The Effect of Dialogue System Output Style Variation on Users' Evaluation Judgments and Input Style

Ivana Kruijff-Korbayová and **Olga Kukina**
Department of Computational Linguistics
Saarland University, Germany
{korbay|olgak}@coli.uni-sb.de

Abstract

A dialogue system can present itself and/or address the user as an active agent by means of linguistic constructions in personal style, or suppress agentivity by using impersonal style. We compare system evaluation judgments and input style alignment of users interacting with an in-car dialogue system generating output in personal vs. impersonal style. Although our results are consistent with earlier findings obtained with simulated systems, the effects are weaker.

1 Introduction

One of the goals in developing dialogue systems that users find appealing and natural is to endow the systems with natural and contextually appropriate output. This encompasses a broad range of research issues. The one we address in this paper pertains to *style in the interpersonal dimension*: does using personal vs. impersonal style of system output have an effect on dialogue system users, in particular, on their judgments about the system and on the way they formulate their input to the system?

We define the personal/impersonal style dichotomy as reflecting primarily a distinction with respect to agentivity: personal style involves the explicit realization of an agent, whereas impersonal style avoids it. In the simplest way it is manifested by the presence of explicit reference to the dialogue participants (typically by means of personal pronouns) vs. its absence, respectively. More generally, active voice and finite verb forms are typical for personal style, whereas impersonal style often, though not exclusively, employs passive constructions or infinite verb forms:

(1) Typical personal style constructions:
   a. I found 20 albums.
   b. You have 20 albums.
   c. Please search for albums by The Beatles.

(2) Typical impersonal style constructions:
   a. 20 albums have been found.
   b. There are 20 albums.
   c. The database contains 20 albums.
   d. 20 albums found.

The designer of a dialogue system has the choice to make it manifest (its own and the user's) agentivity linguistically through the use of personal constructions or not.

Previous experiments with simulated systems have shown that a natural language interface with a synthesized voice should not say "I" (Nass and Brave, 2005) and that users align the style of their input to that of the system output (Brennan and Ohaeri, 1994). (See Section 2 for more detail.)

The dialogue system SAMMIE developed in the TALK project (Becker et al., 2007) can use either personal or impersonal output style. In personal style, it generates constructions making explicit reference to the agent (both the user and the system itself), such as (1a–1c); in impersonal style, it avoids explicit reference to any agent, as in (2a–2d). The system can be set either to use one style consistently throughout a dialogue session, or to align to the user's style, i.e., mimic the user's style on a turn-by-turn basis.

Inspired by the earlier results obtained with simulated systems (Nass and Brave, 2005; Brennan and

Proceedings of the 9th SIGdial Workshop on Discourse and Dialogue, pages 190–197,
Columbus, June 2008.

Ohaeri, 1994), we ran an experiment to test the effects of style manipulation in the SAMMIE system. In this paper, we compare two versions of the system, one using consistently the personal output style and the other the impersonal style. We designed our experiment to test (i) whether the users' judgments of the system's usability and performance differ among the system versions using the personal vs. impersonal style, and (ii) whether users align to the system style.

In Section 2 we review previous experiments concerning the effect of system output style on users' judgments and style. We describe our own experiment in Section 3, present the results in Section 4, and provide a discussion and conclusions in Section 5.

2 Previous Work

(Nass and Brave, 2005) address the issue whether a voice interface should say "I" by investigating several dimensions of user attitudes to their simulated system with a synthetic vs. recorded voice. Generally, agents that use "I" are perceived more like a person than those that do not. However, systems tend to be more positively rated when consistent with respect to such parameters as personality, gender, ontology (human vs. machine), etc. A system with a recorded voice is perceived as more human-like and thus entitled to the use of "I", whereas a synthetic-voice interface is not perceived as human enough to use "I" to refer to itself (Nass et al., 2006).

Another question is whether system output style influences users' input formulation, as would be expected due to the phenomenon of *alignment*, which is generally considered a basic principle in natural language dialogue (Garrod and Pickering, 2004).[1]

Experiments targeting human-human conversation show that in spite of the variety of linguistic expressions available, speakers in spontaneous dialogues tend to express themselves in similar ways at lexical and syntactic levels. For example, the surface form of a question can affect the format of the answer: the question "What time do you close?" will more likely get the response "Five o'clock" than"At five o'clock". On the other hand, "At five o'clock" is a more probable answer to "At what time do you close?" (Levelt and Kelter, 1982). There is evidence that alignment happens automatically as a result of priming, e.g., (Hadelich et al., 2004) for lexical alignment.

Lexical and syntactic alignment is present in human-computer interaction, too. (Brennan, 1996) suggested that users adopt system's terms to avoid errors, expecting the system to be inflexible. However, recent experiments show that alignment in human-computer interaction is also automatic and its strength is comparable to that in human-human communication (Branigan et al., 2003; Pearson et al., 2006).

Early results concerning users' alignment to system output style in the interpersonal dimension are reported in (Brennan and Ohaeri, 1994): They distinguish three styles: anthropomorphic (the system refers to itself using first person pronouns, like in (1a) above, fluent (complete sentences, but no self-reference) and telegraphic, like (2d). They found no difference in users' perception of the system's intelligence across the different conditions. However, they observed that the anthropomorphic group was more than twice as likely to refer to the computer using the second person pronoun "you" and it used more indirect requests and conventional politeness then the other groups. They concluded that the anthropomorphic style is undesirable for dialogue systems because it encourages more complex user input which is harder to recognize and interpret.

The described experiments used either the Wizard-of-Oz paradigm (Brennan, 1996) or preprogrammed system output (Branigan et al., 2003; Nass and Brave, 2005) and involved written communication. Such methods allow one to test assumptions about idealized human-computer interaction. The purpose of our experiment was to test whether similar effects arise in an interaction with an actual dialogue system, which may be plagued, among other factors, by speech recognition problems.

3 Experiment

Dialogue System We used the SAMMIE in-car system developed in the TALK project (Becker et al., 2006; Becker et al., 2007). SAMMIE provides a mul-

[1] This dialogue phenomenon goes under a variety of terms in the literature, besides alignment, e.g., accommodation, adaptation, convergence, entrainment or shaping (used, e.g., by (Brennan and Ohaeri, 1994)).

timodal interface to an MP3 player through speech and haptic input with a button which can be turned, pushed down and pushed sideways in four directions. System output is by speech and a graphical display. The user can perform a range of tasks: control the MP3 player (play/stop/pause playing song, next/previous/go-to track, turn shuffle mode on/off), search and browse by looking for various fields in the MP3 database (song, artist, album, etc.), search and select playlists, edit them or construct new ones. The SAMMIE system was designed with the aim to support natural, intuitive mixed-initiative interaction. Input can be given through any modality at any point and is not restricted to answers to system queries: the user can initiate new tasks as well as give any information relevant to the current task at any time. A sample interaction is shown below (Becker et al., 2006).

(3) U: Show me the Beatles albums.

S: I have these four Beatles albums. [shows a list of album names]

U: Which songs are on this one? [selects the Red Album]

S: The Red Album contains these songs [shows a list of the songs]

U: Play the third one.

S: [song "From Me To You" plays]

The SAMMIE system has a German and an English version which both provide the same functionality. The experiment employed the German version. See (Kruijff-Korbayová et al., 2008) for a description of the natural language generation module.

Setup Figure 1 shows a picture of the experiment setup. To simulate the driving situation, we used the "3D-Fahrschule" software.[2] The driving simulator visuals were projected on a wall-sized back-projection screen. The graphical interface of the SAMMIE system was shown on a display next to the steering wheel. Participants wore headphones with a microphone for the spoken input and output. The button for manual input was positioned to the right of their chair. The experimenter was sitting in an adjacent room and could see and hear everything happening in the experiment lab. The subjects could not see the experimenter, but heard her instructions, including the task assignments, from loudspeakers. If necessary, the subjects were able to talk to the experimenter.

[2]http://www.3d-fahrschule.de/index.htm

Figure 1: Experiment setup

Participants A total of 28 participants were paid to take part in the experiment. All were native German speakers, 22 female and 6 male, 22 students of the Saarland University and 6 employees. All but two participants had a driver's license and 20 participants reported driving more than 500km a year. 10 participants had previous experience with a driving simulation and 6 had used a dialogue system before. Each participant was assigned to one style condition, 14 to personal and 14 to impersonal style. To ensure as even a distribution as possible, there were 11 female and 3 male participants in each style condition, one of whom was a non-driver. There were 4 employees in impersonal style condition and 2 in the personal one.

Procedure Each participant was welcomed by the experimenter, seated in the experiment lab, and given brief written instructions concerning the driving simulator, the SAMMIE system and the evaluation procedure. The participants were instructed to use mainly spoken input to accomplish the tasks, although they were allowed to use manual input, too.

The participants first made a ca. 2-minute drive to get familiar with the driving simulator. Then they were asked to chose a destination city (Amsterdam, Madrid or London) and drive there on a highway. During the driving, the experimenter successively read to the participant 2 trial tasks and 11 experimental tasks to be solved using the SAMMIE system.

The tasks involved exploring the contents of a database of about 25 music albums and were of four types: (1) finding some specified title(s); (2) selecting some title(s) satisfying certain constraints; (3) manipulating the playlists by adding or removing songs and (4) free-use of the system.

The experimental tasks were presented to each participant in randomized order apart from the free use of the system, which was always the last task. To avoid priming by the style of the task formulation, and to help the participants memorize the task, the experimenter (E) repeated each task assignment twice to the participant, once in personal and once in impersonal style, as shown in the example below.

(4) E: *Bitte frage das System nach den Liedern von "Pur". Du willst also wissen welche Lieder von "Pur" es gibt.*
E: Please ask the the system about the songs by "Pur". You would like to know which songs by "Pur" there are.

The time the participants spent completing the individual tasks was not constrained. It took them about 40 minutes to complete all the tasks.

Afterwards, each participant was asked to fill in a questionnaire about their attitudes towards the system, consisting of questions with a 6-point scale ranging from 1 (low grade) to 6 (high grade). The questions were a subset of those used in (Nass and Brave, 2005) and (Mutschler et al., 2007), for example: *How do you assess the system in general:* technical (1) – human-like (6); *Communication with the system seemed to you:* boring (1) – exciting (6); *In terms of usability, the system is:* inefficient (1) —efficient(6).

Upon completing the questionnaire, the participant was paid and discharged.

Collected data The questionnaire responses have been tabulated and the dialogues of the subjects with the system have been recorded and transcribed.[3] The utterances of the participants (on average 95 per session) were subsequently manually annotated with the following features for further analysis:

- Construction type:
  - **Personal** (+/-) Is the utterance a complete sentence in active voice or imperative form
  - **Impersonal** (+/-) Is the utterance expressed by passive voice, infinite verb form (e.g., "Lied abspielen" (*lit.* "song play")), or expletive "es-gibt" ("there-is") construction
  - **Telegraphic** (+/-) Is the utterance expressed by a phrase, e.g., "weiter" ("next")
- Personal pronouns: (+/-) Does the utterance contain a first or second person pronoun
- Politeness marking: (+/-) Does the utterance contain a politeness marker, such as "bitte" ("please"), "danke" ("thanks") and verbs in subjunctive mood (eg. "ich hätte gerne")

4 Results

4.1 Style and Users' Attitudes

The first issue addressed in the experiment was whether the users have different judgments of the personal vs. impersonal version of the system. Since the system used a synthetic voice, the judgments were expected to be more positive in the impersonal style condition (Nass and Brave, 2005). Based on factor analysis performed on attitudinal data from the user questionnaires we created the six indices listed below. All indices were meaningful and reliable

1. General satisfaction with the communication with the system was composed of 3 pairs of adjectives describing communication with the system: disappointing/motivating, uninteresting/interesting and boring/exciting (Cronbach's α=0.86; t(26)=0.29, p=0.39 (one-tailed))
2. Ease of communication with the system comprised 5 parameters: naturalness of the communication with the system, formality/informality and indifference/sympathy of the system's communicative style, participants feelings during the conversation: tensed/relaxed and pleasant/unpleasant (α=0.83; t(26)=0.00, p=0.5 (one-tailed))
3. Usability of the system consisted of 1 pair of adjectives referring to the success

[3] We did not record the data from the driving simulator.

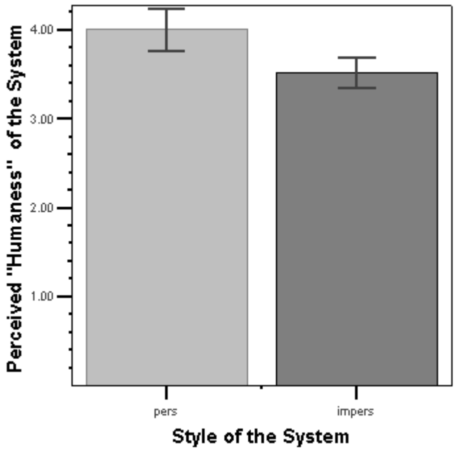

Figure 2: Perceived humanness of the system depending on system output style

of communication with the system: unsuccessful/successful, and 4 pairs of adjectives describing the usability of the system: unpractical/practical, inefficient/efficient, complicated/simple, inconvenient/convenient (α=0.76; t(26)=0.08, p=0.47 (one-tailed))

4. Clarity of the system's speech comprised 2 pairs of adjectives describing the system's speech: unpredictable/predictable and confusing/clear (α=0.88; t(25)=0.87, p=0.2 (one-tailed))
5. Perceived "humanness" of the system was composed of 3 parameters: perceived technicality/humanness, perceived unfriendliness/friendliness and attributed conservatism/innovation (α=0.69; t(25)=1.64, p=0.06 (one-tailed))
6. System's perceived flexibility and creativity comprised 3 parameters: rigidness/flexibility of system's speech, perceived creativity of the system and intelligence attributed to the system (α=0.78; t(26)=0.40, p=0.35 (one-tailed))

We did not find any significant influence of system output style on users' attitudes. The only index with a weak tendency in the predicted direction is *perceived humanness of the system* (t(25)=1.64, p=.06 (one-tailed); see Figure 2). This goes in line with the earlier observation that an interface that refers to itself by means of a personal pronoun is perceived to be more human-like than one that does

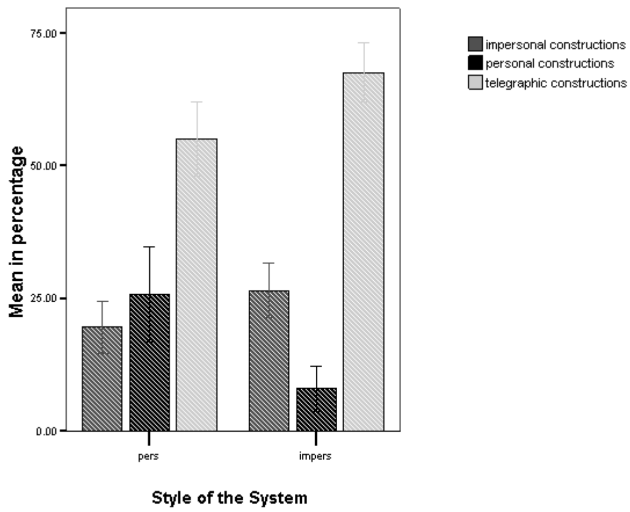

Figure 3: Distribution chart for syntactic construction types in user utterances depending on system output style

not (Nass and Brave, 2005).

4.2 Style and Alignment

The next issue we investigated was whether the users formulated their input differently with the personal vs. impersonal system version. For each dialogue session, we calculated the percentage of utterances containing the feature of interest relative to the total number of user utterances in the session.

First we analyzed the distribution of personal, impersonal and telegraphic constructions across the personal and impersonal style conditions. (The reason we separated telegraphic constructions is because they seem to be neutral with respect to style.) We compared the means of the obtained numbers between the two style conditions. Figure 3 shows the distribution of the types of syntactic constructions across the system output style conditions.

1. We expected the participants to use more personal constructions with the personal style version of the system. Independent samples t-test showed a significant result in the predicted direction (t(19)=1.8, p=0.05 (one-tailed); see Figure 3).

2. We expected to find the reverse effect with regard to the proportion of impersonal verb forms: participants using the personal style

version of the system were expected to have less infinite, passive and "es-gibt" forms than those in the impersonal style condition. However, we did not find any significant difference between the two style conditions (t(26)=1.0, p=0.17 (one-tailed)).

3. According to expectation we also did not find any significant difference in the proportion of telegraphic constructions per style condition (t(26)=1.4, p=0.09 (one-tailed)).

4. In the impersonal style condition we found a significantly lower proportion of verb-containing utterances than utterances in telegraphic form (t(13)=3.5, p=0.00 (one-tailed)). But in the personal style condition there was no statistically significant difference (t(13)=0.7, p=0.25 (one-tailed)).

Next we analyzed the distribution of first and second person pronouns across style conditions. We expected to find more personal pronouns in personal than in impersonal style condition (Brennan and Ohaeri, 1994). However, the results showed no statistically significant difference (t(26)=0.67, p=0.25 (one-tailed)).

Another prediction based on (Brennan and Ohaeri, 1994) was to find more politeness markers in the personal style. However, the analysis showed that participants in the personal style condition did not use significantly more politeness markers than those in the impersonal style condition (t(20)=1.06, p=0.15 (one-tailed)).

Finally, (Brennan and Ohaeri, 1994) predicted that personal style, being more flexible, might cause more speech recognition problems than input in impersonal style. We checked whether participants in the personal style condition had a higher rate of unrecognized utterances than those in the impersonal style condition and found no significant difference (t(26)=0.60, p = 0.28 (one-tailed)).

To summarize, we observed a significant difference in the number of personal constructions across style conditions, in accordance with the expectation based on style alignment in terms of agentivity. But we did not find a significant difference in the distribution of impersonal constructions across style conditions. Not surprisingly, there was also no significant difference in the distribution of telegraphic constructions. An unexpected finding was the higher proportion of telegraphic constructions than verb-containing ones within the impersonal style condition. However, the personal style condition showed no significant effect. Contrary to expectations, we also did not find any significant effect of style-manipulation on the number of personal pronouns, nor on the number of politeness markers.

4.3 Style Alignment over Time

Since alignment can also be seen as a process of gradual adjustment among dialogue participants in the course of their interaction, we were interested in whether participants tended to converge to using particular constructions as their session with the system progressed. For each participant we divided the transcribed conversation in two halves. Using paired samples t-test, we compared the proportion of personal, impersonal and telegraphic constructions in the first and second halves of the conversations for both style conditions.

In the personal style condition, we found no significant change in the usage of construction types between the first and the second half of the dialogue. In the impersonal style condition, we did not find any significant difference in the distribution of impersonal and telegraphic constructions either. However, we found a significant change in the number of personal constructions (t(13)=2.5, p=0.02 (one-tailed)): The participants cut down on the use of personal constructions in the second half.

5 Discussion and Conclusions

We presented the results of an experiment with the in-car multimodal dialogue system SAMMIE, aimed to test whether we obtain effects similar to earlier findings concerning the influence of system output style in the interpersonal dimension on the users' subjective judgments of a system (Nass and Brave, 2005) as well as their formulation of input (Brennan and Ohaeri, 1994). Although our results are not conclusive, they point at a range of issues for further research.

Regarding users' attitudes to the system, we found no significant difference among the styles. This is similar to (Brennan and Ohaeri, 1994) who found no difference in intelligence attributed to the system by the users, but it is at odds with the earlier finding that a synthetic voice interface was judged to be more useful when avoiding self-reference by personal pronouns (Nass and Brave, 2005).

Whereas (Brennan and Ohaeri, 1994) used a flight reservation dialogue system, (Nass and Brave, 2005) used a phone-based auction system which read out an introduction and five object descriptions. There are two points to note: First, the subjects were exposed to system output that was a read out continuous text rather than turns in an interaction. This may have reinforced the activation of particular style features. Second, the auction task may have sensibilized the subjects to the distinction between subjective (the system's) vs. objective information presentation, and thus make them more sensitive to whether the system presents itself as an active agent or not.

Regarding the question whether users align their style to that of the system, where previous experiments showed strong effects of alignment (Brennan and Ohaeri, 1994), our experiment shows some effects, but some of the results seem conflicting. On the one hand, subjects interacting with the personal style version of the system used more personal constructions than those interacting with the impersonal style version. However, subjects in either condition did not show any significant difference with respect to the use of impersonal constructions or telegraphic forms. We also found a higher proportion of telegraphic constructions than verb-containing ones within the impersonal style condition, but no such difference in the personal style. Finally, when we consider alignment over time, we find no change in construction usage in the personal style, whereas we find a decrease in the use of personal constructions in the impersonal style.

That there is no difference in the use of telegraphic constructions across conditions is not surprising. Being just phrasal sentence fragments, these constructions are neutral with respect to style. But why does there seem to be an alignment effect for personal constructions and not for others? One way of explaining this is that (some of) the constructions that we counted as impersonal are common in both styles. Besides their deliberate use as means to avoid explicit reference to oneself, the constructions typical for impersonal style also have their normal, neutral usage, and therefore, some of the utterances that we have classified as impersonal style might just be neutral formulations, rather than cases of distancing or "de-agentivization". However, we could not test this hypothesis, because we have not found a way to reliably distinguish between neutral and marked, truly impersonal utterances. This is an issue requiring further work.

The difference between our results concerning alignment and those of (Brennan and Ohaeri, 1994) is not likely to be due to a difference in the degree of interactivity (as with (Nass and Brave, 2005)). We now comment on other differences between our systems, which might have contributed to the differences in results.

One aspect where we differ concerns our distinction between personal and impersonal style, both in the implementation of the SAMMIE system and in the experiment: We include the presence/absence of agentivity not only in the system's reference to itself (akin to (Nass and Brave, 2005) and (Brennan and Ohaeri, 1994)), but also in addressing the user. This concept of the personal/impersonal distinction was inspired by such differences observed in a study of instructional texts in several languages (Kruijff et al., 1999), where the latter dimension is predominant. The present experiment results make it pertinent that more research into the motives behind expressing or suppressing agentivity in both dimensions is needed.

Apart from the linguistic design of the system's output, other factors influence users' behavior and perception of the system, and thus might confound experiment results, e.g., functionality, design, ergonomics, speech synthesis and speech recognition.

Earlier experiments reported in (Nass and Brave, 2005) suggest that a system with synthesized speech should be more positively rated when it does not refer to itself as an active agent by personal constructions. Whereas the system used by (Brennan and Ohaeri, 1994) used written interaction, we used the MARY text-to-speech synthesis system (Schröder and Trouvain, 2003) with an MBROLA diphone synthesizer, which produces an acceptable though not outstanding output quality. But as discussed earlier, contrary to (Nass and Brave, 2005) we have not observed a difference in the users' attitudes depending on style. It thus remains an open issue what effect speech output quality has on on the users' attitudes and alignment behavior.

Regarding a possible influence of speech recognition on our results, we performed a post-hoc analysis (Kruijff-Korbayová et al., 2008), which did not reveal significant differences in user attitudes or alignment behavior depending on better or worse speech recognition performance experienced by the users. A future experiment should address the possibility of an interaction between system style and speech recognition performance as both factors might be influencing the user simultaneously.

One radical difference between our experiment and the earlier ones is that the users of our system are occupied by the driving task, and therefore only have a limited cognitive capacity left to devote to the interaction with the system. This may make them less susceptible to the subtleties of style manipulation than would be the case if they were free of other tasks. A possible future experiment could address this issue by including a non-driving condition.

Finally, as we pointed out in the introduction, the SAMMIE system can also be used in an style-alignment mode, where it mimics the user's style on turn-to-turn basis. We plan to present experimental results comparing the alignment-mode with the fixed personal/impersonal style in a future publication.

Acknowledgments

This work was carried out in the TALK project (www.talk-project.org) funded by the EU as project No. IST-507802 within the 6^{th} Framework Program.

References

T. Becker, N. Blaylock, C. Gerstenberger, I. Kruijff-Korbayová, A. Korthauer, M. Pinkal, M. Pitz, P. Poller, and J. Schehl. 2006. Natural and intuitive multimodal dialogue for in-car applications: The SAMMIE system. In *Proceedings of ECAI, PAIS Special section*.

T. Becker, N. Blaylock, C. Gerstenberger, A. Korthauer, M. Pitz, P. Poller, J. Schehl, F. Steffens, R. Stegmann, and J. Steigner. 2007. Deliverable D5.3: In-car showcase based on TALK libraries. Technical report, TALK Project, EU FP6, IST-507802.

H. Branigan, M. Pickering, J. Pearson, J. F. McLean, and C. Nass. 2003. Syntactic alignment between computer and people: the role of belief about mental states. In *Proceedings of the Annual Conference of the Cognitive Science Society*.

S. Brennan and J.O. Ohaeri. 1994. Effects of message style on user's attribution toward agents. In *Proceedings of CHI'94 Conference Companion Human Factors in Computing Systems*, pages 281–282. ACM Press.

S. Brennan. 1996. Lexical entrainment in spontaneous dialogue. In *Proceedings of the International Symposium on Spoken Dialogue (ISSD-96)*, pages 41–44.

S. Garrod and M. Pickering. 2004. Why is conversation so easy? *TRENDS in Cognitive Sciences*, 8.

K. Hadelich, H. Branigan, M. Pickering, and M. Crocker. 2004. Alignment in dialogue: Effects of feedback on lexical overlap within and between participants. In *Proceedings of the AMLaP Conference*. Aix en Provence, France.

G.J.M. Kruijff, I. Kruijff-Korbayová, J. Bateman, D. Dochev, N. Gromova, T. Hartley, E. Teich, S. Sharoff, L. Sokolova, and K. Staykova. 1999. Deliverable TEXS2: Specification of elaborated text structures. Technical report, AGILE Project, EU INCO COPERNICUS PL961104.

I. Kruijff-Korbayová, C. Gerstenberger, O. Kukina, and J. Schehl. 2008. Generation of output style variation in the SAMMIE dialogue system. In *Proceedings of INLG'08*, Salt Fork Resort, Ohio.

W.J.M. Levelt and S. Kelter. 1982. Surface form and memory in question answering. *Cognitive Psychology*, 14:78–106.

H. Mutschler, F. Steffens, and A. Korthauer. 2007. Deliverable D6.4: Final report on multimodal experiments Part I: Evaluation of the SAMMIE system. Technical report, TALK Project, EU FP6, IST-507802.

C. Nass and S. Brave, 2005. *Should voice interfaces say "I"? Recorded and synthetic voice interfaces' claims to humanity*, chapter 10, pages 113–124. The MIT Press, Cambridge.

C. Nass, S. Brave, and L. Takayama. 2006. Socializing consistency: from technical homogeneity to human epitome. In *P. Zhang & D. Galletta (Eds.), Human-computer interaction in management information systems: Foundations*, pages 373–390. Armonk, NY: M. E. Sharpe.

J. Pearson, J. Hu, H. Branigan, M. J. Pickering, and C. I. Nass. 2006. Adaptive language behavior in HCI: how expectations and beliefs about a system affect users' word choice. In *CHI '06: Proceedings of the SIGCHI conference on Human Factors in computing systems*, pages 1177–1180, New York, NY, USA. ACM.

M. Schröder and J. Trouvain. 2003. The German text-to-speech synthesis system MARY: A tool for research, development and teaching. *International Journal of Speech Technology*, 6:365–377.