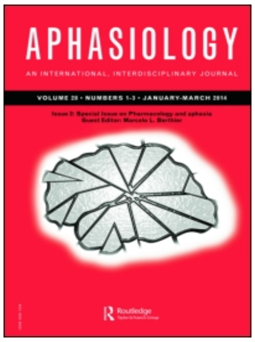

Aphasiology

ISSN: 0268-7038 (Print) 1464-5041 (Online) Journal homepage: http://www.tandfonline.com/loi/paph20

A comparison of two working memory tasks in aphasia

Maria V. Ivanova, Svetlana V. Kuptsova & Nina F. Dronkers

APHASIOLOGY, 2017
VOL. 31, NO. 3, 265–281
http://dx.doi.org/10.1080/02687038.2016.1172699

Routledge
Taylor & Francis Group

A comparison of two working memory tasks in aphasia

Maria V. Ivanova[a], Svetlana V. Kuptsova[a,b] and Nina F. Dronkers[a,c,d]

[a]Neurolinguistics Laboratory, National Research University Higher School of Economics, Moscow, Russia; [b]Center for Speech Pathology and Neurorehabilitation, Moscow, Russia; [c]Center for Aphasia and Related Disorders, VA Northern California Health Care System, Martinez, CA, USA; [d]Department of Neurology, University of California, Davis, CA, USA

ABSTRACT

Background: Overall, there is growing consensus that working memory (WM) should be routinely assessed in individuals with aphasia as it can contribute significantly to their level of language impairment and be an important factor in treatment planning. However, there is still no consensus in the field as to which tasks should be used to assess WM in aphasia. The two main alternatives are adapted complex span tasks and N-back tasks. Both have been used interchangeably in previous studies of WM in aphasia, even though the correspondence between the two tasks has not been properly established.

Aims: The current study investigates the relationship between two WM tasks—complex span and N-back tasks—in a large sample of individuals with aphasia. The relationships of these tasks to measures of language comprehension are also explored, as well as differences in performance patterns between individuals with non-fluent and fluent aphasia.

Methods & Resources: Forty-four participants with aphasia (non-fluent: n = 27; fluent: n = 13; mixed: n = 4) were examined with a modified listening span task (Ivanova & Hallowell, 2014), an auditory verbal 2-back task, and a standardised Russian language comprehension test.

Outcomes & Results: Results revealed a moderate relationship between the two WM measures, but demonstrated a divergence in terms of their relationship to language comprehension. Performance on the modified listening span task was related to language comprehension abilities, but performance on the 2-back task was not, suggesting that the two tasks primarily index different underlying cognitive mechanisms. Furthermore, the relationship between the modified listening span task and language comprehension was significant for individuals with non-fluent aphasia, but not for those with fluent aphasia.

Conclusions: Overall, the data demonstrate that while performance of individuals with aphasia was related on the two tasks, the two tasks cannot be substituted for one another without further inquiries into their underlying differences.

ARTICLE HISTORY
Received 6 November 2015
Accepted 24 March 2016

KEYWORDS
Working memory; complex span tasks; N-back tasks; fluent aphasia; non-fluent aphasia; language comprehension

CONTACT Maria V. Ivanova mivanova@hse.ru

Introduction

Deficits in working memory (WM), the capacity to temporarily hold and manipulate task relevant information, are amongst the most widely acknowledged cognitive impairments in aphasia (Salis, Kelly, & Code, 2015). This is not surprising given the strong association between language and WM (Daneman & Merikle, 1996; Just & Carpenter, 1992) and that WM engages brain areas often damaged in aphasia (Chein, Moore, & Conway, 2011; D'Esposito & Postle, 2015; Ricker, AuBuchon, & Cowan, 2010; Rottschy et al., 2012; Sreenivasan, Curtis, & D'Esposito, 2014). Numerous previous investigations of WM in aphasia have reported decreased WM capacity in individuals with aphasia compared to age-matched neurologically healthy control groups (Christensen & Wright, 2010; DeDe, Ricca, Knilans, & Trubl, 2014; Ivanova & Hallowell, 2014; Tompkins, Bloise, Timko, & Baumgaertner, 1994; for a review see; Salis et al., 2015; Wright & Fergadiotis, 2012). Many studies have established strong relationships between decreased WM capacity and language comprehension (Ivanova, Dragoy, Kuptsova, Ulicheva, & Laurinavichyute, 2015; Sung et al., 2009; Wright, Downey, Gravier, Love, & Shapiro, 2007). Further, an emerging body of research suggests that, in individuals with aphasia, treatment of specific cognitive deficits, including WM, could lead to improvements not only in these cognitive non-linguistic domains, but also in various language comprehension and production abilities (Berthier et al., 2014; Francis, Clark, & Humphreys, 2003; Helm-Estabrooks, Connor, & Albert, 2000; Mayer & Murray, 2002; Salis, 2012; for a review see Salis et al., 2015).

Overall, there is a growing appreciation that WM should be routinely assessed in individuals with aphasia as it can contribute significantly to their level of language impairment and be an important factor in treatment planning. However, there is still no consensus in the field as to which tasks should be used to assess WM in aphasia, whether in aphasia research or in clinical practice (DeDe et al., 2014; Ivanova & Hallowell, 2014; Salis et al., 2015; Wright & Fergadiotis, 2012). The two primary paradigms used in aphasia research are simplified complex span tasks—specifically adapted for individuals with language impairment—and N-back tasks. In a typical complex span task, a processing task (e.g., sentence reading) is given along with a set of stimuli (e.g., words) to be remembered for later recall or recognition. In N-back tasks, participants are presented with a continuous string of items and are instructed to judge whether an item matches a previous one that was presented N-items before. Proponents of complex span tasks state that these tasks are the gold standard for assessing WM capacity in cognitive psychology (Conway et al., 2005) and that variations of these tasks are a valid means of indexing WM capacity within different theoretical frameworks (Caspari, Parkinson, LaPointe, & Katz, 1998; Ivanova & Hallowell, 2014; Sung et al., 2009; Waters & Caplan, 2003). On the other hand, researchers using N-back tasks assert that since these tasks are more language-free in nature, they are more appropriate for indexing cognitive non-linguistic abilities in language-impaired populations (Christensen & Wright, 2010; Mayer & Murray, 2012). Apart from these two tasks, backward span tasks, that require participants to repeat items in the reverse order to which they were presented, have also been used to assess WM in aphasia in a few studies (DeDe et al., 2014; Laures-Gore, Marshall, & Verner, 2011). However, in aphasia research, the adapted complex span and N-back tasks remain the two most commonly used tools to investigate WM capacity (Wright & Fergadiotis, 2012) and were hence utilised for the current study.

A key question regarding these two tasks is whether they can be used interchangeably as a measure of WM. From a conceptual standpoint, the two tasks involve different cognitive processes. The complex span task requires shifting rapidly between several processes: storing and constantly rehearsing incoming items along with performing some type of a parallel processing task. This is radically different from an N-back task that simply engages updating of WM contents. Previous research on executive functions has indicated that these mechanisms—shifting and updating—are clearly separable (Miyake & Friedman, 2012; Miyake et al., 2000). Also, distinctions between the two tasks can be made within existing models of WM. According to Cowan's embedded processes WM model, a prominent and widely acknowledged state-based model of WM (Cowan, 1999; D'Esposito & Postle, 2015), performance on the N-back task merely requires constant updating of the focus of attention, while successful execution of a complex span task also depends on efficient retrieval of items from activated long-term memory.

In accordance with these conceptual considerations, several empirical investigations comparing the two tasks in healthy controls have demonstrated no relationship between them (Jaeggi, Buschkuehl, Perrig, & Meier, 2010; Kane, Conway, Miura, & Colflesh, 2007) while others have indicated a significant relationship between tasks (Schmiedek, Hildebrandt, Lövdén, Wilhelm, & Lindenberger, 2009). Surprisingly no studies to date have directly investigated the relationship between these two types of tasks in individuals with aphasia, a population in which these tasks are often used. Friedmann and Gvion (2003) administered both tasks to six individuals with aphasia, but were only able to perform qualitative analysis of the WM data, and could not establish a clear correspondence between performance on the two types of WM tasks and between performance on N-back and processing of different sentence types. DeDe et al. (2014) presented both tasks to 12 individuals with different types and severity of aphasia, but did not report the correlations between their WM tasks. However, analysis of the raw data reported in their paper indicated no relationship between listening span (an adaptation of a complex span task for individuals with aphasia) and 1-back (r_s (n = 12) = −.243, p = .472) and 2-back tasks (r_s (n = 12) = −.109, p = .751).

In other studies, performance on the complex span task in aphasia studies has repeatedly been related to performance on standardised language tests (Caspari et al., 1998; Ivanova et al., 2015; Ivanova & Hallowell, 2014; Sung et al., 2009). Similar correlations have not been found for N-back tasks (Christensen & Wright, 2010; Wright et al., 2007). Only in the Mayer and Murray (2012) study did the authors obtain a significant correlation between non-nameable (face) N-back task performance and aphasia severity, while the other two verbal N-back tasks were not found to be related to standardised language scores. The authors acknowledge that these results require further investigation and explanation. Thus, it appears that all published studies using complex span tasks demonstrated a relationship between WM and standardised language test performance, while N-back tasks show an almost complete lack of significant correlations with language measures. This discrepancy in reported associations with language measures between the two WM tasks cannot be simply accounted by somewhat less frequent use of N-back tasks in published studies. Additionally, a recent study (Ivanova et al., 2015) suggests that the relationship between WM and language comprehension might depend on the type of aphasia. WM seems to exert more influence on language comprehension abilities of individuals with non-fluent aphasia, suggesting that general cognitive mechanisms might be underlying their language comprehension deficits (Ardila, 2010). These suppositions also require further exploration.

Taken together the earlier findings suggest that while complex span and N-back tasks have been used interchangeably in previous studies of WM in aphasia and similar conclusions regarding WM capacity have been drawn from them, they may, in fact, index different underlying processes. Thus, the relationship between these two WM tasks needs to be explored further, specifically in the aphasia population. The principal aim of the present study was to directly investigate the relationship between these two most commonly used WM tasks in aphasia research—an adapted version of a listening span task and a 2-back task—in a large sample of individuals with different types of aphasia, and to relate performance on these tasks to measures of language comprehension. Such a comparison has not been achieved before in a single study with large numbers of participants receiving both tests. In addition, the large sample allowed us to explore variations in performance patterns between individuals with non-fluent and fluent aphasia that could indicate differences in underlying mechanisms of language comprehension impairment.

In order to make performance on the two tasks comparable, we chose specific versions of these tasks with maximally similar perceptual processing requirements: modified listening span tasks with words and an auditory word 2-back task. In accordance with the literature reviewed earlier, we anticipated little or no relationship between performance on the complex span and N-back tasks, similar to what has been previously demonstrated in healthy controls (Jaeggi et al., 2010; Kane et al., 2007) and in individuals with aphasia (DeDe et al., 2014). Such a result would support the suggestion that the two tasks involve different cognitive mechanisms (Miyake et al., 2000). In addition, we expected that performance on the complex span task, but not the N-back task, would be related to a standardised language comprehension measure, as only performance on complex span tasks has been consistently related to language abilities (Caspari et al., 1998; Ivanova & Hallowell, 2014; Sung et al., 2009). Finally, based on our previous research in this area (Ivanova et al., 2015) we expected that performance on the complex span task would only be significantly related to language comprehension in individuals with non-fluent type of aphasia.

Method

Participants

Individuals with aphasia following stroke were recruited at the Center for Speech Pathology and Neurorehabilitation in Moscow, Russia. All participants were right-handed and native speakers of Russian. None of the participants had diagnosed neurodegenerative disorders, epilepsy, other psychiatric disorders, such as depression (as diagnosed by a certified psychiatrist), or history of alcohol or drug abuse. This information was collected from medical health histories. All participants passed vision and hearing screening prior to administration of experimental tasks. Forty-four individuals participated in the study (21 male, 23 female; M_{age} = 53.39 years, SD = 9.38, age range: 33–73 years). Participants had education levels ranging from completing secondary school to a university degree ($M_{years\ of\ education}$ = 13.32, *SD* = 2.11, years of education range: 10–15 years). They had various types and severity of aphasia resulting from single or multiple left-hemisphere strokes involving perisylvian areas and underlying white matter, the latest being no earlier than 2 months prior to testing ($M_{post\text{-}onset}$ = 29.43, *SD* = 27.99 months, post-onset range: 2–111 months).

Each person with aphasia was examined by a speech-language pathologist and a neuropsychologist of the Center (the second author), and their language deficit was classified according to Luria's system (Luria, 1980). Efferent motor aphasia (most similar to Broca's aphasia) is distinguished by perseverations of syllables and words, non-fluent, effortful and agrammatic language output accompanied by relatively spared comprehension abilities. Dynamic aphasia (with features of Broca's aphasia and possibly Transcortical Motor aphasia) is characterised by disrupted utterance planning and thus difficulties in producing grammatically correct sentences and coherent connected speech. Sensory aphasia (corresponds closely to Wernicke's aphasia) is differentiated by fluent speech with abundant phonological and semantic paraphasias complemented by pervasive difficulties in phoneme perception and selection. Acoustic-mnestic aphasia (somewhat analogous to Anomic aphasia) represents a continuation of sensory aphasia at the word level and is primarily characterised by word retrieval difficulties and reduction of auditory verbal memory. It should be noted that within Luria's approach an individual can be classified as having several types of aphasia simultaneously, for instance, efferent motor aphasia and dynamic aphasia. While there is no accepted one-to-one correspondence between Luria's classification and the western multidimensional approach (see Akhutina, 2015 for more on this), the general distinction made between individuals with non-fluent and fluent aphasias is shared and accepted within both approaches (Ardila, 2010; Ivanova et al., 2015).

For the purposes of this paper, individuals with efferent motor and/or dynamic aphasia were grouped into a "non-fluent" group, while individuals with sensory and/or acoustic-mnestic aphasia were included in a "fluent" group. Note that the assignment of the general categories "non-fluent" and "fluent" was not based on the performance on a single fluency task, but on the qualitative division between two major aphasia syndromes—non-fluent and fluent—suggested by Benson and Ardila (1996) and Ardila (2010). Only individuals who were unanimously classified as having exclusively non-fluent or fluent types of aphasia by both the speech-language pathologist and the neuropsychologist were included in the respective groups. Those individuals for whom specific aphasia types could not be unambiguously determined were classified as having mixed aphasia. Accordingly, aphasia subtypes were distributed among patients as follows: 27 non-fluent, 13 fluent, 4 mixed. Aphasia severity, as indexed by the overall score on the Russian standardised aphasia battery—Assessment of Speech in Aphasia test (ASA; Tsvetkova, Akhutina, & Pylaeva, 1981; see Language assessment section for more information on this test)—ranged from mild to moderate ($M_{\text{aphasia severity}}$ = 81.33 ± 13.61%). There were no significant differences in age, years of education or overall aphasia severity between individuals with non-fluent (n = 27; M_{age} = 53.56 ± 8.74 years; $M_{\text{years of education}}$ = 12.96 ± 2.26 years; $M_{\text{aphasia severity}}$ = 82.66 ± 12.65%) and fluent aphasia (n = 13; M_{age} = 54 ± 11.6 years; $M_{\text{years of education}}$ = 13.69 ± 1.93 years; $M_{\text{aphasia severity}}$ = 76.87 ± 16.5%). Individual participant data are presented in Appendix.

Language assessment

As mentioned earlier, all participants were administered the ASA test (Tsvetkova et al., 1981). This is a traditional Russian language battery for aphasia that includes production and comprehension subtests and a rating of conversational speech (see Ivanova et al. (2015) for a detailed description of the test). The overall score on the test in clinical practice is routinely used as a measure for aphasia severity and not for purposes of classification. The comprehension subtest examines single-word auditory comprehension (matching single nouns (n_{items} = 30) and verbs (n_{items} = 30) to pictures), sentence comprehension (matching sentences of varying complexity to pictures (n_{items} = 15)), following commands (performing manipulations with objects following oral instructions (n_{items} = 10)) and question comprehension in a dialogue (understanding a series of basic everyday questions (n_{items} = 10)). For each subtest a maximum score of 30 can be given. The different comprehension tasks included within the comprehension subtest of the ASA test were designed to be considered in combination. In the current study, the comprehension subtest average score was thus used as the dependent variable for language comprehension. Interested readers can find scores on individual comprehension tasks for each participant in Appendix.

Working memory assessment

Modified listening span (MLS) task

This is a simplified version of the complex span task adapted specifically for individuals with aphasia. In this task, participants have to listen to spoken sentences and match them to the target picture in an array of four pictures and simultaneously remember a word presented aurally for subsequent recognition. The theoretical and methodological rationale for this task and its features are described in more detail in our previous publication (see Ivanova & Hallowell, 2014). In the current study, we used a Russian version of the task that is different from the original English version in that: (a) only the condition with short and simple sentences was used, as the initial study verified its sufficient difficulty for individuals with aphasia to engage in effortful processing and yet maintain an acceptable level of performance on the processing component of the task; (b) high-frequency disyllabic words were used for to-be-remembered items as there is a shortage of high-frequency monosyllabic words in Russian; (c) three sets instead of one for each set-size were presented to participants to increase reliability of assessment. As in the original version, sets ranging in size from two to six items were presented to participants in ascending order. At the end of each set participants had to recall the to-be-remembered words by pointing to matching pictures amongst distractors in a visual array. This format of the processing and recall components minimised the influence of

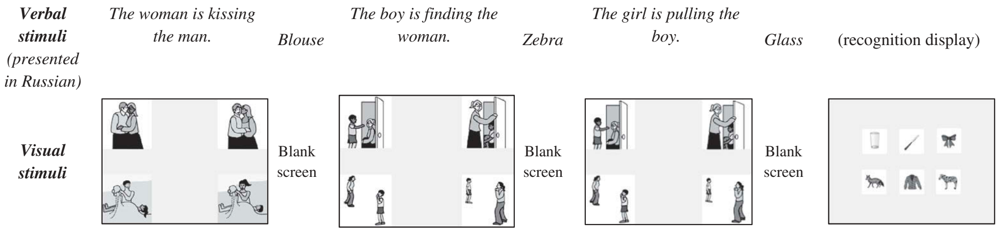

Figure 1. Example of a set from the modified listening span task (set size three). All words for recognition are disyllabic in Russian.

expressive language and motor speech deficits on performance. See Figure 1 for an example of a set from the Russian version of the MLS task. Performance on the task was scored based on a partial credit unit scoring scheme of the recall component (Conway et al., 2005). Storage items (separately presented words) were scored as a proportion of correctly recognised elements per set; for the final MLS score a mean of these proportions was calculated. Thus, in calculating the MLS score, only performance on the recall component of the task was taken into account.

2-back task

In this task, participants were aurally presented with a continuous string of words and were instructed to judge whether a word matched a previous one they had heard 2 items before. The 2-back task contained 20 different stimuli consisting of high-frequency disyllabic concrete words similar but not identical to the words used for to-be-remembered items in the MLS task. The auditory presentation of stimuli and the use of words with comparable psycholinguistic characteristics allowed us to make the perceptual processing requirements between the two WM tasks considerably similar. Altogether, 150 items containing 36 targets (24%) were presented. Items were presented in 3 blocks of 50 items each with an interstimulus interval of 2 seconds. The participants responded with their non-dominant hand by pressing the spacebar on a keyboard to indicate that the current item was the same as the one presented two stimuli before. The percentage of targets and the length of the task were selected to be similar to existing tasks in the literature (Christensen & Wright, 2010; Mayer & Murray, 2012; Wright et al., 2007). Performance on the task was indexed via d′ (d-prime), as it does not depend on a person-specific response criterion (Christensen & Wright, 2010; Lachman, Lachman, & Butterfield, 1979). D′ is computed by subtracting the Z-scored false positive rate from the Z-scored hit rate.

Individual scores for both WM tasks are presented in Appendix.

Results

Descriptive statistics of WM measures for the aphasia group overall, as well as non-fluent, fluent and mixed subgroups are presented in Table 1. To provide an indication of overall accuracy of performance on the 2-back task, descriptive statistics for hit rate and false positive rate are also provided in Table 1.

Table 1. Descriptive statistics for working memory and language tasks.

	Participants with aphasia (n = 44)		Participants with non-fluent aphasia (n = 27)		Participants with fluent aphasia (n = 13)		Participants with mixed aphasia (n = 4)	
Tasks	*M* (SD)	Range	*M* (SD)	Range	*M* (SD)	Range	*M* (SD)	Range
MLS task (proportion of correctly recognised elements per set)	.76 (.14)	.34–.90	.76 (.12)	.39–.90	.73 (.17)	.34–.90	.82 (.08)	.71–.90
2-back (d-prime)	0 (1.4)	−4.04–1.49	.14 (1.4)	−4.04–1.49	−.42 (1.6)	−3.19–1.49	.4 (.34)	−.03–.74
2-back (hit rate)	.90 (.14)	.36–1.00	.93 (.07)	.72–1.00	.82 (.21)	.36–1.00	.94 (.06)	.86–1.00
2-back (false positive rate)	.03 (.04)	.00–.17	.03 (.04)	.00–.17	.02 (.02)	.00–.06	.02 (.01)	.02–.04
ASA comprehension (% correct)	88.8 (12.8)	39–100	91.3 (9.6)	61–100	82.1 (17.6)	39–97	94.1 (3.8)	90–99

Table 2. Correlations between working memory and language variables.

	Spearman correlation
Individuals with different types of aphasia (n = 44)	
• MLS with 2-back	.444** (p = .003)
• MLS with ASA comprehension	.439** (p = .003)
• 2-back with ASA comprehension	.216 (p = .159)
Individuals with non-fluent aphasia (n = 27)	
• MLS with 2-back	.458* (p = .016)
• MLS with ASA comprehension	.625** (p < .001)
• 2-back with ASA comprehension	.151 (p = .454)
Individuals with fluent aphasia (n = 13)	
• MLS with 2-back	.569* (p = .042)
• MLS with ASA comprehension	.270 (p = .373)
• 2-back with ASA comprehension	.192 (p = .529)

* $p < .05$; ** $p < .01$.

As the WM and language scores overall and within each subtype (non-fluent and fluent) of aphasia were not distributed normally due to negatively skewed distribution of scores, non-parametric tests were used. Spearman correlations between different WM and language variables are presented in Table 2. As somewhat expected, only a moderate correlation was observed between the two WM tasks for the aphasia group overall and within the non-fluent and fluent subgroups. There were too few participants with "mixed" aphasia to perform separate analyses within this group. Furthermore, each WM task related differently to language processing, with MLS showing a significant relationship to language comprehension abilities and the 2-back showing no significant relationship.

We also compared the pattern of performance of individuals with different general types of aphasia. Individuals with fluent aphasia performed slightly worse on language comprehension measures compared to individuals with non-fluent aphasia (Z = 2.02, p = .043). There were no significant differences between groups on either of the WM tasks (MLS: Z = 0.14, p = .885; 2-back: Z = 1.06, p = .291). However, when we investigated the relationship between language comprehension and WM in fluent and non-fluent aphasia, a different pattern of performance was observed for the two groups (see also Table 2), with the MLS being related to language comprehension only in the non-fluent group. Performance on the 2-back task was not related to language comprehension scores for either of the groups.

Discussion

The present study aimed to investigate the relationship between performance on two routine WM tasks: a modified version of a complex span task and an N-back task—in individuals with various types of aphasia. Previous studies have yielded conflicting results as to whether these tasks index similar underlying cognitive processes, with the majority of studies of healthy controls showing a minimal relationship between the two tasks (Jaeggi et al., 2010; Kane et al., 2007). However, despite lack of evidence for their similarity, the two tasks have been used interchangeably in aphasia research and the relationship between them has never before been directly explored in individuals with aphasia.

Our investigation of the relationship between the MLS and 2-back tasks when performed by individuals with aphasia demonstrated that the two tasks were moderately related. Further, the relationship between these two tasks remained significant within two different subtypes of aphasia: fluent and non-fluent. While significant, the observed association accounted for only 20–32% variance, suggesting that the two tasks largely target different underlying processes. Tasks that index the same underlying cognitive abilities are predicted to share substantially more common variance. This is consistent with the cognitive literature on executive functions that states that these tasks target different components or submechanisms of executive abilities. Complex span tasks involve rehearsal, shifting flexibly between tasks or mental sets, while N-back tasks primarily rely on updating, that is rapid addition/deletion of WM contents (Conway et al., 2005; Miyake & Friedman, 2012). Previous research investigating these aspects of executive functions indicates that the processes of shifting and updating, while moderately related, are clearly separable (Miyake et al., 2000). Another perspective on the construct of WM is that only tasks that require both active maintenance of presented materials and controlled cue-dependent search of memory when maintenance becomes impeded are indicative of WM abilities (Unsworth & Engle, 2007). Indeed, complex span tasks engage these processes to a greater extent compared to N-back tasks that rely primarily on recognition mechanisms. In addition, the two tasks target distinctive cognitive processes according to Cowan's embedded processes WM model (Cowan, 1999), with N-back involving only retrieval from the central focus of attention and complex span depending on processing of items in both the focus of attention and activated long-term memory.

Still, it should be acknowledged that in our study the two tasks were not completely unrelated to each other. This is possibly because in the current modification of the complex span task the recall component was changed to a recognition task to accommodate potential motor speech deficits of individuals with aphasia. Yet, a traditional complex span task necessitates recall, when participants depend only on their own cues to retrieve items from temporary storage, while N-back tasks principally require recognition to discern items from foils. Additionally, previous studies in healthy controls employed a letter N-back task (Jaeggi et al., 2010; Kane et al., 2007). It is plausible to assume that a visual letter N-back task possess somewhat different processing requirements compared to the auditory word N-back task used in this study to maximise perceptual similarity to the complex span task. Thus, the partial overlap in task demands in the current study may have led to the observed moderate relationship in performance on the two tasks purportedly taxing different aspects of WM. Here we would like to address the issue of generalisability of our results to other versions of these WM tasks. Mayer and Murray (2012) previously demonstrated that performance accuracy on verbal and nonverbal modifications of 2-back tasks in a small sample of individuals with aphasia were highly inter-correlated. Thus, we hypothesise that while visual and nonverbal versions of N-back tasks might show a slightly weaker relationship to verbal complex span tasks due to differing perceptual processing requirements, the overall pattern of results and correspondence to language measures for various types of N-back tasks will be similar. As for generalisation of findings to other types of complex span tasks, it is also reasonable to expect that similar patterns will hold for other verbal, and potentially even nonverbal, modifications of complex span tasks. Previous studies of young healthy controls demonstrated large and significant intercorrelations between various types of verbal and nonverbal complex span tasks (Engle, Tuholski, Laughlin, & Conway, 1999; Kane et al., 2004). Indeed, these suppositions require further exploration and experimental validation. Future investigation will need to determine to what extent patterns observed in individuals without neurological and cognitive impairments hold for the aphasia population.

Additionally, the observed relationship between the two tasks, that supposedly index different aspects of WM or executive abilities, might be due to the fact that brain injury rarely affects only one cognitive system. It is very probable that neural networks supporting various aspects of executive functions are damaged synchronously, particularly in cases of middle cerebral artery stroke. This is especially plausible given that functional neuroimaging studies do show overlapping regions of activation during execution of complex span and N-back tasks (Chatham et al., 2011; Chein et al., 2011; Owen, McMillan, Laird, & Bullmore, 2005; Rottschy et al., 2012). So while in healthy controls the two processes of cue-dependent search and recognition (or switching versus updating) can and do dissociate, in persons with brain injuries these processes (while still separable) can often be damaged concurrently, leading to a detectable relationship between tasks that index these submechanisms. Future lesion studies will help to further clarify the neural underpinnings of the cognitive processes involved in execution of these tasks.

However, what further points to the conceptual differences in the two tasks is their strikingly different relationship to language comprehension. The complex span task was moderately related to a standardised language comprehension measure, while the N-back task was completely unrelated. These findings are in line with prior research, as earlier studies in aphasia also consistently failed to demonstrate a significant relationship between performance on N-back tasks and measures of language comprehension (Christensen & Wright, 2010; Wright et al., 2007), while with various adapted and simplified versions of the complex span task, this relationship was consistently demonstrated (Caspari et al., 1998; Ivanova & Hallowell, 2014; Sung et al., 2009). Possibly performance on the MLS task was related to language processing because of its more complex cognitive nature compared to N-back tasks and because the submechanism of shifting in combination with verbal rehearsal is more critical for language comprehension than updating verbal information. If the current findings are considered within Cowan's embedded processes model (Cowan, 1999), then performance on the N-back task requires constant updating of the capacity-limited focus of attention, while successful execution of a complex span task also necessitates retrieval of items from activated long-term memory that are susceptible to interference and temporal decay. It is evident that language comprehension also engages both levels of the system: focus of attention and activated memory. Thus, complex span tasks capture a wider range of cognitive mechanisms relevant for language processing.

In the current study, the MLS task performance was moderately related to language comprehension in the overall sample and strongly related in the group of non-fluent aphasia, while for individuals with fluent aphasia this relationship was not significant. The differential relationship between this task and language comprehension measures for individuals with non-fluent versus fluent aphasia mirrors our earlier findings in a different sample of individuals with aphasia and with a slightly different WM task (Ivanova et al., 2015). As we have argued in our previous work, this distinct pattern of performance in fluent and non-fluent aphasia, especially given comparable severity, is intriguing. We hypothesise that it points to different underlying mechanisms of impairment (Akhutina, 2015; Ardila, 2010). In non-fluent aphasia, potential cognitive limitations underlie their language difficulties in sequencing linguistic elements, leading to stronger relationships between cognitive tasks and language, while in fluent aphasia perhaps fundamental lexical-semantic language impairment leads to difficulties in performance of even basic language-mediated cognitive tasks. In other words, word level comprehension deficits may transcend the cognitive impairments that individuals with fluent aphasia are experiencing. While the smaller sample size of the fluent group could potentially account for the lack of a significant relationship in this group, we did have .78 power to detect a relationship of similar magnitude to the one detected in the non-fluent group.

One common criticism of listening span tasks is that they overlap in their processing requirements much more closely with the auditory comprehension measures (i.e., both involve listening to sentences and pointing to a pictures) than N-back measures (Wright & Fergadiotis, 2012). This leads to speculation that the observed relationship between listening span and the comprehension measure reflects language demands of both tasks, rather than WM per se (MacDonald & Christiansen, 2002). We believe this is not the case in the present study for the following reasons. First, in calculating the MLS storage score, only performance on the recall component of the task was taken into account (i.e., accuracy of matching sentences to pictures did not contribute to the final score) and still a significant relationship with language measures was observed. Second, if the similarity between tasks rather than common cognitive mechanisms is behind the relationship, than a similar pattern should have been observed for all types of aphasia. However, we did not observe a relationship between recall performance on the MLS task and language comprehension in individuals with fluent aphasia. Taken together this leads us to conclude that the MLS task is not simply another language task, but a cognitively demanding task that taps into specific WM mechanisms that are critical for successful language comprehension (see also Just & Varma, 2002 for conceptual arguments in favour of this position).

The present study demonstrated that although N-back tasks seem similar to traditional complex span measures and also index abilities related to executive functions, they tap into different underlying cognitive faculties. This was demonstrated through a merely moderate intercorrelation and a differential relationship to language comprehension; performance on the MLS task provided a much stronger predictor of auditory language comprehension than the 2-back task. From an important practical standpoint, this means that these tasks are not interchangeable as measures of WM, as the tasks seem to tap into different aspects of this process. Other authors have raised concerns with the construct validity of N-back tasks and their unstable psychometric properties (DeDe et al., 2014; Jaeggi et al., 2010). Thus, we advise that the two tasks—adapted versions of the complex span and N-back tasks—should not be used interchangeably in aphasia research or clinical practice, as findings with these different WM tasks were not the same and cannot be generalised without further inquiry into their underlying differences and potential commonalities. Similarly, findings with other WM tasks, such as forward and backward span, that are more commonly used in clinical practice, should be carefully evaluated and probed for concurrent validity, prior to broadly interpreting results obtained with them as suggestive of WM limitations. In all, the current study highlights the importance of carefully considering task demands and existing empirical evidence supporting validity of a chosen task when assessing cognitive abilities of individuals with aphasia.

Acknowledgements

We extend gratitude to Prof. Victor Shklovsky and the staff of the Center for Speech Pathology and Neurorehabilitation for assistance with participant recruitment. We thank Ekaterina Iskra, Yulia Akinina and Anastasia Kobzeva for their help with data collection and transcription. We are sincerely grateful to all individuals with aphasia who took part in the study.

Disclosure statement

No potential conflict of interest was reported by the authors.

Funding

The article was prepared within the framework of the Basic Research Program at the National Research University Higher School of Economics (HSE) and supported within the framework of a subsidy granted to the HSE by the Government of the Russian Federation for the implementation of the Global Competitiveness Program. This project was also supported by a Research Career Scientist Award to the senior author by the Department of Veterans Affairs Clinical Science Research and Development Program. The contents do not represent the views of the US Department of Veterans Affairs or the United States Government.

References

Akhutina, T. (2015). Luria's classification of aphasias and its theoretical basis. *Aphasiology*, 1–20. doi:10.1080/02687038.2015.1070950

Ardila, A. (2010). A proposed reinterpretation and reclassification of aphasic syndromes. *Aphasiology*, *24*, 363–394. doi:10.1080/02687030802553704

Benson, D. F., & Ardila, A. (1996). *Aphasia: A clinical perspective*. New York, NY: Oxford University Press.

Berthier, M. L., Dávila, G., Green-Heredia, C., Moreno Torres, I., Juárez Y Ruiz De Mier, R., De-Torres, I., & Ruiz-Cruces, R. (2014). Massed sentence repetition training can augment and speed up recovery of speech production deficits in patients with chronic conduction aphasia receiving donepezil treatment. *Aphasiology*, *28*, 188–218. doi:10.1080/02687038.2013.861057

Caspari, I., Parkinson, S. R., LaPointe, L. L., & Katz, R. C. (1998). Working memory and aphasia. *Brain and Cognition*, *37*, 205–223. doi:10.1006/brcg.1997.0970

Chatham, C. H., Herd, S. A., Brant, A. M., Hazy, T. E., Miyake, A., O'Reilly, R., & Friedman, N. P. (2011). From an executive network to executive control: A computational model of the n-back task. *Journal of Cognitive Neuroscience*, *23*, 3598–3619. doi:10.1162/jocn_a_00047

Chein, J. M., Moore, A. B., & Conway, A. R. A. (2011). Domain-general mechanisms of complex working memory span. *NeuroImage*, *54*, 550–559. doi:10.1016/j.neuroimage.2010.07.067

Christensen, S. C., & Wright, H. H. (2010). Verbal and non-verbal working memory in aphasia: What three n-back tasks reveal. *Aphasiology*, *24*, 752–762. doi:10.1080/02687030903437690

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

Cowan, N. (1999). An embedded-process model of working memory. In A. Miyake & P. Shah (Eds.), *Models of working memory. Mechanisms of active maintenance and executive control* (pp. 62–101). Cambridge: Cambridge University Press.

Daneman, M., & Merikle, P. M. (1996). Working memory and language comprehension: A meta-analysis. *Psychonomic Bulletin & Review*, *3*, 422–433. doi:10.3758/BF03214546

DeDe, G., Ricca, M., Knilans, J., & Trubl, B. (2014). Construct validity and reliability of working memory tasks for people with aphasia for people with aphasia. *Aphasiology*, *28*, 692–712. doi:10.1080/02687038.2014.895973

D'Esposito, M., & Postle, B. R. (2015). The cognitive neuroscience of working memory. *Annual Review of Psychology*, *66*, 115–142. doi:10.1146/annurev-psych-010814-015031

Engle, R. W., Tuholski, S. W., Laughlin, J. E., & Conway, A. R. (1999). Working memory, short-term memory, and general fluid intelligence: A latent-variable approach. *Journal of Experimental Psychology. General*, *128*, 309–331. doi:10.1037/0096-3445.128.3.309

Francis, D., Clark, N., & Humphreys, G. (2003). The treatment of an auditory working memory deficit and the implications for sentence comprehension abilities in mild "receptive" aphasia. *Aphasiology*, *17*, 723–750. doi:10.1080/02687030344000201

Friedmann, N., & Gvion, A. (2003). Sentence comprehension and working memory limitation in aphasia: A dissociation between semantic-syntactic and phonological reactivation. *Brain and Language*, *86*, 23–39. doi:10.1016/S0093-934X(02)00530-8

Helm-Estabrooks, N., Connor, L. T., & Albert, M. L. (2000). Treating attention to improve auditory comprehension in aphasia. *Brain and Language*, *74*, 469–472.

Ivanova, M. V., Dragoy, O. V., Kuptsova, S. V., Ulicheva, A. S., & Laurinavichyute, A. K. (2015). The contribution of working memory to language comprehension: Differential effect of aphasia type. *Aphasiology*, *29*, 645–664. doi:10.1080/02687038.2014.975182

Ivanova, M. V., & Hallowell, B. (2014). A new modified listening span task to enhance validity of working memory assessment for people with and without aphasia. *Journal of Communication Disorders*, *52*, 78–98. doi:10.1016/j.jcomdis.2014.06.001

Jaeggi, S. M., Buschkuehl, M., Perrig, W. J., & Meier, B. (2010). The concurrent validity of the N-back task as a working memory measure. *Memory*, *18*, 394–412. doi:10.1080/09658211003702171

Just, M. A., & Carpenter, P. A. (1992). A capacity theory of comprehension: Individual differences in working memory. *Psychological Review*, *99*, 122–149. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/1546114

Just, M. A., & Varma, S. (2002). A hybrid architecture for working memory: Reply to MacDonald and Christiansen (2002). *Psychological Review*, *109*, 55–65. doi:10.1037/0033-295X.109.1.55

Kane, M. J., Conway, A. R. A., Miura, T. K., & Colflesh, G. J. H. (2007). Working memory, attention control, and the N-back task: A question of construct validity. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *33*, 615–622. doi:10.1037/0278-7393.33.3.615

Kane, M. J., Hambrick, D. Z., Tuholski, S. W., Wilhelm, O., Payne, T. W., & Engle, R. W. (2004). The generality of working memory capacity: A latent-variable approach to verbal and visuospatial memory span and reasoning. *Journal of Experimental Psychology. General*, *133*, 189–217. doi:10.1037/0096-3445.133.2.189

Lachman, R., Lachman, J., & Butterfield, E. (1979). *Cognitive psychology and information processing: An introduction*. Hillsdale, NJ: Lawrence Erlbaum Associates Inc.

Laures-Gore, J., Marshall, R. S., & Verner, E. (2011). Performance of individuals with left-hemisphere stroke and aphasia and individuals with right brain damage on forward and backward digit span tasks. *Aphasiology*, *25*, 43–56. doi:10.1080/02687031003714426

Luria, A. R. (1980). *Higher cortical functions in man* (2nd ed.). New York, NY: Basic Books.

MacDonald, M. C., & Christiansen, M. H. (2002). Reassessing working memory: Comment on Just and Carpenter (1992) and Waters and Caplan (1996). *Psychological Review*, *109*, 35–54. doi:10.1037/0033-295X.109.1.35

Mayer, J. F., & Murray, L. L. (2002). Approaches to the treatment of alexia in chronic aphasia. *Aphasiology*, *16*, 727–743. doi:10.1080/02687030143000870

Mayer, J. F., & Murray, L. L. (2012). Measuring working memory deficits in aphasia. *Journal of Communication Disorders*, *45*, 325–339. doi:10.1016/j.jcomdis.2012.06.002

Miyake, A., & Friedman, N. P. (2012). The nature and organization of individual differences in executive functions: Four general conclusions. *Current Directions in Psychological Science*, *21*, 8–14. doi:10.1177/0963721411429458

Miyake, A., Friedman, N. P., Emerson, M. J., Witzki, A. H., Howerter, A., & Wager, T. D. (2000). The unity and diversity of executive functions and their contributions to complex "Frontal Lobe" tasks: A latent variable analysis. *Cognitive Psychology*, *41*, 49–100. doi:10.1006/cogp.1999.0734

Owen, A. M., McMillan, K. M., Laird, A. R., & Bullmore, E. (2005). N-back working memory paradigm: A meta-analysis of normative functional neuroimaging studies. *Human Brain Mapping*, *25*, 46–59. doi:10.1002/hbm.20131

Ricker, T. J., AuBuchon, A. M., & Cowan, N. (2010). Working memory. *Wiley Interdisciplinary Reviews: Cognitive Science*, *1*, 573–585. doi:10.1002/wcs.50

Rottschy, C., Langner, R., Dogan, I., Reetz, K., Laird, A. R., Schulz, J. B., & Eickhoff, S. B. (2012). Modelling neural correlates of working memory: A coordinate-based meta-analysis. *NeuroImage*, *60*, 830–846. doi:10.1016/j.neuroimage.2011.11.050

Salis, C. (2012). Short-term memory treatment: Patterns of learning and generalisation to sentence comprehension in a person with aphasia. *Neuropsychological Rehabilitation*, *22*, 428–448. doi:10.1080/09602011.2012.656460

Salis, C., Kelly, H., & Code, C. (2015). Assessment and treatment of short-term and working memory impairments in stroke aphasia: A practical tutorial. *International Journal of Language & Communication Disorders*, *50*, 721–736. doi:10.1111/1460-6984.12172

Schmiedek, F., Hildebrandt, A., Lövdén, M., Wilhelm, O., & Lindenberger, U. (2009). Complex span versus updating tasks of working memory: The gap is not that deep. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *35*, 1089–1096. doi:10.1037/a0015730

Sreenivasan, K. K., Curtis, C. E., & D'Esposito, M. (2014). Revisiting the role of persistent neural activity during working memory. *Trends in Cognitive Sciences*, *18*, 82–89. doi:10.1016/j.tics.2013.12.001

Sung, J. E., McNeil, M. R., Pratt, S. R., Dickey, M. W., Hula, W. D., Szuminsky, N. J., & Doyle, P. J. (2009). Verbal working memory and its relationship to sentence-level reading and listening comprehension in persons with aphasia. *Aphasiology*, *23*, 1040–1052. doi:10.1080/02687030802592884

Tompkins, C. A., Bloise, C. G., Timko, M. L., & Baumgaertner, A. (1994). Working memory and inference revision in brain-damaged and normally aging adults. *Journal of Speech and Hearing Research*, *37*, 896–912.

Tsvetkova, L., Akhutina, T., & Pylaeva, N. (1981). *Kolichestvennaya otsenka rechi u bol'nykh s aphasiej*. Moscow: Izdatelstvo Moskovskogo Gosudarstvennogo Universiteta.

Unsworth, N., & Engle, R. W. (2007). The nature of individual differences in working memory capacity: Active maintenance in primary memory and controlled search from secondary memory. *Psychological Review*, *114*, 104–132. doi:10.1037/0033-295X.114.1.104

Waters, G. S., & Caplan, D. (2003). The reliability and stability of verbal working memory measures. *Behavior Research Methods, Instruments, & Computers*, *35*, 550–564. doi:10.3758/BF03195534

Wright, H. H., Downey, R. A., Gravier, M., Love, T., & Shapiro, L. P. (2007). Processing distinct linguistic information types in working memory in aphasia. *Aphasiology*, *21*, 802–813. doi:10.1080/02687030701192414

Wright, H. H., & Fergadiotis, G. (2012). Conceptualising and measuring working memory and its relationship to aphasia. *Aphasiology*, *26*, 258–278. doi:10.1080/02687038.2011.604304

Appendix. Individual characteristics of participants with aphasia

No	Age	Sex	Years of education	Months post-onset	Type of aphasia	MLS storage score	MLS processing score	2-back d-prime	2-back hit rate	2-back false positive rate	ASA overall score	ASA comprehension subtest score	ASA comprehension tasks: Dialogue comprehension	Object comprehension	Action comprehension	Sentence comprehension	Instruction comprehension
1	55	f	15	100	non-fluent	.68	.92	−0.23	.83	.02	269.5	138.5	30	26	24.5	28	30
2	47	f	12	19	fluent	.85	.95	1.49	1	.00	251	135	27	30	25.5	27	25.5
3	70	f	15	16	fluent	.44	.77	−2.11	.61	.03	181.5	99.5	30	18	11	21	19.5
4	40	f	15	14	fluent	.84	.98	−0.64	.78	.02	288.5	143.5	30	30	28	27	28.5
5	49	m	15	85	non-fluent	.68	.88	−1.43	.81	.05	240.5	125.5	30	29	26.5	22	18
6	61	f	15	63	non-fluent	.86	.97	1.24	1	.01	267.5	145	30	30	27.5	29	28.5
7	62	f	15	2	fluent	.59	.93	0.54	.97	.03	254	145	30	29	29.5	28	28.5
8	48	m	10	2	non-fluent	.83	.97	−3.25	.86	.13	273	150	30	30	30	30	30
9	50	m	13	59	non-fluent	.9	.98	1.28	.97	.00	270	149	30	30	30	29	30
10	57	f	13	23	non-fluent	.83	1	0.83	.94	.01	291.5	150	30	30	30	30	30
11	56	f	13	10	non-fluent	.9	.97	0.83	.94	.01	265.5	147.5	30	28.5	29.5	30	29.5
12	50	m	10	5	fluent	.89	.95	1.03	.97	.01	269	142	30	30	30	22	30
13	39	f	13	48	non-fluent	.9	.98	0.99	1	.02	285	143	30	29	28.5	27	28.5
14	48	m	15	49	non-fluent	.89	.95	1.49	1	.00	225.5	146.5	30	28.5	29	29	30
15	59	f	15	36	fluent	.85	.88	−3.1	.61	.06	283	143	30	27.5	26.5	29	30
16	67	f	10	16	non-fluent	.39	.73	0.54	.97	.03	237	122	30	26	18	18	30
17	66	f	10	36	non-fluent	.79	.92	−0.99	1	.09	278	147.5	30	29.5	29	29	30
18	36	m	15	12	fluent	.79	.97	0.25	1	.04	252.5	138.5	30	27.5	24	27	30
19	50	m	15	51	mixed	.87	1	0.29	.97	.04	236	135	30	26.5	24.5	24	30
20	63	f	10	17	non-fluent	.8	1	0.26	.83	.00	271.5	146	30	29.5	27.5	29	30
21	50	m	15	5	non-fluent	.76	1	1.24	1	.01	265	141.5	30	28.5	27	26	30
22	68	f	15	4	fluent	.7	.78	−0.57	.89	.04	242.5	128.5	30	29	22.5	23	24
23	54	f	13	12	mixed	.9	.97	−0.03	.86	.02	264	138.5	28.5	29.5	27.5	26	27
24	50	m	13	29	non-fluent	.69	.97	−4.04	.89	.17	232	130	19	30	25	26	30
25	59	m	15	17	non-fluent	.76	.95	0.88	.92	.00	253	137.5	30	29.5	24	27	27
26	48	m	10	4	non-fluent	.85	.65	−0.12	.92	.04	126.5	91.5	24	20.5	14	15	18
27	54	f	10	65	non-fluent	.74	.98	−1.44	.97	.10	259	142.5	30	26.5	27	29	30
28	41	f	15	28	mixed	.71	.98	0.74	1	.03	264.5	143	30	30	24.5	30	28.5
29	65	f	15	9	non-fluent	.86	.95	1.49	1	.00	262	141	30	30	28	23	30
30	67	m	15	12	non-fluent	.68	.95	0.63	.92	.01	209	121.5	30	24.5	21	20	26
31	42	f	15	14	non-fluent	.83	.95	0.54	.97	.03	297	147	30	30	30	27	30
32	51	f	15	52	non-fluent	.72	.73	0.25	1	.04	183	125	30	27	21	23	24
33	64	m	10	96	non-fluent	.75	.98	0.49	1	.04	252	140.5	30	29	26.5	28	27
34	56	m	15	10	mixed	.8	.98	0.58	.94	.02	278	148	30	30	30	28	30

(*Continued*)

(Continued).

No	Age	Sex	Years of education	Months post-onset	Type of aphasia	MLS storage score	MLS processing score	2-back d-prime	2-back hit rate	2-back false positive rate	ASA overall score	ASA comprehension subtest score	ASA comprehension tasks: Dialogue comprehension	Object comprehension	Action comprehension	Sentence comprehension	Instruction comprehension
35	51	m	15	37	fluent	.82	.92	0.99	1	.02	216.5	129.5	28.5	26	21.5	28	25.5
36	54	m	10	6	non-fluent	.52	.93	1.49	1	.00	217	131.5	28.5	28	28	26	21
37	73	f	15	4	fluent	.87	.93	0.63	.92	.01	141.5	59	22.5	7	6.5	14	9
38	50	m	13	26	fluent	.83	.97	0.83	.94	.01	249.5	136	30	28.5	25	27	25.5
39	52	m	10	13	fluent	.34	.62	−3.19	.36	.00	134	86.5	22.5	17	12	20	15
40	44	f	15	9	non-fluent	.8	.93	0.67	.89	.00	231	142	30	30	28	27	27
41	33	f	15	26	non-fluent	.75	1	−0.12	.92	.04	278.5	150	30	30	30	30	30
42	56	m	15	111	non-fluent	.88	.97	0.88	.92	.00	267	142	30	29	26	27	30
43	44	m	13	19	fluent	.72	.73	−1.57	.58	.00	234.5	114	30	24	19.5	18	22.5
44	50	m	10	24	non-fluent	.54	.93	−0.55	.72	.00	189	103.5	27	23.5	11	18	24