Examining Prosody in Spoken Navigation Instructions for People with Disabilities

Cathy Jiao[♠] **Aaron Steinfeld**[♢] **Maxine Eskenazi**[♠]
[♠]Language Technologies Institute, Carnegie Mellon University
[♢]Robotics Institute, Carnegie Mellon University
{cljiao, steinfeld, max}cs.cmu.edu

Abstract

The introduction of conversational systems have made synthesized speech technologies common tools for daily activities. However, not all synthetic speech systems are designed with the needs of people with disabilities in mind. This paper describes a study in which 198 people – 80 participants with self-reported disabilities and 118 participants without – were recruited to listen to navigation instructions from a spoken dialogue system with different prosodic features. Results showed that slowing down speech rate aids in participants' number recall, but not in noun recall. From our results, we provide suggestions for developers for building accessible synthetic speech systems.

1 Introduction

The introduction of conversational systems such as Apple's Siri and Amazon's Alexa have made synthesized speech technologies common tools for daily activities. However, people with disabilities still struggle to interact with synthetic speech systems (Vieira et al., 2022). On the NLP side, current work in accessibility focuses on enhancing system or model functionality. Examples include finding appropriate data to train deep learning models that can be tailored to people with disabilities (Yaneva et al., 2017), or examining negative language model biases against disabilities (Venkit et al., 2022). Yet, user evaluation of these models deployed in systems are limited. On the other hand, past work in HCI and speech systems have noted the importance of using appropriate prosody to achieve better user understanding and recall of the system outputs. (Duffy and Pisoni, 1992; Mirenda and Beukelman, 1987; Paris et al., 2000; Wester et al., 2016; Rodero, 2017). To tie together perspectives from both NLP and HCI, we conducted a study in which user recall of instructions from an existing spoken dialogue system was evaluated in order to (1) determine what speech features are most beneficial for user recall of information, and (2) decide on future features to implement in the system. We recruited 198 people (with and without disabilities) and asked them to recall information under speech conditions which had either (1) *slowed down* speech or (2) *pauses before keywords* in the instruction. Furthermore, we grounded our study to the realm of understanding navigation instructions, which is a challenging setting because it requires users to remember exact *numeric entities* (e.g., departure times, building numbers) and *noun entities* (e.g., unfamiliar street names) in order to navigate effectively.

Our results showed that across all participants (with and without disabilities), slowing down speech rate aided in recalling numeric entities, even as the number of numeric entities in an instruction increased, but was less effective for noun entity recall. Furthermore, in speech conditions where breaks were inserted before numeric and noun entities, participants in general had lower performance in information recall. Our findings suggest that developers may be able to make adjustments that promote overall accessibly. In addition, appropriate cues for different types of information (i.e., numeric vs. noun) should be considered.

2 Related Work

Appropriate prosody in synthetic speech systems is important since synthetic speech often lacks natural cues and pacing found in human voices. Previous works have shown that inappropriate or absent prosodic cues led to large performance gaps in recall tasks from natural and synthetic speech (Paris et al., 2000), and that human voices were preferred

Proceedings of the Third Workshop on Bridging Human–Computer Interaction and Natural Language Processing (3rd HCI+NLP Workshop), pages 1–12

over synthetic voices as task complexity increases (Rodero, 2017). On the other hand, appropriate prosodic cues can improve peoples' abilities to disambiguate information – for example, distinguishing between similar spoken mathematical expressions (Gellenbeck and Stefik, 2009). Furthermore, adding appropriate prosodic cues can decrease performance gaps between different age groups in understanding speech (Stine and Wingfield, 1987; Wingfield et al., 1992; Langner and Black, 2005; Wolters et al., 2007; Roring et al., 2007). Some existing NLP systems have taken this information into account. For instance, CMU GetGoing (Mehri et al., 2019) is a trip-planning dialogue that introduced attention-grabbing prefixes and allowed "barge-in" to provide easier interaction for seniors. However, many modern spoken dialogue systems still fail to capture nuances of interaction for people with disabilities (Vieira et al., 2022). While previous work such as Koul (2003) showed that synthetic speech comprehension among people with disabilities were influenced by the complexity of task and acoustic-phonetic features, our study differs from previous work in that we grounded our study in a trip-planning task. This task is challenging to users (i.e., recalling unfamiliar street names), yet is important for independent daily travel, especially for people who are dependent on public transit.

3 Overview

To analyze the effects of speech prosody on information retention, we recruited participants with 108 and without self-reported disabilities to listen to 109 audio clips containing navigation instructions. The outline of our paper is as follows: in Section 4 we describe the audio clip collection process for the study. In Section 5 we explain the study design and participant recruitment. Finally, in Sections 6, 7 and 8 we provide study results and further discussion.

4 Audio Clips Collection

Audio clips were collected from GetGoing, a spoken dialog system that provides step-by-step navigation instructions to aid senior users in trip planning. GetGoing was previously deployed in Southwestern Pennsylvania and could be accessed by users through telephone. It follows a traditional pipeline dialog system architecture with natural language understanding and speech modules. Notably, the system uses Google Maps API[1] to curate sets of directions between two locations. Furthermore, the system uses Vonage's text-to-speech API [2] to output synthesized speech to the user.

The audio clips collection process involved three stages. In the *first stage*, GetGoing was used to generate navigation instructions using random start and end destination points around Southwestern Pennsylvania. In the *second stage*, a subset of the generated instructions was selected according to two information parameters defined in Section 4.1 to ensure our instruction set was well-balanced. This resulted in 48 unique instructions. In the *third stage*, speech conditions (described in Section 4.2) were created by either slowing down the speech or inserting breaks before keywords using the Vonage API. We recorded the speech output of GetGoing for each of the 48 instructions under each condition, and ended up with a total of 192 audio clips.

4.1 Navigation Instruction Parameters

To ensure the curation of a balanced set of navigation instructions from GetGoing, we defined each instruction by two information parameters. The first parameter is the *number of nouns entities* in an instruction. Noun entities included street and building names (e.g., *"Frew St"*, *"UPMC Presby"*). The second parameter is the *number of numeric entities* in an instruction. Numeric entities included any word that contained a number (e.g., *"9:30pm"*, *"7th Street"*). Certain items in an instruction were counted as both a noun entity and numeric entity. For instance, in the instruction *"Take the bus to Liberty avenue and 7th Street."*, the name *"7th Street"* was counted as a noun entity and numeric entity. We used noun and numeric entities as information parameters for two reasons. First, when navigating it is important to remember names, streets, and times. Second, our instructions are not multi-step and do not contain any other major pieces of information.

Next, we defined eight unique parameter combinations, or groups, which are listed in Table 1. Each parameter group contained 1-2 noun enti-

[1] https://developers.google.com/maps/documentation/

[2] https://www.vonage.com/developer-center/

Group	Numeric Entities	Noun Entities	Words *mean*	*sd*
1	3	2	17.3	3.6
2	2	2	18.7	3.6
3	1	2	19.2	3.9
4	0	2	14.2	3.6
5	3	1	16.2	3.4
6	2	1	15.8	3.0
7	1	1	11.8	2.5
8	0	1	11.3	2.4

Table 1: Navigation instruction parameter groups. Each group contains a set number of noun and numeric entities per instruction. The average number of words per instruction in each group is listed on the right column.

ties and 0-3 numeric entities. For each parameter group, six instructions were randomly selected from the inital set of instructions collected from GetGoing, resulting 48 unique instructions. Table 7 in the Appendix lists all instructions and their parameter values.

4.2 Audio Clip Conditions

Using Vonage's text-to-speech API, we created four speech conditions:

1. **Default**: The default Vonage API voice.
2. **Default-slow**: The Vonage API voice with the speech rate set to “slow".
3. **Break-short**: The default Vonage API voice with a 5ms break before every noun and numeric entity.
4. **Break-long**: The default Vonage API voice with a 15ms break before every noun and numeric entity.

We selected these conditions since previous work has shown that inserting pauses can aid in speech understanding, especially for seniors (Langner and Black, 2005; Wolters et al., 2007). While other prosodic cues and its effects on information retention can be explored, we decided to focus on these four conditions for the scope of our experiments.

We applied each of the four conditions to each of the 48 instructions to create 192 audio clips. Table 2 shows the duration of clips according to each condition. Clips in the *Default* condition had the shortest average duration, while clips in *Default-slow* had the longest average duration.

Condition	Clip Duration (s) *mean*	*sd*
Default	7.6	1.9
Default-slow	8.8	2.0
Break-short	7.8	1.9
Break-long	8.0	2.0

Table 2: Average audio clip lengths within the four speech conditions.

5 Study Design

In this section, we describe the study design, procedure, and participants. All participants were recruited on Prolific.co. From there, they were directed to a Qualtrics page to consent to the study. After consenting, participants were redirected to a website to do the tasks. An example of the task interface is provided in Figure 4 in the Appendix. Each participant listened to 24 audio clips and entered what they heard in a text box on the website. Participants listened to six clips from each of the four conditions. Each of these six clips were randomly sampled from instructions from different parameter groups. All clips were presented to the participant in random order, and the website prevented participants from listening to the clip more than once in order to test recall. Once the participants finished the tasks, they were redirected back the the Qualtrics page where they answered a short questionnaire, which is included in Table 6 in the Appendix.

5.1 Participants

We recruited 198 adult participants from the U.S. through Prolific.co (two additional participants did not complete the study). The participants ages ranged from 18 to 78 years (*mean = 37.3, sd = 14.3*), and the age breakdown is shown Table 4. The participants reported their genders as follows: 105 participants self-identified as female and 82 as male, 10 as other genders, and 1 did not disclose their gender. The majority of participants reported English as their native language *(n = 192)*. Furthermore, the majority of participants reported that they used a computer daily on a scale from “1 (Never)" to “7 (Often - daily)" *(mean = 6.85, sd =*

Self-Reported Disabilities	Participants (n=198)
cognitive disability	12
communicative disability	3
dexterity disability	9
hearing impairment	13
mental disability	39
mobility disability	17
vision impairment	23
other	8
none	118

Table 3: Participants' self-reported disabilities. Some participants have multiple self-reported disabilities

Age Group	Participants (n=198)
18-24	35
25-34	70
35-44	41
45-54	21
55-64	19
65+	12

Table 4: Participants' ages.

0.51, min = 4, max = 7) and occasionally used a voice assistant *(mean = 3.58, sd = 2.09, min = 1, max =7)*.

We used filters from Prolific.co to recruit participants with and without disabilities. We first recruited 98 participants without any selection criteria. We then used five different Prolific filters to recruit people with varying disabilities. More specifically, we recruited participants who indicated they have (1) *vision*, (2) *hearing*, (3) *mobility*, (4) *chronic conditions*, and (5) *cognitive* disabilities on Prolific. Twenty participants were recruited using each filter. In addition, we asked participants to self-report any disabilities during the post-study questionnaire (shown in Table 6 in the Appendix) to handle discrepancies and ambiguity from the Prolific filter categories. The breakdown of participant self-reported disabilities is reported in Table 3.

6 Results

We grouped participants in two categories: those who had self-reported disabilities and those who did not, and analyzed recall of noun and numeric entities across the four speech conditions. While it is possible to examine the subgroups of self-reported disabilities individually (e.g., all participants with a mobility disability), we avoided this since the subgroups are imbalanced with respect to the number of participants. In addition many participants reported more than one disability.

6.1 Noun Retention Accuracy

Annotations for nouns were done manually, and each noun was assigned as "correct" or "incorrect" by phonetic similarity since participants transcribed text from audio. For example, *"Knight St"* was considered to be the same as *"Nite Street"*. Figure 1 shows the results.

In clips with *one noun*, little difference in noun accuracy (averaged across all conditions) was observed among people with self-reported disabilities (*mean = 94.27*) and those without (*mean = 90.61*. In clips with *two nouns*, the noun accuracy dropped in both groups, but the difference still remained small between people with self-reported disabilities (*mean = 46.35*) and those without (*mean = 43.08*). Large changes were not observed in noun accuracy across different conditions among both groups.

6.2 Number Retention Accuracy

In clips with *one number*, little difference in number accuracy (averaged across all conditions) was observed between people with self-reported disabilities (*mean = 79.38*) and those without (*mean = 78.25*). In clips with *two numbers*, the number accuracy dropped in both groups. However, there was greater decrease in performance for people with self-reported disabilities (*mean = 36.46*) than those without (*mean = 43.08*). In clips with *three numbers*, little difference was observed in the number accuracy between people with self-reported disabilities (*mean = 34.17*) and those without (*mean = 35.03*).

Unlike noun accuracy, differences in number accuracy were noticeable across conditions. For instructions with ≥ 2 numbers, participants (those with and without self-reported disabilities) in general performed the highest in number recall the *default-slow*. Meanwhile participants performed the worst in *break-short*, followed by *break-long*.

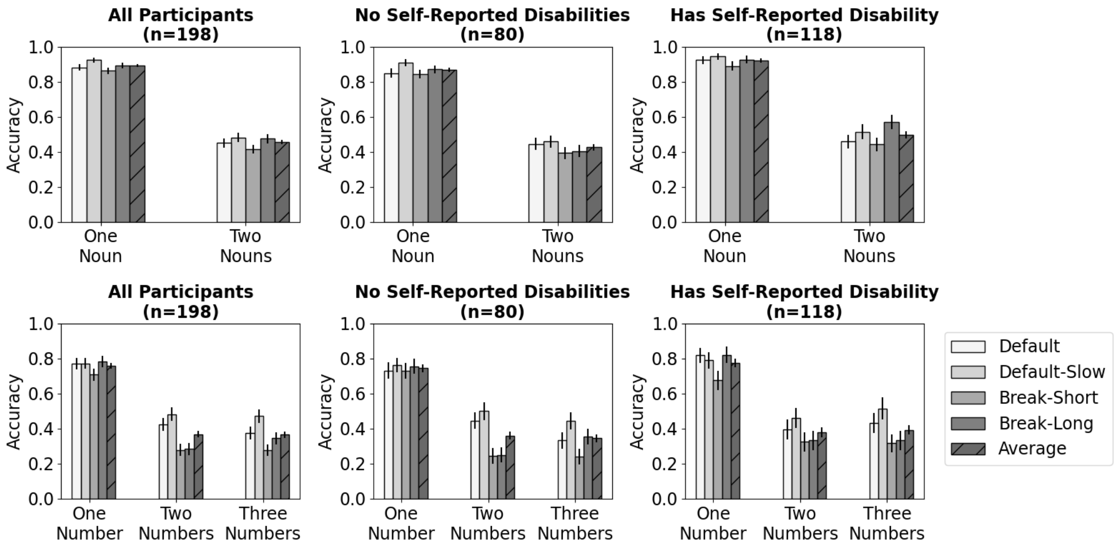


Figure 1: Noun and number recall accuracy of participants across conditions. The x-axis reflects the number of noun entities and numeric entities in a clip.

Notably, *default-slow* was the only condition which obtained higher accuracy than *default*, which suggested an overall slower speech rate was more beneficial in number recall than inserted pauses before numeric information.

7 Further Analysis

7.1 Parameter Lengths

Since the content of the information parameters values can differ – for instance, recalling a short and common street name (e.g., *"Main St"*) versus a long street name (e.g., *"Presto-Sygan Rd"*) – we further analyzed participants' retention across different conditions with respect noun and number lengths.

7.1.1 Noun Length

To evaluate noun recall with respect to its length (i.e., the number of characters in a noun) we used a finer resolution metric: first, given the participant's transcription, $\hat{t}$, and the true transcription, t, we applied Metaphone (Philips, 1990) which converted the transcriptions to standardized string representations of its English pronunciations. Using these representations, $\hat{m} \leftarrow \text{Metaphone}(\hat{t})$ and $m \leftarrow \text{Metaphone}(t)$, we then measured noun retention by taking the Levenshtein distance between $\hat{m}$ and m. Figure 3 shows the results. Little difference in noun retention was seen between the conditions even as noun length increased across all participants. Furthermore, little difference in performance was noticed between participants with and without self-reported disabilities as noun length increased. However, noun retention rapidly decreased starting from nouns with ≥ 16 characters (Figure 3). This suggested that very long nouns should be kept to a minimum in order to help users retain information.

7.1.2 Number Length

Figure 2 shows number accuracy across different conditions as number length – *the number of digits in a number* – increased. Across all participants, number recall was highest in the *default-slow* condition as number length increased for both participants with and without self-reported disabilities. In addition, *default-slow* was also the only condition that had higher number accuracy than the *default* condition.

7.2 Retention by Age Group

Since our study was motivated by previous work in improving accessibility in dialogue system for seniors, we examined the effects of the four conditions on all participants with respect to their ages. As shown in Table 5, the participants' noun recall (averaged across all conditions) decreased slightly as age increased: from *mean acc = 68.0* for participants aged 25-35 to *mean acc = 65.3* for par-

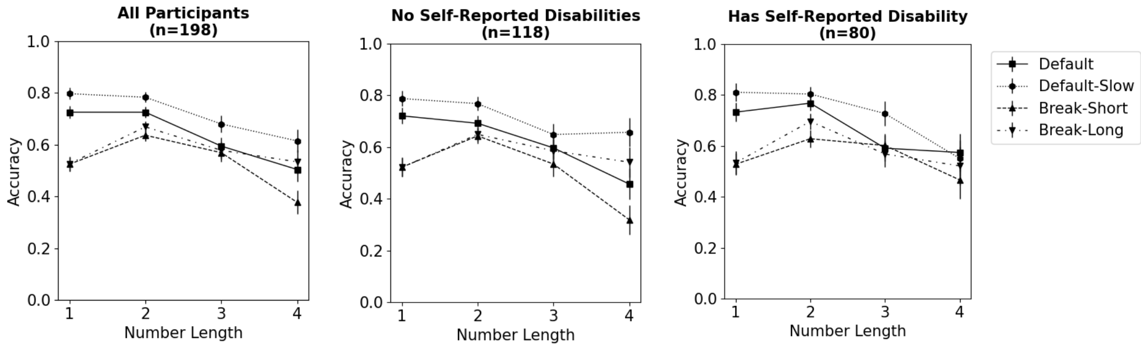


Figure 2: Retention accuracy of participants as the number of digits in a number increases.

Figure 3: Participants' recall as proper noun length increase. Proper noun recall is measured by Levenshtein distance between the Metaphone (Philips, 1990) representation of the user responses (a higher value indicates lower performance).

Age	Noun Acc.		Number Acc.	
	mean	*sd*	*mean*	*sd*
18-24	69.0	46.3	47.5	49.9
25-34	68.0	46.7	52.4	49.9
35-44	68.1	46.6	52.8	49.9
45-54	63.7	48.1	48.0	50.0
55-64	65.3	47.6	44.4	49.7
65+	65.5	47.5	42.1	49.4

Table 5: Noun and number accuracy by age group (averaged across all condition)

ticipants aged 65+. As shown in Figure 5 in the Appendix, when *one noun* was present in a clip, were no large changes in performance across age groups among the different conditions. However, when *two nouns* were present, performance fluctuated in different conditions as age increased.

Unlike noun recall, performance in number recall noticeably decayed as age increased: from *mean acc = 52.4* for participants aged 25-35 to *mean acc = 42.1* for participants aged 65+ as shown in Table 5. Figure 5, shows when *one number* was present in a clip, participants' performance in the conditions were similar in younger age groups, but differed in older age groups. Notably, when there were ≥ 2 numbers in an instruction, performance in *default-slow* was often better than the other conditions (similar to results in Section 6.2). These findings supported previous work which showed that slower speech is preferred among older populations (Langner and Black, 2005; Wolters et al., 2007; Mehri et al., 2019).

8 Discussion

There are several implications that can be drawn from our findings. First, from our experiments in Section 6.2, we observed that speech conditions did not lead to large differences on noun retention between (1) individuals with and without self-reported disabilities, and (2) across age groups. In addition, speech conditions had little effect on noun retention even as noun length increased across all participants as highlighted in

Section 7.1.2. This suggests that system designers may have some flexibility when presenting nouns to users, as long as the nouns (i.e., names, streets, places) are not long. However, additional care may be taken into consideration when some nouns are more important to remember than others, and this is left as a direction for future research

Next, with respect to number retention, in Section 6.1 we observed that having a slower overall speech rate was helpful for number recall, while inserting pauses before numbers had negative effects. Furthermore, results in Section 7.1.1 showed that a slower overall speech rate also aided in number retention as length of numbers (i.e., number of digits in a number) increased. However, a follow up question is why participants in general performed worse in *break-short* and *break-long* conditions compared to *default-slow* and *default* conditions with respect to number retention. A possible explanation may be that added pauses before every key piece of information caused participants to focus on too many cues. This can overload the participants with information, and cause forgetting. Further investigation on where to place appropriate pauses is beneficial. Overall, these findings suggest that system developers should take into account speech prosody when communicating numbers, regardless of whether the user has a disability.

Interestingly, our study showed little difference in noun and number recall between participants with and without self-reported disabilities in general. A possible explanation is that participants did our study online, and therefore had their environment and computing device set up for good listening. For example, people with hearing loss may have turned up the volume on their speakers, worn headphones, or enabled a Bluetooth connection to hearing aids. Hence, future research should explore use in less optimal conditions (e.g., using a phone from a city street).

Based on our work, we have several recommended directions for future research. One point to consider is limiting prosodic cues based on information priority. For instance, rather than inserting pauses before every noun and number, only inserting pauses before long numbers and uncommon nouns may lead to a positive effect on recall. Furthermore, we acknowledge that our study is limited to noun and number recall, and that our analysis considered these parameters to be independent of each other. In realistic settings, other important pieces of information may also be present in navigation directions. For instance, instructing the user to "*turn right* on Frew St and *go up the ramp* to the bus station" adds additional load to remember specific remember actions they must take. Also, some pieces of information are more important for the user to understand and recall than others. For example, a user may prioritize recalling bus arrival times over the name of their destination stop street.

Finally, finding effective prosodic features for information retention could be explored as a problem of personalization (e.g., blind user preferences for screenreaders). For example, allowing users to select voice styles alongside information presentation styles can allow for easier usage for individuals with different disabilities. Finding successful ways to achieve this goal also requires further investigation on how people with disabilities interact with spoken dialogue systems.

9 Conclusion

In this study, we recruited people with and without disabilities and evaluated their information retention in different speech conditions. We found that having an overall slow speech rate was useful for number retention across all participants (with and without self-reported disabilities), but was less effective in improving noun retention. We also showed that inserting breaks before nouns and numbers did not improve in information retention. Thus, finding appropriate prosodic cues for different pieces of spoken information is an interesting direction to explore.

Ethics and Limitations

Our study was approved by our Institutional Review Board. Each participant received $7.50 USD for participating in the study, and took on average 19 minutes (SD = 12) to complete the study. While investigating information retention from dialogue systems across different languages and cultures is important, we note our recruitment was limited to participants from the US and participants were mostly native English speakers.

Acknowledgements

This work was supported by a grant from the National Institute on Disability, Independent Living, and Rehabilitation Research (NIDILRR grant number 90REGE0007).

References

Susan A Duffy and David B Pisoni. 1992. Comprehension of synthetic speech produced by rule: A review and theoretical interpretation. *Language and Speech*, 35(4):351–389.

Ed Gellenbeck and Andreas Stefik. 2009. Evaluating prosodic cues as a means to disambiguate algebraic expressions: An empirical study. In *Proceedings of the 11th International ACM SIGACCESS Conference on Computers and Accessibility*, Assets '09, page 139–146, New York, NY, USA. Association for Computing Machinery.

Rajinder Koul. 2003. Synthetic speech perception in individuals with and without disabilities. *Augmentative and Alternative Communication*, 19(1):49–58.

Brian Langner and Alan W Black. 2005. Using speech in noise to improve understandability for elderly listeners. In *IEEE Workshop on Automatic Speech Recognition and Understanding, 2005.*, pages 392–396. IEEE.

Shikib Mehri, Alan W Black, and Maxine Eskenazi. 2019. Cmu getgoing: An understandable and memorable dialog system for seniors.

Pat Mirenda and David Beukelman. 1987. A comparison of speech synthesis intelligibility with listeners from three age groups. *Augmentative and Alternative Communication*, 3(3):120–128.

Carol R. Paris, Margaret H. Thomas, Richard D. Gilson, and J. Peter Kincaid. 2000. Linguistic cues and memory for synthetic and natural speech. *Human Factors*, 42(3):421–431. PMID: 11132803.

Lawrence Philips. 1990. Hanging on the metaphone.

Emma Rodero. 2017. Effectiveness, attention, and recall of human and artificial voices in an advertising story. prosody influence and functions of voices. *Computers in Human Behavior*, 77:336–346.

Roy W Roring, Franklin G Hines, and Neil Charness. 2007. Age differences in identifying words in synthetic speech. *Human factors*, 49(1):25–31.

Elizabeth L Stine and Arthur Wingfield. 1987. Process and strategy in memory for speech among younger and older adults. *Psychology and aging*, 2(3):272.

Pranav Narayanan Venkit, Mukund Srinath, and Shomir Wilson. 2022. A study of implicit bias in pretrained language models against people with disabilities. In *Proceedings of the 29th International Conference on Computational Linguistics*, pages 1324–1332, Gyeongju, Republic of Korea. International Committee on Computational Linguistics.

Alessandro Diogo Vieira, Higor Leite, and Ana Vitória Lachowski Volochtchuk. 2022. The impact of voice assistant home devices on people with disabilities: A longitudinal study. *Technological Forecasting and Social Change*, 184:121961.

Mirjam Wester, Oliver Watts, and Gustav Eje Henter. 2016. Evaluating comprehension of natural and synthetic conversational speech. In *Proc. speech prosody*, volume 8, pages 736–740.

Arthur Wingfield, Sarah C Wayland, and Elizabeth AL Stine. 1992. Adult age differences in the use of prosody for syntactic parsing and recall of spoken sentences. *Journal of Gerontology*, 47(5):P350–P356.

Maria Wolters, Pauline Campbell, Christine DePlacido, Amy Liddell, and David Owens. 2007. Making speech synthesis more accessible to older people. In *6th ISCA Workshops on Speech Synthesis (SSW-6)*.

Victoria Yaneva, Constantin Orăsan, Richard Evans, and Omid Rohanian. 2017. Combining multiple corpora for readability assessment for people with cognitive disabilities. In *Proceedings of the 12th Workshop on Innovative Use of NLP for Building Educational Applications*, pages 121–132, Copenhagen, Denmark. Association for Computational Linguistics.

Appendix

1. What is your age?
2. What is your native language?
3. What other languages do you speak?
4. What is your gender?
5. Are there other aspects of your identity that are important to you (racial, ethnic, or otherwise)?
6. How often do you use a computer? Answer on a scale from 1 to 7.
   1 = Never. 7 = Often (daily).
7. How often do you use a smart voice assistant (Siri, Alexa, etc)? Answer on a scale from 1 to 7.
   1 = Never. 7 = Often (daily).
8. Do you have a disability that you would like to disclose? (Select as many as you like.)
   (a) I have a mobility device/disability
   (b) I have a dexterity disability
   (c) I have a vision impairment
   (d) I have a hearing impairment
   (e) I have a communication impairment
   (f) I have a cognitive impairment
   (g) I have a mental issue
   (h) Other (please describe)
   (i) I prefer not to disclose
   (j) None

Table 6: Post-Study Questionnaire

Clip	Noun 1	Noun 2	Number 1	Number 2	Number 3
Next, you should Walk for about 14 mins to Fifth avenue + Beechwood.	Fifth avenue	Beechwood	14	Fifth	-
Next, you should Take the 53L to Fourth avenue at Wood Street It will depart at 9:51pm.	Fourth avenue	Wood Street	53	9:51	-
Next, you should Walk for about 4 mins to 4200 Fifth avenue.	Fifth avenue	-	4	4200	Fifth
Next, you should Walk for about 4 mins to Baum Blvd + Liberty avenue.	Baum Blvd	Liberty avenue	4	-	-
Next, you should Walk to 8th Street + 6th StreetNS at 9:37pm.	8th Street	6th StreetNS	9:37	8th	6th
The final step is to Take the 51 to Brownsville Rd. It will depart at 9:58pm.	Brownsville Rd	-	51	9:58 PM	-
The final step is to Walk for 15 mins to UPMC Presby.	UPMC Presby	-	15 mins	-	-
The final step is to Take the bus to Liberty avenue + 7th Street.	Liberty avenue	7th Street	7th	-	-
The first thing that you want to do is to Walk for about 18 mins to South Busway + Pioneer avenue Ramp far side.	South Busway	Pioneer avenue Ramp far side	18 mins	-	-
Next, you should Turn left onto Panther Hollow Road	Panther Hollow Road	-	-	-	-
The final step is to Take the bus to Fifth avenue + University Place It will depart at 6:51pm.	Fifth avenue	University Place	Fifth	6:51 PM	-
Next, you should Walk for about 25 mins to 400 Presto-Sygan Rd	Presto-Sygan Rd	-	25 mins	400	-
The first thing that you want to do is to Take the 88 to 7th Street+ Penn avenue. It will depart at 9:19pm.	7th Street	Penn avenue	88	7th	9:19 PM
The final step is to Take the 89 to Frick Park	Frick Park	-	89	-	-
The first thing that you want to do is to Walk for about 5 mins to Wood Street+ Third avenue far side.	Wood Street	Third avenue far side	5 mins	-	-
The first thing that you want to do is to Walk to Island avenue + Chartiers near side .	Island avenue	Chartiers near side	-	-	-
Next, you should Take the 21 to Stanwix Street. It will depart at 9:16pm.	Stanwix Street	-	21	9:16 PM	-
The final step is to Turn right on Forest avenue	Forest avenue	-	-	-	-
The first thing that you want to do is to Walk for about 4 mins to 46 18th Street.	18th Street	-	4 mins	46	18th
Next, you should Take the 88 to Liberty avenue + 17th Street. It will depart at 7:01pm.	Liberty avenue	17th Street	88	17th	7:01 PM
The first thing that you want to do is to Take the 71D to Hamilton avenue + Lang. It will depart at 9:45pm.	Hamilton avenue	Lang	71D	9:45 PM	-
The first thing that you want to Walk for 16 mins on Frew Street.	Frew Street	-	16 mins	-	-
The final step is to Walk for about 5 mins to 7101 Frankstown avenue.	Frankstown avenue	-	5 mins	7101	-
The final step is to Walk for about 1 min to 2900 7th Street.	7th Street	-	1 min	2900	7th
The final step is to Take the 70D to Stanwix Street.	Stanwix Street	-	70D	-	-
Next, you should Walk to Sarah Street+ East Carson Street.	Sarah Street	East Carson Street	-	-	-
The final step is to Take the 31 to Washington avenue + James Streetfar side . It will depart at 10:26pm.	Washington avenue	James Streetfar side	31	10:26 PM	-
The first thing that you want to do is to Take the 71B for about 15 mins to 395 Jacks Run Road.	Jacks Run Road	-	71B	15 mins	395
Next, you should Take the bus to Freeport Rd + Butler.	Freeport Rd	Butler	-	-	-
The first thing that you want to do is to Take the 13 to Forest avenue.	Forest avenue	-	13	-	-
Next, you should Walk for 10 mins McKnight Rd.	McKnight Rd	-	10 mins	-	-
The first thing that you want to do is to Take the 56 to Brownsville	Brownsville	-	56	-	-
The final step is to Walk for about 2 mins to Liberty avenue + Fifth avenue.	Liberty avenue	Fifth avenue	2 mins	Fifth	-
The first thing that you want to do is to Walk to 5th Street + 17th Street at 10:45pm.	5th Street	17th Street	5th	17th	10:45 PM
The final step is to Take bus 61 for about 2 mins to 5235 Clairton Boulevard.	Clairton Boulevard	-	61	2 mins	5235
The first thing that you want to do is to Take the 28X to Forbes avenue. It will depart at 9:27pm.	Forbes avenue	-	28X	9:27 PM	-
The first thing that you want to do is to Walk for 16 mins to 300 Monongahela avenue.	Monongahela avenue	-	16 mins	300	-
Next, you should Take the 34 to Shadyside Village	Shadyside Village	-	34	-	-
Next, you should Take the bus to Cambronne Street+ Winhurst. It will depart at 9:58pm.	Cambronne Street	Winhurst	9:58 PM	-	-
The final step is to Walk to 51th Street + 19th Street at 9:45pm.	51th Street	19th Street	9:45 PM	51th	19th
Next, you should Take the 88 to Halket Street	Halket Street	-	88	-	-
The final step is to Walk for about 2 mins to Penn avenue + Village of Eastside Shpg Ctr	Penn avenue	Village of Eastside Shpg Ctr	2 mins	-	-
The final step is to Walk to Liberty avenue at Wood Street.	Liberty avenue	Wood Street	-	-	-
Next, you should Take bus 19 for about 9 mins to 7034 Blackhawk Street.	Blackhawk Street	-	19	9 mins	7034
The first thing that you want to do is Take the bus to Giant Eagle Drive + Iggle Video.	Giant Eagle Drive	Iggle Video	-	-	-
The final step is to Take the 75 to 5th avenue / Halket Street It will depart at 10:09pm.	5th avenue	Halket Street	75	5th	10:09 PM
The first thing that you want to do is to Take the bus to Sandusky Street+ General Robinson Street. It will depart at 9:41pm.	Sandusky Street	General Robinson Street	9:41 PM	-	-
The first thing that you want to do is to Walk to Main Street	Main Street	-	-	-	-

Table 7: Instructions

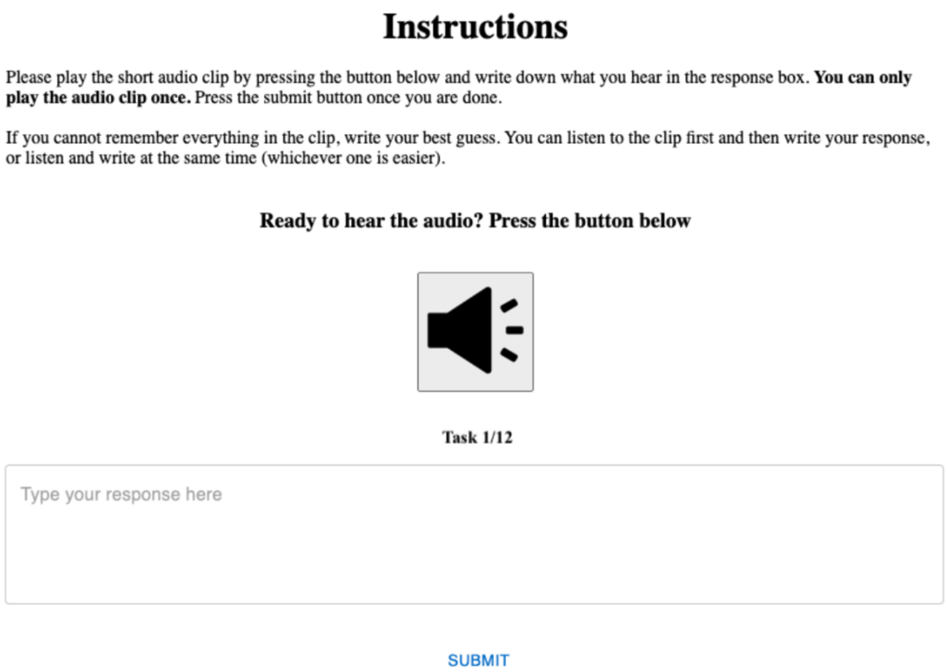
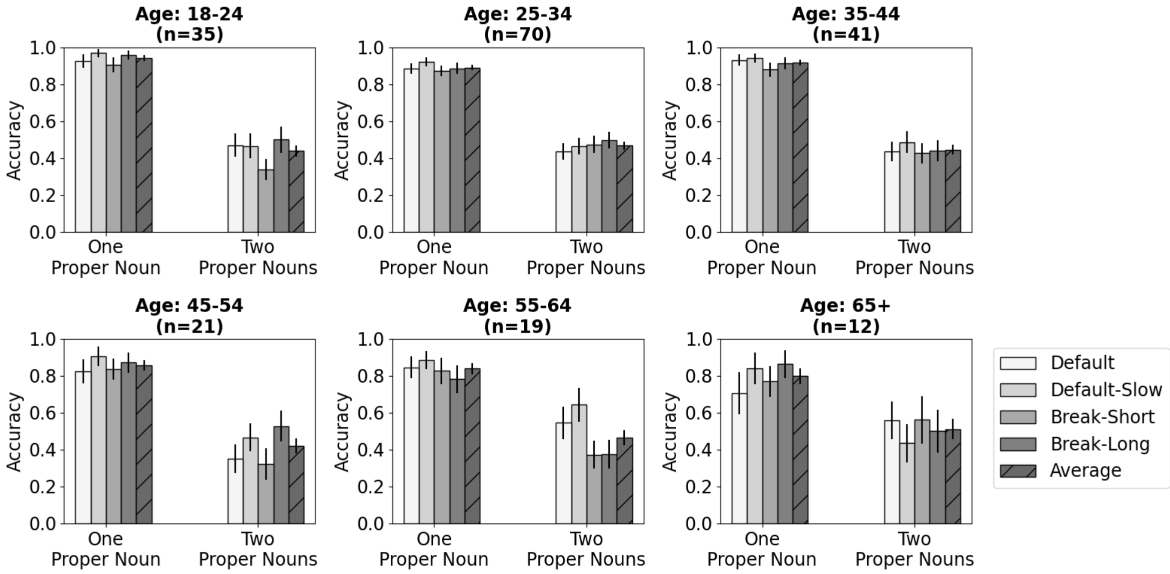
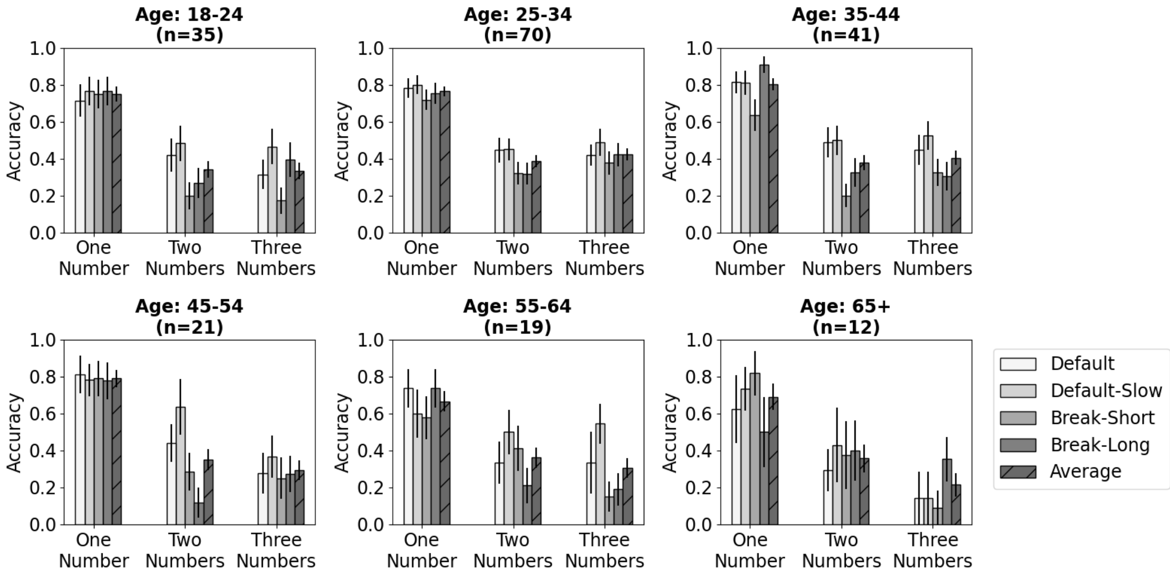


Figure 4: Task interface



(a) Noun recall accuracy of age groups across conditions.



(b) Number recall accuracy of age groups across conditions.

Figure 5: Noun and number recall accuracy of participants across conditions with respect to different age groups.