capable of performing the task, even if the user is less efficient than the system. The advantages of this strategy is that it avoids reliance on the system and deterioration of human skills through nonuse. This heuristic plays a central role in dialogue control strategies, and will influence initiative and turn-taking behaviours. Necessary future research in dialogue management should compare the effects of mixed-initiative systems vs. system-driven dialogue systems vs. user-controlled systems and their respective turn-taking strategies in terms of their effects on maintaining autonomy on the one hand while providing timely help on the other.

To sum up the desiderata for dialogue systems assisting elderly users, at the prosodic level, prior research indicates that adjusting volume and intonation and modelling elderly speech in speech recognition software will help, along with following a strategy of prosodic convergence. Likewise employing normal stress patterns which use contrastive stress to distinguish opposed or compared entities and slightly slowing the presentation rate can be beneficial for this user group. At the syntax level, simple sentence structure without embedded clauses are recommended. At the semantic level, speech technologies should accommodate to their users' needs, taking into account the findings on positive and negative aspects of elderspeak, employing semantic elaborations like repetition, expansion and comprehension checks which are tailored to the addressee's expertise. Lastly, at the discourse level an adjustable autonomy strategy of dialogue management which fosters independent behaviours and reduces system dependence is recommended.

Summary and Areas for Future Work

So far we have seen that some features of elderspeak are beneficial while others are detrimental for older listeners, providing guidelines for tailoring dialogue systems' information-giving communication to older addressees which maximise their comprehension and minimise loss of self-confidence and other negative effects of elderspeak. These guidelines include dialogue strategies like semantic elaboration via varying lexical choice, repetition and paraphrasing 50% of instructions and frequent comprehension checks. They also prescribe adopting simple grammatical structure with explicit content and low propositional density while not shortening sentences and using a varied vocabulary with a high type to token ratio.

Additionally while there is a fine balance between modelling individual users' needs and accounting for some general characteristics of elderly listeners, there are also balances to be made along the axes of control and helpfulness in order to tailor dialogue systems to older users in the most beneficial way. While most elderly people have more difficulties than younger adults with spatial tasks, individuals also vary in the degree of their difficulties and the extent to which they are cognitively impaired.

Recommendations for dialogue system development suggest that dialogue systems should avoid aiming for a nurturing tone and instead employ prosodic convergence, gradually adjusting nonverbal features to mirror features of the addressee's speech to facilitate social cohesion. Additionally, semantic elaborations need to be tailored to the addressee's expertise where appropriate. Research in elderspeak recommends that dialogue assistance systems provide help when it's needed, and not before, following an adjustable autonomy strategy, where assistance software lets the user do as much of the task as they are capable of, and only provides assistance when the user can't perform some action.

There are several areas for future work building dialogue systems in the spatial instruction domain given the current state of research on elderspeak.

In terms of psycholinguistic research, there is a gap where spatial strategies are concerned. Future research should investigate what spatial strategies best facilitate comprehension with elderly listeners in various task types and situations. Furthermore, feedback from elderly listeners in response to the use of particular spatial strategies and also semantic elaboration which is too limited or excessive should be studied in order for dialogue systems to be able to detect overaccommodation and underaccommodation and adjust dialogue strategies accordingly for this user group, creating more responsive behaviour which avoids some of the pitfalls of elderspeak.

In terms of computational linguistics research, determining how systems can automatically classify users' needs and abilities would be particularly useful for this user group in order to adaptively tailor system behaviour to an individual user's needs. Following along this line of research, dialogue systems should be evaluated for optimizing initiative and turn-taking strategies for older users in a given type of task, comparing mixed vs user-initiated vs system-controlled initiative models and the corresponding turn-taking strategies. Further work is also needed in adapting semantic elaborations and spatial strategies used to account for growing expertise with long-term users in familiar domains. Also, from a more computational perspective, speech recognition systems should be optimized for older users' voices.

Acknowledgements

Thanks to Thora Tenbrink for proofreading this document and for providing much helpful feedback.

References

Baltes, M. M. and Wahl, H.W. (1992). The Dependency- Support Script in Institutions: Generalization to Community Settings. *Psychology and Aging* 7(3): 409-418.

Brown, P. and Levinson, S. (1987). *Politeness: Some universals in language usage*. Cambridge: Cambridge University Press.

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

Cohen, G. and Faulkner, D. (1986). Does Elderspeak Work? The Effect of Intonation and Stress on Comprehension and Recall of Spoken Discourse in Old Age. *Language and Communication* 6: 91-98.

Giles, H., Coupland, N. and Coupland, J. (1991). Accommodation theory: Communication, context, and consequence. Contexts of accommodation: developments in applied sociolinguistics. In H. Giles, J. Coupland and N. Coupland (Eds.), *Cambridge: University Press*, pp. 1-68.

Goschler, J., Andonova, E., and Ross, R. (2008). Perspective Use and Perspective Shift in Spatial Dialogue. In C. Freksa, N. Newcombe, P. Gärdenfors, and S. Wölfl (Eds.), *Spatial Cognition VI: Learning, Reasoning, and Talking about Space*. Berlin: Springer, pp. 250-265.

Hummert, M. L. and Ryan, E. B. (1996). Toward understanding variations in patronizing talk addressed to older adults: Psycholinguistic features of care and control. *International Journal of Psycholinguistics* 12(2): 149-169.

Kemper, S. 2001. Over-Accommodations and Under-Accommodations to Aging. In N. Charness, D. Parks and B. Sabel Eds.), *Communication, Technology and Aging: opportunities and challenges for the future*. New York: Springer, pp. 30-46.

Kemper, S. and Harden, T. (1999). Experimentally disentangling what's beneficial about elderspeak from what's not. *Psychology and Aging*, 14, 656-670.

Kemper, S., McDowd, J., Metcalf, K. and Liu, C. Young and Older Adults' Reading of Distractors. (2008). *Educational Gerontology*, 34(6): 489-502.

Kemper, S., Othick, M., Gerhing, H., Gubarchuk, J., and Billington, C. (1998). The effects of practicing speech accommodations to older adults. *Applied Psycholinguistics*, 19: 175-192.

Kemper, S., Vandeputte, D., Rice, K., Cheung, H., and Gubarchuk, J. (1995). Speech adjustments to aging during referential communication task. *Journal of Language and Social Psychology*, 14: 40-59.

Levy, B. (2003) Conscious Versus Unconscious Levels of Aging Self-Stereotypes. *The Journals of Gerontology Series B: Psychological Sciences and Social Sciences*, 58: 215-216.

- Moffat, S., Zonderman, A. and Resnick, S. (2001). Age Differences in Spatial Memory in a Virtual Environment Navigation Task. *Neurobiology of Aging*, 22.
- Mueller, C. and Wasinger, R. (2002). Adapting Multimodal Input for the Elderly. In *ABIS-02, Adaptivity and User Modelling in Interactive Software Systems*.
- O'Connor, B.P., and Rigby, H. (1996). Perceptions of baby talk, frequency of receiving baby talk, and self-esteem among community and nursing home residents. *Psychology and Aging*, 11: 147-154.
- Parmelee, P. A. and Lawton, M. P. (1990). The design of special environments for the aged. *Handbook of the psychology of aging*. J. E. Birren and K. W. Schaie. San Diego, CA, Academic Press: 464-488.
- Ryan, E.B., Bourhis, R.Y., and Knops, U. (1991). Evaluative perceptions of patronizing speech addressed to elders. *Psychology and Aging*, 6, 442-450.
- Rodin, J. and Langer, E.(1980). Aging Labels: The Decline of Control and the Fall of Self-Esteem. *Journal of Social Issues*, 36: 12-29.
- Sanders, R. Holtzer, R. Lipton, C. Hall and J. Verghese. (2008). Egocentric and Exocentric Navigation Skills in Older Adults, A. *Journal of Gerontology*, 63A(12).
- Stine-Morrow, E., Gagne, D., Morrow, D. and DeWall, B. (2004). Age Differences in Rereading. *Memory and Cognition*, 5: 696-710.
- Tenbrink, T., Ross, R., Thomas, K. and Andonova, E. (2010). Efficient Negotiation of Routes in Map-based Dialogues Between Humans and Dialogue Systems. *To appear in the Journal of Visual Languages and Computation*.
- Torrey, C., Fussell, S. and Kiesler, S. (2005). Appropriate Accommodations: Speech Technologies and the Needs of Older Adults. *Proceedings of the American Association of Artificial Intelligence*, 2005.
- Torrey, C., Powers, A., Marge, M., Fussell, S. and Kiesler, S. (2006). Effects of Adaptive Robot Dialogue on Information Exchange and Social Relations. *Proceedings of the International Conference on Human-Robot Interaction* , HRI2006.
- Wahl, H. W. (1991). Dependence in the Elderly from an Interactional Point of View: Verbal and Observational Data. *Psychology and Aging*, 6: 238-246.
- Wingfield, A., Lahar, C. J. and Stine, E. L. (1989). Age and decision strategies in running memory for speech: Effects of prosody and linguistic structure. *Journal of Gerontology: Psychological Sciences*, 44: 106-113.
- Wingfield, A., and Stine-Morrow, E. L. (2000). Language and Speech. In *Handbook*

of Aging and Cognition, 2nd ed. Edited by F. I. M. Craik and T. A. Salthouse. Erlbaum Associates, 359–416.

Wolters, M., Vipperla, R., and Renals, S. (2009). Age Recognition for Spoken Dialogue Systems: Do we need it? In *Proc. of Interspeech*.

Wolters, M., Georgila, K., MacPherson, S., and Moore, J. (2009b). Being Old Doesn't Mean Acting Old: Older users' interaction with spoken dialogue systems. *ACM Transactions on Accessible Computing*, 2(1): 1-39.

Wolters, M., Georgila, K., Logie, R., MacPherson, S., Moore, J., and Watson, M. Reducing Working Memory Load in Spoken Dialogue Systems. *Interacting with Computers*, 21(4): 276-287.