CAN OLDER PEOPLE REMEMBER MEDICATION REMINDERS PRESENTED USING SYNTHETIC SPEECH?

Maria K. Wolters¹, Christine Johnson², Pauline E. Campbell³, Christine G. DePlacido²,
Brian McKinstry⁴

1 School of Informatics and School of Philosophy, Psychology, and Language Sciences,
University of Edinburgh

2 Speech and Hearing Sciences, Queen Margaret University

3 NMAHP Research Unit, Glasgow Caledonian University

4 Centre for Population Health Sciences, University of Edinburgh

Funder: Chief Scientist Office Scotland, UK EPSRC/BBSRC

Word Count: 1967 excluding Tables, Figures, Captions, and References
Reminders are often part of interventions to help older people adhere to complicated medication regimes. Computer-generated (synthetic) speech is ideal for tailoring reminders to different medication regimes. Since synthetic speech may be less intelligible than human speech, in particular under difficult listening conditions, we assessed how well older people can recall synthetic speech reminders for medications. 44 participants aged 50-80 with no cognitive impairment recalled reminders for one or four medications after a short distraction. We varied background noise, speech quality, and message design. Reminders were presented using a human voice and two synthetic voices. Data were analyzed using generalized linear mixed models. Reminder recall was satisfactory if reminders were restricted to one familiar medication, regardless of the voice used. Repeating medication names supported recall of lists of medications. We conclude that spoken reminders should build on familiar information and be integrated with other adherence support measures.

(Abstract: 148 words)
BACKGROUND AND SIGNIFICANCE

As more people live longer, levels of multimorbidity have increased, leading to more complex medication regimes. Non-adherence can have serious consequences, including hospitalization and death [1,2]. Forgetting what to take when is a common cause of non-adherence [3]. Reminders that help users remember those details are an important part of successful adherence interventions [4-6].

Spoken reminder messages can be presented through many channels such as interactive voice response systems, a digital TV, or electronic pill boxes. Using a computer-generated (synthetic) voice, spoken messages can be tailored quickly and cost-effectively to different patients and medication regimes, which may make them more effective at changing medication behavior.[7]

However, despite major advances in synthetic speech technology, synthetic speech is still not as intelligible as human speech,[8] and the perceived listening effort is significantly higher.[9] In this study, we assess whether the current intelligibility levels of synthetic speech are sufficient for older people, who are likely to have hearing loss, to recall medication names, which are difficult to remember.[10]

Hearing loss starts early. About one in four adults aged 50-59 has a hearing loss in the frequencies covered by speech sounds;[11] incidence rises sharply with age.[12] As more cognitive and perceptual effort is required to process auditory stimuli, fewer cognitive resources may be available for understanding and remembering what was said.[13]
Older people find synthetic speech more difficult to understand than younger people.[14,15] However, most related work [14-19] has focused on older speech synthesis technology and may therefore overestimate intelligibility problems. In this study, we used publicly available implementations of the two main current speech synthesis approaches, statistical parametric synthesis (SPS [20,21]) and unit selection (USEL [22,23]), shown in Figure 1.

FIGURE 1 ABOUT HERE

Reminders are often heard against background noise, such as radio, TV, or road traffic, or they can be transmitted through noisy telephone lines, which may make them more difficult to understand. Listeners may also be distracted by other tasks or people after hearing a reminder. Therefore, we varied listening conditions in our study design and included a brief distraction before recall.

Often, negative effects of listening conditions can be counteracted by improving the phrasing and organization of the message. One option is to adapt the message to listeners’ expectations and ways of thinking about medication, for example by linking each medication to the reason for taking it,[24,25] or to repeat key information.[26] Both options were tested in our study.
OBJECTIVES

This study was designed to determine if older people with a range of hearing ability (from clinically normal to some age-related hearing loss) find medication reminders more difficult to recall when these are presented using a synthetic as opposed to a human voice. We included two factors that might impair recall, memory load (one versus four medications per reminder) and listening conditions (difficult versus acceptable). Finally, we tested whether message design might be able to counteract recall problems. Two alternatives were tested: explaining medication indications (Explanation) and repeating medication names (Repetitions).

MATERIALS AND METHODS

STIMULI

We created 12 one-medication reminders, 24 baseline 4-medication reminders, 12 4-medication reminders where the medications were repeated, and 12 4-medication reminders where the indication of the medications was explained (c.f. Figure 2).

FIGURE 2 ABOUT HERE

We chose 12 names of commonly used UK medications (6 over-the-counter / 6 prescription).

Auditory stimuli were generated using three male voices with a standard British English accent, a human voice; a SPS synthetic voice [21]; and a USEL synthetic voice [23].
Participants heard stimuli in two levels of background noise (multi-speaker speech shaped babble [27], signal-to-noise ratios 0 and +10) and two levels of signal transmission quality (noisy / clear telephone line).

To reduce duration, we only tested the Explanation and Repetition message designs for high-memory load reminders with four medications. Participants heard both designs under two listening conditions, clear phone line/soft background noise (acceptable) and noisy phone line/loud background noise (difficult).

PARTICIPANTS

Ethical approval for this study was granted by the South East Scotland NHS Ethics Board, Reference number 10/S1103/43. Participants were recruited from four Family Practices in Edinburgh through the Scottish Primary Care Research Network. We contacted native speakers of English aged 50+ with no prescriptions for hearing aids, no cognitive impairment, and no neurodegenerative disorder and sufficient mobility to attend testing. 56 participants attended for testing. 12 were found not to meet inclusion criteria, leaving 44 participants.

PROCEDURE

In a questionnaire, participants indicated whether they were familiar with the twelve medications used in the task. For each familiar medication, they were asked to describe its use.
Participants were screened for cognitive impairment using the Addenbrooke’s Cognitive Examination-Revised [28]. Working Memory Capacity was assessed by a reading span test [29]. Working memory scores were recoded into three levels, first quartile, second/third quartile, and fourth quartile.

Hearing levels were measured using pure-tone audiometry [30,31]; severity was categorized following [32]. Participants were considered to have a hearing loss if they had at least a mild hearing loss on their poorer ear. We excluded people with a severe hearing loss on at least one ear or a conductive hearing loss (following [33]).

The intelligibility test consisted of a total of 8 training and 72 test trials. Participants adjusted loudness to a comfortable level at the start of the experiment.

In each trial, participants heard a medication reminder, followed by a two-second pause and a simple five-word distractor sentence [34]. After repeating the distractor, participants had to recall the medications in the reminder. Figure 3 shows a sample trial.

FIGURE 3 ABOUT HERE

To support recall, participants were shown a screen with a list of 24 possible medications and their key indication, 12 were medications that occurred in the reminders and 12 were distractor medications matched for indication. Responses were audio-recorded and
transcribed verbatim by the experimenter. The reminder score was the number of correctly remembered medications.

All participants heard the same reminder/distractor pairs in the same, randomized order. The assignment of voices to sentences was balanced using a Latin Square design.

Participants were debriefed in a semi-structured exit interview.

STATISTICAL ANALYSIS
Intelligibility scores were analyzed using generalized linear mixed models [35,36] and two levels of analysis, individual reminder responses versus responses grouped by participant. Fisher’s Exact test and pairwise t-tests corrected for multiple comparisons as implemented in R (Version 3.0.2) were used to illustrate effects of single predictors.

RESULTS
Demographic and baseline data are summarized in Table 1. The decline in working memory capacity and the prevalence of hearing loss in the sample is consistent with normal ageing [11,37].
Table 1 Demographics of participants

<table>
<thead>
<tr>
<th>Gender</th>
<th>female</th>
<th>N=24</th>
<th>55%</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>male</td>
<td>N=20</td>
<td>45%</td>
</tr>
<tr>
<td>Age</td>
<td>Mean (SD)</td>
<td>M=64 (9)</td>
<td></td>
</tr>
<tr>
<td>50-59</td>
<td>N=15</td>
<td>34%</td>
<td></td>
</tr>
<tr>
<td>60-69</td>
<td>N=14</td>
<td>32%</td>
<td></td>
</tr>
<tr>
<td>70+</td>
<td>N=14</td>
<td>32%</td>
<td></td>
</tr>
<tr>
<td>not given</td>
<td>N=1</td>
<td>2%</td>
<td></td>
</tr>
<tr>
<td>Hearing Loss</td>
<td>none</td>
<td>N=26</td>
<td>59%</td>
</tr>
<tr>
<td>(on worse ear)</td>
<td>mild, monaural</td>
<td>N=7</td>
<td>16%</td>
</tr>
<tr>
<td></td>
<td>mild, binaural</td>
<td>N=8</td>
<td>18%</td>
</tr>
<tr>
<td></td>
<td>Moderate on worse ear,</td>
<td>N=2</td>
<td>4.5%</td>
</tr>
<tr>
<td></td>
<td>mild on better ear</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>moderate, binaural</td>
<td>N=1</td>
<td>2%</td>
</tr>
<tr>
<td>Working</td>
<td>Mean (SD)</td>
<td>43 (13)</td>
<td></td>
</tr>
<tr>
<td>Memory</td>
<td>age 50-59</td>
<td>51 (12)</td>
<td></td>
</tr>
<tr>
<td>Capacity^</td>
<td>age 60-69</td>
<td>42 (15)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>age 70+</td>
<td>38 (11)</td>
<td></td>
</tr>
</tbody>
</table>

^The score is the sum of all items recalled correctly in the correct order. Scores range from 0 to 70.
On average, participants knew five of the over-the-counter and two of the prescription medications.

Recall was perfect for 74.8% (790 of 1056) of the one-medication reminders, but only for 4.5% (95 of 2112) of the four medication reminders. When participants made a mistake, they mostly forgot a medication name or confused the medication with another target medication.

Table 2 shows the results of mixed model analysis for one medication, four medications (baseline message design only), and four medications (message design varied).

Table 2 Coefficients for Individual-Level Predictors in Mixed Models

<table>
<thead>
<tr>
<th>Model A</th>
<th>Predictor</th>
<th>Beta</th>
<th>SD</th>
<th>Z value</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Intercept</td>
<td>1.20</td>
<td>0.51</td>
<td>2.34</td>
<td>P&lt;0.02*</td>
</tr>
<tr>
<td>1 Medication</td>
<td>Hearing Loss</td>
<td>-0.73</td>
<td>0.29</td>
<td>-2.55</td>
<td>P&lt;0.02*</td>
</tr>
<tr>
<td></td>
<td>WM Group</td>
<td>0.38</td>
<td>0.20</td>
<td>1.86</td>
<td>P&lt;0.07</td>
</tr>
<tr>
<td></td>
<td>Phone Line</td>
<td>-0.19</td>
<td>0.18</td>
<td>-1.03</td>
<td>P&lt;0.3</td>
</tr>
<tr>
<td></td>
<td>Noise Level</td>
<td>-1.37</td>
<td>0.18</td>
<td>-7.8</td>
<td>P&lt;0.001***</td>
</tr>
<tr>
<td></td>
<td>Known</td>
<td>1.95</td>
<td>0.2</td>
<td>10.0</td>
<td>P&lt;0.001***</td>
</tr>
<tr>
<td></td>
<td>HTS Voice</td>
<td>-0.39</td>
<td>0.21</td>
<td>-1.88</td>
<td>P&lt;0.06</td>
</tr>
<tr>
<td></td>
<td>USEl Voice</td>
<td>-0.46</td>
<td>0.21</td>
<td>-2.25</td>
<td>P&lt;0.03*</td>
</tr>
<tr>
<td>Model B</td>
<td>Intercept</td>
<td>-0.02</td>
<td>0.15</td>
<td>-0.11</td>
<td>P&lt;0.95</td>
</tr>
<tr>
<td>---------</td>
<td>------------</td>
<td>-------</td>
<td>------</td>
<td>-------</td>
<td>--------</td>
</tr>
<tr>
<td></td>
<td>Hearing Loss</td>
<td>-0.19</td>
<td>0.08</td>
<td>-2.33</td>
<td>P&lt;0.05*</td>
</tr>
<tr>
<td></td>
<td>WM Group</td>
<td>0.07</td>
<td>0.06</td>
<td>1.28</td>
<td>P&lt;0.2</td>
</tr>
<tr>
<td>Basic</td>
<td>Phone Line</td>
<td>-0.04</td>
<td>0.05</td>
<td>-0.72</td>
<td>P&lt;0.5</td>
</tr>
<tr>
<td></td>
<td>Noise Level</td>
<td>-0.11</td>
<td>0.05</td>
<td>-2.25</td>
<td>P&lt;0.03*</td>
</tr>
<tr>
<td></td>
<td>Known</td>
<td>0.29</td>
<td>0.02</td>
<td>12.32</td>
<td>P&lt;0.001***</td>
</tr>
<tr>
<td></td>
<td>SPS Voice</td>
<td>-0.20</td>
<td>0.06</td>
<td>-3.48</td>
<td>P&lt;0.001***</td>
</tr>
<tr>
<td></td>
<td>USel Voice</td>
<td>-0.14</td>
<td>0.06</td>
<td>-2.36</td>
<td>P&lt;0.02</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Model C</th>
<th>Intercept</th>
<th>0.02</th>
<th>0.14</th>
<th>0.18</th>
<th>P&lt;0.9</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Hearing Loss</td>
<td>-0.17</td>
<td>0.07</td>
<td>-2.37</td>
<td>P&lt;0.02*</td>
</tr>
<tr>
<td></td>
<td>WM Quantile</td>
<td>0.08</td>
<td>0.05</td>
<td>1.55</td>
<td>P&lt;0.2</td>
</tr>
<tr>
<td>Varied</td>
<td>Listening</td>
<td>-0.05</td>
<td>0.04</td>
<td>-1.33</td>
<td>P&lt;0.2</td>
</tr>
<tr>
<td>Message</td>
<td>Difficult</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Design</td>
<td>Known</td>
<td>0.23</td>
<td>0.02</td>
<td>11.49</td>
<td>P&lt;0.001***</td>
</tr>
<tr>
<td></td>
<td>SPS Voice</td>
<td>-0.19</td>
<td>0.05</td>
<td>-4.12</td>
<td>P&lt;0.001***</td>
</tr>
<tr>
<td></td>
<td>USel Voice</td>
<td>-0.10</td>
<td>0.05</td>
<td>-2.24</td>
<td>P&lt;0.03*</td>
</tr>
<tr>
<td></td>
<td>Explain</td>
<td>0.05</td>
<td>0.05</td>
<td>1.09</td>
<td>P&lt;0.3</td>
</tr>
<tr>
<td></td>
<td>Repeat</td>
<td>0.17</td>
<td>0.05</td>
<td>3.57</td>
<td>P&lt;0.001***</td>
</tr>
</tbody>
</table>

*: significant at p<0.05, **: significant at p<0.01, ***: significant at p<0.001. The table only shows fixed effects (individual response level). On the individual response level, we used a default intercept and the predictors Hearing Loss, Memory, phone line quality, background noise level, medication familiarity, and voice type. For the Message Design
model, we added predictors corresponding to message type. On the participant level, we used a default intercept that captured inter-individual variation and modeled listener-specific effects of phone line quality and background noise level.

As expected, listeners with a hearing loss found it significantly more difficult to recall the reminders than listeners with normal hearing. Working Memory Capacity does not have a significant effect after controlling for hearing. (Table 2, all three models). Signal quality did not affect recall; the effect of background noise varied depending on the number of medications and message design.

Looking at one-medication reminders, we find known medications are typically recalled correctly (recall for known 89.3% (484 of 541), for unknown 59.5% (306 of 512), p<0.001, OR 5.7, 95% CI [4.1,8.0]).

Although voice type affected recall, this difference was almost entirely due to unknown medications. For known medications, participants recalled 89% (human voice), 90% (SPS voice), and 88% (unit selection voice), for unknown medications, they recalled 64.8% (human voice) and 52.2% (both synthetic voices).

We found similar patterns in the four-medication case. Listeners knew an average of 1.9 medications, and remembered 1.6. The average number of medications recalled for the
human voice was 1.86 (SD 1.1), 1.5 (SD 1.1) for the SPS voice, and 1.6 (SD 1) for the unit selection voice (Table 2).

Message design improved recall. With the basic design, listeners recalled an average of 1.6 medications (SD 1.0), with explanations, 1.8 (SD 1.1), and with repetitions, 2 (SD 1.0). Only the effect of repeating medication names was significant after controlling for all other relevant variables (cf Table 2).

Repeating medications particularly helped listeners understand the synthetic voices. In the difficult listening condition (loud noise, bad signal), listeners recalled an average of 1.7 medications (SD=1.1) in the human voice, but only 1.4 (SD=1.1) when medications were presented in the SPS voice, and 1.5 (SD=1.0) when they heard the unit selection voice. Using repetitions, recall increased to a mean of 2.1 (human, SD=1.0), 1.9 (SPS, SD=1.1) and 2.1 (USEL, SD=1.1) respectively. The USEL voice performed as well as the human voice (pairwise comparisons using t-test with adjustments, p<0.6) and outperformed the SPS voice (p<0.02).

**DISCUSSION AND CONCLUSION**

When presenting short reminders for known medications, a high-quality synthetic voice works as well as a human voice. However, for longer lists of medication names or unknown medication names, recall is not satisfactory, regardless of voice.
Therefore, spoken medication reminders will only work if they build on what users know. In practice, this means that spoken reminders have to be integrated into complex interventions that allow for patient input.

The choice of speech synthesis method matters. While the USEL and the SPS voice were similar overall, the USEL voice, which preserves more acoustic information, was more robust when lists of medication names were repeated, and the SPS voice performed better for one-medication reminders. Reminders could also be sent directly to users’ hearing aids, bypassing much of the ambient distortion.

Repeating key information is significantly better at improving recall than explanations, even though in the exit interviews, participants thought that the explanations were particularly helpful. While repetition alone is usually not enough to help people recognize or recall information,[38,39] in this case, it provides a second chance to catch an auditory glimpse of the speech signal,[40] building a more robust percept that is in turn more likely to be recognized later.

LIMITATIONS

The distractor task was designed to prevent participants from rehearsing the medication name(s) subvocally, a common strategy for remembering auditory information.[41] We need to examine the effect of more realistic distractors that use modalities other than hearing.
It is not clear why working memory had little effect. This could be due to the high proportion of participants with relatively good working memory, or it could be that working memory is less important than other cognitive abilities, such as speed of processing.

In line with best practice in designing health IT,[42] our study was conducted in a laboratory environment. In deployment, reminders will need to be adjusted further to fit in with a person’s concepts of their medication and their routine.[43,44] For example, to aid recall, the names of all medications were clearly visible on a single page; real medicine cabinets are often far more messy.[45,46] We also used actual medication names instead of phrases such as “morning medication” that patients typically use to describe medications. This requires clinicians and pharmacists to cooperate with patients and caregivers. Otherwise, there is a risk that patients may not remember what pills are included in the “morning medication”.

CONCLUSION

In order to assess the usefulness and usability of computer-generated reminders in practice, they need to be integrated into a multimodal medication management or home care system that also provides tailored information about other aspects of the medication such as dosage and side effects.
ACKNOWLEDGEMENTS

We thank Elaine Niven and Dina van der Hulst for training in the administration and scoring of cognitive tests, and Martin Cooke, Elaine Niven and Catherine Mayo for valuable discussion. Simulations of signal transmission quality were kindly provided by the German Telekom Usability Lab, Berlin.

COMPETING INTERESTS

No competing interests declared.

FUNDING

This project was funded by the Chief Scientist Office Scotland, grant number CZG/2/494, and the EPSRC/BBSRC initiative SPARC.

CONTRIBUTION STATEMENT

MW, BM, and CDP designed the study, MW, CDP, CJ, and PC piloted the design, MW conducted the statistical analysis, MW prepared a first draft of the paper, and all authors discussed the interpretation of results and read and commented on several drafts of the paper.
REFERENCES


Figure Legends

Figure 1. Speech synthesis technologies used.

Figure 2. Sample trial.

Figure 3. Examples for each reminder type. In the basic design, which is used for the one-medication reminders and the baseline four-medication reminders, participants hear an introduction (“please remember the following (four) medication(s)”), followed by a pause and the medication names. In the repetition condition, the list of medications is repeated, in the explanation condition, each medication is followed by its indication.
Can Older People Remember Medication Reminders Presented Using Computer-Generated Speech?
Figure 2.
System: Please remember to take the following four medications: Aspirin, Corsodyl, Amoxicillin, and Piriton. [2 second pause] Rachel wins three old mugs.
Experimenter: [Switches to sentence screen]
User: Rachel wins two old mugs.
Experimenter: [Types “Rachel wins two old mugs.”]
Experimenter: [Switches to medication screen which asks user to recall medications]
User: Aspirin, Piriton, and something else.
Experimenter: [ticks Aspirin and Piriton on screen, switches to next task]

Figure 3.
Can Older People Remember Medication Reminders Presented Using Computer-Generated Speech?

<table>
<thead>
<tr>
<th>Reminder</th>
<th>Example</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 medication</td>
<td></td>
</tr>
<tr>
<td><strong>Baseline</strong></td>
<td>Please remember to take the following medication: Aspirin.</td>
</tr>
<tr>
<td>4 medications</td>
<td></td>
</tr>
<tr>
<td><strong>Baseline</strong></td>
<td>Please remember to take the following four medications: Aspirin, Corsodyl, Amoxicillin, and Dulcolax.</td>
</tr>
<tr>
<td><strong>Repetition</strong></td>
<td>Please remember to take the following four medications: Aspirin, Corsodyl, Amoxicillin, and Dulcolax. I repeat: Aspirin, Corsodyl, Amoxicillin, and Dulcolax.</td>
</tr>
<tr>
<td><strong>Explanation</strong></td>
<td>Please remember to take the following four medications: Aspirin, to thin your blood, Corsodyl, for your mouth ulcer, Amoxicillin, for your infection, and Dulcolax, for your constipation.</td>
</tr>
</tbody>
</table>