Working memory capacity and L2 speech production in a picture description task with repetition

Memória de trabalho e a produção oral em L2 numa tarefa de descrição de figura com repetição

Kyria Finardi
UFSC/CNPQ

Abstract

This study investigates the relationship between working memory capacity and gains in L2 oral performance in a picture description task with repetition. It departs from Fortkamp (2000) who found significant correlations between working memory capacity and measures of L2 speech performance and Bygate´s (2001b) findings that when repeating an oral task participants gained in complexity of L2 speech at the expense of accuracy and fluency, to propose that there might be a correlation between gains in complexity of L2 speech and working memory capacity. Results show there are gains in complexity of L2 speech in the second trial of the same task, but these gains do not correlate with working memory capacity. The lack of correlations is explained by the small number of participants, which did not allow for variation in terms of working memory capacity.

Keywords

Working memory capacity; Picture description with repetition; Complexity of L2 speech.

Resumo

O estudo investiga a relação entre a capacidade de memória de trabalho e os ganhos na produção oral em L2 durante uma tarefa de descrição de figura com repetição. Partindo de Fortkamp (2000) que encontrou correlações entre a capacidade de memória de trabalho e medidas de produção oral em L2 e Bygate (2001b) que mostrou que na repetição de tarefas os participantes ganham em complexidade de fala em detrimento, principalmente da acurácia e da fluência, este estudo propõe que pode haver correlações entre os ganhos na complexidade da fala em L2 e a capacidade de memória de trabalho. Resultados do experimento mostram ganhos na complexidade da fala em L2, mas os mesmos não correlacionam com a capacidade de memória de trabalho. A falta de correlações é explicada em razão do pequeno número de participantes que não permitiu variação na capacidade de memória de trabalho.

Palavras-chave

Capacidade de memória de trabalho; Descrição de figura com repetição; Complexidade da fala em L2.

Introduction

Speaking a foreign language is a cognitive skill that involves many complex sub-processes for its execution. As with other skills, most of these sub-processes take place without our awareness and/or control and must be automatized so as to free mental capacity to execute other processes and sub-processes (ANDERSON, 1995).

One way to approach the sub-processes involved in L2 speaking is to adopt an information processing perspective that conceptualizes human beings as autonomous, active, and limited-capacity processors (ASHCRAFT, 1994) with a working memory system responsible for online processing and temporary maintenance of information in the performance of complex tasks, such as problem solving, reading, and speaking among others (BADDELEY & LOGIE, 1999). The mental processes involved in the performance of complex tasks compete for the limited attention capacity of the working memory, which has to be shared between on-line processing and storage of relevant information (DANEMAN, 1991).

Most studies on L2 speech production agree that the mastery of a foreign language involves speaking it with complexity, fluency, and accuracy (BYGATE, 2001; D'ELY, 2003; FORTKAMP, 2000; SKEHAN, 1998, to mention but a few). Studies on task effects and speech production show that there are trade-off effects among these three competing goals of oral production. There seem to be, in particular, trade-off effects between complexity and accuracy (BYGATE, 2001; FORTKAMP, 2000; D'ELY, 2003, 2004; SKEHAN, 1996; SKEHAN, 1998).

Bygate (2001b) studied the effects of task familiarity on speech performance. His assumption was that when learners had the opportunity to perform the task for the second time, their attention would be targeted to different aspects of the oral production process at each new practice opportunity, thus allowing learners to improve their performance gradually and differentially across speech dimensions. He claimed that task repetition could influence learners' oral performance by relocating their focus of attention. He hypothesized that the

performance in the repeated task would be better than the first trial in terms of fluency, accuracy, and complexity. His hypothesis was not fully confirmed, as there were trade-off effects among the three competing goals of oral performance; that is, accuracy, fluency, and complexity. However, he found that, overall, complexity seemed to improve in the repeated task. He therefore concluded that speech performance lost in accuracy and fluency so as to gain in complexity in the repetition condition.

Similarly to the trade-off effects found in L2 speech production, studies on working memory (MW) show evidence the trade-off effects between its two main functions, namely, the storage and processing of information during the execution of complex tasks. Most research to date acknowledges the fact that working memory capacity (WMC) may be seen as a possible independent variable in the processes involved in both L1 (DANEMAN & GREEN, 1986; DANEMAN, 1991) and L2 speaking (FORTKAMP, 1999, 2000; FINARDI & PREBIANCA, 2006; WEISSHEIMER, 2007). These studies have shown that individuals with more working memory capacity (WMC) tend to outperform those with less WMC in fluency, accuracy, complexity, and lexical density.

Bearing this panorama in mind, the aim of this paper is to analyze the role of working memory capacity in the L2 speech production of learners exposed to a picture description task with repetition. The assumption underlying this paper is that learners who have more working memory capacity will be able to allocate more attention to the processes involved in speaking an L2, retrieving more information from long-term memory, and, as a consequence, benefiting more from the repetition condition than those participants with less working memory capacity.

The Study

The main assumption supporting the present study is that L2 speaking is a complex cognitive task that is carried out within the constraints of a limited-capacity system, namely, working memory. In this system, there are trade-off effects between the storage and processing functions of working memory, just as in L2 speaking there seems to be now sufficient evidence for the trade-off effects among fluency, accuracy and complexity when L2 learners perform under processing pressure (FORTKAMP, 2000; BYGATE, 2001; WEISSHEIMER, 2007). Aiming at investigating this relationship, the following research question is put forward:

Is there a relationship between working memory capacity and gains in performance in L2 speech production measures in a picture description task with repetition?

Method

The main objective of this study is to investigate the relationship between individual differences in working memory capacity and L2 learners' gains in performance in the second trial of the repetition condition. The study departs from the assumption that individuals with more working memory capacity will be able to allocate more attention to the processes involved in performing a task under the repetition condition, thus, benefiting more from this condition than individuals with less working memory capacity, and as a consequence, showing more gains in performance. With that aim in mind, the following hypotheses were proposed:

1. There are no gains in performance in terms of complexity/ accuracy/ fluency of L2 speech in the second trial of the repetition condition.
2. There are gains in performance in terms of complexity/fluency/accuracy of L2 speech in the second trial of the repetition condition and such gains correlate with individual differences in working memory capacity.
3. There are gains in performance in terms of complexity/fluency/accuracy of L2 speech in the second trial of the repetition condition but such gains do not correlate with individual differences in working memory capacity.

Participants

This study was conducted in an intact class environment where students attended English classes for free in exchange for participating in research. Different researchers were collecting data on this group in a collaborative enterprise, but all of them did so as part of the class routine. The total number of students in the group was 24, but only twelve of them (6 male and 6 female) were used for this specific study. Although the rest of the group was not used for this particular study (for some of them had missed class or not taken all the tests), all the students in class would follow the same procedure, doing the same tasks and tests. The researcher was teaching the group throughout the whole semester. All the participants in this experimental class were pre-tested with an

oral interview to take part in participacion in this group to ensure that all participants had the same L2 oral proficiency level (in this case, intermediate).

Data collection

Two tools for data collection were used in this study, a working memory capacity test and a speech production test, both in L2. The working memory capacity test used was Weissheimer's (2006) speaking span test (SST) which was constructed following Daneman's (1991) and consisted of 120 unrelated words organized in six sets of 2, 3, 4, 5 and 6 words (Appendix 1). Each word was presented individually in the middle of a computer screen for one second. Participants were instructed to read the words silently. After ten milliseconds, the next word in the set appeared in the same position and this procedure was followed until the set ended and a black screen with question marks (the same number as the words presented in that set) appeared on the screen. Participants were informed that these question marks signaled the number of words presented and the number of sentences they should try to make. Participants were also instructed to try to use the words in the correct order in a grammatically correct way to form sentences and say them aloud in English.

A training phase (60 words, the first three trials) preceded the testing phase (60 words, trials 4, 5 and 6, in Appendix 1) and the actual test did not start until the participants reported feeling comfortable to perform the test. Participants' speaking span was defined as the maximum number of words (out of 60) for which they could generate sentences. Two scores were calculated for this test, a strict and a lenient one. In the strict score, only sentences which were grammatically correct and used the target word in the correct order were given one point. In the lenient score, half a point was also given for sentences which were partially correct (e.g.: The girl live on the farm) or when the sentence was correct but the target word was in a different order.

Speech production was elicited through a picture description task. The picture was an advertisement for clothes in a magazine showing many people on a busy street (Appendix 2). The reason for using a description task was two-fold: The first can be traced back to Fortkamp (2000), who, through linear regression analysis, found that working memory capacity (WMC) was a better predictor of speech performance in the description task than in the narrative task. In the narrative task, the performance was only linearly related to working memory

capacity. Among the different measures she used for speech performance, complexity earned the highest scores, thus, showing that, WMC was a good predictor of speech performance among the many measures, especially of complexity. The second reason is that, according to Robinson (1995), picture description is a *here and now* type of task, which is less cognitively demanding than *there and then* tasks such as narratives. Since this group was an intermediate level group, the researcher thought that using a less demanding task would probably elicit fewer mistakes, thus showing better effects on accuracy during the second trial. This assumption was not supported by the data, for the only improvement found in the second trial of the description task was in terms of complexity, possibly because of the interaction between type of task and task familiarity (BYGATE, 2001b).

Participants were shown the picture during individual sessions with the researcher and were then asked to describe it without looking at notes. These sessions were recorded and transcribed. During the classes, participants were shown the transcriptions for their description in the first trial and had the opportunity to check vocabulary with the teacher. After a one-month interval, in which participants had regular EFL classes and saw their transcriptions, participants received the same instruction as in the first trial and were shown the same picture again and asked to describe it in individual sessions that were also recorded.

Four measures of speech performance were calculated following Fortkamp (2000) for the transcriptions at the first and second trial of the picture description task. Complexity (Comp) of speech was measured in terms of number of dependent clauses per minute and was calculated by dividing the total number of dependent clauses by the time taken to accomplish the task in seconds and then multiplied by 60 to express the number in minutes. Fluency (Fl) was measured in terms of unpruned speech rate and was calculated by, dividing the total number of semantic units produced, including repetitions, by the total time, including pause time, and expressed in the number of seconds that the subject took to complete the task. Accuracy (Acc) was calculated by the total number of errors divided by the number of semantic units produced and the resulting figure multiplied by 100 to express the number of errors per 100 words. Finally, lexical density (LD) was calculated by the total number of weighted lexical items divided by the total number of weighed linguistic items and multiplied by 100 so as to obtain the percentage of weighted lexical items over the total number of weighted linguistic items in the speech sample.

In this study, gains in performance were operationalized as speech with more complexity and/or fluency and/or lexical density and/or accuracy in the second trial of the description task and were measured through a comparison of means. To investigate the relationship between individual differences in working memory capacity and gains in speech production in the second trial of the repetition condition a correlation was run between the speaking span test and gains in speech production measures.

Results

Raw scores of the four measures of speech performance in Table 1 show a complex pattern. There seems to be an improvement in terms of complexity of L2 speech in the second trial but these gains are paid for by a loss in the other measures of speech performance, especially of accuracy, which seems to be the most penalized dimension of speech performance in the second trial. Recall that accuracy scores have to be interpreted in a different way from the other measures, since the higher the score, the more mistakes the participant produced.

TABLE 1
Scores speech production measures at first and second trials

Participant	Comp 1	Comp 2	Acc 1	Acc 2	SR 1	SR 2	LD 1	LD 2
1	,57	,90	2,61	3,01	87,4	111,3	61,72	55,71
2	0	,39	1,13	6,54	80,0	67,2	60,14	57,5
3	0	,69	1.78	3,19	103	65.58	62.22	58.27
4	2,27	1,24	,51	2,48	147	67,08	58,38	58,92
5	0	,79	1,42	4,39	54,5	72,79	67,46	58,67
6	1,20	2,52	4,16	9,17	67,2	68,84	60,60	57,40
7	1,24	2,64	3,28	6,41	114	82,50	54,88	55,33
8	0	1,71	4,76	3,06	48,5	84	78,26	60,74
9	,25	,39	3,47	4,68	54	48	57,55	56,74
10	,82	1,31	1,86	3,75	87,9	87,69	60,18	55,35
11	0	.90	1.66	6.21	60	36.31	70.68	57.67
12	2,68	3,06	3,51	3,06	115	99,79	57,5	58,76

In a preliminary analysis of the raw data, independent sample *t*-tests (Appendix 3) were run for all the speech production measures (Complexity-Comp, Accuracy- Acc, Speech Rate- SR and Lexical Density-LD) but no significant gains in performance were found in the second trials except for the complexity measure. The *t*-test for the complexity measures showed that in fact there were gains in complexity in the performance of the second trial of the repetition condition and these gains were significant at $p< 0.05$ as can be seen in Table 2.

TABLE. 2
Paired Samples *t*-Test for complexity measures at first and second trials

	Mean	Std.Dev.	St.error	Lower	Upper	*t*	Df	Sig.
Comp 1								
Comp 2	-,6258	,7111	,2053	-1,0777	-,1740	-3,049	11	,011

However, the result of a Pearson Product Moment Correlation Analysis did not show significant correlations between gains in complexity of speech in the second trial and individual differences in working memory capacity. The only significant correlation found was between complexity in the first and second trials (r = .702*). Thus, only hypothesis 3 was confirmed, that is, there were significant gains in terms of complexity of speech in the second trial of the repetition condition but these gains did not correlate with individual differences in working memory capacity. Individuals with a higher working memory capacity were therefore not the ones to profit more in terms of complexity of speech in the second trial of the repetition condition.

Discussion

A plausible explanation for the lack of correlation between gains in complexity of L2 speech at second trial and individual differences in working memory capacity is that the sample size used in this study was too small to allow statistical variation. Ideally, so as to see differences in terms of working memory capacity, the group should be split into high and low spans. Unfortunately, this group varied very little in terms of working memory capacity, since most of the participants were categorized as medium spans and so the number of higher and lower spans was too small to see differences in the performance of this condition.

Another variable that may have hindered the results of this study is the fact that the group selected was an experimental group, where students had to do many tasks and tests as part of the program. The same group was used in four different studies and so the participants may have felt tired and overexposed to the applied tests. The second trial of the description, for instance, was done on the same day as the speaking span test and, although the data collection session took no longer than 30 minutes per participant, the students may have been unwilling to take their time and show their best performance in the second trial of the description, since they were doing the same task for the second time.

Another aspect that must be taken into consideration when analyzing the results of this study is the fact that some students may have perceived the tasks as tests and so behaved accordingly. As Iwashita et al (2002) suggest, performance on tests differs from performance in class and so has to be analyzed differently and with caution. Whereas some of the students may have perceived the tasks as tests and felt stressed during its performance, others may have simply regarded the tasks as repetitions and so were not willing to do their best. Whatever the case at hand, task implementation for research purposes must be carried out with care and consideration of these issues.

Conclusion

The main goal of this study was to analyze individual differences in working memory capacity and their relationship with gains in L2 speech performance of learners exposed to the repetition condition. As can be seen from the analysis, there were significant gains in performance in terms of complexity of L2 speech in the second trial of the repetition condition, but these gains did not correlate with individual differences in working memory capacity, probably due to the limited sample size used in this study.

The theory supporting the repetition condition is not only logical but also appealing and aligned with the information processing paradigm that sees repetition as an important condition to help in the automatization of procedures. Moreover, if we agree that by automatizing procedures there will be more attentional resources to devote to other components of the task execution, then it makes sense to think that depending on an individual's working memory capacity, he/she will pay more or less attention to the speech production task at a second trial, focusing on different aspects of this activity in each encounter with the task.

Unfortunately, this study could not produce evidence for this assumption, and so it remains to be seen whether individual differences in working memory capacity play a crucial role in the benefits advocated by the theory that sees repetition as an important pedagogical aid to L2 speech development. However, in order to be fair, it is important to view the lack of significant correlations between gains in L2 speech production measures at second trial and working memory capacity in this study as caused by methodological (limited number of participants) rather than theoretical problems. Future studies should still be able to produce evidence for this claim once the methodological limitations are overcome and larger samples sizes are used.

Since the working memory is at the crux of human cognition, it can not be left behind in studies aiming at making claims for the beneficial effects of task designs. There is evidence for the role of both task manipulation and individual differences in working memory capacity in human cognition. Nevertheless, these studies are usually carried out in isolation and so the results have to be integrated later on if we are to draw a precise map of what affects human cognition in general. To fill in this gap, more studies are needed to scrutinize the effects of task type and task conditions and working memory capacity on human cognition in general, and on L2 in particular.

Notwithstanding the limitations of this study, it succeeded in providing further support for the benefits advocated by the repetition condition (Bygate, 2001b), at least to the extent that they relate to the complexity of L2 speech. Given this panorama, it remains to be seen how durable these effects are and how they transfer to other types of task. Studies with different types of task and task conditions integrated with individual differences in working memory capacity should provide important insights to inform future pedagogical practices.

Finally, if we are to have a clear picture of the effects of the repetition condition on speech performance and its relationship to individual differences in working memory capacity, more measures of speech performance would have to be analyzed to see how they interact under this specific condition. The same holds true for the construct of working memory capacity, which, as a latent variable, requires indirect techniques of measurement and analysis. One possibility to safeguard against expected shortcomings in measuring latent variables would be the use of multiple measures of working memory capacity and complex factor analysis, which, unfortunately, were beyond the scope of this paper.

References

ANDERSON, J.R. *Learning and Memory.* Singapore: Wiley, 1995. Chapter 9.

BADDELEY, A.D.; LOGIE, R.H. The multiple-component model. In: Akira MIYAKE, Akira; SHAH, Priti (Ed.). *Models of working memory: Mechanisms of active maintenance and executive control.* New York, NY: Cambridge University Press, 1999. p. 28-61.

BYGATE, M. Theoretical perspectives on speaking. *Annual Review of Applied Linguistics*, 18, p. 20-42, 1998.

BYGATE, M. Speaking. In: CARTER, R.; NUNAN, D. (Ed.). *The Cambridge Guide to Teaching English to Speakers of Other Languages.* Cambridge: Cambridge University Press, 2001. Chapter 2, p. 14-20.

BYGATE, M. Effects of task repetition on the structure and control of oral language. In: BYGATE, M.; SKEHAN, P.; SWAIN, M. (Ed.). *Researching pedagogic tasks – second language learning and testing.* Longman, 2001b.

DANEMAN, M.; GREEN, I. Individual differences in producing and comprehending and producing words in context. *Journal of Memory and Language,* 25, p. 1-18, 1986.

DANEMAN, M. Working memory as a predictor of verbal fluency. *Journal of Psycholinguistic Research,* 20, p. 445-464, 1991.

DE BOT, K. A bilingual production model: Levelt's speaking model adapted. *Applied Linguistics,* 13, p. 1-24, 1992.

D'ELY, R. Do task planning and task repetition enhance learners' second language performance? An empirical study. Trabalho de pesquisa não publicado. UFSC, 2003.

D'ELY, R. A focus on learners' metacognitive processes: strategic planning, repetition and planning for repetition as catalysts of interlanguage development. *Trabalho de pesquisa não publicado.* UFSC*,* 2004.

D'ELY; R.; FONTANINI, I.; WEISSHEIMER, J.; BERGSLEITHNER, J.; PERUCCI, M. (2004). Memória de trabalho e desempenho em tarefas L2. *Revista Brasileira de Linguística Aplicada,* v. 5, n.2, 2005.

FINARDI, K.; PREBIANCA, G. Working memory capacity and speech production in L2: evidence from a picture description task. *Revista de Estudos da Linguagem,* Faculdade de Letras da UFMG, v. 14, n.1, p. 231-260, 2006.

FORTKAMP, M.B.M.Working memory capacity and aspects of L2 speech production. *Communication and Cognition,* v.32, p. 259-296, 1999.

FORTKAMP, M.B.M. *Working memory capacity and L2 speech production: an exploratory study.* 2000. Tese (Doutorado) - UFSC, Pós-Graduação em Inglês e Literatura Correspondente, Florianópolis, 2000

FOSTER, P.; SKEHAN, P. The influence of planning and task type on second language performance. *Studies in Second language acquisition,* v. 18, p.299-323, 1996.

IWASHITA, N.; MCNAMARA, T; ELDER, C. Can we predict task difficulty in an oral proficiency test? Exploring the potential of an information processing approach to task design. *Language Learning,* v. 51, n. 3, p.401-436, 2002.

LEVELT, W. J. M. *Speaking: From intention to articulation*. Cambridge, M.A.: Bradford/MIT Press, 1989.

MIYAKE, A.; SHAH, P. Toward a unified theory of working memory: emerging general consensus, unresolved theoretical issues, and future research directions. In: MIYAKE, A.; SHAH, P. (Ed.). *Models of working memory: mechanisms of active maintenance and executive control,* 1999. p.442-481.

ROBINSON, P. Aptitude, awareness, and the fundamental similarity of implicit and explicit second language learning. In: SCHMIDT, R. (Ed.). *Attention and awareness in foreign language learning*. Honolulu, Hawaii: University of Hawaii, Second Language Teaching & Curriculum Center, 1995. p.303-357.

SKEHAN, P. A framework for the implementation of task-based instruction. *Applied Linguistics,* v. 17, p.38-6, 1996.

SKEHAN, P. *A Cognitive Approach to Language Learning*. Oxford: Oxford University Press, 1998.

WEISSHEIMER, J.; FORTKAMP, M.B.M. *The effects of practice and strategy use on L2 working memory capacity and fluency*. Unpublished research paper. UFSC, 2004.

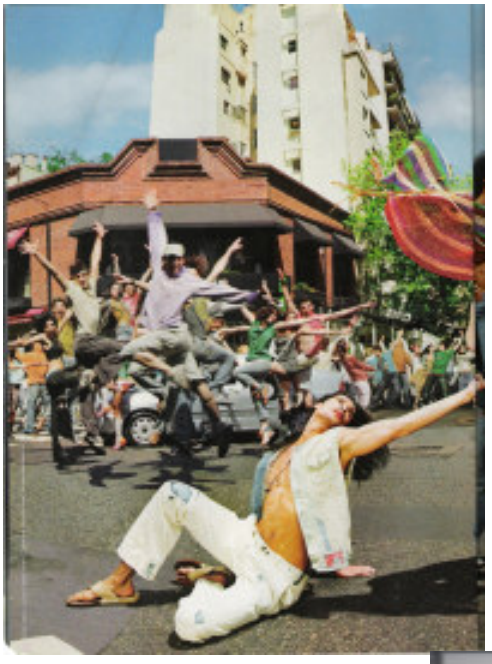
WEISSHEIMER, J. *Working memory capacity and the development of L2 speech production – an exploratory study*. Unpublished research paper. UFSC, 2006.

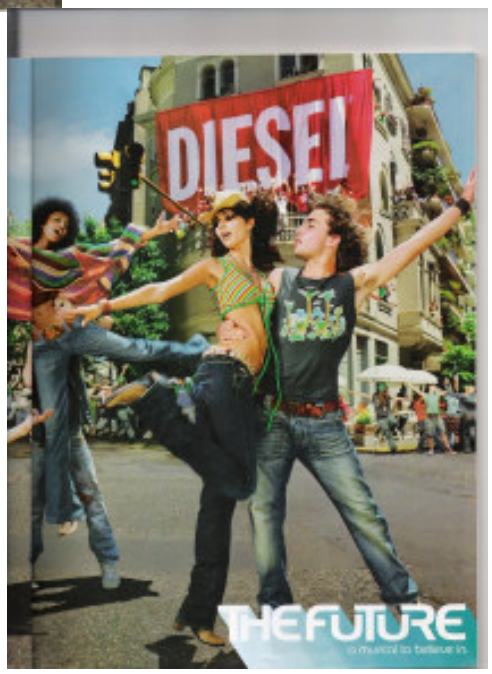
WEISSHEIMER, J. *Working memory capacity and the development of L2 speech production: an exploratory study*. Unpublished Doctoral Dissertation, Florianopolis, UFSC, 2007.

APPENDIX 1

Words included in the L2 Speaking Span Test

Trial 1	Trial 2	Trial 3	Trial 4	Trial 5	Trial 6
HOUSE	PEOPLE	BOSS	ARM	SPOON	BALL
BEACH	EARTH	ISLAND	COURSE	BANK	TOOL
SCHOOL	WIFE	TEA	GUY	DATE	ICE
HOBBY	SOCCER	MOUTH	POINT	GAS	BREAD
FAMILY	POWER	SPORT	TRAIN	SKY	SEA
TEAM	WORLD	BABY	COW	CAR	BAG
MUSIC	SUMMER	IDEA	FIRE	DOOR	YEAR
NIGHT	OCEAN	MOVIE	SHOE	PEN	KING
FRIEND	APPLE	SPACE	KEY	DISK	BAND
SNACK	ROOM	TAXI	SNOW	BIRD	FLAG
DRUG	BALL	GIFT	OIL	SEAT	JOB
HONEY	NURSE	CLOCK	DOOR	BATH	AIR
LIGHT	TRUCK	WOMAN	BOAT	GIRL	BRAIN
FACE	ACTRESS	FISH	TOY	CLUB	BOY
MOTHER	MOON	MILK	ART	STREET	CLASS
COFFEE	WORKER	LUNCH	BOX	BED	FARM
PRISON	HEAD	WINDOW	FLOOR	MIND	BUS
NUMBER	CITY	MONEY	ROCK	MAIL	TV
POEM	DRESS	PROBLEM	COAT	BEER	FILE
	PLANT	PARTY	BOOK	PAIR	CROWD




DIESEL
THE FUTURE

APPENDIX 3

T-tests between trials 1 and 2 of the speech production measures

Paired Samples Test

		Paired Differences					t	df	Sig. (2-tailed)
		Mean	Std. Deviation	Std. Error Mean	95% Confidence Interval of the Difference				
					Lower	Upper			
Pair 1	accuracy 1 - accuracy 2	-2.1500	2.1821	.6299	-3.5364	-.7636	-3.413	11	.006
Pair 2	speech rate 1 - speech rate 2	10.6108	31.0887	8.9745	-9.1420	30.3636	1.182	11	.262
Pair 3	complexity 1 - complexity 2	-.6258	.7111	.2053	-1.0777	-.1740	-3.049	11	.011
Pair 4	weighed lexical density description 1 - weighed lexical density description 2	4.8758	5.7461	1.6587	1.2250	8.5267	2.939	11	.013

N=12