The effects of disfluent repetitions and speech rate on recall accuracy in a discourse listening task

Jillian Donahue, Christine Schoepfer and Robin Lickley
Clinical Speech Audiology and Language Research Centre, Queen Margaret University, Edinburgh, UK

Abstract

While many studies have examined the effects of disfluency on word recognition and local syntactic or semantic issues, fewer have addressed the impact on comprehension at a discourse level. In this work, we ask what effects features typical in the pathological condition of cluttering (essentially, rapid, disfluent and unintelligible speech) have on our ability to retain the information conveyed in speech. Specifically, we manipulate repetition disfluencies and speech rate in passages of running speech. Forty participants listened to four recordings of passages presented in four conditions: Control, Rapid, Disfluent, Rapid + Disfluent. They were asked to recall details of the passages and rate their speed, fluency and comprehensibility. Both repetition disfluencies and increased speech rate significantly reduced recall of information from discourse. Though no relationship was found between the working memory span of individuals and information recall, we argue that the cognitive load of these features of cluttered speech significantly affects intelligibility and thus recall of speech.

Introduction

There is increasing interest in the effects of disfluencies of various kinds on speech comprehension. Many studies focus on the effects of filled pauses (uh, um). While these hesitation markers may often prepare the listener for unexpected words (Arnold et al., 2004; Corley, MacGregor & Donaldson, 2007; Maxfield, Lyon & Silliman, 2009), others types, including repetitions, may not have the same effect (MacGregor, Corley & Donaldson, 2009).

Most studies on comprehension of disfluent speech focus on effects on word recognition (as above) or on local syntactic interpretations (e.g., Ferreira & Bailey, 2004), but few have examined the broader impact on discourse comprehension. An exception to this is Fraundorf and Watson (2011), who found that filled pauses inserted into spoken narratives enhanced recall of plot points in the narratives, while coughs did not. The effect of disfluent repetition on recall of discourse has not been tested.

Models of discourse comprehension emphasise the involvement of working memory in recall (Kintsch & Van Dijk, 1978), but it is unknown how this interacts with the effects of repetition disfluencies and speech rate.

Some studies have focused on the local effects of typical repetition disfluencies within single sentences. In some cases, listeners fail to recall brief repetitions in spontaneous speech (Bard & Lickley, 1998). In a word monitoring task, repetitions have been found to associate with faster monitoring latencies when they disrupt a phonological phrase, but not otherwise (Fox Tree, 1995).

A study measuring elicited neural responses to words in repetition disfluencies suggested that typical repetitions may have an impact on processing of subsequent words (MacGregor, Corley & Donaldson, 2009).

In this study, rather than examining effects within isolated sentences, we focus on the effect on recall of discourse of word-onset repetitions typical of stuttering and cluttering. One reason for this is that we were interested in the intelligibility of speech that exhibits characteristics of the disorder of fluency known as Cluttering. Cluttering is an under-researched disorder, characterised by disfluencies and rapid speech rate (St Louis et al., 2007).

Word onset (fragment) repetitions are frequent in typical speech, but also in cluttered (and stuttered) speech. They occur most often on stressed syllables of content words. Since content words are crucial in conveying meaning in speech, breakdown in fluency on these words may be expected to have an impact on a listener's comprehension.

The other major aspect of cluttered speech is an elevated speech rate. Faster speech presents the listener with a somewhat degraded signal and less time to process it, so it is only to be expected that there would be effects on comprehension. Such effects have been found in studies with both younger and older adults, where speech rate was systematically varied (Wingfield et al. 1999; Wingfield, Peelle & Grossman, 2003).

Finally, comprehension depends to some extent on the listener. In older people, a clear relationship has been found between an individual's working memory capacity and their recall of information in degraded speech (Ward et al., 2016).

We ask whether any effects of disfluency or speech rate on recall of speech will vary with working memory capacity in younger adults.

In our experiment, we follow the method used by Fraundorf and Watson (2011) whereby participants listen to retold versions of readings abridged from Lewis Carroll's *Alice's Adventures in Wonderland* (1865). We used 4 passages, one for each experimental condition.

There 2 main hypotheses concerning effects on recall: (1) that recall of the passages will be adversely affected by the presence of repetition disfluencies; (2) that recall will be adversely affected by an elevated rate of speech.

We also hypothesised that participants with poorer scores for working memory would recall less of the passages than other participants, with greater effects in degraded passages.

Method

Participants

Forty individuals participated in the study, all having learnt English before the age of six, and all high-school qualified and aged between 18 and 33 years (mean age: 26.4 years and median age: 27 years; male to female ratio 23:17). Participants were excluded if they reported having a history of hearing or language comprehension difficulties, dyslexia or fluency disorder. They were also excluded if they had significant experience of interacting with individuals with stuttered or cluttered speech.

Materials

Four passages (mean word count 302) were used. Three of the passages were those used by Fraundorf and Watson (2011), and a fourth one was selected (also from Carroll, 1865) to fit in with the design of our study. The stories were paraphrased from the original story, scripted and then retold by a 26 year old female with a North American accent, using as natural a delivery as possible, and a typical rate of speech. Each story contained 14 key plot points, and the speaker retold the story one plot point at a time. If any disfluencies occurred in the retelling, the section was re-recorded until it was fluent. The recordings of the passages were created using a TASCAM DR-100 digital audio recorder in the Queen Margaret University speech laboratory.

Each passage was then edited using Audacity® 2.1.0 so that there was a total of four versions. A control version contained no further edits.

A disfluent version of each passage was created by cutting and splicing word onsets on 10% of the content words in each passage.

A fast version of the control and the disfluent passages was created Audacity's "Change Tempo" tool, keeping fundamental frequency the same as the original. A percentage increase in speech rate (51%) was determined by calculating the difference in number of syllables per minute between three conversational samples of cluttered speech and two 'typical' speech samples, as well as figures for typical speech rate found in Tauroza and Allison (1990) and Wood (2001).

Working memory (WM) span was measured using the Forward Digit Span, Reverse Digit Span and Non-Word Repetition tasks (FDS, RDS, NWR), adapted from the *Comprehensive Test of Phonological Processing* (Wagner, Torgesen & Rashotte, 1999). Three tests were chosen since a clear definition of WM is still evasive and it was felt that a score based on a range of tests would provide a more representative measure of WM capacity (Conway et al., 2005).

Procedure

Participants were seated in a quiet room and equipped with headphones (Sennheiser HD201), with one of the experimenters (first two authors) present.

All tasks were preceded by short practice tests.

First, the Working Memory assessments, FDS, RDS, NWR were administered. The digit span tasks began with three strings of two digits and ended with three strings of nine digits. A maximum score of 24 was possible. The non-word repetition task consisted of 27 test items. In each case, the test was terminated after a participant had produced three errors or reached ceiling. The overall WM score had a range of 0–25.

Next, participants heard the four passages in conditions and orders determined by Latin Square, such that each participant was tested in each condition and each passage was presented in all four conditions across the experiment. After hearing each passage, the participant was asked to retell the story that they had just heard as accurately as possible. There was a maximum score of 28 for each passage, with points given on the basis of whether full (2 points) or partial (1 point) information had been provided for each plot point. After each passage, participants also gave subjective ratings of speed, fluency and intelligibility on a 1–5 Likert scale.

Ethical permission was granted by the QMU Ethics panel, and all participants gave informed consent.

Results

Since one passage was found to have lower recall scores across all conditions, recall scores were normalized and z scores used.

We expected that repetition disfluencies would have a negative impact on recall and they did. We also hypothesized that an increased rate of speech would result in poorer recall and it did. There was also a combined effect of disfluency and increased rate, such that passages with both disfluency and increased speech rate attracted the lowest recall scores (Table 1).

Table 1. Means and Standard Deviations of Raw and Normalised Passage Recall Scores under each Condition. Maximum possible raw score is 28.

Condition	N	Raw Mean	Raw SD	Normalised Mean	Normalised SD
Control	40	15.65	5.53	0.33	0.84
Rapid	39	12.82	5.65	– 0.14	0.96
Disfluent	40	13.70	6.99	0.04	1.03
Combined	39	11.79	7.11	– 0.25	1.06

Scores for control passages were higher than all other conditions. There was a significant difference in the scores between Control and Rapid; *t(38) = 3.96, p < .001*, between Control and Disfluent; *t(39) = 2.53, p = .016* and between Control and Combined; *t(38) = 5.82, p < .001.* Significant differences were also found between Disfluent and Combined; *t(38) = 2.58, p = .014.* The difference between the Rapid and Disfluent conditions approached significance; *t(38) = _ 1.79, p = .082*, but Rapid and Combined conditions were not significantly different; *t(37) = 0.668, p = .508.* These results are illustrated in Figure 1.

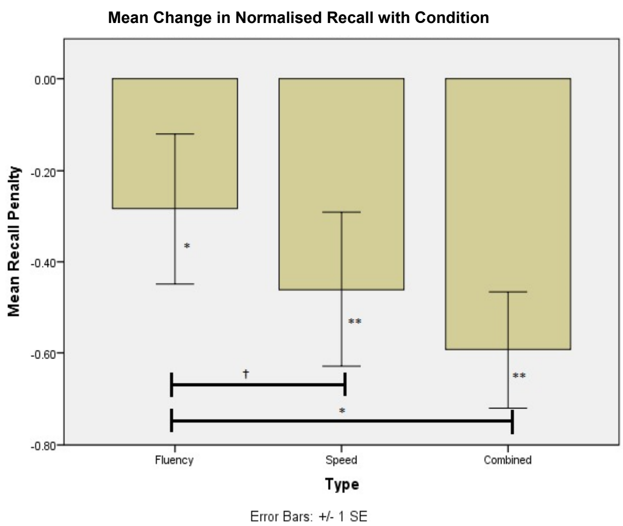


Figure 1. The Effect of Disfluencies, Rapid Speech Rate and Combined Conditions on Passage Recall Scores. Boxes show the difference between control recall scores and recall scores in each test condition.
** p < .05, ** p < .01*

Subjective judgements of disfluency, speed and intelligibility reflected the manipulations in the passages. Ratings for disfluency and speed are summarised in Table 2.

For intelligibility, the subjective ratings correlated strongly with recall scores (r_s between –41 and –.59) in the manipulated conditions, but no correlation was found in the control condition, since intelligibility scores were at ceiling.

Working memory scores are displayed in Table 3. Contrary to our expectation, there was no significant correlation between the composite WM score and recall scores in any condition. Similarly, no association was found between any of the components of the WM tests and recall performance in any conditions, except for the NWR assessment where a significant positive correlation was found with normalised Recall Scores during the Rapid condition ($r_s(39) = +.329, p = .041$).

Table 2. Means of Participant Subjective Ratings of Speed and Fluency for Each Condition (Scale 1–5, where 5 means fast/disfluent).

	Speed Mean	SD	Fluency Mean	SD
Control Condition	2.60	0.81	1.50	0.78
Rapid Condition	4.51	0.56	2.56	1.05
Disfluent Condition	2.58	0.81	3.75	0.70
Combined Condition	4.23	0.71	3.97	0.78

Table 3. Means and Standard Deviations (SD) of the individual WM assessments (FDS, RDS, NWR) and Comprehensive WM Score

	N	Mean	SD
Forward Digit Span	40	16.55	2.94
Non-Word Repetition	40	17.30	2.55
Reverse Digit Span	40	10.88	3.06
Comprehensive WM Score (Mean Score)	40	14.92	2.32

Discussion

In our experiment, both repetition disfluencies and increased rate of speech adversely affected recall of the content of what was heard. We did not find a clear relationship between listeners' working memory scores and their ability to recall details of what they had heard.

Our stuttering-like disfluency types and the overall methodology mean that it is hard to draw comparisons with previous work that examined the effects of typical repetitions in isolated sentences. The repetitions that we used were similar to stuttered repetitions, so we assume that they would be more disruptive to phonological phrases and have some effect on the processing of subsequent words (as per Fox Tree (1995) and MacGregor, Corley & Donaldson (2009)). It seems that this disruption has an effect on recall for the content of the speech.

Similarly, for speech rate, previous work has examined response latency to single sentences (Wingfield et al. 1999; Wingfield, Peelle & Grossman, 2003). Our findings suggest that the difficulty observed for single sentences translates into negative effects on recall of the information in longer passages of speech.

Taken together, the findings confirm that the faster speech and stutter-like disfluencies that characterise pathologically cluttered speech are likely to have a significant effect on how much listeners can understand of what is being said.

The speech manipulations also affected subjective judgements of disfluency, speech rate and intelligibility of speech, lending some support to the use of such judgements in clinical assessment of cluttering (e.g., the *Cluttering Severity Instrument*, Bakker & Myers (2011)).

That the recall results did not vary significantly with individuals' working memory scores was unexpected. However, the participants were from a relatively homogeneous group, of young adults with a minimum of a full high school level of education, and the range of their working memory scores was fairly compact.

Future work in recall may take various directions. First, it is interesting to ask to what extent typical repetition disfluencies, involving single repetitions of short function words or word onsets, hinder recall of information, if at all and whether this varies with prosodic disturbance. It is also of interest to know at what point speech becomes too fast for full intelligibility. In both cases, there may be a habituation effect, which may have clinical implications for assessment of cluttered speech. Finally, despite our null findings, past work suggests that a listener's working memory capacity should have an impact on recall of speech affected by disfluency and increased rate, so work with populations with more varied working memory would be of interest.

Acknowledgements

This work was completed as part of a Masters dissertation project in Speech and Language Therapy at Queen Margaret University, Edinburgh, carried out by the first two authors, supervised by the third author.

References

Arnold, J. E., M. K. Tanenhaus, R. J. Altmann & M. Fagnano. 2004. The Old and Thee, uh, New Disfluency and Reference Resolution. *Psychological Science,* 15(9):578–582.

Bakker, K. & F. Myers. 2011. Cluttering Severity Instrument. *Computer software and manual].* http://associations.missouristate.edu/ICA/

Bard, E. G. & R. Lickley. 1998. Disfluency Deafness: Graceful Failure in the Recognition of Running Speech. *Proceedings of the Twentieth Annual Conference of the Cognitive Science Society*, University of Wisconsin, Madison, Minnesota, USA, Lawrence Erlbaum, 108–113.

Carroll, L. 1865. *Alice's Adventures in Wonderland*. New York: MacMillan.

Conway, A. R. A., M. J. Kane, M. F. Bunting, D. Z. Hambrick, O. Wilhelm & R. W. Engle. 2005. Working memory span tasks: A methodological review and user's guide. *Psychonomic Bulletin and Review* 12(5):769–786.

Corley, M., L .J. MacGregor & D. I. Donaldson. 2007. It's the way that you, er, say it: Hesitations in speech affect language comprehension. *Cognition* 105(3):658–668.

Ferreira, F. & K. G. Bailey. 2004. Disfluencies and human language comprehension. *Trends in cognitive sciences 8*(5):231–237.

Fox Tree, Jean, E. 1995. The effects of false starts and repetitions on the processing of subsequent words in spontaneous speech. *Journal of Memory and Language* 34:709–738.

Fraundorf, S. H. & D. G. Watson. 2011. The disfluent discourse: Effects of filled pauses on recall. *Journal of Memory and Language* 65(2):161–175.

Kintsch, W. & T. A. Van Dijk. 1978. Toward a model of text comprehension and production. *Psychological Review* 85(5):363–394.

MacGregor, L. J., M. Corley & D. I. Donaldson. 2009. Not all disfluencies are are equal: The effects of disfluent repetitions on language comprehension. *Brain and Language* 111(1):36–45.

Maxfield, N. D., J. M Lyon & E. R. Silliman. 2009. Disfluencies along the garden path: Brain electrophysiological evidence of disrupted sentence processing. *Brain and Language* 111(2):86–100.

St Louis, K.O., F. Myers, K. Bakker & R. Lawrence 2007. Understanding and Treating Cluttering. In E. G. Conture & R. F. Curlee (eds.), *Stuttering and Related Disorders of Fluency* (3rd ed). New York: Thieme, 297–325.

Tauroza, S. & D. Allison. 1990. Speech rates in British English. *Applied Linguistics* 11(1):90–105.

Wagner, R. K., J. K. Torgesen & C. A. Rashotte. 1999. *Comprehensive test of phonological processing.* Texas: Pro-Ed.

Ward, C. M., C. S. Rogers, K. J. Van Engen & J. E. Peelle. 2016. Effects of age, acoustic challenge, and verbal working memory on recall of narrative speech. *Experimental aging research* 42(1):97–111.

Wingfield, A., J. E. Peelle & M. Grossman. 2003. Speech rate and syntactic complexity as multiplicative factors in speech comprehension by young and older adults. *Aging, Neuropsychology, and Cognition,* 10(4):310–322.

Wingfield, A., P. A. Tun, C. K. Koh & M. J. Rosen. 1999. Regaining lost time: adult aging and the effect of time restoration on recall of time-compressed speech. *Psychology and Aging* 14(3):380

Wood, D. 2001. In search of fluency: What is it and how can we teach it? *Canadian Modern Language Review* 57(4):573–589.

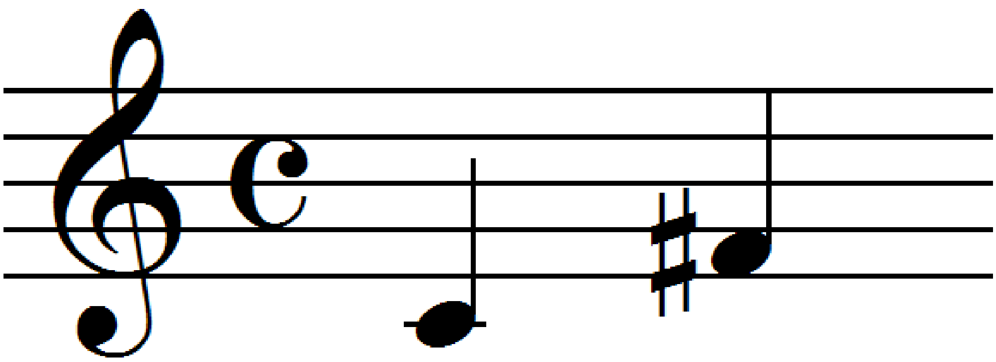
Proceedings of

DiSS 2017

The 8th Workshop on Disfluency in Spontaneous Speech

KTH Royal Institute of Technology
Stockholm, Sweden
18–19 August 2017

TMH-QPSR
Volume 58(1)



Edited by
Robert Eklund & Ralph Rose

Conference website: http://www.diss2017.org
Proceedings also available at: http://roberteklund.info/conferences/diss2017

Cover design by Robert Eklund
Graphics and photographs by Robert Eklund (except ISCA and KTH logotypes)

Proceedings of DiSS 2017, Disfluency in Spontaneous Speech
Workshop held at the Royal Institute of Technology (KTH), Stockholm, Sweden, 18–19 August 2017
TMH-QPSR volume 58(1)
Editors: Robert Eklund & Ralph Rose
Department of Speech, Music and Hearing
Royal Institute of Technology (KTH)
Lindstedtsvägen 24
SE-100 44 Stockholm, Sweden

ISSN 1104-5787
ISRN KTH/CSC/TMH–17/01-SE