Proceedings of Meetings on Acoustics

Volume 19, 2013 http://acousticalsociety.org/

ICA 2013 Montreal
Montreal, Canada
2 - 7 June 2013

Speech Communication
Session 4pSCb: Production and Perception I: Beyond the Speech Segment (Poster Session)

4pSCb12. The influence of multiple narrators on adults' listening comprehension

Brittan A. Barker* and Cornetta Mosely

***Corresponding author's address: Louisiana State University, Baton Rouge, LA 70806, barkerb@lsu.edu**

Research has demonstrated that variable, talker information-such as the number of talkers-affects listeners' perception and processing of linguistic information during various laboratory tasks. In particular, the detrimental effects of multiple talkers are highlighted during online speech perception tasks with little contextual support (isolated word recognition; e.g., Mullennix et al., 1989; Ryalls & Pisoni, 1997; Sommers & Barcroft, 2011). Nonetheless, it is unclear how multiple talkers might affect listeners' perception of linguistic information in more complex spoken language tasks utilizing real-time, fluent speech. The present experiments were conducted determine if information contributed by multiple talkers influences adults' auditory story comprehension in the presence of both quiet and background noise. The accuracy and reaction time data did not support the hypothesis that talker information directly affects the perception of linguistic information during auditory story comprehension. Thus these data bring to light theoretical perspectives that emphasize the importance of looking across experimental tasks to better understand talker-specific information's pattern of influence on spoken language processing (e.g., Sommers & Barcroft, 2006; Werker & Curtin, 2005).

Published by the Acoustical Society of America through the American Institute of Physics

 [DOI: 10.1121/1.4800954]
Received 22 Jan 2013; published 2 Jun 2013

Research has demonstrated that variable, talker information—such as the number of talkers—affects listeners' perception and processing of linguistic information during assorted laboratory tasks. In particular, the detrimental effects of multiple talkers are highlighted during online speech perception tasks with little contextual support, such as isolated word identification[1-6]. However despite this growing body of research examining talker-specific speech perception, more research is needed. More information is specifically needed to determine if and when talker-specific information is mandatory or supplemental and when it is facilitative or detrimental, especially in complex spoken language tasks utilizing real-time, fluent speech[7]. Understanding the role of talker-specific information in linguistic processing is not only important for refining and/or supporting predictions of theoretical models of speech perception and processing (e.g., PRIMER[8] or the Representation Quality hypothesis[9]), but it may also be useful for thinking about the types of added barriers hearing impaired listeners face when listening to spoken language via the restricted signals of their listening devices (e.g., cochlear implants or hearing aids) with reduced audibility.

Currently there are no data available examining the role of multiple talkers in complex, ecologically valid spoken language tasks, such as auditory story comprehension. Some researchers, however, have extended their experimental paradigms beyond low-level word identification and examined the role of multiple talkers in sentence recognition. These data give some insight into how adult listeners may perceive, process, encode, and recall talker-specific information in a more real-life spoken language task. For example in Karl and Pisoni's seminal study[5], they compared listeners' sentence transcription and word recall when sentences were presented by single and multiple talkers. Their results echoed previous talker-specific word identification work showing that adults listening to single talkers transcribed sentences more accurately than those listening to sentences spoken by multiple talkers, however the negative effect of multiple talkers was *not* evident when Karl and Pisoni examined cued recall performance for correctly transcribed sentences. Their data suggested that the processing of talker-specific information accompanying spoken sentences may be quite different from isolated words, thus highlighting the need to determine talker-specific information's role in spoken language processing. The present experiments were conducted to expand on the previous study's manipulation of talker-specific information and determine whether or not multiple narrators influence adults' auditory story comprehension in the presence of both quiet and background noise.

In Experiment 1, we examined the effect of multiple narrators (i.e. talkers) on adults' story comprehension in a quiet listening environment; a high-level complex listening task. In Experiment 2 we replicated the story comprehension task of Experiment 1 in the presence of multitalker babble. We chose story comprehension as our task because if we are to ultimately determine whether talker-specific information is crucial for the analysis of spoken language, we need to understand the relationship between this talker-specific and linguistic information in complex, natural spoken language tasks that parallel the real-life requirements. We included noise in Experiment 2 to enhance the task's face validity. For both Experiments, we predicted the listeners would be negatively affected by the presence of multiple narrators.

EXPERIMENT 1: THE EFFECT OF TALKER VARIABILITY ON ADULTS' AUDITORY STORY COMPREHENSION IN QUIET

In this study adults listened to stories narrated by either single or multiple talkers in a quiet listening environment. Following the stories, each participant answered multiple-choice questions aimed to thoroughly assess auditory comprehension. Experiment 1 aimed to answer the following questions: 1) Does listening to a story relayed from multiple narrators affect a listener's ability to answer comprehension questions about said story? 2) Does the talker information contributed from multiple narrators differentially affect various components of the story (e.g., its saliency and type of knowledge being conveyed)?

Methods

Experimental Design

This experiment utilized a mixed 2 X 2 X 2 factorial design. The between-subjects variable was *narrator type* (single narrator, multiple narrators) and the within-subject variables were *story saliency* (main idea, details) and *story knowledge* (stated, implied).

Participants

A total of 32 adults participated in this study. The participants were all native, American-English undergraduate students who attended a public southern university in the United States. Each participant was screened for normal-hearing thresholds and reported no history of language impairment at the time of testing. Each participant received extra credit in his or her communication sciences and disorders class for participating.

Stimuli

Phonetically balanced stories. Stimuli consisted of 10 phonetically balanced short stories and their accompanying multiple-choice questions from the Resource Allocation Paradigms of Pittsburgh (RAPP)[10]. Each story's 10 corresponding questions (N = 100) assessed all levels of comprehension knowledge. Two questions assessed stated knowledge and 2 questions assessed implied knowledge of the stories' main ideas; 3 questions assessed stated knowledge and 3 questions assessed implied knowledge of the stories' details.

Narrator recordings. 12 different talkers (6 females) recorded the 10 aforementioned stories. Similar to Goh's study [11], our talkers were selected to serve as narrators based on objective measures assessing 20 adult participants' perceived similarity of the voices using multidimensional scaling (MDS) analyses [12]. Of these 12 talkers, 1 male and 1 female served as the narrator for the single talker condition. The remaining 5 males and 5 females served as the narrators for the multiple talker condition.

All 10 stories were recorded and edited using *Adobe Audition 2.0* (2004) sound editing software. Ultimately, for the stories heard by participants in the single narrator condition, the same male or female talker spoke all of the story utterances. Narrator gender was counterbalanced across the participants in the single narrator condition. For the stories heard by participants in the multiple narrator condition, each of the 10 talkers spoke 1 of 10 utterances per story. Narrator order for the multiple narrator condition was chosen semi-randomly so that no one narrator presented the same utterance in more than one block.

Apparatus

We used *E-Prime 2* (2012) experiment building software on a *Dell Optiplex 745* personal computer, monitor, and loudspeaker setup to execute the auditory story comprehension task. The experimental setup was located in a double-walled sound booth.

Procedure

Participants were pseudo-randomly assigned to either the *single* or *multiple narrator* condition. The experiment was administered in a double-walled sound booth at a personal computer equipped with loudspeakers. All participants were tested individually. All of the speech stimuli were presented at a comfortable listening level via the computer's two loudspeakers located at approximately ±45º azimuth.

Once the participant was seated in front of the computer she read the instructions for the task on the computer's monitor—listen to a series of short stories and answer the accompanying multiple-choice questions. The experiment then began with a practice story and 10 accompanying questions. During practice, the text *LISTEN* was centered on the monitor while a short story narrated by a single talker played over the computer's loudspeakers. After the participant heard the story, she was presented with the 10 corresponding multiple-choice questions. Each question appeared individually on the screen. When the participant finished reading the question, she was instructed to press the computer's space bar to reveal the 5 multiple-choice answers on the screen. The participant responded to the question by pressing the corresponding letter of the selected answer on the keyboard. *E-Prime* recorded the participant's accuracy and RT for each question. After the participant answered the question, the next question appeared on the screen following the same format. The test phase immediately followed and consisted of two testing blocks; each block consisted for 5 stories followed by their corresponding questions. The question sets' order followed the order of the stories' presentation. After an optional break the participant listened to the 5 remaining stories and answered their corresponding questions. Story order was counterbalanced across all participants.

After the final comprehension questions, *Did you recognize any of the narrators? Hit A for 'yes' or B for 'no'* appeared on the monitor. If the participant indicated that she recognized a narrator, the experimenter gave her a form assessing her familiarity with the possible narrator(s). None of the 32 participants were familiar with the narrators in the experiment. The experiment lasted approximately 30-45 min in duration.

Results

Accuracy and latency served as the dependent variables in the experiment (see Table 1). Accuracy was calculated as the total percent of comprehension questions answered correctly. We treated latency data as follows: 1) latencies for trials receiving a score of 0 on the accuracy measure were excluded and 2) latencies exceeding 3 SD of the mean of each condition were considered outliers and excluded. Mean accuracy and latency data were each submitted to a split-plot, 3-factor ANOVA (See Figure 1). The between-subjects variable was *narrator type* (single, multiple); the within-subjects variables were *story saliency* (main idea, details) and *story knowledge* (stated, implied).

Accuracy

The analysis of the accuracy data showed significant main effects for *story saliency* [F (1, 30) = 189.90, p = .000] and *story knowledge* [F (1, 30) = 25.73, p = .000] with large effect sizes (*partial* η^2 = .86 and *partial* η^2 = .46, respectively). There was no significant main effect of *narrator type* [F (1, 30) = .146, p = .71], nor any significant 2- or 3-way interactions.

Latency

The analysis of the latency data showed a significant main effect for *story knowledge* [F (1, 30) = 59.61, p = .000] with a large effect size, *partial* η^2 = .67. The analysis also showed a showed a significant 2-way interaction for *story saliency* * *story knowledge* [F (1, 30) = 21.89; p = .000] with a large effect size, *partial* η^2 = .42. There was no significant main effect of *narrator type* [F (1, 30) = 1.28, p = .27], nor any additional significant 2- or 3-way interactions.

TABLE 1. Mean accuracy and latency data for Experiment 1, story comprehension in quiet.

measure	narrator	saliency	knowledge	M (SD)
accuracy (% correct)	single	main ideas	stated	87.2 (12.8)
			implied	79.7 (16.5)
		details	stated	61.9 (17.0)
			implied	58.7 (14.6)
	multiple	main ideas	stated	89.4 (7.5)
			implied	79.7 (10.9)
		details	stated	57.5 (16.1)
			implied	54.9 (10.5)
latency (RT ms)	single	main ideas	stated	4288.06 (577.37)
			implied	5629.43 (903.40)
		details	stated	4665.49 (1260.05)
			implied	4897.52 (983.50)
	multiple	main ideas	stated	4434.70 (848.92)
			implied	5991.44 (1219.43)
		details	stated	4717.39 (1035.58)
			implied	5779.94 (1427.54)

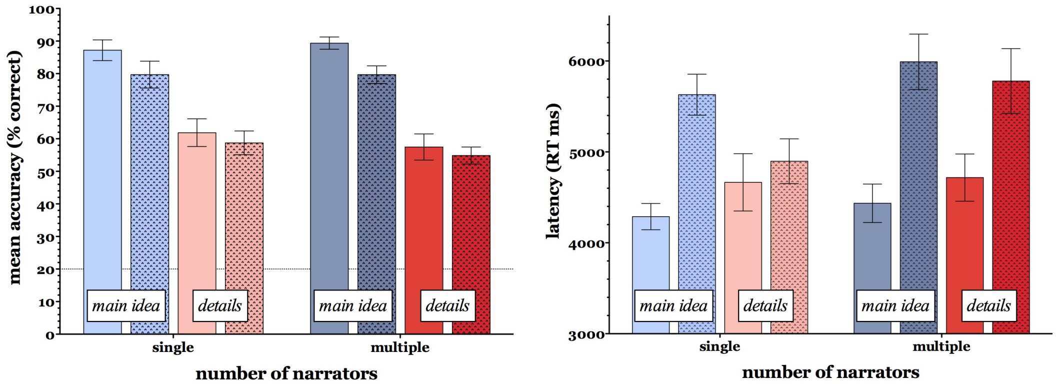

FIGURE 1. Group mean accuracy (left panel) and latency (right panel) for participants in Experiment 1. The blue-toned bars represent questions assessing the stories' *main ideas*; the red-toned bars represent questions assessing the stories' *details*. The unpatterned bars represent questions assessing *stated* knowledge; the patterned bars represent questions assessing *implied* knowledge. Error bars represent S.E.M.

EXPERIMENT 2: THE EFFECT OF TALKER VARIABILITY ON ADULTS' AUDITORY STORY COMPREHENSION IN MULTITALKER BABBLE

Experiment 2 expected to answer similar questions to those of Experiment 1, but with the added variable of background noise (0 dB SNR). Thus, Experiment 2 aimed to answer the following questions: 1) Does listening to a story relayed from multiple narrators affect a listener's ability to answer comprehension questions about said story? 2) Does the talker information contributed from multiple narrators differentially affect various components of the story (e.g., its saliency and type of knowledge being conveyed)? 3) Do the effects of the talker-information contributed by multiple narrators differ when the story comprehension task is completed in the presence of background noise, as opposed to quiet? We predicted, that while adding face validity, the added background noise would heighten the listening complexity of the story comprehension task and subsequently yield the predicted negative effects resulting from the presence of multiple narrators.

Methods

Experiment 2's methods were the same as those of Experiment 1, with the following exceptions: 1) **32 adults** (who did not participate in Experiment 1) participated in Experiment 2 and 2) audio stimuli included **8-talker babble** which played simultaneously in the background during the story narrations at 0 dB SNR. All story comprehension **questions were completed in quiet**. Thus when one block of story narrations was complete, the background babble would cease and the participant would complete the story comprehension questions.

Results

Again, accuracy and latency served as the dependent variables in the experiment (see Table 2). Mean accuracy and latency data were each submitted to a split-plot, 3-factor ANOVA (See Figure 2). The between-subjects variable was *narrator type* (single, multiple); the within-subjects variables were *story saliency* (main idea, details) and *story knowledge* (stated, implied).

Accuracy

Accuracy data analyses showed significant main effects with large effect sizes of *story saliency* [F (1, 30) = 282.789; p = .000; *partial* η^2 = .90] and *story knowledge* [F (1, 30) = 12.52; p = .001; *partial* η^2 = .29]. The analysis also showed significant 2-way interactions with large effect sizes: *story saliency* * *story knowledge* [F (1, 30) = 19.51; p = .000 *partial* η^2 = .39]; *story saliency* * *narrator* [F (1, 30) = 6.41; p = .02 *partial* η^2 = .18]; and *story knowledge* * *narrator* [F (1, 30) = 6.41; p = .02 *partial* η^2 = .18]. There was no significant main effect of *narrator type* [F (1, 30) = 1.23; p = .28], nor was there a significant 3-way interaction [F (1, 30) = 2.05; p = .16].

The ANOVA was followed by pairwise comparisons utilizing independent *t*-tests in an attempt to isolate the effects of narrator. A *t*-test comparing the mean accuracy for the questions assessing main idea saliency + implied knowledge between the two groups of participants showed a significant difference in mean accuracy. In other words, the story comprehension of the participants in the single narrator group (M = 82.19%, SD = 1.30%) was significantly more accurate than that of the participants in the multiple narrator group (M =69.38%, SD = 1.38%); t (30) = 2.70, p = .01, with a large effect size r = .98.

Latency

The analysis of the latency data showed a significant 2-way interaction of *story saliency* * *story knowledge* [F (1, 30) = 7.36; p = .01] with a large effect size *partial* η^2 = .20. Additionally, the analysis revealed significant main effects for *story saliency* [F (1, 30) = 182.46, p = .000] and *story knowledge* [F (1, 30) = 21.30, p = .000] with large effect sizes (*partial* η^2 = .86 and *partial* η^2 = .42, respectively). There was no significant main effect of *narrator type* [F (1, 30) = 2.15, p = .15], nor any additional significant 2- or 3-way interactions.

TABLE 2. Mean accuracy and latency data for Experiment 2, story comprehension in quiet.

measure	narrator	saliency	knowledge	M (SD)
accuracy (% correct)	single	main ideas	stated	87.2 (11.7)
			implied	82.2 (13.0)
		details	stated	56.5 (14.7)
			implied	59.2 (13.2)
	multiple	main ideas	stated	84.1 (13.6)
			implied	69.4 (13.8)
		details	stated	56.7 (13.3)
			implied	57.1 (10.1)
latency (RT ms)	single	main ideas	stated	4390.85 (666.68)
			implied	6050.99 (1089.59)
		details	stated	4192.66 (883.81)
			implied	5270.80 (870.69)
	multiple	main ideas	stated	4020.66 (882.30)
			implied	5715.33 (1011.69)
		details	stated	3772.01 (521.57)
			implied	4971.54 (766.12)

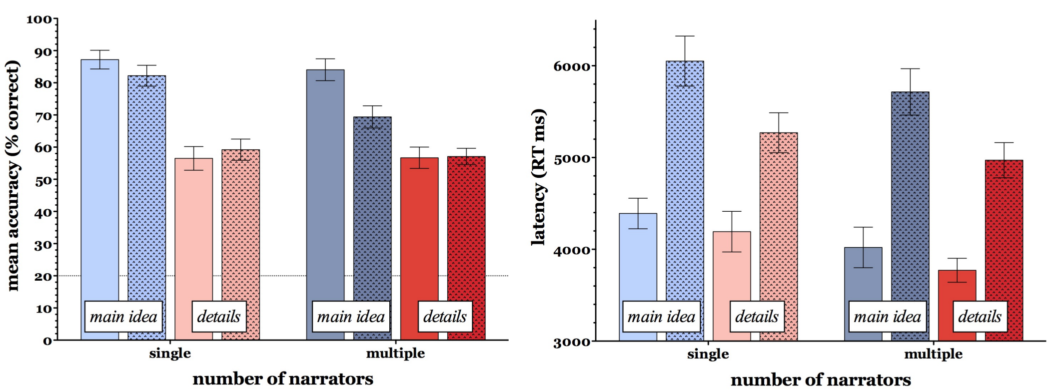

FIGURE 1. Group mean accuracy (left panel) and latency (right panel) for participants in Experiment 2. The blue-toned bars represent questions assessing the stories' *main ideas*; the red-toned bars represent questions assessing the stories' *details*. The unpatterned bars represent questions assessing *stated* knowledge; the patterned bars represent questions assessing *implied* knowledge. Error bars represent S.E.M.

SUMMARY & CONCLUSIONS

The two experiments described within are the first to examine how variable talker information (contributed by multiple talkers) influences adults' auditory story comprehension. Overall the results were contrary to both our predictions and previous studies exploring the effect of multiple talkers on lower-level perceptual tasks such as word identification[4, 11, 15]. The results of Experiment 1 and 2 indicated *no* effect of *narrator*—as measured by accuracy and latency. Participants hearing the stories narrated by a single talker in quiet and noise performed as accurately and quickly on the 100 story comprehension questions as did the participants hearing stories narrated by 10 different talkers. When attempting to answer specific questions about the affect of multiple narrators in a noisy listening environment, the adults' data from Experiment 2 suggested that the added background noise *did* heighten the listening complexity of the task. The accuracy and latency analyses indicated significant 2-way interactions and follow-up pairwise comparisons suggested that multiple narrators indeed interacted with *story saliency* and *knowledge* in such a way that they negatively affected listeners' comprehension when the task was completed in the presence of background noise. Together these data indicated that multiple talkers do not affect story comprehension accuracy and latency in the relatively straightforward manner that they seem to affect the low-level perceptual task of word identification. Finally, both Experiments revealed significant main effects of *story saliency* and *knowledge*, suggesting that listeners answered questions about the stories' main ideas and stated knowledge more accurately and quickly than the questions about the stories' details and implied knowledge respectively. These results were predictable and robust given what is known about adults' learning, memory, cognition, and story comprehension in general[16].

Our findings overall indicate that information contributed by multiple talkers influences listeners' spoken language processing in different ways, in different listening tasks. Although portions of our results corroborate past work highlighting the listeners' vulnerability to talker-effects, the specific role of such variable talker information in complex spoken language tasks utilizing real-time, fluent speech remains elusive. Nonetheless, our results do bring to light the theoretical perspective shared with the PRIMIR framework[8] emphasizing the importance of looking across experimental tasks to better understand talker-specific information's pattern and influence on spoken language processing. Our results also underscore theoretical perspectives that incorporate hierarchical predictive processing[17] and emphasize the differences in cognitive processing at work during low-level perceptual tasks (e.g., word identification) and high-level learning tasks (e.g., auditory story comprehension). Pursuing additional research that explores talker-specific spoken language processing from such perspectives will help the field understand the role of talker-specific information in the speech signal, especially in complex real-life spoken language tasks utilizing fluent speech.

ACKNOWLEDGMENTS

We want to thank all of the individuals who recorded stimuli for the present studies. A big thank you to Lauren Biamonte for the thoughtful organization and research she contributed to the stimuli development process. We are grateful to Mary Elizabeth Dilday, Lindsay Meyer Turner, Annie Schubert, Lauren Biamonte, and the members of LSU's Spoken Language Processing lab for helping collect data. We are very grateful to Emily Elliott for her collaboration on data analysis and interpretation. Finally, we thank Janna Oetting for feedback on this paper.

REFERENCES

1. J. W. Mullennix, D. B. Pisoni and C. S. Martin, "Some effects of talker variability on recall of spoken word recognition," Journal of the Acoustical Society of America **85**, 365-378 (1989).
2. B. O. Ryalls and D. B. Pisoni, "The effect of talker variability on word recognition in preschool children," Developmental Psychology **33** (3), 441-452 (1997).
3. M. S. Sommers and J. Barcroft, "Indexical information, encoding difficulty, and second language vocabulary learning," Applied Psycholinguistics **32** (2), 417-434 (2011).
4. S. D. Goldinger, D. B. Pisoni and J. S. Logan, "On the nature of talker variability effects on recall of spoken word lists," Journal of Experimental Psychology: Learning, Memory, and Cognition **17** (1), 152-162 (1991).
5. J. R. Karl and D. B. Pisoni, "The role of talker-specific information in memory for spoken sentences," The Journal of the Acoustical Society of America **95**, 2873 (1994).
6. C. S. Martin, J. W. Mullennix, D. B. Pisoni and W. V. Summers, "Effects of talker variability on recall of spoken word lists," Journal of Experimental Psychology: Learning, Memory, and Cognition **15** (4), 676-684 (1989).
7. R. N. Aslin and L. B. Smith, "Perceptual development," Annual Review of Psychology **39**, 435-473 (1988).

8. J. F. Werker and S. Curtin, "Primir: A developmental framework of infant speech processing," Language, Learning, & Development **1** (2), 197-234 (2005).
9. M. S. Sommers and J. Barcroft, "Stimulus variability and the phonetic relevance hypothesis: Effects of variability in speaking style, fundamental frequency, and speaking rate on spoken word identification," Journal of the Acoustic Scociety of America **119**, 2406-2416 (2006).
10. P. J. Doyle, M. R. McNeil, G. H. Park, A. J. Goda, E. Rubenstein, K. Spencer, B. Carroll, A. Lustig and L. Szwarc, "Linguistic validation of four parallel forms of a story retelling procedure," Aphasiology **14**, 537-549 (2000).
11. W. D. Goh, "Talker variability and recognition memory: Instance-specific and voice-specific effects," Journal of Experimental Psychology: Learning, Memory, & Cognition **31** (1), 40-53 (2005).
12. J. B. Kruskal and M. Wish, *Multidimensional scaling*. (SAGE Publications, Inc., Beverly Hills, 1978).
13. (Adobe Systems Incorporated, San Jose, CA, 2004).
14. (Psychology Software Tools, Inc., 2012).
15. C. A. Yonan and M. S. Sommers, "The effects of talker familiarity on spoken word identification in younger and older listeners," Psychology & Aging **15** (1), 88-99 (2000).
16. T. R. D. Fossett, M. R. McNeil, P. J. Doyle, H. Rubinsky, S. Nixon, W. Hula and J. Brady, "Assessing the validity of multiple-choice questions for rapp story comprehension," Aphasiology **18**, 439-519 (2004).
17. A. Clark, "Whatever next? Predictive brains, situated agents, and the future of cognitive science.," Behavioral and Brain Sciences (in press).