Can older people remember medication reminders presented using synthetic speech?

Citation for published version:

Digital Object Identifier (DOI):
10.1136/amiajnl-2014-002820

Link:
Link to publication record in Edinburgh Research Explorer

Document Version:
Peer reviewed version

Published in:
Journal of the American Medical Informatics Association : JAMIA

Publisher Rights Statement:
This is a pre-copyedited, author-produced PDF of an article accepted for publication in Journal of the American Medical Informatics Association following peer review. The version of record Wolters, M. K., Johnson, C., Campbell, P. E., Deplacido, C. G., & Mckinstry, B. (2015). Can older people remember medication reminders presented using synthetic speech?. Journal of the American Medical Informatics Association : JAMIA, 22(1), 35-42. 10.1136/amiajnl-2014-002820 is available online at: http://jamia.oxfordjournals.org/content/22/1/35.

General rights
Copyright for the publications made accessible via the Edinburgh Research Explorer is retained by the author(s) and / or other copyright owners and it is a condition of accessing these publications that users recognise and abide by the legal requirements associated with these rights.

Take down policy
The University of Edinburgh has made every reasonable effort to ensure that Edinburgh Research Explorer content complies with UK legislation. If you believe that the public display of this file breaches copyright please contact openaccess@ed.ac.uk providing details, and we will remove access to the work immediately and investigate your claim.
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: 1996 excluding Tables, Figures, Captions, References, and indications of Figure locations

Corresponding author:
Maria K Wolters, School of Informatics, University of Edinburgh
Informatics Forum IF 3.46
10 Crichton Street
Edinburgh EH8 9AB
United Kingdom
maria.wolters@ed.ac.uk

Skype: mkwolters
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 speech synthesis 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 synthetic speech is sufficiently intelligible to allow 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.\cite{14,15} However, most related work \cite{14-19} has focused on older speech synthesis technology and may therefore overestimate intelligibility problems. Here, we used publicly available implementations of the two main current speech synthesis approaches, statistical parametric synthesis \cite{20,21} and unit selection \cite{22,23} (cf Figure 1). Both methods start with a large database of speech. Given a message, unit selection systems search the database for pieces of speech that fit the message and concatenate them. Statistical parametric systems use the database to build a set of statistical models of all core speech units, which are then used to generate arbitrary new messages.

![FIGURE 1 ABOUT HERE](image)

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.

Better message phrasing and organization can counteract unfavourable listening conditions. In our study, we tested two options: adapting the message to how listeners think about medication, for example by linking each medication to the reason for taking it;\cite{24,25} and repeating key information.\cite{26}
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

Based on 12 frequently taken UK medications (6 over-the-counter / 6 prescription), we created 12 one-medication reminders, 24 baseline 4-medications reminders, 12 4-medications reminders where the medications were repeated, and 12 4-medications reminders where indications were explained (c.f. Table 1).
Table 1: Examples of the medication reminders used in the study

<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>

Auditory stimuli were generated using three male voices with a standard British English accent: a human voice; a statistical parametric synthesis voice [21]; and a unit selection 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 hearing aids, no cognitive impairment, no neurodegenerative disorder, and sufficient mobility to attend testing. Of 56 participants who attended, 44 met the inclusion criteria.

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 in their poorer ear. We excluded people with a severe
hearing loss in 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 2 shows a sample trial.

**FIGURE 2 ABOUT HERE**

To support recall, participants were shown a list of 24 possible medications and their key
indication. 12 targets that occurred in the reminders and 12 alternatives 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 (GLMM,[35,36]). Following established best practice for GLMMs,[36] the effect of each predictor is illustrated with a graph showing the 95% confidence interval for that predictor. 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 2. The decline in working memory capacity and the prevalence of hearing loss in the sample is consistent with normal ageing [11,37].

Table 2 Demographics of participants

<table>
<thead>
<tr>
<th>Gender</th>
<th>female</th>
<th>N=24</th>
<th>55%</th>
<th>male</th>
<th>N=20</th>
<th>45%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age</td>
<td>Mean (SD)</td>
<td>M=64 (9)</td>
<td></td>
<td>50-59</td>
<td>N=15</td>
<td>34%</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>60-69</td>
<td>N=14</td>
<td>32%</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>70+</td>
<td>N=14</td>
<td>32%</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>No information</td>
<td>N=1</td>
<td>2%</td>
</tr>
<tr>
<td>Education (Highest Level Achieved)</td>
<td>Postgraduate</td>
<td>N=11</td>
<td>25%</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>-----------------------------------</td>
<td>--------------</td>
<td>------</td>
<td>-----</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Graduate</td>
<td>N=9</td>
<td>20%</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Vocational</td>
<td>N=6</td>
<td>14%</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Secondary School</td>
<td>N=14</td>
<td>32%</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>None</td>
<td>N=1</td>
<td>2%</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No information</td>
<td>N=3</td>
<td>7%</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hearing Loss (on worse ear)</td>
<td>none</td>
<td>N=26</td>
<td>59%</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>mild, monaural</td>
<td>N=7</td>
<td>16%</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>mild, binaural</td>
<td>N=8</td>
<td>18%</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Moderate on worse ear, mild on better ear</td>
<td>N=2</td>
<td>4.5%</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>moderate, binaural</td>
<td>N=1</td>
<td>2%</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Working</td>
<td>Mean (SD)</td>
<td>43 (13)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Memory age 50-59</td>
<td>51 (12)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Capacity^ age 60-69</td>
<td>42 (15)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>age 70+</td>
<td>38 (11)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

^ Secondary school qualifications include GCSE, Leaving Certificate, and A-Levels.

^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.

Figure 3 shows the results of the GLMM analyses for one medication, four medications (baseline message design only), and four medications (message design varied).

FIGURE 3 ABOUT HERE

As expected, listeners with a hearing loss found it more difficult to recall the reminders than listeners with normal hearing. Working Memory Capacity does not have a significant effect after controlling for hearing. Signal quality did not affect recall; the effect of background noise varied depending on the number of medications and message design. Figure 3 illustrates the corresponding parameter values.

Looking at one-medication reminders, we find that 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% (statistical parametric synthesis), and 88% (unit selection), 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 statistical parametric synthesis, and 1.6 (SD 1) for the unit selection voice (Figure 3b).

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 (Figure 3c).

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 using statistical parametric synthesis, 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 (statistical parametric synthesis, SD=1.1) and 2.1 (unit selection, SD=1.1) respectively. Unit selection performed as well as the human voice (pairwise comparisons
using t-test with adjustments, p<0.6) and outperformed statistical parametric synthesis (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. Knowledge can come from users’ own experience of taking medications, from friends and family, or from the media. In practice, this means that spoken reminders have to be integrated into complex interventions that allow for patient input.

Speech synthesis method matters. While both synthetic voices performed similarly overall, unit selection, which preserves more acoustic information, was more robust when lists of medication names were repeated, and statistical parametric synthesis 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 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. We also did not probe why participants were familiar with a given medication.

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. Working memory might also be 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 tailored to fit with a person’s concepts of their medication and their routine and to reduce the potential for unintended disclosure of illness.[43,44]

The medication list provided an overview of possible medications similar to a list provided by a clinician, or to a medicine cabinet where people can see all their medications at once. However, real medicine cabinets (and medication lists) are often far more messy.[45,46] Reading level may also have affected participants’ ability to use the list.
Finally, we used actual medication names instead of phrases such as “morning medication” that patients typically use to describe medications. However, such phrases require clinicians and pharmacists to cooperate with patients and carers. Otherwise, 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 system that also provides tailored information about other aspects of the medication such as dosage and side effects.
ACKNOWLEDGEMENTS

We thank Vasilis Karaiskos for programming and conducting the listening tests, 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, Technical University of 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 1

The figure shows the process of generating synthetic speech for pronouncing a word, specifically "aspirin." The process involves:

1. **Unit Selection**
   - Search database for appropriate units
   - Select units, smooth over links

2. **Stress on first syllable, pronunciation: ‘aspirin’**
   - Stress on the first syllable
   - Pronunciation: ‘as@r-in’

3. **Statistical Parametric Synthesis**
   - Choose appropriate speech models for text from database
   - Generate speech from statistical models
   - Add intonation and statistical models

The diagram illustrates the steps from selecting units to generating the actual speech sound, highlighting the pronunciation and stress on the first syllable.
FIGURE 2
FIGURE 3
Can Older People Remember Medication Reminders Presented Using Synthetic Speech?
Can Older People Remember Medication Reminders Presented Using Computer-Generated Speech?