



Phonological working memory and FOXP2

Katrin Schulze^{a,b,*}, Faraneh Vargha-Khadem^{a,c}, Mortimer Mishkin^d

^a UCL Great Ormond Street Institute of Child Health, 30 Guilford Street, London WC1N 1EH, UK

^b Clinical Psychology and Psychotherapy Unit, Department of Psychology, Heidelberg University, Hauptstraße 47-51, 69117 Heidelberg, Germany

^c Great Ormond Street Hospital for Children NHS Foundation Trust, London, UK

^d Laboratory of Neuropsychology, National Institute of Mental Health, 49 Convent Drive, Bethesda, MD 20892, USA



ARTICLE INFO

Keywords:

Working memory
Phonological memory
FOXP2
Speech impairment
KE family

ABSTRACT

The discovery and description of the affected members of the KE family (aKE) initiated research on how genes enable the unique human trait of speech and language. Many aspects of this genetic influence on speech-related cognitive mechanisms are still elusive, e.g. if and how cognitive processes not directly involved in speech production are affected. In the current study we investigated the effect of the FOXP2 mutation on Working Memory (WM).

Half the members of the multigenerational KE family have an inherited speech-language disorder, characterised as a verbal and orofacial dyspraxia caused by a mutation of the FOXP2 gene. The core phenotype of the affected KE members (aKE) is a deficiency in repeating words, especially complex non-words, and in coordinating oromotor sequences generally. Execution of oromotor sequences and repetition of phonological sequences both require WM, but to date the aKE's memory ability in this domain has not been examined in detail.

To do so we used a test series based on the Baddeley and Hitch WM model, which posits that the central executive (CE), important for planning and manipulating information, works in conjunction with two modality-specific components: The phonological loop (PL), specialized for processing speech-based information; and the visuospatial sketchpad (VSSP), dedicated to processing visual and spatial information. We compared WM performance related to CE, PL, and VSSP function in five aKE and 15 healthy controls (including three unaffected members of the KE family who do not have the FOXP2 mutation).

The aKE scored significantly below this control group on the PL component, but not on the VSSP or CE components. Further, the aKE were impaired relative to the controls not only in motor (i.e. articulatory) output but also on the recognition-based PL subtest (word-list matching), which does not require speech production. These results suggest that the aKE's impaired phonological WM may be due to a defect in subvocal rehearsal of speech-based material, and that this defect may be due in turn to compromised speech-based representations.

1. Introduction

Half the members of the KE family suffer from an inherited disorder that severely impairs their speech and language function (Vargha-Khadem et al., 2005). This impairment is caused by the mutation of one copy of the FOXP2 gene and is inherited in an autosomal dominant pattern (Lai et al., 2001; Vargha-Khadem et al., 2005). The unaffected members of the family (uKE) have inherited the intact copy of the gene and do not suffer from any speech and language impairments.

The affected members of the KE family (aKE) were shown to be so impaired phonologically that every affected member's performance on a word and nonword repetition test containing complex articulation patterns discriminated them successfully from every unaffected member (uKE; Alcock et al., 2000; Watkins et al., 2002a). Furthermore,

group comparison of the aKE and the uKE demonstrated that the former were impaired on expressive language tasks, including naming, verbal fluency, past tense production, non-word reading, and spelling (Watkins et al., 2002a). In addition, the aKE were deficient not only in the domains of speech and language, but also in the execution of oromotor sequences that do not involve speech (Watkins et al., 2002a), indicating that their core problem is an orofacial dyspraxia that interferes most strikingly with speech.

Many tasks in which the aKE are deficient, such as word and non-word repetition, verbal fluency, and mimicking orofacial movement sequences, involve working memory (WM). For example, the impairment of the aKE becomes more pronounced if they are asked to imitate (i) multisyllabic compared to monosyllabic consonant-vowel combinations and (ii) parallel and sequential orofacial movements compared to

* Corresponding author at: Clinical Psychology and Psychotherapy Unit, Department of Psychology, Heidelberg University, Hauptstraße 47-51, 69117 Heidelberg, Germany.
E-mail addresses: katrin.schulze@psychologie.uni-heidelberg.de (K. Schulze), f.vargha-khadem@ucl.ac.uk (F. Vargha-Khadem), mishkinm@mail.nih.gov (M. Mishkin).

single orofacial movements (Vargha-Khadem et al., 2005), both of which are characterised by an increased WM load. This could indicate that problems with WM contribute to the aKEs' core deficit, a possibility supported by neuroimaging findings: Brain structures that differ anatomically or functionally in the aKE compared to the uKE and normal controls, such as Broca's area, premotor cortex, and the cerebellum (Belton et al., 2003; Liegeois et al., 2011; Vargha-Khadem et al., 1998; Watkins et al., 2002b), are known to be critical for speech production (Hickok, 2012) as well as for verbal WM (Baddeley, 2003; Koelsch et al., 2009; Schulze et al., 2011).

WM is a limited capacity system that enables the temporary maintenance and manipulation of units of information. The influential WM model of Baddeley and Hitch (1974), which has been supported by many research studies, assumes an attentional control system, the *Central Executive (CE)*, that operates in conjunction with two slave systems, namely, the *Visuospatial Sketchpad (VSSP)*, which stores both visual and spatial information, and the *Phonological Loop (PL)*, which processes speech-based material. It has been suggested that verbal material in WM is maintained and rehearsed in a manner comparable to that used in subvocal speech (Baddeley, 2012; Schulze and Koelsch, 2012).

This hypothesis that WM problems contribute to the aKEs' core deficit is supported by findings in patients with apraxia of speech (AOS), an acquired impairment due to damage of neural structures "responsible for planning and programming motor movements for speech" (American Speech-Language-Hearing Association; ASHA, 2007), resulting in, among other symptoms, a decreased speech rate and phoneme distortions. Patients with acquired apraxia or dyspraxia of speech may also display an abnormally low WM span for verbal material (Ortiz and Martins, 2010; Rochon et al., 1990; Waters et al., 1991, 1992). These findings indicate that speech apraxia is associated with an impairment of phonological WM.

The above findings, together with the evidence that a core deficit in the aKE is a developmental verbal and orofacial dyspraxia, raise the question of whether the aKE are impaired in verbal WM. To examine this question, we compared the performance of the aKE with that of the controls (both uKE and normal controls) on each of the three WM components: CE, VSSP, and PL.

2. Material and methods

2.1. Participants

Five aKE (two female, mean age: 38.40 yrs., SD = 12.1 yrs.) and 15 controls (seven female, mean age: 36.47 yrs., SD = 9.13 yrs.) took part in this study; three of the controls were uKE. Unless otherwise specified, the term "controls" in this paper refers to this group of 15 participants, including three uKE. Of the aKE, one female was from the second generation of the KE family, and the four other participants were from the third generation. The aKE and controls did not differ either in age ($t_{[18]} = 0.380, p = 0.709$) or in gender ($\chi^2_{[1]} = 0.067; p = 0.795$). The study was approved by the NHS Research Ethics Service Committee London, Bloomsbury, and informed consent was obtained from all the participants.

2.2. Verbal and non-verbal IQ

Participants were given the Wechsler Abbreviated Scale of Intelligence (WASI™), except for three of the controls, who received the Wechsler Adult Intelligence Scale-third edition (WAIS-III). Results are reported using Standard Scores. The Vocabulary and Similarities subtests were administered to assess verbal performance [(Standard Scores Vocabulary subtest + Standard Scores Similarities subtest)/2], and the Block Design and Matrix Reasoning subtests were used to evaluate performance in the non-verbal domain [(Standard Scores Matrix subtest + Standard Scores Block subtest)/2].

2.3. Working memory

WM was investigated using the Working Memory Test Battery for Children (WMBTC; Pickering and Gathercole, 2001) based on the Baddeley and Hitch WM model (Baddeley, 2003, 2012; Baddeley and Hitch, 1974). In the WM test battery, the PL was analysed with digit recall, word-list matching, word-list recall, and nonword-list recall. For digit recall, the participants were instructed to repeat sequences of digits in the order spoken by the experimenter (forward span range: 1–9). For the word-list matching task, the participants were presented with two sequences of words and were asked to indicate whether the words in the second sequence were in exactly the same order as the words in the first sequence (same/different judgement, span range: 2–9). For word-list recall, the participants were presented with sequences of words and asked to recall them immediately and in the same order that the experimenter had spoken them (forward span range: 1–7). The nonword-list recall task had the same requirements as the word-list recall task (forward span range: 1–7).

The VSSP was investigated using block recall and maze memory. For the block-recall task, participants were asked to repeat a sequence that the experimenter tapped out on the blocks (forward span range: 1–9). For the maze-memory task, participants were presented with two-dimensional mazes and asked to memorise the routes traced by the experimenter from the center of the maze to its outside boundary. Successive mazes increased in size and, consequently, in difficulty (span range: 2–8).

The CE was investigated using listening recall, counting recall, and recall of digits backwards. Thus, these tasks required the participants not only to copy what they had seen or heard but to manipulate and reorder the information as well. During listening recall, participants were presented with sentences, and they were asked to indicate after each one whether the statement was true or false (e.g., "Pineapples play football"); after each set of sentences, they were also asked to recall the last word of each sentence (span range: 1–6). For the counting test, participants were asked to count aloud the dots presented on each of several pages; once all of the pages had been presented, the participants were asked to recall the number of dots counted on each page (span range: 1–7). The backwards digit recall task required participants to recall sequences of digits in the reverse of the order in which they had been presented (backward span range: 2–7).

Before each subtest, participants practiced the task using several examples, starting with a span of 1 or 2 and increasing, usually, to a span of 3. The subtest began with the highest span the participant had achieved in this practice session. The participants were then given up to six trials at each span, but if they performed four trials correctly (as defined in the manual, this counted as six successfully completed trials), the experimenter moved to the next span level.

The aKE's nonword test results are not included here, because, unlike their performance on the other tasks, the aKE had already shown severe problems with the examples due to their verbal and orofacial dyspraxia. In the nonword examples, 3 trials (span range: 1–3) were presented. The aKE performed only 47% of these examples correctly, four out of the five aKE having been unable to repeat the rest (average span: 1.4). In contrast, the controls performed 90% of the examples correctly (average span: 2.7). On recall of the nonword list, the aKE performed an average of 8.80 trials (SD, 5.45), whereas the controls performed an average of 14.87 trials (SD, 2.23).

2.4. Statistical analysis

Standard Scores were used to compare the two groups' performance on both the WASI and the WAIS-III. To analyse the WM performance in the WMBTC, we used two measures: (i) the raw scores, i.e., how many trials participants completed correctly for each subtest (referred to as WM correct responses); and (ii) the span scores, i.e. the highest span for which participants performed at least four trials correctly on each

subtest (referred to as WM span). Both scores were z-transformed [individual z-score = (individual score – control mean)/control SD]. Only these z-scores were analysed statistically to measure WM performance.

Statistical analyses were conducted with version 23 of the IBM SPSS Statistics software package. The nonparametric Mann-Whitney (*U*) test was performed if the assumptions of normality for a *t*-test, as tested with the Shapiro-Wilk test, were not fulfilled, or when only the aKE and uKE were compared. For the mixed ANCOVA, Mauchly's Tests indicated that the assumption of Sphericity had not been violated ($p < 0.05$), and the Shapiro-Wilk test indicated that the data (WM correct responses for PL, CE, VSSP) were normally distributed. The homogeneity of variance was analysed with the Levene's test. If significant, the corrected F-ratio is reported (Welch's *F*). The α -levels ($p = 0.05$) were Bonferroni-corrected to account for multiple comparisons.

Performance on the nonverbal subtests of the WAIS-III/WASI (averaged Standard Scores) was included as a covariate in the ANCOVA to examine the possibility that an overall speech-unrelated impairment might be influencing the observed WM performance.

3. Results

3.1. WASI/WAIS-III subtests

The results of the WASI/WAIS-III subtests are listed in Table 1. To compare the performance between the groups, univariate ANOVAs were conducted with the standard scores in the Vocabulary, Similarities, Block Design, and Matrix Reasoning subtests as the dependent variables. Compared to the control group ($n = 15$), the aKE were significantly impaired in the Vocabulary subtest (Welch-Test; $F_{[1,4.56]} = 19.616, p = 0.008$) and in the Similarities subtest ($F_{[1,18]} = 11.275, p = 0.004$). No difference was observed for either the Block Design subtest ($F_{[1,18]} = 0.029, p = 0.867$) or the Matrix Reasoning subtest ($F_{[1,18]} = 0.365, p = 0.553$) (corrected alpha-level = 0.0125). The aKE performed more poorly than the controls in the verbal subtests ($F_{[1,18]} = 27.995, p < 0.001$), but not in the nonverbal subtests ($F_{[1,18]} = 0.158, p = 0.696$).

When comparing the standard scores for Vocabulary, Similarities, Block Design, and Matrix Reasoning subtests between the aKE and the uKE, no significant difference was observed ($ps > 0.099$), presumably due to the small sample size.

Watkins et al. (2002a) had reported the following results on the WAIS subtests, using Scaled Scores: Vocabulary: uKE 7.4 (0.8) and aKE 5.0 (1.8); Similarities: uKE 9.4 (2.6) and aKE 5.9 (2.4); and Block Design: uKE 9.7 (3.4) and aKE 9.1 (2.7). The authors had included thirteen aKE in their study (age range: 9–75 years). Using Scaled Scores, we obtained the following results in the present study: Vocabulary: uKE 9.7 (1.2) and aKE 4.2 (3.6); Similarities: uKE 10.3 (1.5) and aKE 7.4 (2.6); Block Design: uKE 12.0 (2.6) and aKE 11.4 (1.8), and Matrix Reasoning: uKE 10.7 (1.5) and aKE 11.4 (0.9).

3.2. Components of working memory

3.2.1. Comparison between aKE and controls ($n = 15$, including 3 uKE)

3.2.1.1. *WM correct responses.* To analyse the z-scores of the raw scores, a mixed ANCOVA was calculated with Group as a between-subjects

Table 1
Mean (and Standard Deviation) for the Standard Scores on these subtests of the WASI or the WAIS-III.

	Control group (including uKE), $n = 15$	uKE, $n = 3$	aKE, $n = 5$
Vocabulary	106.7 (7.9)	98.3 (6.1)	71.0 (17.4)
Similarities	105.0 (10.1)	102.3 (7.8)	86.4 (12.8)
Block	109.0 (17.4)	110.7 (14.6)	107.6 (9.0)
Matrix	110.7 (11.5)	103.7 (6.1)	107.4 (5.2)

factor (aKE, controls), WM (PL, VSSP, and CE) as a within-subjects factor, and performance on the nonverbal subtests of the WAIS-III/WASI (averaged Standard Scores) included as a covariate. There was a significant main effect of Group ($F_{[1,17]} = 8.410, p = 0.010$) due to the poorer performance of the aKE compared to that of the controls. We also observed a significant interaction between Group and WM ($F_{[2,34]} = 10.417, p < 0.001$). Univariate ANCOVAs indicated that this interaction was caused by the selective impairment of the aKE compared to the controls on the PL component ($F_{[1,17]} = 22.928, p < 0.001$), but not on the VSSP ($F_{[1,17]} = 0.152, p = 0.702$) or the CE ($F_{[1,17]} = 2.419, p = 0.138$) (corrected alpha-level = 0.017). The interaction between WM and the averaged performance on the two nonverbal subtests entered as a covariate was not significant ($F_{[2,34]} = 2.018, p = 0.149$).

As indicated in Table 2, there were no group differences (corrected alpha-level = 0.006) on either the VSSP-related subtests (Block Recall, Maze Memory) or the CE-related subtests (Listening Recall, Counting Recall or Backwards Digit Recall; the latter fell short of significance after the alpha-level correction). On the PL-related subtests, however, the controls scored significantly higher than the aKE on Digit Recall and Word-List matching, and their scores fell just short of significance on Word List Recall (see Figs. 1 and 2).

3.2.1.2. *WM span.* Similar results were obtained for the WM span scores (see Table 2). There were no group differences (corrected alpha-level = 0.017) for the VSSP or CE spans, but the groups did differ on the PL span, where the aKE performed more poorly than the controls. Again, examination of the individual subtests indicated that the groups did not differ on either the VSSP-related subtests (Block Recall, Maze Memory) or on the CE-related subtests (Listening Recall, Counting Recall, and Backwards Digit Recall). However, the aKE performed significantly more poorly than the controls on two of the three PL-related subtests (Digit Recall and Word-List Matching).

3.2.2. Comparison between aKE and uKE ($n = 3$)

3.2.2.1. *WM correct responses.* In addition, WM correct responses were compared between unaffected and affected KE family members (uKE; $n = 3$ vs. aKE; $n = 5$). Although results comparing these small sample sizes did not reach significance when alpha-levels were corrected, the pattern is similar to the above reported results. The uKE demonstrated a better performance for PL at an uncorrected p-value ($U = 0, z = 2.249, p = 0.036$), but no group differences emerged for VSSP or CE ($ps > 0.250$). Again, examination of the individual subtests indicated that groups did not differ on either the VSSP-related subtests or on the CE-related subtests ($ps > 0.250$), but the aKE performed more poorly at an uncorrected p-value than the uKE on two of the three PL-related subtests [Digit Recall ($U = 1, z = 1.950, p = 0.071$); Word-List Matching ($U = 0, z = 2.277, p = 0.036$); and Word-List Recall ($U = 0, z = 2.263, p = 0.036$)].

3.2.2.2. *WM span.* Very similar results were obtained when the WM span scores were compared between aKE and uKE. We observed no differences between the groups for CE and VSSP related tasks ($ps > 0.250$), but the aKE showed a trend for impaired performance for PL ($U = 0, z = 2.277, p = 0.036$). This was reflected in impaired performance for Word-List Matching at an uncorrected p-value ($U = 0, z = 2.366, p = 0.036$) and a tendency towards a worse performance for Digit Recall ($U = 1.5, z = 1.919, p = 0.071$). Identical to our previous results, no difference was observed for the Word-List Recall ($p = 0.143$) between the aKE and uKE.

4. Discussion

We used a test (WMBTC; Pickering et al., 2001) based on the Baddeley and Hitch WM model that allowed us to analyse the different components of WM separately: The central executive (CE), the

Table 2

Comparison between controls (n = 15) and aKE (n = 5). Nonparametric tests [Mann-Whitney (U)] were performed if the assumption of normality for a t-test was not fulfilled. Significant results are indicated in bold font (¹corrected alpha-level = 0.017; ²corrected alpha-level = 0.006).

WM components/tests	WM correct responses	WM span
Phonological loop (PL)¹	$t_{[18]} = 4.964, p < 0.001$	$U = 1.50, z = 3.161, p < 0.001$
Digit Recall ²	$U = 3.00, z = 3.023, p = 0.001$	$U = 6.50, z = 2.799, p = 0.004$
Word-List Matching ²	$U = 0.50, z = 3.237, p < 0.001$	$U = 1.50, z = 3.223, p < 0.001$
Word-List Recall ²	$t_{[18]} = 1.928, p = 0.070$	$U = 15.00, z = 2.276, p = 0.053$
Visuospatial Sketchpad (VSSP)¹	$t_{[18]} = 0.550, p = 0.589$	$U = 32.50, z = 0.449, p = 0.672$
Block Recall ²	$t_{[18]} = 0.123, p = 0.903$	$U = 36.50, z = 0.093, p = 0.933$
Maze Memory ²	$t_{[18]} = 0.709, p = 0.488$	$U = 33.00, z = 0.410, p = 0.735$
Central Executive (CE)¹	$t_{[18]} = 1.627, p = 0.121$	$U = 16.00, z = 1.913, p = 0.066$
Listening Recall ²	$t_{[18]} = 1.449, p = 0.165$	$U = 17.00, z = 1.985, p = 0.081$
Counting Recall ²	$t_{[18]} = 0.109, p = 0.914$	$U = 36.50, z = 0.103, p = 0.933$
Backwards Digit Recall ²	$t_{[17.31]} = 2.634, p = 0.017$	$U = 18.00, z = 1.769, p = 0.098$

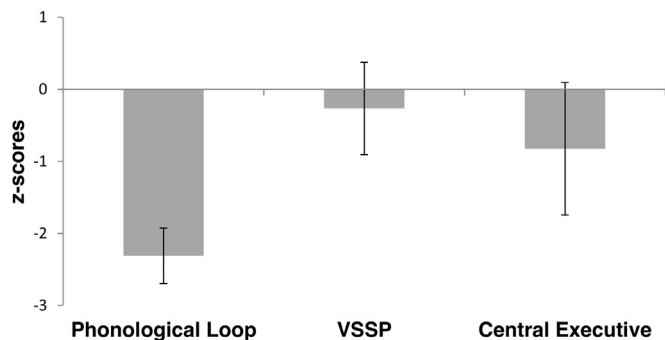


Fig. 1. Performance on the different WM components (z-scores of WM correct responses) of the aKE group compared to that of the controls (n = 15, including the uKE), negative values indicate poorer performance of the aKE.

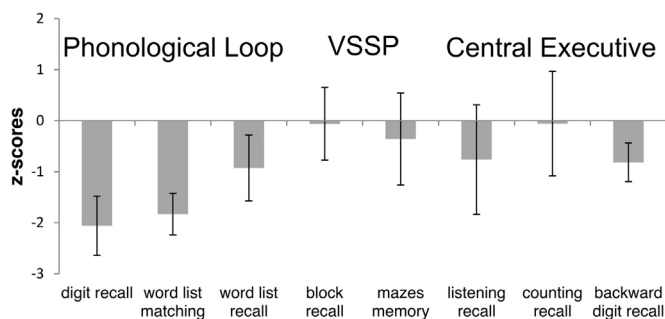


Fig. 2. Performance on the different WM subtests (z-scores of WM correct responses) of the aKE group compared to that of the controls (n = 15, including the uKE), negative values indicate poorer performance of the aKE.

visuospatial sketchpad (VSSP), and the phonological loop (PL). Compared to controls (n = 15, including the uKE), the aKE were significantly impaired only on the tasks related to PL. Importantly, the aKE were also markedly impaired in the recognition-based, word-list matching subtest of the PL, in which repetition (i.e. motor output) of the speech-based material is not required. In addition, although no significant difference was indicated between the groups for the CE (potentially due to the small sample size of the aKE), the affected members appear to show a tendency towards an inferior performance for these tasks (see Figs. 1 and 2). Some CE-related tasks involve the manipulation of verbally presented material (e.g., Listening Recall and Backwards Digit Recall). Therefore, the observed WM deficit for the PL in aKE, requiring the maintenance of verbal material, might have impaired their performance for the CE tasks as well. Our findings thus corroborate previous results indicating that the *FOXP2* mutation appears to affect predominantly oromotor and speech-related processes (Vargha-Khadem et al., 1995, 2005).

Several characteristics of the PL suggest that the codes used to store,

maintain, rehearse, and manipulate information in the PL of WM resemble subvocal speech. First, the *phonological similarity effect* (Conrad, 1964) describes the phenomenon that (i) memory span is decreased for visually presented verbal items that are acoustically similar compared to those that are acoustically dissimilar and (ii) errors in this task are typically phonological rather than visual. These findings indicate that visually presented verbal material is recoded into subvocal auditory codes in order to store it in WM. Second, *articulatory suppression* of overt or covert movement of the articulators (for example Baddeley et al., 2002, 1975; Larsen and Baddeley, 2003; Surprenant et al., 1999) disrupts maintenance and rehearsal of stored material in the articulatory loop (Baddeley, 1992, 2003). Finally, the *word-length effect*, the phenomenon that memory spans increase (Baddeley et al., 1975) and are more accurate (Baddeley et al., 2002) for lists of short words than for lists of long words, suggests that subvocal rehearsal within the PL occurs in real time.

The view that verbal WM relies on the assistance of the oromotor system is further supported by studies of patients. The American Speech-Language-Hearing Association (ASHA) defines Childhood Apraxia of Speech (CAS) as a speech disorder "in which the precision and consistency of movements underlying speech are impaired in the absence of neuromuscular deficits". The core impairment of CAS is viewed as a deficit in planning and/or programming movement sequences (ASHA, 2007). Importantly, children with CAS show impairments in non-word repetition, which are presumably due to deficits related to verbal memory processes (Shriberg et al., 2012). By contrast, apraxia of speech (AOS) is acquired later in life following damage to neural structures "responsible for planning and programming motor movements for speech" (ASHA, 2007), resulting in decreased speech rate and phoneme distortions, among other symptoms (ASHA, 2007). Regarding WM processes, patients with acquired apraxia or dyspraxia of speech have been observed to display, among others: (i) a reduced WM span for words and digits (Ortiz et al., 2010; Rochon et al., 1990; Waters et al., 1991, 1992); (ii) a lack of or reduced word-length and phonological similarity effects (Rochon et al., 1990; Waters et al., 1991, 1992); (iii) chance performance in recognition when asked to indicate whether two probes had been presented previously in the same order (Waters et al., 1991); (iv) a direct relationship between lowered articulation rate and lowered WM span (Waters et al., 1992); and (v) a direct relationship between lowered articulation rate and lowered phonological similarity and word-length effects (lower articulation rate was related to smaller effects; Waters et al., 1992). The pattern of WM deficits described here in patients with apraxia (reduced span and no word-length or phonological similarity effects) can also be observed in normal controls under the condition of articulatory suppression, i.e., when articulatory rehearsal is diminished or unavailable (Baddeley, 1992, 2012; Baddeley et al., 1975). This indicates that speech apraxia is associated with an impairment of the articulatory rehearsal process, an explanation that could also account for the PL deficits observed in the aKE in our study.

Interestingly, it seems that vocal learning – the ability to learn new sounds through imitation – may well require auditory WM. Monkeys, who appear to lack the ability to mimic auditory stimuli, and thus are not vocal learners, seem to possess, at best, a passive form of auditory short-term memory. Scott et al. (2012) presented macaques with sounds from different categories (pure tones, environmental sounds, monkey calls, etc.) and found that the animals performed extremely poorly on an auditory serial delayed match-to-sample task. They demonstrated an ‘overwriting’ effect for auditory stimuli (including monkey vocalisations) that was far greater than that shown for visual stimuli, suggesting that monkeys’ auditory retention depends on a passive form of short-term memory (STM) and not on working memory, which requires active manipulation of the stimuli. When humans are tested for the retention of auditory stimuli that they cannot mimic, they too seem to rely on passive auditory STM (McKeown et al., 2011; Mercer and McKeown, 2010).

Auditory long-term memory (LTM) in vocal learners, like auditory WM, appears to rely on the assistance of the oromotor system. Schulze et al. (2012) reported that participants could not store long-lasting representations of sounds for subsequent recognition if those sounds (e.g. reversed words, i.e., words played backwards) could not be mimicked or labeled. Importantly, failure to recognize the stimuli was not attributable to a perceptual deficit, inasmuch as different reversed words presented with an intrapair interval of 500 ms were easily discriminated from each other (Schulze et al., 2012). The authors concluded that a sound’s pronounceability, that is, the potential to activate the speech production system subvocally, was both necessary and sufficient for storing that sound in auditory LTM.

As noted earlier, the PL has been shown to rely on cortical structures that participate in oromotor control, such as Broca’s area, premotor cortex and insular cortex (e.g., Baddeley, 2003; Bamiou et al., 2003; Koelsch et al., 2009; Schulze et al., 2011), the supplementary motor area (Paulesu et al., 1993; Schulze et al., 2011), and the cerebellum (Baddeley, 2003; Koelsch et al., 2009; Schulze and Koelsch, 2012; Schulze et al., 2011). This same brain network was found to be deficient in the affected members of the KE family (aKE): Structural abnormalities were identified in both the cerebellum and the inferior frontal gyrus bilaterally (Belton et al., 2003; Vargha-Khadem et al., 1998; Watkins et al., 2002b); and a functional deficiency was documented not only in Broca’s area (Liegeois et al., 2003) and the cerebellum (Liegeois et al., 2011), but also in the premotor, supplementary motor, and primary motor cortices (Liegeois et al., 2011). This network, including the insula, is part of the dorsal stream of the dual-stream model of speech processing (Hickok, 2009; Hickok et al., 2011; Hickok and Poeppel, 2007). This model proposes that there are two routes to speech processing: A ventral pathway that operates as a lexical interface in the superior temporal gyrus, and a dorsal pathway responsible for sensory-motor integration, i.e., for mapping the speech signals onto articulatory representations in the frontal lobe. On the basis of this proposal, our findings suggest that, in the aKE, it is predominantly the dorsal stream that is functionally deficient.

Additional support for the above hypothesis comes from studies of the arcuate fasciculus, which forms part of the dorsal auditory stream (Saur et al., 2008). A study by (Lopez-Barroso et al., 2013) suggests that the direct segment of the arcuate fasciculus, which connects the posterior temporoparietal cortices with the inferior frontal gyrus, is essential for word learning: The strength of the functional connectivity between these two cortical areas (often referred to as “Wernicke’s” and “Broca’s” territories, respectively) is correlated with recognition memory in a task in which participants are asked to indicate whether a nonsense word had been presented before. The density and complexity of the human arcuate fasciculus increased dramatically during human evolution and so differs substantially from that in nonhuman primates, including apes (Petrides and Pandya, 2009; Rilling et al., 2011, 2008; Thiebaut de Schotten et al., 2012). This difference provides a possible explanation for the inability of nonhuman primates to store auditory

information in either LTM (Fritz et al., 2005) or WM (Scott et al., 2012).

The proposal that the primary function of the dorsal route is one of sensory-motor integration (Hickok, 2009; Hickok et al., 2011; Hickok et al., 2007), is in line with accumulating evidence that motor-related cortical areas, such as the premotor and primary motor cortices, are activated not only during speech production, but also during speech perception (Fadiga et al., 2002; Mottonen and Watkins, 2009; Pulvermuller et al., 2006; Watkins et al., 2003; Wilson et al., 2004). Such findings strongly suggest that speech processing and speech production share some of the same neural circuitry.

In summary, the aKE’s impairment in articulation: (i) extends to WM, presumably including the internal rehearsal of speech-based material, thereby implicating the PL component of WM, selectively; and (ii) this extension of the aKE’s speech-related difficulty appears to be due to the same structural abnormalities that cause the articulatory disorder, namely, a compromised dorsal speech-processing stream.

Funding

This work was supported by the Royal Society e-Gap Grant 2006/R1 and the Intramural Research Program of the National Institute of Mental Health, National Institutes of Health, Department of Health and Human Services. This research was also supported by the NIHR Great Ormond Street Hospital Biomedical Research Centre. The views expressed are those of the authors and not necessarily those of the NHS, the NIHR or the Department of Health.

Conflict of interest statement

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

References

- Alcock, K.J., Passingham, R.E., Watkins, K.E., Vargha-Khadem, F., 2000. Oral dyspraxia in inherited speech and language impairment and acquired dysphasia. *Brain Lang.* 75 (1), 17–33. <http://dx.doi.org/10.1006/brln.2000.2322>.
- American Speech-Language-Hearing Association. (2007). Childhood apraxia of speech [Technical Report]. Available from www.asha.org/policy.
- Baddeley, A.D., 1992. Working memory. *Science* 255 (5044), 556–559. <http://dx.doi.org/10.1126/science.1736359>.
- Baddeley, A.D., 2003. Working memory: looking back and looking forward. *Nat. Rev. Neurosci.* 4 (10), 829–839. <http://dx.doi.org/10.1038/nrn1201>.
- Baddeley, A.D., 2012. Working memory: theories, models, and controversies. *Annu. Rev. Psychol.* 63, 1–29. <http://dx.doi.org/10.1146/annurev-psych-120710-100422>.
- Baddeley, A.D., Chincotta, D., Stafford, L., Turk, D., 2002. Is the word length effect in STM entirely attributable to output delay? Evidence from serial recognition. *Q. J. Exp. Psychol.-A* 55 (2), 353–369. <http://dx.doi.org/10.1080/02724980143000523>.
- Baddeley, A.D., Hitch, G.J., 1974. Working memory. In: *Recent Advances in Learning and Motivation VIII*. Academic Press, New York, pp. 47–89.
- Baddeley, A.D., Thomson, N., Buchanan, L., 1975. Word length and the structure of short-term memory. *J. Verbal Learn. Verbal Behav.* 14 (6), 575–589. [http://dx.doi.org/10.1016/S0022-5371\(75\)80045-4](http://dx.doi.org/10.1016/S0022-5371(75)80045-4).
- Bamiou, D.E., Musiek, F.E., Luxon, L.M., 2003. The insula (Island of Reil) and its role in auditory processing. Literature review. *Brain Res. Brain Res. Rev.* 42 (2), 143–154. [http://dx.doi.org/10.1016/S0165-0173\(03\)00172-3](http://dx.doi.org/10.1016/S0165-0173(03)00172-3).
- Belton, E., Salmond, C.H., Watkins, K.E., Vargha-Khadem, F., Gadian, D.G., 2003. Bilateral brain abnormalities associated with dominantly inherited verbal and orofacial dyspraxia. *Hum. Brain Mapp.* 18 (3), 194–200. <http://dx.doi.org/10.1002/hbm.10093>.
- Conrad, R., 1964. Acoustic confusions in immediate memory. *Br. J. Psychol.* 55, 75–84. <http://dx.doi.org/10.1111/j.2044-8295.1964.tb00899.x>.
- Fadiga, L., Craighero, L., Buccino, G., Rizzolatti, G., 2002. Speech listening specifically modulates the excitability of tongue muscles: a TMS study. *Eur. J. Neurosci.* 15 (2), 399–402. <http://dx.doi.org/10.1046/j.0953-816x.2001.01874.x>.
- Fritz, J., Mishkin, M., Saunders, R.C., 2005. In search of an auditory engram. *Proc. Natl. Acad. Sci. USA* 102 (26), 9359–9364. <http://dx.doi.org/10.1073/pnas.0503998102>.
- Hickok, G., 2009. The functional neuroanatomy of language. *Phys. Life Rev.* 6 (3), 121–143. <http://dx.doi.org/10.1016/j.plrev.2009.06.001>.
- Hickok, G., 2012. Computational neuroanatomy of speech production. *Nat. Rev. Neurosci.* 13 (2), 135–145. <http://dx.doi.org/10.1038/nrn3158>.
- Hickok, G., Houde, J., Rong, F., 2011. Sensorimotor integration in speech processing: computational basis and neural organization. *Neuron* 69 (3), 407–422. <http://dx.doi.org/10.1016/j.neuron.2011.01.019>.

- Hickok, G., Poeppel, D., 2007. The cortical organization of speech processing. *Nat. Rev. Neurosci.* 8 (5), 393–402. <http://dx.doi.org/10.1038/nrn2113>.
- Koelsch, S., Schulze, K., Sammler, D., Fritz, T., Muller, K., Gruber, O., 2009. Functional architecture of verbal and tonal working memory: an fMRI study. *Hum. Brain Mapp.* 30 (3), 859–873. <http://dx.doi.org/10.1002/hbm.20550>.
- Lai, C.S.L., Fisher, S.E., Hurst, J.A., Vargha-Khadem, F., Monaco, A.P., 2001. A forkhead-domain gene is mutated in a severe speech and language disorder. *Nature* 413 (6855), 519–523. <http://dx.doi.org/10.1038/35097076>.
- Larsen, J.D., Baddeley, A.D., 2003. Disruption of verbal STM by irrelevant speech, articulatory suppression, and manual tapping: do they have a common source? *Q. J. Exp. Psychol.-A* 56 (8), 1249–1268. <http://dx.doi.org/10.1080/02724980244000765>.
- Liegeois, F., Baldeweg, T., Connelly, A., Gadian, D.G., Mishkin, M., Vargha-Khadem, F., 2003. Language fMRI abnormalities associated with FOXP2 gene mutation. *Nat. Neurosci.* 6 (11), 1230–1237. <http://dx.doi.org/10.1038/nn1138>.
- Liegeois, F., Morgan, A.T., Connelly, A., Vargha-Khadem, F., 2011. Endophenotypes of FOXP2: dysfunction within the human articulatory network. *Eur. J. Paediatr. Neurol.* 15 (4), 283–288. <http://dx.doi.org/10.1016/j.ejpn.2011.04.006>.
- Lopez-Barroso, D., Catani, M., Ripolles, P., Dell'Acqua, F., Rodriguez-Fornells, A., de Diego-Balaguer, R., 2013. Word learning is mediated by the left arcuate fasciculus. *Proc. Natl. Acad. Sci. USA* 110 (32), 13168–13173. <http://dx.doi.org/10.1073/pnas.1301696110>.
- McKeown, D., Mills, R., Mercer, T., 2011. Comparisons of complex sounds across extended retention intervals survives reading aloud. *Perception* 40 (10), 1193–1205. <http://dx.doi.org/10.1068/p6988>.
- Mercer, T., McKeown, D., 2010. Updating and feature overwriting in short-term memory for timbre. *Atten. Percept. Psychophys.* 72 (8), 2289–2303. <http://dx.doi.org/10.3758/APP.72.8.2289>.
- Mottron, R., Watkins, K.E., 2009. Motor representations of articulators contribute to categorical perception of speech sounds. *J. Neurosci.* 29 (31), 9819–9825. <http://dx.doi.org/10.1523/JNEUROSCI.6018-08.2009>.
- Ortiz, K.Z., Martins, F.C., 2010. The relationship between severity of apraxia of speech and working memory. *Dement. Neuropsychol.* 4 (1), 63–68. <http://dx.doi.org/10.1590/S1980-57642010DN40100011>.
- Paulesu, E., Frith, C.D., Frackowiak, R.S., 1993. The neural correlates of the verbal component of working memory. *Nature* 362 (6418), 342–345. <http://dx.doi.org/10.1038/362342a0>.
- Petrides, M., Pandya, D.N., 2009. Distinct parietal and temporal pathways to the homologues of Broca's area in the monkey. *PLoS Biol.* 7 (8), e1000170. <http://dx.doi.org/10.1371/journal.pbio.1000170>.
- Pickering, S.J., Gathercole, S.E., 2001. Working Memory Test Battery for Children (WMTB-C). Pearson.
- Pulvermuller, F., Huss, M., Kherif, F., Martin, F.M.D.P., Hauk, O., Shtyrov, Y., 2006. Motor cortex maps articulatory features of speech sounds. *Proc. Natl. Acad. Sci. USA* 103 (20), 7865–7870. <http://dx.doi.org/10.1073/pnas.0509989103>.
- Rilling, J.K., Glasser, M.F., Jbabdi, S., Andersson, J., Preuss, T.M., 2011. Continuity, divergence, and the evolution of brain language pathways. *Front. Evolut. Neurosci.* 3, 11. <http://dx.doi.org/10.3389/fnevo.2011.00011>.
- Rilling, J.K., Glasser, M.F., Preuss, T.M., Ma, X.Y., Zhao, T.J., Hu, X.P., Behrens, T.E.J., 2008. The evolution of the arcuate fasciculus revealed with comparative DTI. *Nat. Neurosci.* 11 (4), 426–428. <http://dx.doi.org/10.1038/Nn2072>.
- Rochon, E., Caplan, D., Waters, G.S., 1990. Short-term memory processes in patients with apraxia of speech: implications for the nature and structure of the auditory verbal short-term memory system. *J. Neurolinguist.* 5 (2/3), 237–264. [http://dx.doi.org/10.1016/0911-6044\(90\)90013-0](http://dx.doi.org/10.1016/0911-6044(90)90013-0).
- Saur, D., Kreher, B.W., Schnell, S., Kummerer, D., Kellmeyer, P., Vry, M.S., Weiller, C., 2008. Ventral and dorsal pathways for language. *Proc. Natl. Acad. Sci. USA* 105 (46), 18035–18040. <http://dx.doi.org/10.1073/pnas.0805234105>.
- Schulze, K., Koelsch, S., 2012. Working memory for speech and music. *Ann. N. Y. Acad. Sci.* 1252, 229–236. <http://dx.doi.org/10.1111/j.1749-6632.2012.06447.x>.
- Schulze, K., Vargha-Khadem, F., Mishkin, M., 2012. Test of a motor theory of long-term auditory memory. *Proc. Natl. Acad. Sci. USA*. <http://dx.doi.org/10.1073/pnas.12047171109>.
- Schulze, K., Zysset, S., Mueller, K., Friederici, A.D., Koelsch, S., 2011. Neuroarchitecture of verbal and tonal working memory in nonmusicians and musicians. *Hum. Brain Mapp.* 32, 771–783. <http://dx.doi.org/10.1002/hbm.21060>.
- Scott, B.H., Mishkin, M., Yin, P., 2012. Monkeys have a limited form of short-term memory in audition. *Proc. Natl. Acad. Sci. USA* 109 (30), 12237–12241. <http://dx.doi.org/10.1073/pnas.1209685109>.
- Shriberg, L.D., Lohmeier, H.L., Strand, E.A., Jakielski, K.J., 2012. Encoding, memory, and transcoding deficits in childhood apraxia of speech. *Clin. Linguist. Phon.* 26 (5), 445–482. <http://dx.doi.org/10.3109/02699206.2012.655841>.
- Surprenant, A.M., Neath, I., LeCompte, D.C., 1999. Irrelevant speech, phonological similarity, and presentation modality. *Memory* 7 (4), 405–420. <http://dx.doi.org/10.1080/741944920>.
- Thiebaut de Schotten, M., Dell'Acqua, F., Valabregue, R., Catani, M., 2012. Monkey to human comparative anatomy of the frontal lobe association tracts. *Cortex* 48 (1), 82–96. <http://dx.doi.org/10.1016/j.cortex.2011.10.001>.
- Vargha-Khadem, F., Gadian, D.G., Copp, A., Mishkin, M., 2005. FOXP2 and the neuroanatomy of speech and language. *Nat. Rev. Neurosci.* 6 (2), 131–138. <http://dx.doi.org/10.1038/Nrn1605>.
- Vargha-Khadem, F., Watkins, K., Alcock, K., Fletcher, P., Passingham, R., 1995. Praxic and nonverbal cognitive deficits in a large family with a genetically transmitted speech and language disorder. *Proc. Natl. Acad. Sci.* 92(3), 930–933.
- Vargha-Khadem, F., Watkins, K.E., Price, C.J., Ashburner, J., Alcock, K.J., Connelly, A., Passingham, R.E., 1998. Neural basis of an inherited speech and language disorder. *Proc. Natl. Acad. Sci. USA* 95 (21), 12695–12700. <http://dx.doi.org/10.1073/pnas.95.21.12695>.
- Waters, G.S., Caplan, D., Hildebrandt, N., 1991. On the structure of verbal short-term memory and its functional-role in sentence comprehension – evidence from neuropsychology. *Cogn. Neuropsychol.* 8 (2), 81–126. <http://dx.doi.org/10.1080/02643299108253368>.
- Waters, G.S., Rochon, E., Caplan, D., 1992. The role of high-level speech planning in rehearsal – evidence from patients with apraxia of speech. *J. Mem. Lang.* 31 (1), 54–73. [http://dx.doi.org/10.1016/0749-596X\(92\)90005-1](http://dx.doi.org/10.1016/0749-596X(92)90005-1).
- Watkins, K.E., Dronkers, N.F., Vargha-Khadem, F., 2002a. Behavioural analysis of an inherited speech and language disorder: comparison with acquired aphasia. *Brain* 125, 452–464. <http://dx.doi.org/10.1093/brain/awf058>.
- Watkins, K.E., Strafella, A.P., Paus, T., 2003. Seeing and hearing speech excites the motor system involved in speech production. *Neuropsychologia* 41 (8), 989–994. [http://dx.doi.org/10.1016/S0028-3932\(02\)00316-0](http://dx.doi.org/10.1016/S0028-3932(02)00316-0).
- Watkins, K.E., Vargha-Khadem, F., Ashburner, J., Passingham, R.E., Connelly, A., Friston, K.J., Gadian, D.G., 2002b. MRI analysis of an inherited speech and language disorder: structural brain abnormalities. *Brain* 125, 465–478. <http://dx.doi.org/10.1093/brain/awf057>.
- Wilson, S.M., Saygin, A.P., Sereno, M.I., Iacoboni, M., 2004. Listening to speech activates motor areas involved in speech production. *Nat. Neurosci.* 7 (7), 701–702. <http://dx.doi.org/10.1038/nn1263>.