Flipping the 4-3-2 Activity

Benjamin Thanyawatpokin
Ritsumeikan University
Carl Vollmer
Ritsumeikan Uji Junior and Senior High School

Reference Data:

Thanyawatpokin, B., Vollmer, C. (2018). Flipping the 4-3-2 activity. In P. Clements, A. Krause, & P. Bennett (Eds.), *Language teaching in a global Age: Shaping the classroom, shaping the world*. Tokyo: JALT.

Building on the research of Boers (2014) and Thai and Boers (2015), this still-in-progress study explores the effect of time conditions on learner monologue output. The effects of a shrinking time condition (4-3-2), constant time condition (2-2-2), and expanding time condition (2-3-4) are explored and analyzed based on fluency, complexity, and accuracy. By exploring these conditions, this research places a different lens on the 4-3-2 activity argued for by Maurice (1983), Nation (1989), and others. Through the exploration of a variant of the 4-3-2 activity, we gained a better understanding of the effect of time conditions on in-class monologue activities. Results of the study show that both shrinking and constant time conditions result in improved fluency, and the expanding time condition leads to more complex output at the expense of fluency and accuracy.

Boers（2014）とThai and Boers（2015）の研究に基づいて、この研究では、時間的プレッシャーで学習者にモノローグアウトプット（monologueoutput）をさせる活動の検証をする。時間を縮める「4/3/2」、時間制限を変えない「2/2/2」と、時間を延ばす「2/3/4」活動の効果を三つの項目（流暢さ、複雑さ、精度さ）で分析する。この三つの項目の内、Maurice（1983）、Nation（1989）その他が紹介した「4/3/2」メソッドをさらに深く調べる。これらの調査で「4/3/2」をより客観的な視点で見ることができたこの研究の結果、時間を縮めた活動と時間が変わらない活動は生徒の流暢さを上達させた。その一方、時間を延ばす活動は流暢さと精度さを劣化させたが、複雑さ（接続詞と関係代名詞の使用）において能力が上がったことがわかった。

In an ideal world, the language learning classroom should facilitate output that is more fluent, accurate, and complex. Swain (1995) argued that pushed output from learners leads to marked improvements and thus activities that result in additional output are beneficial in the language classroom. One way to push learner output is by altering the conditions of the task or activity being used in class. One such activity, often used to improve fluency through a sequence of monologues with a shrinking time condition, was first described by Maurice (1983), was expanded upon by Nation (1989), and came to be known as the 4-3-2 activity or technique. In this activity, learners deliver a monologue on the same topic to a partner under increasing time pressure. The first monologue is delivered for 4 minutes, the second monologue with a new partner for 3 minutes, and the third monologue again with a new partner for 2 minutes. By repeating the same task under a decreasing time condition, output has been shown to improve both fluency and accuracy (Boers, 2014; Thai & Boers, 2015).

Considering the 4-3-2 Activity

Although 4-3-2 has become a commonly adapted and used activity in the language classroom, it should be able to withstand more critical investigation into how it functions in the classroom. For example, at the 2016 Vocab@Tokyo Conference, Nation (2016) argued that there were assumptions being made about many language activities including 4-3-2, and that they should be subject to further exploration and research. Through such exploration, teachers would be able to test what they think they know, and thus come to a greater understanding of the language classroom. Given that Nation has been a strong advocate for 4-3-2, bringing attention to this issue should give credence to the need to explore 4-3-2 and alternative variations of the activity.

Research by Boers (2014) has already looked more closely at 4-3-2, and an alternate version 3-3-3, to see if repetition leads to gains in fluency and accuracy instead of time constraints. Complexity in the monologues was also monitored to see if time constraints would affect the quality of the output. Boers found that when comparing the shrinking

time activity (4-3-2) with a constant time activity (3-3-3), dysfluencies decreased in both versions of the activity but that the constant time condition led to fewer dysfluencies. In addition, the speech rate of the shrinking time condition increased, but the constant time condition also saw improvements in speech rate. Nation (1989) argued that 4-3-2 would lead to improvements in accuracy, but Boers found the opposite to be true. Boers found that errors increased with the shrinking time condition and decreased in the constant time condition. Thus, the results of this study counter the expected results of 4-3-2 and provide a challenge to what we had previously thought. The results for complexity were not as conclusive, as the results did not demonstrate statistical significance, but the shrinking time condition led to a slight drop in subordinate clauses, and the constant time condition led to a slight gain. These results were not enough for Boers to reach a strong conclusion on the effect on complexity. Consequently, the study by Boers shows that some assumptions made about 4-3-2 were incorrect when applied in the classroom.

In a similar study, Thai and Boers (2015) reported results that show what can occur during a 4-3-2 activity. Monologues for fluency, complexity, and accuracy were analyzed under a shrinking time condition and a constant time condition. Based on the level of the students in this study, shorter versions of 4-3-2 and 3-3-3 were needed, thus this study utilized a 3-2-1 shrinking time condition and a 2-2-2 constant time condition. The results indicated that the shrinking time condition improved fluency but came at the expense of complexity and accuracy, both of which decreased by the final monologue. In the constant time condition, fluency improved but not to the extent that it did in the shrinking time condition. Although the shrinking time condition did not result in increases in accuracy, the constant time condition resulted in increases that reached significance. Improvements in complexity were not observed in either the constant time condition or the shrinking time condition. Thus, the results of this study showed that a constant time condition can lead to improvements in fluency and accuracy, supporting Boers (2014). Additionally, a shrinking time condition has only been shown to lead to improvements in fluency; the data does not support improvements in complexity and accuracy.

All variations of 4-3-2 display the concept of pushed output (Swain, 1995). Studies have shown that pushed output can be beneficial for improving oral production. Izumi (2002) found that learners who were required to produce output in addition to receiving foreign language input had higher rates of grammatical comprehension than their classmates who did not produce output. Additionally, Sadeghi Beniss and Edalati Bazzaz (2014) found that pushed output led to improvements in accuracy. Thus, it is clear that there are benefits to implementing activities that involve pushed output. A variation of 4-3-2, which allows additional time for output, may also prove beneficial and will be explored further in this study.

Current Study

Building on the momentum of these findings, the aim of this study is to further our understanding of 4-3-2. Two iterations—a shrinking time condition and a constant time condition—have been conducted (Boers, 2014; Thai & Boers, 2015), but an expanding time condition has yet to be explored. Three monologue activities that use 4-3-2 were explored in this study in an iteration that had not previously been investigated.

Methods

Participants

The present study presents a preliminary data set of 10 students but will eventually include data from 17 first-year high school students and 17 first-year university students for a total of 34 subjects. The participants were English majors taking an English communication course and had TOEFL levels from 375-450 and can be considered in the beginner-to-intermediate range.

Procedures

This study was conducted over 3 consecutive weeks in a once-a-week English communication class. Only one session was conducted each week to prevent learner fatigue. Similar to previous studies (see Ogawa, 2016; Thai & Boers, 2015), the level of the learners demanded a shorter version of 4-3-2. The three versions of the activity used in this study were 3-2-1, 2-2-2, and 1-2-3. Each session used a different conversation topic so that learners would not be able to rely on what they had said the previous week. The chosen topics were considered common enough for all learners to talk about but broad enough that filling longer monologue times would still be possible, for example, "What is the best place you have visited?" Learners were rotated between rounds so they would have a new partner for each round of their monologue.

The level of the students necessitated a preplanning of 5 minutes before starting any of the monologues. Thai and Boers (2015) also allowed for preplanning in their study, arguing that it would help learners organize their monologues. Because the 1-2-3 version was expected to require extensive additional content to be created, supporting learner

organization on this new content was considered necessary. In order to keep planning times consistent across all versions of the monologues, the planning time was made a constant variable. Therefore, a 2-minute planning period was given at the end of each round of monologues. In total, learners received 5 minutes of preplanning, followed by a 2-minute planning time after the first and second monologues. During these planning periods, learners were allowed to use a piece of paper to write their ideas for their monologues. Students were only allowed to write on their paper during the planning time. No writing was allowed during the monologues, as this would have impacted performance during the task.

Each pair was recorded; a digital voice recorder was placed on the desk of each pair and the recordings were then transcribed. Students that were present for and participated in all monologue sessions were included in the data. If a student was absent for one day, he or she was not included in this study. Monologues were fully transcribed and sorted into two different versions: an original transcription and a pruned transcription. The original transcriptions include all hesitations, false starts, and pauses longer than 0.3 seconds. The pruned transcriptions exclude dysfluencies and pauses, leaving only the meaning-based messages from the monologue. The same transcription procedure was used in Boers (2014), and both types of transcription were used to analyze different aspects of language output during the monologues.

Monologues were analyzed for fluency, complexity, and accuracy and were based on the analysis of previous studies (Boers, 2015; Thai & Boers, 2015). These areas of oral production are also frequently targeted in task performance and oral production (e.g., Ellis, 2009; Foster & Tavakoli, 2009; Yuan & Ellis, 2003). Fluency was measured by collecting data related to syllables per minute, total syllable count, total length of pauses, average length of pauses, and false starts (dysfluencies). Complexity was measured by counting subordinate clauses used in the monologue. Lastly, accuracy was gauged by counting language errors (verb, tense, syntactic errors, vocabulary) and the number of times Japanese was used. In total, there were eight different measurements taken for the data in this study. In order to better convey how these measurements were applied to fluency, complexity, and accuracy, these concepts will be explored further.

Fluency is often thought to relate to the rate of production in a certain amount of time (Skehan & Foster, 1999). To measure fluency, both the rate of speech and dysfluencies must be considered to account for production and hindrances to production. In measuring speech rates, the pruned transcriptions were used for calculating syllables per minute and an overall syllable count. Dysfluencies were also considered in the analysis for fluency using the original transcription. Measured dysfluencies were the total length of pauses, average length of pauses, and false starts. False starts included speech at the beginning of a sentence that was repeated or replaced by additional speech that quickly followed.

Complexity is the ability to utilize and control advanced language effectively (Skehan & Foster, 1999). The measure of complexity used by Boers (2014), which was the number of subordinate clauses, was used in this study. Although subordinate clauses are a commonly used measure of complexity, there may in fact be more ways to measure complexity than only considering subordinate clauses.

Accuracy is the avoidance of errors in language production (Skehan & Foster, 1999). Accuracy was measured by grammatical errors such as subject-verb nonagreement and tense errors. Japanese use was also considered under accuracy, as occasionally learners would use Japanese words to fill lexical gaps or give brief explanations in Japanese. Students were considered to have lower accuracy if they resorted to using Japanese in their monologues. Although accuracy is a much more complex concept to discuss, for the purposes of this study we have grouped all common errors previously mentioned. Further study could explore the grammatical nuances of speech during expanding time monologues.

Results

Patterns can be observed between deliveries of the monologues (see Table 1). Differences between the first and third monologue were calculated for significance ($p < 0.01$) based on a t test. Results for the shrinking time condition and the constant time condition reflect similar results to those found in previous research into those conditions (Boers, 2014; Thai & Boers, 2015).

Table 1: Mean Change Between First and Last Monologue

Measure	3-2-1	2-2-2	1-2-3
Syllables per minute	59.43*	0.6	-36.3*
Total syllable count	-74.7*	21.6	133.3*
Total length of pauses	-24.32*	-5.6	19.5*
Average length of pause	-0.78	-0.2	0.52
False Starts	-2.9*	-0.3	0.6*
Subordinate clauses	-1.6	1	5.2*
Grammatical errors	-2.5*	1.2	2.8*
Japanese usage	-1.9*	-0.4	0

*Significant ($p < 0.01$) change between first and third monologue based on *t* test.

Fluency

Just as designed, the shrinking time condition resulted in significantly faster speech with fewer dysfluencies. However, it is not possible to say that all of these gains are solely related to time as the constant time condition also saw improvements in fluency. Determining if it is repetition or time that is influencing these changes will require further research. It is also noteworthy that the constant time condition overall produced a small amount of extra content with a shorter length of pauses. However, when the changes are looked at on a per-minute basis, fluency marked by syllables per minute did not experience a significant rise. This increase in speed could have been due to the fact that some of the content of the third monologue may have been planned after the second monologue. When we look at total mean values, the expanding time condition shows significantly more production but with more dysfluencies and at a slower rate. However, when we look at the monologue at a per-minute basis, syllables per minute and total syllable count show a significant decrease. This suggests that the participants were speaking more slowly at the end of the three repetitions. This could have been due to the demand to create more content and speak for longer periods of time. Whether the decreases in fluency were caused by the expanding time condition is a point that should be investigated in further research.

Table 2: Mean Change in Fluency of First and Last Monologue per Minute of Output

Measure	3-2-1		2-2-2		1-2-3	
	M	*SD*	M	SD	M	SD
Syllables per minute	104.14*	27.31	0.30	8.02	-92.83*	14.79
Total syllable count	59.43*	21.82	10.80*	9.52	-36.30*	9.19
Total length of pauses	-5.04	8.14	-2.88*	3.73	-5.47	5.41
Average length of pause	0.20	0.49	-0.10	0.15	-0.33*	0.39
False starts	-0.43	0.85	-0.20	1.03	-0.43	1.01

*Significant ($p < 0.05$) change between first and third monologue based on *t* test.

Complexity

When we look at the mean total in Table 1, results show similarities to findings by Boers (2014). Students were observed to use markedly fewer subordinate clauses in the shrinking time condition. However, when the results are broken down to per-minute figures, it can be seen that students were in fact using more subordinate clauses in the shrinking time condition. The constant time condition did show slight gains in complexity, but because the results did not reach significance and the data sample was small, it is difficult to reach much of a conclusion on this result as of yet. In the increasing time condition, total means show an increase in usage of subordinate clauses, however when the results are looked at per minute, there was no change and in fact a slight decrease in the number over time. Based on the results, it seems that the expanding time condition is the best for pushing learners to use more complex language (for an example of an expanding time condition, pruned transcription, see the Appendix). However, this is only when the dialogue is looked at as a whole and not minute-by-minute. Having the additional time allowed learners the opportunity to add extra details, examples, and explanation to their monologue.

Table 3: Mean Change in Complexity Between First and Last Monologue per Minute of Output

Measure	3-2-1		2-2-2		1-2-3	
	M	*SD*	M	SD	M	SD
Subordinate clauses	1.13*	1.03	0.45	0.90	-0.80	1.51

*Significant ($p < 0.05$) change between first and third monologue based on *t* test.

Accuracy

When we look at the mean totals, the shrinking time condition showed a significant increase in accuracy (observed by the change in grammatical errors); however, when the results are broken down to per-minute figures, decreasing accuracy is shown as grammatical errors increased significantly. The minute-by-minute results closely mirror findings by Thai and Boers (2015). As the study is still awaiting more participants to be added to the data pool, this point of contention will need to be more closely investigated. On the other hand, when we look at the expanding and constant time conditions as a whole, grammatical errors appeared to increase. If we look at the data on a per-minute basis, the increasing time condition group showed more accuracy gains, resulting in fewer grammatical errors. In terms of Japanese usage, a significant change was not observed. As mentioned previously, the need to produce new output without the benefit of repetition and potentially with a short period of planning might be the cause of this result. Interestingly, the use of Japanese decreased or stayed moderately the same in all three versions of the monologues.

Table 4: Mean Change in Accuracy Between First and Last Monologue per Minute of Output

Measure	3-2-1		2-2-2		1-2-3	
	M	*SD*	M	SD	M	SD
Grammatical errors	3.23*	2.24	0.50	1.55	-2.60*	2.46
Japanese usage	-0.23	0.47	-0.20	0.54	-0.27	0.58

*Significant ($p < 0.05$) change between first and third monologue based on 1 test.

Classroom Implications

The two current sets of results (one based on the total mean changes and one on per-minute changes) show vastly contradictory results, which make drawing definitive conclusions quite difficult. However, some useful insights for the classroom have emerged. The shrinking time condition exhibited signs of causing students to speak faster as a whole. This suggests that teachers who utilize this activity in their classrooms should expect positive results when looking at fluency in the shrinking time condition. Whether the expanding time condition aids accuracy is still in question and further research is needed to determine if such an activity also improves accuracy. In the constant time condition of 2-2-2, fluency was improved slightly and additional content was also added because learners were able to finish the content of previous monologues faster. As such, a constant time condition can aid fluency while also pushing output (Swain, 1995). Finally, the possibility that an expanding time condition may aid in improvements in complexity has been supported in this study. However, this is only when scores are looked at as a whole and not minute-by-minute production.

Based on these results, teachers can select the version that best suits the needs of the learners or targets a goal of a course or lesson. Besides in conventional discussion, incorporation of 1-2-3 into the classroom could include discussion games in which students must continue talking until the end of a round or else they are "out." Teachers can also adjust the time limit to meet the needs of their learners. For higher level learners, the activity could be expanded to debates during which teams must continue a debate for longer amounts of time such as 4-6-8 or 5-7-9. For lower level learners, the time limit can be decreased, so that discussions follow a 30 second-45 second-1 minute pattern. As such, any of the three versions observed in this study can prove beneficial and adaptable in the classroom.

Limitations

The biggest limitation in this study is that it is still ongoing, additional subjects to be included and further analysis to be done. Additionally, there may be weaknesses in the analysis that still need to be addressed: for example, only considering subordinate clauses as complexity and only considering grammatical errors, without breaking these errors into subcategories. These issues of analysis still can be adjusted, as the analysis does not require any changes to the data. The use of planning time in this study also may have impacted the data and further investigation into the impact of planning time, especially on the expanding time condition, may prove beneficial.

Conclusion

The effect of different time conditions on a series of three repeated monologues was explored in this study. The first of these series, argued for by Maurice (1983) as well as Nation (1989), involved a shrinking time condition and thus increased pressure on the learner to speak faster. The study indicated that the time pressure did in fact increase fluency and might also have some influence on accuracy as well. The constant time condition, introduced in previous studies (Boers, 2014; Thai & Boers, 2015), also showed improvements in fluency, although not to the extent of the shrinking time condition. Extra content was also added in the constant time condition, leading to pushed output (Swain, 1995). The new element that this study added to previous research was an expanding time condition series of monologues. The results indicated less fluent, less accurate, but more complex monologues. Overall, this study provides a closer look at the way repetition and time influence classroom language production for students.

Bio Data

Benjamin Thanyawatpokin is currently a teacher and researcher at Ritsumeikan University. His research interests include CALL and the application of video games to the language learning classroom. In addition to this, he also does some research into student output teaching methods. <btpokin@gmail.com>

Carl Vollmer is an English and social studies teacher at Ritsumeikan Uji Junior and Senior High School. He holds an MA in English education from Ritsumeikan University. His research interests include task-based learning, conversation analysis, and written corrective feedback. <carlvollmer34@gmail.com>

References

Boers, F. (2014). A reappraisal of the 4/3/2 activity. *RELC Journal, 45*(3), 221-235. https://doi.org/10.1177/0033688214546964

Ellis, R. (2009). The differential effects of three types of task planning on the fluency, complexity, and accuracy in L2 oral production. *Applied Linguistics, 30*(4), 474-509. https://doi.org/10.1093/applin/amp042

Foster, P., & Tavakoli, P. (2009). Native speakers and task performance: Comparing effects on complexity, fluency, and lexical diversity. *Language Learning, 59*(4), 866-896. https://doi.org/10.1111/j.1467-9922.2009.00528.x

Izumi, S. (2002). Output, input enhancement, and the noticing hypothesis: An experimental study on ESL relativization. *Studies in Second Language Acquisition, 24*, 541-577. https://doi.org/10.1017/S0272263102004023

Maurice, K. (1983). The fluency workshop. *The TESOL Newsletter, 17*(4), 29.

Nation, P. (1989). Improving speaking fluency. *System, 17*(3), 377-384.

Nation, P. (2016, September). *Unanswered questions in L2 vocabulary teaching and learning*. Paper presented at the 2016 Vocab@Tokyo Conference, Tokyo, Japan.

Ogawa, C. (2016). Examining the effects of types of pretask planning on oral performances. *JALT Journal, 38*(2), 97-118. Retrieved from http://jalt-publications.org/node/3/articles/5540-examining-effects-types-pretask-planning-oral-performances

Sadeghi Beniss, A. R., & Edalati Bazzaz, V. (2014). The impact of pushed output on accuracy and fluency of Iranian EFL learners' speaking. *Iranian Journal of Language Teaching Research, 2*(2), 51-72.

Skehan, P., & Foster, P. (1999). The influence of task structure and processing conditions on narrative retellings. *Language Learning, 49*(1), 93-120. https://doi.org/10.1177/136216889700100302

Swain, M. (1995). Three functions of output in second language learning. In G. Cook & B. Seidlehofer (Eds.), *Principles and practice in applied linguistics* (pp. 245-256). Oxford, England: Oxford University Press.

Thai, C., & Boers, F. (2015). Repeating a monologue increasing time pressure: Effects on fluency, complexity, and accuracy. *TESOL Quarterly,* 1-25. https://doi.org/10.1002/tesq.232

Yuan, F., & Ellis, R. (2003). The effects of pre-task planning and on-line planning on fluency, complexity and accuracy in L2 monologic oral production. *Applied Linguistics, 24*(1), 1-27. https://doi.org/10.1093/applin/24.1.1

Appendix

Example Pruned Transcription for 1-2-3 Activity

Transcription key:
Underline – Sections of grammatically improved output from previous monologues
Italics – Additional content, not in previous monologues
Bold – Japanese use

1 Minute

- I want to shopping at least once a week
- before starting university life
- I often to Aeon mall in Himeji because its the biggest shopping mall shopping mall
- near my house and my home town's supermarket isn't good
- It's very near my house
- It's only five minutes walk from my house but, vegetables and fruits and flesh and they supermarket doesn't have a very good selection of goods
- I haven't been there lately because my university life is very busy
- But, I like shopping just looking clothes department clothes department vegetables and

2 Minutes

- I went shopping at least once a week before starting University life
- I often went to Aeon mall in Himeji because its the biggest shopping mall near my house and my home town's shopping isn't good
- It's very near my house
- It's only 5 minutes walk from my home, but vegetables, tables, fruits, and meat every day and the supermarket doesn't have a very good selection of goods
- I haven't been there lately because, my university life is very busy
- But, I like shopping just looking at cosmetic and clothes department or vegetables, *fruits, corner in a supermarket or is very fun for me*
- *relive relief relief university life stress by shopping*
- *and I like eat out than home I don't like cooking and I don't like to have dishes that they aren't cooked by my family*

3 Minutes

- I want to shopping at least once a week before starting University life. I often to Aeon Aeon Mall in Himeji because its the biggest shopping mall near my house
- And my home town's supermarket and ***fuku***. It's they are very uh it is very near my house
- It's only 5 minutes walk from my house
- But, the supermarket's vegetables and fruits and meat doesn't have a very good selection of goods
- And everytime there are no people. I haven't been there lately because my University life is very busy, but I like shopping
- Just looking at cosmetic, clothes department or vegetables fruits corner in a supermarket is very fun for me
- And, it's relieves my university life by shopping
- And I like better to eat out than at home. I don't like cooking and I don't like to have dishes that are cooked by my family
- *because my family doesn't good for cooking*
- *My family sometime buy expensive brand items for example, clothes*
- *My mother bought brand clothes and cosmetic, but it's rare*
- I think I think if you don't have expensive or brand goods, you don't need it
- Cheap goods is I think cheap goods is good
- For example Daiso's items is easy to broke but, I think I don't need