The Effects of Read-aloud Assistance on Second Language Oral Fluency in Text Summary Speech

Shungo Suzuki and Judit Kormos [1]

[1]*Department of Linguistics and English Language, Lancaster University, UK*

Abstract

Focusing on text summary speaking tasks, the present study investigated the effects of the activation of phonological representations during text comprehension (operationalized by read-aloud assistance) on the subsequent retelling speech. A total of 24 Japanese learners of English completed text summary speaking tasks under two conditions: (a) reading without read-aloud assistance and (b) reading with read-aloud assistance. Their speech data were analyzed by lexical overlap indices (i.e., the ratio of characteristic single-words and multiword sequences) and by fluency measures capturing three major dimensions of fluency–speed, breakdown, and repair fluency. The results showed that read-aloud assistance directly facilitated lexical overlaps with source texts and indirectly improved speed and repair fluency. Furthermore, read-aloud assistance was found to affect the interrelationship between lexical overlaps and utterance fluency. The findings suggested that read-aloud assistance might help second language learners to store multiword sequences as a single unit (i.e., chunking) during text comprehension.

1. Background

Previous studies suggest that the activation of speech content and linguistic forms prior to speech should contribute to efficient speech processing, leading to fluent speech (Skehan, 2014). From the perspective of L2 speech production, the processes of specifying speech content and linguistic forms (i.e., *conceptualization* and *formulation*) mainly rely on controlled processing (Kormos, 2006). Thus, the activation of underlying speech processing can be assisted at the following different levels: content specification, lemma selection, syntactic structure, and phonological representations.

One of the effective fluency enhancement strategies is the use of multiword sequences (MWS) including formulaic sequences and n-grams (Stengers, Boers, Housen, & Eyckmans, 2011). L2 speech production model assumes MWS are stored as a whole and are retrieved as a single unit. This direct single-step retrieval can save attentional resources for other speech processing such as conceptualization or syntac-tic and phonological encoding (Skehan, 2014)

Considering the advantages of the activation of linguistic forms as well as the use of MWS, L2 speakers tend to be fluent in text summary tasks where they retell what they read. In addition, read-aloud assistance (RAA), in which one reads a text while simultaneously listening to its oral recording, may further enhance their utterance fluency (UF) due to its dual-modal input. For instance, intonation can help L2 learners to segment texts into larger units of grammatical and/or semantic information (cf. Košak-Babuder, Kormos, Ratajczak, & Pižorn, 2019). Therefore, RAA can be hypothesized to lead to the facilitated text comprehension as well as the activation of MWS (i.e., *chunking)*. However, it is still unclear the extent to which such an enhanced reading by dual-modal input can play a supportive role in subsequent speech processing.

Taken together, the present study–as part of a larger project–investigated the effects of RAA on the use of MWS and UF in text summary speech and the relationship between them.

2. Method

2.1 Participants

A total of 24 Japanese-speaking learners of English were recruited at a private university in Japan. Their self-reported scores in English proficiency tests such as TOEFL and IELTS, their proficiency levels ranged from B1 to C1 levels on the CEFR scale.

2.2 Materials

Our text summary speaking task included two elements: source texts and their recordings for RAA. We selected two expository texts from a Dreamreader.net (http://dreamreader.net/). To maximize the comparability of these two texts, we initially pooled multiple texts and analyzed them in terms of text length, lexical complexity, and readability. Regarding lexical complexity, we used the JACET8000 wordlist which are specifically tailored for Japanese learners of English and replaced vocabulary items above Level 5 with synonyms within Level 1-4. As for readability, we used both a Flesh-Kincaid Reading Ease value based on Coh-Metrix (McNamara, Graesser, McCarthy, & Cai, 2014) and an overall complexity score based on the TextEvaluator® (Educational Testing Service, 2013) to select a pair of comparable texts. The textual characteristics of the selected texts are summarized in Table 1. The RAA stimuli were recorded by a L1 Canadian English speaker who had 15-year teaching experience of English at universities in Japan. We also ensured the comparability of the delivery speed of recordings across texts (see Table 1).

Table 1. Characteristics of the source texts.

	Text A	Text B
Topic	US Flag	Red Cross
Flesh-Kincaid value	71.21	64.79
TextEvaluator® score	380	660
Common Core Grade	2–3	5–7
Length in words	324	303
Delivery speed (wpm)	116.4	119.6

2.3 Procedures

Considering the possibility that individual difference factors may affect the effects of RAA (Liu & Todd, 2014), we decided to use a within-subjects design. In other words, our participants completed both conditions (i.e., [+/- RAA]) while the order of conditions and the source texts were counterbalanced.

For each condition, the participants were first instructed to focus on the gist of meaning of texts rather than the details of information such as dates and were then provided with the source text. The students were allowed to read the texts for approximately three minutes (i.e., the same duration as the RAA recording) either under [+RAA] or [-RAA] conditions. After the text comprehension phase, additional three minutes were given as planning time. During this period students could plan or rehearse their speech while looking at the source text. Afterwards, participants were instructed to retell the content of the source text in English without looking at the text.

2.4 Analysis

The present study analyzed participants' speaking performance in terms of lexical overlaps with source texts and UF. All the audio-recorded speech data were transcribed and annotated for dysfluency phenomena such as self-corrections and false starts. The transcripts were

segmented into Analysis of Speech units (AS-units; Foster, Tonkyn, & Wigglesworth, 2000) as well as clauses and were then submitted to subsequent analyses.

Regarding lexical overlaps, a set of n-gram keyword overlap indices were computed by the Tool for the Automatic Analysis of Cohesion (TAACO 2.0; Crossley, Kyle, & Dascalu, 2019): single words, bigrams, trigrams, and quadgrams. These indices tap into the extent to which characteristic words and n-grams from the source text are used in the speech transcript. TAACO 2.0 identifies single- and multi-word keywords using the news and magazine section of the Corpus of Contemporary American English (COCA) as a reference corpus (for a detailed description, see Crossley et al., 2019).

As for UF, we employed a set of fluency measures covering speed, breakdown, and repair fluency. Following prior research, we set the duration of silent pauses as 250 milli-seconds (Bosker, Pinget, Quené, Sanders, & de Jong, 2013). Using *Praat* software (Boersma & Weenink, 2012), we computed *articulation rate* (AR; the mean number of words per minute, divided by total speech duration excluding pauses), *final-* and *mid-clause pause ratio* (FCPR, MCPR; the mean number of pauses between/within clauses per 100 words), and *dysfluency ratio* (DR; the ratio of dysfluencies to the total number of words).

3. Results

3.1 Lexical overlaps across conditions

In order to detect the differences in lexical overlaps between the source text and subsequent text summary speech across conditions, a set of Wilcoxon signed-ranks tests were performed. As summarized in Table 2, no significant difference was found across condi-tions except for single words ($Z = 2.25$, $p = 0.024$, $d = 0.36$), suggesting that RAA during text comprehension enhanced lexical overlaps only at the single-word level.

3.2 Utterance fluency across conditions

Another set of Wilcoxon signed-ranks tests were performed to examine whether participants' UF significantly differed across conditions. As observed in Table 3, the results showed that there were only marginally significant changes in AR ($Z = 1.74$, $p = 0.081$, $d = 0.40$) and DR ($Z = 1.71$, $p = 0.086$, $d = 0.34$). In other words, our participants spoke slightly more fluently under the [+RAA] condition during text comprehension.

Table 2. Descriptive Statistics and Wilcoxon signed-ranks tests for single-word and n-gram keywords

| KW% measure | [- RAA] | | [+ RAA] | | Wilcoxon-signed rank | | |
	M	*SD*	*M*	*SD*	*Z*	*p*	*d*
Single-words	0.115	0.025	0.133	0.041	2.251	0.024	0.355*
Bigrams	0.095	0.057	0.091	0.062	0.200	0.841	0.046
Trigrams	0.034	0.030	0.290	0.033	0.469	0.639	0.097
Quadgrams	0.018	0.021	0.013	0.022	0.719	0.472	0.147

Table 3. Descriptive Statistics and Wilcoxon signed-ranks tests for utterance fluency measure

| Dimension | UF measure | [- RAA] | | [+ RAA] | | Wilcoxon-signed rank | | |
		M	*SD*	*M*	*SD*	*Z*	*p*	*d*
Speed	AR	104.95	20.84	109.49	20.27	1.743	0.081	-0.397†
Breakdown	FCPR	0.108	0.029	0.114	0.029	0.829	0.407	-0.198
	MCPR	0.404	0.172	0.423	0.179	1.057	0.290	-0.160
Repair	DR	0.193	0.090	0.166	0.099	1.714	0.086	0.343†

3.3 Interrelationship between lexical over-laps and fluency across conditions

We also examined how RAA affected the inter-relationship between lexical overlaps and UF by calculating Spearman's rank order correlations (see Table 4). Under the [-RAA] condition, the results revealed that only FCPR was correlated with trigram (rs = -.626, p = 0.001) and quadgram keyword percentages (rs = -.533, p = 0.007), indicating that speakers using keyword trigrams and quadgrams tend to produce fewer pauses at clausal boundaries. On the other hand, under the [+RAA] condition, keyword trigrams were correlated with all the UF measures in a supportive direction while keyword quadgrams were supportively correlated with AR (rs = .599, p = 0.002) and MCPR (rs = -.567, p = 0.004). Table 4. Correlations between lexical overlap indices and utterance fluency measures

UF measure	Single words		Bigrams		Trigrams		Quadgrams	
	[-RAA]	[+RAA]	[-RAA]	[+RAA]	[-RAA]	[+RAA]	[-RAA]	[+RAA]
AR	.119	-.301	.079	.104	.250	.438*	.309	.599**
FCPR	-.206	-.020	-.345†	-.543**	-.626**	-.412*	-.535**	-.219
MCPR	.098	.176	-.065	-.267	-.143	-.587**	-.163	-.567**
DR	.048	-.180	.150	-.253	-.174	-.603**	-.299	-.373†

4. Discussion

4.1 The effects of read-aloud assistance on lexical overlaps

We found that our L2 speakers used more single-word keywords under the [+RAA] condition than the [-RAA] condition. This positive effect of RAA may indicate that the activation of phonological representations can facilitate the use of single presented in the source texts. This might also suggest that RAA facilitates text comprehension and that speakers successfully retell the content of the source text using the characteristic words.

4.2 The effects of read-aloud assistance on utterance fluency

Meanwhile, we found only marginally significant gains in UF under the [+RAA] condition. L2 speakers tended to produce faster speech with fewer dysfluency phenomena with RAA than without RAA. Regarding speed fluency, the activation of phonological representations may lead to smooth articulatory gestures. Meanwhile, the RAA may reduce two types of self-repairs; the enhanced text comprehension could have reduced information or appropriacy repairs whereas the facilitated selection of lexical items from source texts might have reduced error repairs (Kormos, 2006).

4.3 The effects of read-aloud assistance on the interrelationship between lexical overlaps and utterance fluency

The results of correlational analyses revealed that RAA intensified the relationship between n-gram overlaps (mainly, *trigrams* and *quadgrams*) and UF. Although tri- and quad-gram overlaps were correlated only with FCPR under the [-RAA] condition, these two lexical overlap indices were correlated with most of the UF measures under the [+RAA] condition. These correlations under both conditions suggested supportive relationships between multiword lexical overlap and UF measures.

From the perspective of L2 speech production, these results may indicate that the activation of phonological representations, operationalized by RAA, can facilitate *chunking* during text comprehension (Ellis, 2003). The phonological information such as into-nation boundaries can help L2 readers to segment texts into larger units of meaning, and then they can establish the connections among lexical items within the intonation units. As a result, the use of n-gram keywords can positively contribute to UF when L2 learners summarize a text in a dual-mode input condition.

These results may suggest that the activation of phonological representations directly enhances the selection of lexical items and, to a small extent, indirectly facilitates the speed of linguistic encoding processes. In addition, such phonological activation may also play a facilitative role in chunking during text comprehension and in lexical overlaps and fluency of the subsequent text summary speech.

5. Conclusion

The present study investigated the effects of RAA on the subsequent text summary speech in terms of lexical overlaps and UF in the case of 24 Japanese learners of English. To minimize the effects of individual differences such as preferred modality of input processing, the study used a within-subjects design. The results showed that multi-modal input at the text comprehension phase –reading with RAA–had a direct impact on the proportion of single-word keywords and also an indirect impact on UF in the subsequent text summary speech. In addition, the effects of RAA were found in the interrelationship between multiword lexical overlaps and UF, suggesting that the activation of phonological representations by RAA may facilitate L2 readers' chunking during text comprehension and fluency during subsequent speech.

Finally, it should be noted that the number of participants in our study was relatively small. In relation to the small sample size, we could not investigate the interaction effects between texts and conditions, which we instead tried to minimize by carefully preparing comparable texts. Future research is therefore expected to be conducted with a larger sample size and perform more advanced statistical modelling (e.g., mixed-effects modelling) to carefully examine the relationship between texts, conditions, lexical overlaps and fluency.

References

Boersma, P. & D. Weenink. 2012. Praat: doing phonetics by computer [Computer software]. Retrieved from www.praat.org/

Bosker, H. R., A.-F. Pinget, H. Quené, T. Sanders & N. H. de Jong. 2013. What makes speech sound fluent? The contributions of pauses, speed and repairs. *Language Testing, 30*(2), 159–175.

Crossley, S. A., K. Kyle & M. Dascalu. 2019. The Tool for the Automatic Analysis of Cohesion 2.0: Integrating semantic similarity and text overlap. *Behavior Research Methods, 51*(1), 14–27.

Educational Testing Service. 2013. TextEvaluator®. Retrieved from https://textevaluator.ets.org/

Ellis, N. 2003. Constructions, chunking and conne-ctionism: The emergence of second language structure. In C. Doughty & M. H. Long (Eds.), *Handbook of second language acquisition*. Malden, MA: Blackwell, 63–103.

Foster, P., A. Tonkyn & G. Wigglesworth. 2000. Measuring spoken language: A unit for all reasons. *Applied Linguistics, 21*(3), 354–375.

Kormos, J. 2006. *Speech production and second language acquisition*. Mahwah, N.J.: Lawrence Erlbaum Associates.

Košak-Babuder, M., J. Kormos, M. Ratajczak & K. Pižorn. 2019. The effect of read-aloud assistance on the text comprehension of dyslexic and non-dyslexic English language learners. *Language Testing, 36*(1), 51–75.

Liu, Y. T. & A. G. Todd. 2014. Dual-modality input in repeated reading for foreign language learners with different learning styles. *Foreign Language Annals, 47*(4), 684–706.

McNamara, D., A. C. Graesser, P. M. McCarthy & Z. Cai. 2014. *Automated evaluation of text and discourse with Coh-Metrix*. Cambridge University Press.

Skehan, P. 2014. *Processing perspectives on task performance*. Amsterdam: John Benjamins.

Stengers, H., F. Boers, A. Housen & J. Eyckmans. 2011. Formulaic sequences and L2 oral proficiency: Does the type of target language influence the association? *IRAL - International Review of Applied Linguistics in Language Teaching, 49*(4), 321–343.