**Language processing and executive functions in early treated adults with PKU**

Sara De Felice1,2, Cristina Romani1, Tarekegn Geberhiwot 2, Anita MacDonald3, and Liana Palermo1,2,4

**1. School of Life and Health Sciences, Aston University, Birmingham, UK**

**2. Queen Elizabeth Hospital, Birmingham, UK**

**3. Birmingham Children’s Hospital, Birmingham, UK**

**4.** Department of Medical and Surgical Sciences, Magna Graecia University of Catanzaro, Catanzaro, Italy

Address correspondence to:

Cristina Romani, Ph.D.

School of Life and Health Sciences

Aston University, Aston Triangle

Birmingham B4 7ET - England

e-mail C.Romani@Aston.ac.uk

Tel: 00 44 121 – 204 4081

We provide an in-depth analysis of language functions in early-treated adultswith phenylketonuria (AwPKU), an inherited metabolic disease. We tested 15-33 AwPKU and 24-32 age- and education-matched controls on tasks tapping: 1. narrative production (retelling the story of Cinderella), 2. language pragmatics (e.g. comprehension of humour, metaphors, and inferred meaning), 3. prosody discrimination 4. inhibitory control and planning in lexical selection tasks (Blocked Cyclic Naming; Hayling Sentence Completion Test). Verbal narratives were analysed with measures proven sensitive to detect impairments in other populations, including rate of Correct Information Units (CIUs), mean length of utterance (MUL), speech rate (number of words per minute), rate of errors, and coherence of narrative. AwPKU showed normal basic language processing abilities, with no deficits in lexical retrieval, phonological/articulatory processing or sentence construction even in a narrative task. Instead, impairments were seen in language tasks requiring planning and reasoning abilities. In narrative production, rate of CIU was lower than controls; in understanding of metaphors and inferred meaning, responses were slower; in the Hayling task, more errors were made in conditions requiring a focused lexical search. These results are consistent with results on an overlapping group of participants (Palermo et al., 2007). They suggest that higher-order EF deficits in PKU cannot be explained by an accumulation of lower-order deficits in basic language skills and/or a general reduction in speed of processing. More generally, they suggest EF are based on specific neurophysiological resources and cannot simply be conceived has an emergent property of lower level processing.

Key words: PKU, metabolic disorders, language, pragmatics, executive functions, speed of processing.

Phenylketonuria (PKU) is a rare inherited metabolic disease, with an occurrence of 1 in 10.000 live births on average in many European countries (Hardelid et al., 2008). It is caused by a mutation in the phenylalanine hydroxylase (*PAH)* gene on chromosome 12, which leads to a deficiency of PAH, an enzyme converting phenylalanine (Phe) into tyrosine , a precursor of dopamine (Scriver & Kaufman, 2001). Damage to PAH-activity results in increased blood Phe-levels, with toxic concentration in the brain as well as a lack of tyrosine. Untreated, PKU causes severe learning disabilities. The introduction of new-born screening in the late 1960’s together with a low phenylalanine diet has prevented severe mental disability (Blau, van Spronsen, & Levy, 2010). However, despite early intervention, some mild cognitive impairments are well-established in people with PKU, and a low-average IQ is a consistent finding (for a review see Moyle, Fox, Arthur, Bynevelt, & Burnett, 2007). Moreover, studies have often reported poor performance in tasks which rely on executive functions (EFs; for a review see Christ, Huijbregts, de Sonneville, & White, 2010; DeRoche & Welsh, 2008) and a reduction in speed of processing (for a review see Albrecht, Garbade, & Burgard, 2009).

In our previous research, we have assessed the cognitive abilities of adults with PKU (from now on AwPKU) in a variety of cognitive tasks and linked performance to dietary control indexed by blood Phe levels (Palermo et al., 2017; Romani et al., 2017; Romani, McDonald, De Felice, & Palermo, in press, this issue). Consistent with previous research, AwPKU showed impairments in higher level EFs. Thus, they were impaired in tasks involving **thinking** and **reasoning** (e.g., Tower of Hanoi, Wisconsin Card Sorting Test, Verbal Similarities from the WAIS), **planning** and **flexibility** (Tower of Hanoi, Wisconsin Card Sorting Test, Semantic Fluency) and **monitoring** (e.g., sustained attention and short-term memory; see Palermo et al., 2017). Also in align with previous research, we found that AwPKU had difficulties in tasks involving visual search and tapping visuo-spatial attention. In these latter tasks, their performance was slower than controls and differences increased with increasing difficulty of the search conditions.

In contrast to these impairments, we found that our AwPKU sample showed normal word retrieval abilities. PKU participants performed well in tasks tapping mainly lexical access, such as picture naming and spelling (both in terms of speed and accuracy). In reading words, although they were slightly slower than controls, effects of frequency and regularity were normal, with only a fixed delay across conditions. Similarly, in tasks of picture and colour naming, AwPKU showed normal effects of semantic interference (although in the context of a Stroop task, colour naming was slightly slower compared to controls). These results cannot be explained by difficulties in lexical access. Instead, performance may be slower in certain conditions either due to a greater demand on visual attention or because of a tendency to be more caution at the point of returning an answer, which would be reflected in a fixed delay across conditions (see Romani et al., in press, this issue).

In the present study, we want to explore the interplay between EFs and language abilities in more complex linguistic tasks which have received little attention in the PKU literature. Since AwPKU have significant deficits in higher order EFs, possibly mediated by damage to frontal areas, we expect to see deficits in complex language tasks such as those involving narrative production and understanding language pragmatics (Martin & McDonald, 2003). What remains unclear, however, is whether these impairments involve discreet EFs with their own neurological substrates or whether, instead, impairments are wide-spread and not specific. According to this second hypothesis, deficits are more evident or limited to executive tasks because these tasks are more complex. Complex tasks require diverse cognitive skills and minor/subtle impairments may accumulate with catastrophic consequences. Our study wants to provide evidence for these two alternative hypotheses.

If higher order EFs tap their own, independent cognitive resources, they may be impaired even when processing in simple language tasks, requiring limited interactions with other cognitive skills, is completely normal. This should be the case even when basic language skills are measured sensitively by considering RTs, assessing conditions varying for difficulty, and/or by measuring performance in the context of a complex task (narrative production). In contrast, if EF are simply an emergent property of complex tasks, impairments should be always accompanied by impairments in more basic skills when measures which are sensitive enough are used. Performance should also be impaired across complex tasks rather than impairments being restricted to certain tasks/functions.

A better understanding of the relations between different potential impairments in AwPKU would be important for a clinical management of their difficulties, but, also, to increase our understanding of the elusive nature of executive functions (for a debate see Jurado & Rosselli, 2007; Stuss, 2011). Assessing the nature of higher order impairments in populations with inherited metabolic disorders is particularly relevant. These disorders do not result in focal brain damage, but disrupt neuronal metabolisms more generally. In PKU, elevated levels of Phe are toxic for the oligodendroglia which forms the myelin sheets wrapping up axons in the brain (Anderson & Leuzzi, 2010). Therefore, one may expect impairments to occur across the board and, when they are restricted to complex tasks, this to be the consequence of an accumulation of minor impairments in different skills or, within a cognitive module, of delays in the processing operations needed to complete the task (hypothesis 1). If results are not consistent with this interpretation, this would, alternatively, suggest that the neural circuitries which support EFs are specific and particularly sensitive to a metabolic disruption (hypothesis 2). Before providing more information on our study rationale, we will briefly review the literature on language impairments in PKU.

**Review of language studies**. We have conducted a systematic review of the literature through PubMed, Web of Science, and ScienceDirect databases. Key words included *language/linguistic/verbal AND functions/abilities/skills* *AND phenylketonuria/ PKU*. Inclusion criteria were studies investigating at least one form of linguistic ability in a PKU sample (e.g. production, comprehension, spelling/reading). A total of 794 studies were identified. After removing duplicates and filtering of non-English papers, a total of 375 original-research studies were identified. A first screening from titles and abstracts left 14 relevant studies, whose text was assessed for eligibility. Four studies were discarded for not including any language tasks in addition to verbal IQ, while three further studies where identified through snow-balling. Finally, we conducted a further search using the key words *cognition/cognitive functions* *AND phenylketonuria/PKU* to check for potential relevant studies which did not include *language/linguistic/verbal* key words in titles and/or abstract. This procedure led to the identification of one further study. Hence, overall, we have identified a total of 14 studies, eleven focusing on children and/or adolescents (Azen, Koch, Friedman, Wenz, Fishier, 1996; Chang, Gray, & O'Brien, 2000; Faust, Libon & Pueschel, 1987; Fishler, Azen, Henderson, Friedman, & Koch, 1987; Gassió et al., 2005; Gejão, Ferriera, Silva, Anastácio-Pessan, & Lamônica, 2009; Melnick, Michals, & Matalon, 1981; Ozanne, Krimmer, & Murdoch, 1990; Pennington, Van Doornick, McCabe, McCabe, 1985; Soleymani, Keramati, Rohani, & Jalaei, 2015¸ Zartler & Sassaman, 1981) and three on adults (Brumm et al, 2004; Koch et al., 2002; Palermo et al., 2017). Results are summarized in the Appendix where articles are organized based on the linguistic function investigated.

Seven studies examined *expressive language/naming* and five reported deficits (Gejão et al., 2009; Melnick et al., 1981; Soleymani et al., 2015; Brumm et al., 2004; Palermo et al., 2017). In children, Melnick et al. (1981) reported half of their sample to have a linguistic delay. Gejão et al. (2009) reported poor performance on measures of language use. Also, Ozanne et al. (1990) examined two groups of children with PKU that differed in developmental level, so that they were either still in the early stages of language acquisition (group 1;<5 years old) or they were older and with expected better communication skills (group 2; > 5 years old). Although overall the PKU sample did not show significant impairment compared to controls, three participants (27%) in group 1 showed performance <75% of their chronological age, and three participants (16%) in group 2 showed performance below the 10th percentile. In adults, Palermo et al, (2017) found no impairment in picture naming in contrast to Brumm et al. (2004), but they did find impairments in the vocabulary and similarities subtests of the WASI which include a reasoning component (as other tasks investigated in the present paper). Only three studies have analysed *spontaneous speech* and they were all carried out on children (Faust et al., 1987; Melnick et al., 1981; Gassio et al., 2005). Of these, only two used quantifiable measures and they failed to show any deficit.

Seven studies examined *comprehension/receptive language.* Four did not find impairments (Faust et al., 1987; Zartler & Sassaman, 1981; Ozanne et al., 1990; Brumm et al., 2004), while three showed impartments in children on test of picture vocabulary and measurements of linguistic development (Gejão et al., 2009; Melnick et al., 1981; Soleymani et al., 2015), which were particularly severe in late-treated children (Soleymani et al., 2015).

S*pelling and reading* showed a degree of deficit in a few children studies. Seven studies compared a PKU-group’s performance with controls/normative data. Three found some spelling deficits (Faust et al., 1987; Fishler et al., 1987; Chang et al., 2000) and four found some reading deficits (Faust et al., 1987; Fishler et al., 1987; Azen et al., 1996; Palermo et al., 2017). In contrast, the two studies investigating spelling in AwPKU showed good performance (Brumm et al., 2004; Palermo et al., 2017), suggesting that a spelling deficit may reflect a developmental delay, rather than a permanent, condition-related deficit. Regarding reading, studies have generally failed to report separate results for accuracy and speed. When this has been assessed in AwPKU (Palermo et al. 2017), results have shown a reduction in speed, but good accuracy. This has been interpreted as a fixed delay in making a response combined with difficulties in serial allocation of attention which create problems especially in non-word reading and in reading long words (but normal frequency effects; see Romani et al., this issue).

Association between language performance and metabolic control (measured in terms of Phe-concentration) was investigated in a few studies only, with inconsistent results. For *naming and expressive language*, some significant correlations with accuracy measures were reported in late-treated, but not early-treated children (Soleymani et al., 2015) and not in adults (Brumm et al., 2004; Romani et al., 2017). Similarly, for *comprehension,* a high Phe-level was associated with worse performance only in late-treated groups of children (Soleymani et al., 2015)*,* while no significant association was reported in adults (Brumm et al., 2004). For *spelling and reading,* associations were reported in children by some longitudinal studies (Fishler et al., 1987; Azen et al., 1996; Koch et al., 2002), but not in adults (Brumm et al., 2004; Romani et al., 2017).

Overall, our review suggests that language functions are only mildly affected in PKU. However, studies are few and results arevariable. Moreover, there are only few studies examining adults with PKU and limited studies exploring performance beyond single word comprehension and production. Only two studies have examined narrative production in children and, to the best of our knowledge, there is no study which has explored pragmatic linguistic skills, where the wider context in which a sentence is presented must be considered to achieve full understanding.

**Study rationale**. In our study, we rely on a distinction between basic or fundamental language skills and higher-order EFs. With basic/fundamental language skills, we refer to skills (like semantic, phonological, morphological, articulatory, and syntactic processing) which manipulate specific kinds of linguistic representations and whose domain of application is restricted to the comprehension and production of language. With higher-order EFs, we refer to functions which may be deployed in language tasks, but also to a wider range of tasks (e.g., the visuo-spatial tasks). These functions control behaviour on the basis of set goals and for this purpose relay on the coordination of different skills (language, attention, visuo-spatial processing), are engaged in a less automatic way, and are supported by wider brain networks of which the frontal lobes can be the hub (see Aron, 2008; Badre & D’Esposito, 2009; Godefroy, Cabaret, Petit-Chenal, Pruvo, & Rousseaux, 1999; Luria, 1978; Shallice, 1982). These functions include planning, flexibility and monitoring. There is a large conceptual and empirical overlap between the concepts tapped EFs and reasoning or fluid intelligence (e.g., see Decker, Hill, & Dean, 2007). We will consider reasoning as part of EFs. However, for our purposes, we are not dependent on how reasoning is categorized (see Duncan, 2010, but also Friedman & Miyake, 2017). Our working hypothesis, based on previous findings, is that AwPKU may show no deficits on measures of basic language functions, but still be impaired on more complex linguistic tasks requiring reasoning and higher order EFs.

Clearly, human communication requires employing skills that exceed processing the meanings of individual words or the syntax that binds them together. Inferential abilities are required to understand jokes and metaphorical language, but also to understand the implications of a sentence in the context of the wider message in which it is embedded. These abilities involve higher-order EFs such as reasoning, monitoring and flexibility (for understanding of metaphors see Carriedo et al., 2016;for understanding humour/sarcasms see McDonald & Pearce, 1996; Uekermann, Channon, Winkel, Schlebusch, & Daum, 2007). Similarly, connected speech – of which narrative production is an example—requires reasoning and planning to organize ideas into a coherent discourse, where relevant information is delivered in the most efficient way for the aims of the speaker and in the context of the present audience (e.g., see Blain-Brière, Bouchard & Bigras, 2014). The role played by EFs, especially planning and monitoring, in narrative coherence and informativeness, has been highlighted in clinical populations (see for William Syndrome: Marini, Martelli, Gagliardi, Fabbro, & Borgatti, 2010; Schizophrenia: Marini et al., 2008; traumatic-brain injury: Carlomagno, Giannotti, Vorano, & Marini, 2011; frontotemporal dementia: Ash et al., 2006). These higher-order abilities, which allow speech to be flexibly understood or produced, are commonly described as “pragmatic skills”.

The fact that pragmatic skills, used here as a descriptive term, depend on EFs is uncontroversial (see Martin & McDonald, 2003). Pragmatic deficits are associated with syndromes which are typically characterized by EF difficulties, such as attention-deficit hyperactivity disorder (ADHD; Bishop & Baird, 2001; Geurts, Verte, Oosterlaan, Roeyers, & Sergeant, 2004) and autism spectrum disorder (Landa, 2000; Young, Diehl, Morris, Hyman, & Bennetto, 2005; Wang, Lee, Sigman, & Dapretto, 2006). Pragmatic deficits are also present in syndromes where brain-damage is diffuse and/or affects the frontal lobes considered the hub of EF, as in patients with traumatic brain injury (e.g., Angeleri et al., 2008; Bihrle, Brownell, Powelson, & Gardner, 1986; McDonald, 1993), frontotemporal dementia (e.g., Ash et al., 2006; Blair, Marczinski, Davis-Faroque, & Kertesz, 2007; Rapp & Wild, 2011) and a disruption of the dopaminergic system (for Parkinson’s disease see McNamara & Durso, 2003; Hall, Ouyang, Lonnquist, & Newcombe, 2011; for schizophrenia see Mueser, Bellack, Douglas, & Morrison, 1991; Sponheim, Surerus-Johnson, Leskela, & Diep, 2003 ) possibly due to links to the frontal lobes (Tekin & Cummings, 2002). Finally, neuroimaging studies have highlighted how pragmatics rely on frontal areas for understanding humour (Bartolo, Benuzzi, Nocetti, Baraldi, & Nichelli, 2006; Chan et al., 2013; Goel & Dolan, 2001; Moran, Wig, Adams, Janata, & Kelley, 2004), making inferences (i.e. how plausible a sentence is based on world knowledge and social context, Kuperberg et al., 2000), and comprehending and producing metaphors (see Bohrn, Altmann, & Jacobs, 2012; Benedek et al., 2014; Rapp, Leube, Erb, Grodd, & Kircher, 2004).

Given the evidence of executive deficits in PKU (see Palermo et al., 2017 and for reviews DeRoche & Welsh, 2008; Christ et al., 2010), as well the presence of diffuse brain damage, possibly more severe in frontal areas (for an association between reduced dopamine in PKU and frontal functions see Boot et al., 2017;Diamond, & Baddeley, 1996; for an association between frontal lobe, dopamine and EFs see Floresco & Magyar, 2006), we expect that AwPKU will show deficits in complex language tasks which rely on EFs. We want to investigate relationships with more basic language functions since these relationships will shed light on more specific hypotheses both on the nature of PKU and EFs. As mentioned, one possibility is that the metabolic disorder affecting individuals with PKU results in diffuse brain damage which equally affects all kinds of functions –basic language functions included. In this case, deficits will be seen even in basic language functions when conditions are taxing enough and when performance is assessed with measures which are not affected by ceiling/floor effects. Alternatively, the metabolic impairment in PKU may selectively affect EFs because these functions rely particularly on frontal areas and/or on the neuro-transmitter –dopamine—which is hypothesised to be particularly affected in this disorder (Boot et al., 2017). Results consistent with this second possibility will support a view of EFs that sees them as having a specific neurophysiological substrate and not simply as an emergent property of overall brain proficiency. A version of this second view sees EFs as a reflection of speed of processing as a property of whole, integrated brain networks.

What to predict for processing prosody is unclear. Understanding of linguistic prosody (distinguishing a statement from a question) and emotional prosody (understanding the state of mind of the speaker) can be considered basic functions. This information needs to be integrated with other linguistic and contextual information to derive the complete meaning of a sentence (e.g., see Martin & McDonald, 2003), but understanding prosody, per se, does not require reasoning and EFs in the same way as understanding metaphors and jokes does. Consistent with this tenant, Uekermann et al. (2007) showed that processing of prosody was unaffected and uncorrelated with measures of EFs in a group of individuals with depression; Tonks et al. (2008) showed that understanding emotional prosody and executive functions can dissociate in children with brain injury. Nevertheless, processing prosody relies on frontal areas. This has been shown by fMRI studies in healthy participants (Meyer, Alter, Friederici, Lohmann, & von Cramon, 2002; Wildgruber, Ackermann, Kreifelts, & Ethofer, 2006; for a meta-analysis of PET and fMRI studies see Phan, Wager, Taylor & Liberzon, 2002) and by prosody impairments in individuals with damaged frontal areas (see for primary progressive aphasia, Rohrer, Sauter, Scott, Rossor & Warren, 2010; for individuals with frontal-lobe lesions, Breitnestein, Daum, & Ackermann, 1998). If PKU affects frontal areas either selectively or as part of diffuse brain-damage both deficits in processing prosody and EF may be seen. Alternatively, if PKU affects selectively the neurophysiological substrate of higher order executive functions, deficits in processing prosody may not occur.

Finally, the hypothesis that EFs rely on specific neurophysiological substrates will be consistent with impairments where some higher order EFs are affected, while other, more specific EFs are spared. We have already presented some results indicating that inhibitory capacities are spared in AwPKU (Palermo et al., 2017). Here, we want to acquire further evidence using additional word retrieval tasks (Blocked Cyclic Naming, Healy Sentence Completion Test, see later), but also prosody tasks which require to judge prosodic information in conflict with the linguistic message of the sentence. The hypothesis that impairments in PKU are diffuse predicts impairments across the board. The hypothesis that impairments are selective for some higher-order EFs will be consistent with no particular difficulty in conditions which require inhibition.

**Plan of study**. First, we assessed performance in a narrative production task. As common practice, we used a recount of the Cinderella story (see Andreetta & Marini, 2014; Kavé, Leonard, Cupit, & Rochon, 2007; MacWhinney, Fromm, Holland, Forbes, & Wright, 2010; Stark, 2010; Thompson, 2003). Narratives analysis is an effective tool to investigate both basic linguistic skills and pragmatics (e.g., see Botting, 2002). We will assess: 1. Number of semantic, morphological, phonological and articulatory errors; 2. Rate of words produced per unit of time; 3. Sentence length, as a measure of syntactic ability; 4. Number and order of the thematic units/concepts of the story; and 5. Rate of meaningful content words out of total number of words produced (content information units or CIU, tapping informativeness). The first three measures tap basic language skills; the last two tap EFs through the production of a coherent and effective narrative (Andreetta & Marini, 2014). The hypothesis of independence between higher-order and basic linguistic functions predicts that the first three measures will be normal, while the last two will be impaired, with worst performance occurring with the informativeness score (CIUs), as the most sensitive measure of discourse planning and organization (Carlomagno et al., 2011; Marini et al., 2010).

Second, in order to further assess word retrieval abilities, we administered a picture naming task where effects of semantic interference are assessed by asking participants to repeatedly name sets of semantically related or unrelated items (Blocked Cyclic Naming). In these conditions, with semantically related sets, interference builds across repeated cycles, reducing speed of naming (see Abdel Rahman, & Melinger, 2007; Belke, Meyer, & Damian, 2005; Damian, Vigliocco, & Levelt, 2001). This task relies on lexical selection abilities but also on inhibitory EF.

Third, we examined pragmatic skills with tasks tapping understanding of 1. humour, 2. metaphors, 3. inferred meaning, and 4. prosody. The hypothesis of selective deficits to higher-order EF predicts deficits in the first three types of tasks which are strongly dependent on EFs such as reasoning, inhibition, monitoring and flexibility. Understanding prosody which is less related to EFs may be normal or deficits may be limited to tests where emotional prosody and linguistic content are in contrast with one another.

Finally, we examined performance in the Hayling Sentence Completion test which requires participants to complete a sentence with a word which is neither appropriate nor semantically related to any of the other words in the sentence. This is a complex task which requires a combination of skills: fast retrieval of words from the lexicon, inhibition of related words and, crucially, monitoring and planning an efficient lexical search so that words from areas semantically unrelated to the sentence are tapped (see Hornberger, & Bertoux, 2015, and Robinson et al., 2015). Impairments are expected since this task engages a variety of EF. Previous studies have shown an association between the Hayling test and the ability to provide cohesive and coherent verbal narratives (see Barker, Young, & Robinson, 2017). The same association will be investigated here.

**Participants**

Thirty-eight early-treated participants with classical PKU were recruited from a pool of patients currently followed by the Department of Inherited Metabolic Disorders at the University Hospitals Birmingham and who had been continuously treated with a low-phenylalanine diet since birth. Diagnosis was by new-born screening conducted at 5-7 days after birth. At the time of testing seven participants were on an unrestricted diet and 31 on a low-phenylalanine diet. We invited all early-treated patients attending the clinic to participate, plus some who were still contactable, but not attending clinic follow-up appointments. All individuals who responded to the invitation were tested. The PKU participants were compared to a group of 39 healthy control participants matched for age, gender and educational status. Healthy volunteers were recruited through the Aston University volunteering website. Slightly different subgroups carried out the verbal narratives and the pragmatic language tasks.

Table 1 shows demographic data including blood Phe levels for the whole PKU sample. Demographic information for the sub-groups of participants for each task are included with the relevant results to show age and educational matching. Data on historical blood Phe levels were obtained from the PKU database at The Clinical Chemistry Department at Birmingham Children's Hospital. Blood Phe concentrations were measured prior to each testing session to determine current levels. Table 1 shows Phe concentrations reported in three age bands: 0-10 years old, 11-16 years old and 17 years to adulthood, as well as at testing time. The Phe level in each band was determined by calculating a mean blood Phe for each year and, then, averaging the mean values for the years in the time band. Fluctuations were calculated using the Phe SD for each year and then calculating the mean SD across the years in the time band.

------------------------------------

Insert Table 1 about here

----------------------------------

The research was approved by the NHS and Aston University Ethics committees. All participants gave voluntary informed consent to take part.

**2. Experimental Investigation**

**2.1 Narrative Production: Recalling the Cinderella Story**

**Task**. Participants were asked to tell the story of Cinderella, with no time constraints. Before telling the story, each participant was shown a picture book of the story where the text was blocked out. This served as a reminder of the story, reducing differences due to poor recall. Participants were asked to look at the pictures for as long as necessary. The book was taken away when they were ready to recall the story. Audio-files were transcribed using the Express Scribe software (v 5.6, NCH Software), and then transferred into a Word file. A minimal transcription style was used. Longer pauses and hesitations were marked with suspension points. All words produced in the recording were transcribed, including initiation utterances (e.g. *“Ok, Cinderella… how I shall start…?”*) and dysfluencies (*erm, uhm, ehm*). Words which can be written either separately or as one word were transcribed as whole words (e.g. *stepsister*, instead of *step sisters*).

Performance was assessed with 5 measures

1. Errors: Errors were classified into three main categories: (i) semantic/lexical errors: They included word substitutions (e.g. sisters instead of stepsisters, his instead of her) and approximations/circumlocutions (e.g. the king’s son instead of the prince; Meteyard & Patterson, 2009); (ii) grammatical/morphological errors: they included regularisation of irregular plurals (e.g. mices), inconsistent tense forms and inappropriate use of morphemes, as well as omissions (e.g. ‘one of the provisos of being able to go to the ball $ *is that* $ you have to be back by midnight’ the information between the dollar signs was omitted; see also Meteyard & Patterson, 2009; Webster, Franklin, & Howard, 2007) (iii) phonological errors/false starts: they included phonological substitutions (e.g. ‘Cirinderella’ instead of ‘Cinderella’) and false starts (e.g. ‘Cin-Cinderella’) as well as incomplete words (e.g. ‘godmoth-… fairy godmother’). Error scores were calculated for each subcategory individually, and then an error percentage was computed over total Word Count (described below; Errors/tWC\*100).

2. Speech rate: The total Word Count (tWC) was used to compute a measure of speech rate in terms of well-formed word produced per sec.

3. Mean Length of Utterance (MLU): This is a measure of syntactic competence on the basis that more syntactic competence is necessary to organize more words into a single sentence (Marini, Andreetta, del Tin, & Carlomagno, 2011). MLU was calculated by dividing the total number of words by the number of utterances. An utterance was defined as a complete thought expressed by a connected group of words. To define utterance boundaries, we adopted mainly a semantic criteria although grammatical and acoustic criteria were also considered when in doubt (following a semantic criterion, an utterance is a conceptually homogeneous piece of information; following a grammatical criterion, it is a complete sentence which could include subordinate clauses; following an acoustic criterion, it is a group of words delimited by identifiable pauses, which could be either empty or filled by hesitations/filler words; see Marini et al., 2011).

4.Macro structure/concept organization: This assesses the global structure of a narrative by considering the quantity and quality of the main concepts produced which are necessary for a coherent discourse (Halper, & Cherney, 1998). For our Cinderella story we considered 16 main concepts (e.g. “Cinderella tries on slipper, it fits”, see Appendix 2). They were adapted from Bartels-Tobin and Hinckley (2005) to be consistent with the picture book we used. Performance was scored in terms of the accuracy of the information given (0: no or wrong information; 1: partially accurate information; 2: correct information). This provided a *Main Concepts Score (MC)* out of maximum score= 32.

5. Rate of Correct Information Units (CIU): This measures accurate, relevant and informative words out of total number of words produced or *total word count (tWC*; Nicholas & Brookshire, 1993b). Repetitions were included in the count of CIU only if they served as emphasis. Errors and incorrect information (e.g. three sisters – instead of two stepsisters) were excluded. *tWC* included all words, including uninformative repetitions, words filler (e.g. well/ something like that), expressions which were not directly related to the narrative (e.g. I think/ from what I remember/ can’t remember their names) and all error types. Hesitations and false starts were excluded from both the tWC and from the count of CIU.

**Results**. Results are shown in Table 2 and 3. Table 2 shows group performance with statistical differences from controls and N of participants with poor performance (z scores ≥ 2 SD from control mean). Table 3 shows individual scores.

-------------------------------------

Insert Tables 2 and 3 about here

---------------------------------------

As a group, AwPKU do not show differences from controls in terms of rate of errors and speech rate. This is consistent with normal performance in lexical access and subsequent stages of phonological encoding and articulatory programming. Similarly, sentence construction abilities appear normal as assessed in terms of mean utterance length. The ability to construct an efficient narrative is also equivalent to the controls, possibly because we used a non-sensitive measure. Instead, as a group, AwPKU, produced a significantly smaller proportion of correct and topically informative words (CIU) compared to controls. Six participants reported very poor scores with this measure (≥2 z scores from control group) and were significantly different from the controls using the Crawford single case analysis implemented in the program Singlims\_ES.exe (Crawford, Garthwaite, & Porter, 2010) [[1]](#footnote-1). Here, again, however, there is variability, with some PKU participants showing completely normal performance.

**2.2 Blocked Cyclic Naming**

**Task.** The *Blocked cyclic naming* (e.g., Belke 2008; Belke, Brysbaert, Meyer, & Ghyselinck, 2005; Crowther & Martin, 2014; Schnur, Schwartz, Brecher, & Hodgson, 2006) required participants to name, as soon as and as accurately as possible, sets of pictures shown repeatedly on a computer screen*.* The pictures were 72 black and white drawings from twelve semantic categories (professions, food, vegetables, body parts, fruit, animals, clothing, vehicles, insects, birds, food, appliances, and tools). The task included 24 blocks (12 blocks with sets of semantically related pictures and 12 blocks with sets of semantically unrelated pictures). Each block included 4 repetitions of the same set of 6 pictures (4 cycles). In the semantically related sets, items belonged to the same semantic category (e.g., “teacher”, “painter”, “police”, “nurse”, “chef”, “fireman”). In the unrelated sets, they recombined items from the related sets (e.g., “popcorn”, “scissors”, “eggplant”, “nurse”, “tv”, “swan”). The order of the items differed across cycles. Once the participant had named a picture, it disappeared (triggered by a voice response) and after a 100 ms. ISI, the subsequent picture was shown. Prior to the experiment, participants were familiarized with the set of pictures by presenting them in random order on the screen and providing feedback after naming. Reaction times were recorded with a voice key; errors were recorded manually and checked off-line using a tape recording.

What is expected with this task is an interaction between semantic relatedness and cycle. With unrelated cycles, performance should become progressively faster with repeated presentations (cycles) because of an accumulation of practice. With unrelated cycles, instead, performance should flatten out after the second cycle because a practice effect is counter-balanced by an increased semantic interference which builds up with progressive cycles (Abdel Rahman & Melinger, 2007; Belke, Meyer, et al., 2005; Damian et al., 2001).

**Results**. Results for the Blocked Cyclic Naming task are reported in Table 4 and Figure 1. Performance was satisfactory in terms of both accuracy and speed, as well as in terms of semantic interference. Two participants showed a particularly strong increase in semantic interference in terms of errors, but they were too few to derive any conclusion.

---------------------------------------------------

Insert Table 4 and Figure 1 about here

----------------------------------------------------

**2.3. Prosody, Pragmatics and Hayling Sentence Completion Test**

**Tasks*. Prosody.*** Prosody discrimination and comprehension was evaluated using three modified subtests from the Comprehensive Affect Test System (CATS; Froming, Levy, Schaffer, & Ekman, 2006). This battery is a sensitive measure of prosodic processing in a variety of populations (e.g. in healthy older adults, Schaffer, Wisniewski, Dahdah, & Froming, 2009; in Parkinson’s Disease, Ventura et al., 2012; in psychosis, Castagna et al., 2013, Rossell et al., 2013).

1. The *Nonemotional prosody discrimination* test requires listening to two identical sentences with non-emotional content, played one after the other. The sentences can be both declarative, both questions, or one of each. The task is to indicate whether the two sentences have the same prosody (either both declarative or interrogative) or not (being one declarative and one interrogative). There are 12 trials (6 same and 6 different).
2. The *Emotional prosody discrimination* test requires listening to a pair of sentences read by an actor with a tone conveying happiness, sadness, anger, fright or neutrality. Participants must indicate whether the voice in two sentences reflect the same or different emotions. The two sentences contain the same words and the task includes 12 trials (6 same and 6 different).
3. *Conflicting prosody—attend to prosody* test requires listening to a sentence read by an actor conveying happiness, sadness, or neutrality. The emotional tone of the sentence, however, is at odds with its linguistic content. For example, the sentence “The teacher was a monster” was said with a happy tone. The participant is instructed to ignore the meaning, but to focus on the emotional tone to decide whether the speaker is happy, in a neutral mood, or sad. The task includes 12 trials (4 Happy, 4 Neutral, 4 Sad).

***Language Pragmatics****.* Pragmatic language skills in terms of figurative language comprehension (metaphor, humour), and comprehension of inferred meaning were evaluated using three modified subtests from the Right Hemisphere Language Battery (Bryan, 1995). The name of the battery refers to the relationship between difficulties with language pragmatics and right-hemisphere damage reported in the literature (for a review see Mitchell & Crow, 2005). However, this battery provides sensitive measures of pragmatic skills in a variety of cohorts and has been used to demonstrate subtle impairments in other non-brain damaged populations (e.g. dyslexic adults Griffiths, 2007; adults with autistic spectrum disorder, Lewis, Woodyatt, & Murdoch, 2008; individuals with Asperger’s syndrome, Gunter, Ghaziuddin, & Ellis, 2002; healthy young and older adults, Zanini, Bryan, De Luca, & Bava, 2005; for a review see Rad, 2014).

The three sub-tests from the Right Hemisphere Language Battery (Bryan, 1995) were:

1. The *Metaphor Picture* testrequires reading of a sentence containing a common metaphor and then selecting the correct matching picture among four alternatives depicting the correct metaphorical meaning, the literal meaning, and selective aspects of the sentence. For example, participants have to match the sentence “Anyone who can spend £20,000 must have money to burn” with one of four pictures detecting respectively: a) a lot of money, b) a rich man with a luxury car; c) a sparking campfire, d) a man who is burning some money. The test includes 10 items plus a practice item (maximum score = 10). The task was implemented in superLab to record latency times**.**
2. The *Appreciation of Humour* test requires choosing the right ending for a joke (to make it funny) among four alternatives: the actual punchline, a straight ending of neutral content, a straight ending with emotional content and a surprise ending unrelated to the body of the joke. For example, for the story “A judge had just finished telling the prisoner that he was free to go, as the jury found him not guilty of fraud. The prisoner then asked:” there were the following four endings a) “When can I leave sir?”; b) “What about my friends?”; c) “Does that mean I can keep the money?”, d) “What time is it please?”. The test includes 10 items plus a practice item (maximum score = 10). The task was implemented in superLab to record latency times.
3. The *Comprehension of Inferred Meaning* test requires answering four specific questions about a situation described in a short paragraph. The test includes a practice paragraph plus three paragraphs followed each by four questions (maximum score = 12). For example: ‘*“How long have we been going now?” asked Irene. “Oh, just about two hours” replied Tom. “Only one more station and we’ll be there. How long since we last saw your Mother?” said Tom. “It must be about six months, I wonder if she will come to meet us?” “I shouldn’t think so, I doubt if she could walk so far now”’.*  Questions: a) How were they travelling? b) Was their destination near home?; c) Who were they visiting?; d) Did this person have a car?. The total time required to answer each question was recorded with a stopwatch.

***The Hayling Sentence Completion Test***(Burgess & Shallice, 1997). This test has two parts. Each part involves auditory presentation of 15 sentences (read by the examiner) where the last word is missing. The participant must complete the sentence. In the first part of the test, however, the sentence can be completed with the first meaningful word which comes to mind, yielding a measure of lexical retrieval speed. In the second part, instead, the sentence must be completed with a word which does not fit with the sentence and is totally unrelated to its meaning, giving a more complex measure of EFs (suppressing the expected response/planning a new lexical search). Latency times and number of correct answers were recorded for each section*.* Time was measured by the experimenter through a chronometer.In the second part, errors were coded astype-A errors (participant failed to suppress the predicted word)and type-B errors (participant responded with a word semantically related to the predicted word).

**Results**. Performance is reported in Table 5 and 6. Table 5 shows group performance together with statistical comparisons from controls and N of participants with poor scores. Table 6 shows individual z scores, as well as aggregate scores in complex tasks. This measure includes CIU scores from the verbal narrative.

-----------------------------------------

Insert Table 5 and 6 about here

-----------------------------------------

Performance in the Prosody tasks was normal. This was the case even when participants had to respond to the emotional tone of the voice and ignore the contrasting content of the sentence. These results show a good ability to interpret prosody, but also, good inhibitory control, consistent with some of our previous results (e.g., normal interference effect in the Stroop task, see Palermo et al., 2017).

Performance on complex language tasks was satisfactory for accuracy, but not speed. AwPKU needed more time than controls to comprehend metaphors and infer meaning. As in the case of rate of CIU in narrative production, some PKU participants were severely impaired, while many others performed normally. Performance on the second part of the Hayling Sentence Completion test was also impaired both in terms of accuracy and speed. It is only in the appreciation of humour subtest that performance is normal (see later for a discussion). Across the complex language tasks examined here, moreover, AwPKU demonstrated an unusual number of poor scores (>2 z scores from control mean; aggregating numbers across 5 tasks, Fisher exact probability = .047).

Deficient performance in the Hayling Test contrasts with normal interference effects in lexical access (tapped here by both Blocked Cyclic Naming and Continuous Naming, see Romani et al., this issue). This seems surprising given that all these tasks require inhibition of unwanted responses. These tasks, however, also differ in complexity and in the degree to which they require a planned search of the lexicon. The Hayling Test requires a top-down mediated search in areas unrelated to the sentence; in contrast, in picture naming, lexical search is stimulus driven (driven by the picture). Common interpretations of the Hayling Test, in fact, stress the need to generate, implement and monitor a search strategy (see Barker et al., 2017; Hornberger, & Bertoux, 2015; Robinson et al., 2015) and, so, this test has been used to assess not only inhibition, but also strategic planning in different populations (e.g., Attention Deficit Hyperactivity Disorder, see Clark, Prior & Kinsella, 2000; andParkinson’s Disease, see Bouquet, Bonnaud, & Gil, 2003). Consistent with a strong EF component, the Hayling Test is also associated with reasoning in healthy older adults (Bielak, Mansueti, Strauss, & Dixon, 2006).

As shown in Table 6, there was high variability across participants. Some individuals showed significant impairments compared to controls when assessed with the Crawford single case analysis: six PKU participants were impaired for proportion of CIU, five for Hayling time, three for Hayling errors, two for Humour time, four for Metaphor time and four for Inferred Meaning time. Overall, four PKU participants showed an overall significant impairment in the complex language tasks compared to controls (4/38 =11%) and 16% showed poor performance (1.5 SD from control means or more). On the other hand, half of the sample (19/38=50%) showed a completely normal performance, within .05 SD from the control mean. There was also variability across tasks, reflected in the fact that correlations among tasks were generally not significant, except for a significant correlation between the number of semantically-related errors in The Hayling Test and the percentage of CIU in narrative production in the PKU group (Pearson r: AwPKU= -.533 p=.02. vs. Controls= -.146 p>.05). This is consistent with the Hayling Test and the percentage of CIU both tapping higher order planning abilities (for the same finding see Barker et al., 2017).

**General Discussion**

The present study carried an in-depth investigation of language processing in a group of early-treated adults with PKU, with a focus on high-level linguistic abilities and relations with Executive Functions (EFs). First, our results confirmed our previous reports of unimpaired performance in basic language functions. Even in a complex task involving narrative production, where resource limitations may put strain on basic skills, the PKU group showed control-equivalent performance in terms of rate of different types of errors (morphological, semantic, phonological), speech rate, and mean sentence length. Normal speech rate in the narrative task was paralleled by normal performance both in terms of errors and RTs in a naming task (Blocked Cyclic Naming), which is contrary to a single/uniform speed deficit. Normal mean utterance length is also inconsistent with a generalized resource limitation, which should have produced shorter sentences. Understanding of prosody, another basic language skill, was also normal. These results indicate normal basic language skills and are consistent with our previous reports of normal performance in a continuous naming task and excellent performance in spelling (see Palermo et al., 2017). Normal understanding of the emotional tone of sentences and a normal appreciation of humour also demonstrate a normal capacity to appreciate emotions and relate to them. This is consistent with other results showing normal performance in detecting facial emotions and in demonstrating empathy (the last one tested with the Interpersonal Reactivity Index; see Palermo, MacDonald, Limback, Robertson, Boocock, Geberhiwot & Romani, in preparation). Moreover, the fact that performance is normal on tasks supposedly involving frontal functions (understanding prosody), suggest that impairments of EF in AwPKU are not caused by a generalized impairment in frontal functions.

Despite good performance on basic language functions, AwPKU showed deficits in complex language tasks. In narrative production, AwPKU showed a reduced efficacy in conveying information (reduced % CIU). CIU is a measure of communicative informativeness and efficiency (Nicholas & Brookshire, 1993) and is tightly dependent on planning (see Andreetta & Marini, 2014). In tasks involving understanding metaphorical language and inferring meanings from a short text, AwPKU were slower than controls. In the second part of the Hayling test, which requires completing a sentence with an unrelated word, thus, taxing planning and strategic abilities, AwPKU were both slower and less accurate. Because of their complexity, these tasks require a variety of cognitive abilities. However, ***reasoning*** and ***planning*** are a common denominator (for a role of planning and reasoning in connected speech see Marini et al., 2008, understanding metaphors, Amanzio, Geminiani, Leotta, & Cappa, 2008; inferring meaning, Saldert, Fors, Ströberg, & Hartelius, 2010). No deficits, however, were shown in understanding humour. It is possible that it is the emotional valence of the joke that helps AwPKU to return a quick answer.

Our previous results on a sample of AwPKU, overlapping with the one studied here, are consistent in indicating deficits in higher-order EFs involving ***reasoning*** and ***planning***, and ***flexibility*** and ***monitoring***. We have previously shown difficulties in semantic fluency (requiring planning a lexical search), the tower of Hanoi and the WCST (requiring planning, monitoring, flexibility and reasoning) and the similarities of the WAIS (requiring reasoning; see Palermo et al., 2017). In contrast, our results do not indicate difficulties ***inhibitory control,*** another EF. In the present study, AwPKU showed normal interference effects in Blocked Cyclic Naming, a task which requires repeated naming of sets of semantically related words. This task taxes lexical selection abilities (tapping lateral inhibition of lexical competitors), but may also have a top-down executive component if the set of possible targets is kept in working memory with decision mechanisms selecting the right candidate trial by trial (Crowther & Martin, 2014). Moreover, AwPKU performed normally in the conflicting prosody subtest which requires inhibiting the content of the sentence to correctly judge the emotional tone of the voice. These results are consistent with our previous results showing normal interference effects in the Stroop tasks and with results from other groups showing normal performance in a variety of tasks requiring inhibitory control in AwPKU (Brumm et al., 2004; Channon, German, Cassina, & Lee, 2004; Channon, Mockler, & Lee, 2005; Channon, Goodman, Zlotowitz, Mockler, & Lee, 2007; Feldmann, Denecke, Grenzebach, & Weglage, 2005).

The dissociations we have reported a) between basic language functions and higher order EF and b) among different types of EFs and c) with tasks assumed to involve frontal (understanding prosody), have clinical implications for our understanding of PKU, but also for better understanding of the nature of EFs and their interactions with more basic functions. Similar dissociations between basic language functions (e.g. phonology, syntax) and higher-level linguistic abilities (e.g. discourse coherence) have been reported across clinical populations (e.g. brain-damaged patients, McDonald & Pearce, 1996; Pearce, McDonald & Colheart, 1998; Shapiro & Danly, 1985; Shammi & Stuss, 1999; Autistic Spectrum Disorder, Minshew, Goldstein, & Siegel, 1995; Ozonoff & Miller, 1996; dementia, Rousseaux, Sève, Vallet, Pasquier, & Mackowiak-Cordoliani, 2010). The challenge, however, is to understand if they reflect different kinds of neurological impairments or, more simply, differences in the difficulty/sensitivity of the tasks involved.

Difficulties/delays in carrying out more complex tasks, such as planning a narrative or inferring the right meaning of a sentence from context, may be the consequence of summing minor impairments across different cognitive domains and /or of delays in individual operations needed for the task (e.g., retrieving the meaning of individual words in a sentence, integrating them with syntax, integrating the meaning of the current sentence with those of previous sentences and with the social context and general aims of the discourse). Both possibilities (a. summing minor impairments or b. summing minor delays) are versions of the more general hypothesis that sees deficits of EF not as primary, but as the consequence of a generalized disruption of brain processing. EF will *appear* selectively impaired because the tasks used to tap them avoid ceiling effects and are more sensitive to mild impairments. This hypothesis will explain why impairments of EF are common across populations with diffuse brain damage. This hypothesis is also consistent with the view that higher cognitive functions, defined as EF/fluid intelligence/reasoning, only depend on the efficiency of basic functions and/or on the efficiency of speed of processing as a generalized property of the brain (see Salthouse, 1996; Channon, et al., 2005, 2007; Feldmann, et al., 2005; Moyle, Fox, Bynevel, Arthur, & Burnett, 2007). Our results challenge this interpretation since deficits of basic functions were not shown, even in the context of complex tasks and with tasks were RTs were taken. The fact that deficits may vary for different types of EFs is also consistent with the hypothesis that these functions rely on specific neural circuitries which may be selectively affected by different types of damage, and not with the hypothesis that they reflect a general reduction in processing resources (see also Cepeda, Blackwell, & Munakata, 2013; Friedman et al., 2006; Friedman, & Miyake, 2017; Miyake et al., 2000). Our results are particularly consistent with those of Cepeda et al., (2013), who analysed relationships between tasks tapping EFs and simple vs. complex tasks measuring speed of processing in healthy participants across they life span. They found that executive measures were only associated with speed of processing in more complex tasks. These results show that speed of processing in complex tasks may depend on EFs, rather than EFs being a simple reflection of speed of processing.

Our results, of course, do not rule out relationships between EFs, basic skills and speed of processing. EFs and tasks tapping speed of processing were significantly correlated in our sample of AwPKU (see Romani et al., in press, this issue). Similarly, EF and speed of processing show a similar trajectory across the life span, improving from childhood to adulthood, and then declining after midlife (for studies on speed of processing: Cerella & Hale, 1994; Jenkins, Myerson, Joerding, & Hale, 2000; Nettelbeck & Burns, 2010; for studies on EFs: Anderson, 2002; Best & Miller, 2010; Romine & Reynolds, 2005). The fact that there is some over-lap, however, does not exclude a partial independence. This is what is emphasized here.

Clinically, our results show the success of dietary control in early-treated adults with PKU, but also some limitations with poorer performance in higher order functions which, although not specific to language, contribute to good performance in complex language tasks. These deficits may have implications for quality of life and professional achievement. It is important, however, to stress the great variability in performance across tasks, but, especially, among individuals (see Palermo et al., 2017, Manti et al., 2017). While some AwPKU performed poorly, others performed normally even in complex language tasks.

We have tried to relate performance in complex language tasks to level of dietary control as measured by blood Phe. We have subdivided participants according to different metabolic measures (Phe average and Phe fluctuations) taken at different times during the life span. We have only found a marginally significant difference in participants with low average Phe across the life span (Phe< 600 mml; N=17; mean=421; SD=98) and participants with high Phe average across the life span (Phe>600 mml; N=16; mean=930; SD=201) in terms of rate of CIU (z score: -1.45 vs z score: -.29; F (1,31) = 4.1; p=.05). Other differences did not reach statistical significance. This can be due to the limited number of participants as well as to the fact that metabolic control was relatively good across the group with relatively limited variation. For example, we did not find a difference in cognitive performance according to concurrent Phe either when participants were subdivided using a median split or when we compared individuals with Phe >600 and >900. However, we had only 12 and 10 participants, respectively, in the two groups. More power is needed to derive meaningful conclusions. It remains a possibility that some unknown factor is protecting the unaffected individuals from the negative consequences of PKU. This issue is related to the general issue of the pathogenesis of the neuropsychological impairments seen in AwPKU. High executive functions may rely particularly on dopaminergic circuitries which are affected in PKU because of lack of conversion of Phe into tyrosine, a precursor of dopamine. It is also possible that axonal damage due to Phe toxicity is particularly important for some cognitive functions compared to others (e.g., higher order EFs and also visuo-spatial functions relative to language functions, as shown by Romani et al., this issue). These hypotheses, however, are only speculative. Our results highlight the need to understand more about the neurophysiological causes of these impairments in PKU (due to untargeted metabolic control or to the limits of dietary intervention in amending this disease?) and to identify the factors contributing to the individual vulnerability and resilience.

Finally, one should note that the pattern of results shown by AwPKU in the production of narratives (characterized by poor communicative informativeness, but normal speech rate, rate of errors and mean length of utterance) also characterizes the narrative production of healthy older adults (Marini, Boewe, Caltagirone, & Carlomagno, 2005). In Romani et al. (this issue), we have shown that AwPKU also show patterns similar to those of older adults in tasks of visual search. These results, in conjunction with report of similar neuro-degenerative markers in PKU and pathological aging (e.g., Alzheimer Disease, see Adler-Abramovich et al., 2012; Soloway, Soloway, & Warner, 2013), raise the question of whether this population shows pre-mature aging or whether, instead, these are markers of some frontal and/or diffuse brain damage which is not population/disease specific.

**Conclusions.** The present study provides a first in-depth analysis of high-level linguistic functions in AwPKU. Our results contribute to defining more precisely and comprehensively the cognitive profile associated with this disease. The patterns of results showed by AwPKU, however, have also implications for normal cognition by showing that complex EFs rely on specific neurophysiological resources, possibly being particularly reliant on dopaminergic circuitry, and are not simply a reflection of general brain prowess which could spare basic functions.

**References**

Abdel Rahman, R., & Melinger, A. (2007). When bees hamper the production of honey: Lexical interference from associates in speech production. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 33,* 604–614.

Adler-Abramovich, L., Vaks, L., Carny, O., Trudler, D., Magno, A., & Caflisch, A., ... Gazit, E. (2012). Phenylalanine assembly into toxic fibrils suggests amyloid etiology in phenylketonuria. *Nature Chemical Biology, 8(8),* 701-706. Retrieved from http://dx.doi.org/10.1038/nchembio.1002

Albrecht, J., Garbade, S. F., & Burgard, P. (2009). Neuropsychological speed tests and blood phenylalanine levels in patients with phenylketonuria: A meta-analysis. *Neuroscience and Biobehavioral Reviews, 33,* 414–421. Retrieved from http://dx.doi.org/10.1016/j.neubiorev.2008.11.001.

Amanzio, M., Geminiani, G., Leotta, D., & Cappa, S. (2008). Metaphor comprehension in Alzheimer’s disease: Novelty matters. *Brain and Language, 107(1),* 1-10. Retrieved from http://dx.doi.org/10.1016/j.bandl.2007.08.003

Anderson, P. (2002). Assessment and development of Executive Function (EF) during childhood. *Child Neuropsychology, 8(2),* 71-82. Retrieved from http://dx.doi.org/10.1076/chin.8.2.71.8724

Anderson, P., & Leuzzi, V. (2010). White matter pathology in phenylketonuria. *Molecular Genetics and Metabolism, 99,* S3-S9. Retrieved from http://dx.doi.org/10.1016/j.ymgme.2009.10.005.

Andreetta, S., & Marini, A. (2014). The effect of lexical deficits on narrative disturbances in fluent aphasia. *Aphasiology, 29(6),* 705-723. Retrieved from http://dx.doi.org/10.1080/02687038.2014.979394

Angeleri, R., Bosco, F., Zettin, M., Sacco, K., Colle, L. & Bara, B. (2008). Communicative impairment in traumatic brain injury: A complete pragmatic assessment. *Brain and Language, 107(3),* 229-245.

Aron, A. (2008). Progress in Executive-Function Research. *Current Directions in Psychological Science, 17(2),* 124-129. Retrieved from http://dx.doi.org/10.1111/j.1467-8721.2008.00561.x

Ash, S., Moore, P., Antani, S., McCawley, G., Work, M., & Grossman, M. (2006). Trying to tell a tale: Discourse impairments in progressive aphasia and frontotemporal dementia. *Neurology, 66(9),* 1405-1413. Retrieved from http://dx.doi.org/10.1212/01.wnl.0000210435.72614.38

Azen, C., Koch, R., Friedman, E., Wenz, E. & and Fishier, K. (1996). Summary of findings from the United States Collaborative Study of children treated for phenylketonuria. *European Journal of Pediatrics, 155(S1),* S29-S32.

Barker, M., Young, B., & Robinson, G. (2017). Cohesive and coherent connected speech deficits in mild stroke. *Brain and Language, 168,* 23-36. Retrieved from <http://dx.doi.org/10.1016/j.bandl.2017.01.004>

Bartels-Tobin, L., & Hinckley, J. (2005). Cognition and discourse production in right hemisphere disorder. *Journal of Neurolinguistics, 18(6),* 461-477. Retrieved from http://dx.doi.org/10.1016/j.jneuroling.2005.04.001

Bartolo, A., Benuzzi, F., Nocetti, L., Baraldi, P., & Nichelli, P. (2006). Humor Comprehension and Appreciation: An fMRI Study. *Journal of Cognitive Neuroscience, 18(11),* 1789-1798. Retrieved from http://dx.doi.org/10.1162/jocn.2006.18.11.1789

Bedre, D., & D’Esposito, M. (2009). Is the rostro-caudal axis of the frontal lobe hierarchical? *Nature Review Neuroscience*, 10(9), 659-669.

Belke, E. (2008). Effects of working memory load on lexical-semantic encoding in language production. *Psychonomic Bulletin and Review, 15*, 357–363. doi: 10.3758/PBR.15.2.357

Belke, E., Brysbaert, M., Meyer, A. S., & Ghyselinck, M. (2005). Age of acquisition effects in picture naming: Evidence for a lexical-semantic competition hypothesis. *Cognition, 96,* B45–B54.

Belke, E., Meyer, A. S., & Damian, M. F. (2005). Refractory effects in picture naming as assessed in a semantic blocking paradigm. Quarterly. *Journal of Experimental Psychology Human Experimental Psychology 58A,* 667–692. doi: 10.1080/02724980443000142.

Benedek, M., Beaty, R., Jauk, E., Koschutnig, K., Fink, A., & Silvia, P. … Neubauer, A. C. (2014). Creating metaphors: The neural basis of figurative language production. *Neuroimage, 90,* 99-106. Retrieved from http://dx.doi.org/10.1016/j.neuroimage.2013.12.046

Best, J., & Miller, P. (2010). A Developmental Perspective on Executive Function. *Child Development, 81(6),* 1641-1660. Retrieved from http://dx.doi.org/10.1111/j.1467-8624.2010.01499.x

Bielak, A. A. M., Mansueti, L., Strauss, E., & Dixon, R. A. (2006). Performance on the Hayling and Brixton tests in older adults: Norms and correlates. *Archives of Clinical Neuropsychology, 21(2),* 141–149. doi:10.1016/j.acn.2005.08.006

Bihrle, A. M., Brownell, H. H., Powelson, J. A., & Gardner, H. (1986). Comprehension of humorous and nonhumorous materials by left and right brain-damaged patients. *Brain and Cognition, 5(4),* 399–411. doi:10.1016/0278-2626(86)90042-4

Bishop, D., & Baird, G. (2001). Parent and teacher report of pragmatic aspects of communication: use of the Children's Communication Checklist in a clinical setting. *Developmental Medicine and Child Neurology, 43(12),* 809. Retrieved from <http://dx.doi.org/10.1017/s0012162201001475>

Blain-Brière, B., Bouchard, C., & Bigras, N. (2014). The role of executive functions in the pragmatic skills of children age 4-5*. Frontiers in Psychology, 5(240),* 1-14. Retrieved from http://dx.doi.org/10.3389/fpsyg.2014.00240

Blair, M., Marczinski, C., Davis-Faroque, N., & Kertesz, A. (2007). A longitudinal study of language decline in Alzheimer's disease and frontotemporal dementia. *Journal of The International Neuropsychological Society, 13(02).* Retrieved from http://dx.doi.org/10.1017/s1355617707070269

Blau, N., van Spronsen, F. J., & Levy, H. L. (2010). Phenylketonuria. *Lancet, 376,* 1417–1427. Retrieved from http://dx.doi.org/10.1016/S0140-6736(10)60961-0

Bohrn, I., Altmann, U., & Jacobs, A. (2012). Looking at the brains behind figurative language—A quantitative meta-analysis of neuroimaging studies on metaphor, idiom, and irony processing. *Neuropsychologia, 50(11),* 2669-2683. Retrieved from <http://dx.doi.org/10.1016/j.neuropsychologia.2012.07.021>

Boot, E., Hollak, C. E. M., Huijbregts, S. C. J., Jahja, R., van Vliet, D., Nederveen, A. J., ... & Abeling, N. G. G. M. (2017). Cerebral dopamine deficiency, plasma monoamine alterations and neurocognitive deficits in adults with phenylketonuria. *Psychological Medicine*, 1-12.

Botting, N. (2002). Narrative as a tool for the assessment of linguistic and pragmatic impairments. *Child Language Teaching and Therapy, 18(1),* 1-21. Retrieved from http://dx.doi.org/10.1191/0265659002ct224oa.

Bouquet, C. A., Bonnaud, V., & Gil, R. (2003). Investigation of supervisory Attentional system functions in patients with Parkinson?S disease using the Hayling task. *Journal of Clinical and Experimental Neuropsychology (Neuropsychology, Development and Cognition: Section A), 25(6),* 751–760. doi:10.1076/jcen.25.6.751.16478

Breitenstein, C., Daum, I., & Ackermann, H. (1998). Emotional Processing following Cortical and Subcortical Brain Damage: Contribution of the Fronto-Striatal Circuitry. *Behavioural Neurology, 11(1),* 29-42. Retrieved from http://dx.doi.org/10.1155/1998/579029

Brumm, V., Azen, C., Moats, R., Stern, A., Broomand, C., Nelson, M., & Koch, R. (2004). Neuropsychological outcome of subjects participating in the PKU Adult Collaborative Study: A preliminary review. *Journal of Inherited Metabolic Disease, 27(5),* pp.549-566.

Bryan, K.L. (1995). *The Right Hemisphere Language Battery.* (2nd ed). Whurr Publishers, London.

Burgess, P. W. & Shallice, T. (1997). *The Hayling and Brixton Tests.* Thames Valley Test Company, Thurston, Suffolk.

Carlomagno, S., Giannotti, S., Vorano, L., & Marini, A. (2011). Discourse information content in non-aphasic adults with brain injury: A pilot study. *Brain Injury, 25(10),* 1010-1018. Retrieved from http://dx.doi.org/10.3109/02699052.2011.605097

Carriedo, N., Corral, A., Montoro, P., Herrero, L., Ballestrino, P., & Sebastián, I. (2016). The Development of Metaphor Comprehension and Its Relationship with Relational Verbal Reasoning and Executive Function. *PLOS ONE, 11(3),* 1-20.

Castagna, F., Montemagni, C., Maria Milani, A., Rocca, G., Rocca, P., Casacchia, M., & Bogetto, F. (2013). Prosody recognition and audiovisual emotion matching in schizophrenia: The contribution of cognition and psychopathology. *Psychiatry Research, 205(3),* 192-198. Retrieved from http://dx.doi.org/10.1016/j.psychres.2012.08.038

Cepeda, N., Blackwell, K., & Munakata, Y. (2013). Speed isn't everything: complex processing speed measures mask individual differences and developmental changes in executive control. *Developmental Science, 16(2),* 269-286.

Cerella, J., & Hale, S. (1994). The rise and fall in information-processing rates over the life span. *Acta Psychologica, 86(2-3),* 109-197. Retrieved from http://dx.doi.org/10.1016/0001-6918(94)90002-7

Chan, Y., Chou, T., Chen, H., Yeh, Y., Lavallee, J., Liang, K., & Chang, K. (2013). Towards a neural circuit model of verbal humor processing: An fMRI study of the neural substrates of incongruity detection and resolution. *Neuroimage, 66,* 169-176. Retrieved from http://dx.doi.org/10.1016/j.neuroimage.2012.10.019

Chang, P., Gray, R., & O'Brien, L. (2000). Patterns of academic achievement among patients treated early with phenylketonuria. *European Journal of Pediatrics, 159(S2),* S96-S99.

Channon, S., German, E., Cassina, C., & Lee, P. (2004). Executive Functioning, Memory, and Learning in Phenylketonuria. *Neuropsychology, 18(4),* 613-620. Retrieved from http://dx.doi.org/10.1037/0894-4105.18.4.613

Channon, S., Goodman, G., Zlotowitz, S., Mockler, C., & Lee, P. J. (2007). Effects of dietary management of phenylketonuria on long-term cognitive outcome. *Archives of Disease in Childhood, 92,* 213–218. Retrieved from http://dx.doi.org/10.1136/adc.2006.104786

Channon, S., Mockler, C., & Lee, P. (2005). Executive functioning and speed of processing in phenylketonuria. *Neuropsychology, 19,* 679–686. Retrieved from http://dx.doi.org/10.1037/0894-4105.19.5.679

Christ, S. E., Huijbregts, S. C. J., de Sonneville, L. M. J., & White, D. A. (2010). Executive function in early-treated phenylketonuria: Profile and underlying mechanisms. *Molecular Genetics and Metabolism, 99,* 22–32. Retrieved from http://dx.doi.org/10.1016/j.ymgme.2009.10.007

Clark, C., Prior, M., & Kinsella, G. (2000). Do Executive Function Deficits Differentiate Between Adolescents with ADHD and Oppositional Defiant/Conduct Disorder? A Neuropsychological Study Using the Six Elements Test and Hayling Sentence Completion Test. *Journal of Abnormal Child Psychology, 28(5),* 403-414. Retrieved from <http://dx.doi.org/10.1023/a:1005176320912>

Crawford, J. R. , Garthwaite, P. H. , & Porter, S. (2010). 'Point and interval estimates of effect sizes for the case-controls design in neuropsychology: Rationale, methods, implementations, and proposed reporting standards. *Cognitive Neuropsychology*, *27 (3),* 245-260. http://dx.doi.org/10.1080/02643294.2010.513967

Crowther, J., & Martin, R. (2014). Lexical selection in the semantically blocked cyclic naming task: the role of cognitive control and learning. *Frontiers in Human Neuroscience, 8.* Retrieved from http://dx.doi.org/10.3389/fnhum.2014.00009

Damian, M. F., Vigliocco, G., & Levelt, W. J. M. (2001). Effects of semantic context in the naming of pictures and words. *Cognition, 81,* B77–B86.

Decker, S., Hill, S., & Dean, R. (2007). Evidence of Construct Similarity in Executive Functions and Fluid Reasoning Abilities. *International Journal of Neuroscience, 117(6)*, 735-748. Retrieved from http://Dx.Doi.Org/10.1080/00207450600910085

DeRoche, K., & Welsh, M. (2008). Twenty-five years of research on neurocognitive outcomes in early-treated phenylketonuria: Intelligence and executive function. *Developmental Neuropsychology, 33,* 474–504. Retrieved from http://dx.doi.org/10.1080/87565640802101482

Diamond, A., & Baddeley, A. (1996). Evidence for the Importance of Dopamine for Prefrontal Cortex Functions Early in Life [and Discussion]. *Philosophical Transactions of The Royal Society B: Biological Sciences, 351(1346),* 1483-1494. Retrieved from http://dx.doi.org/10.1098/rstb.1996.0134

Duncan, J. (2010). The multiple-demand (MD) system of the primate brain: mental programs for intelligent behaviour. Trends in Cognitive Sciences, 14(4), 172-179. Retrieved from <http://dx.doi.org/10.1016/j.tics.2010.01.004>

Faust, D., Libon, D., & Pueschel, S. (1987). Neuropsychological Functioning in Treated Phenylketonuria. *The International Journal of Psychiatry in Medicine, 16(2),* 169-177.

Feldmann, R., Denecke, J., Grenzebach, M., & Weglage, J. (2005). Frontal lobe-dependent functions in treated phenylketonuria: Blood phenylalanine concentrations and long-term deficits in adolescents young adults. *Journal of Inherited Metabolic Disease, 28(4),* 445-455. Retrieved from http://dx.doi.org/10.1007/s10545-005-0445-7

Fishler, K., Azen, C. G., Henderson, R., Friedman, E. G., Koch, R. (1987) Psychoeducational findings among children treated for phenylketonuria. *American Journal of Mental Deficiency,* *92(1),* 65-73.

Floresco, S., & Magyar, O. (2006). Mesocortical dopamine modulation of executive functions: beyond working memory. *Psychopharmacology, 188(4),* 567-585. Retrieved from <http://dx.doi.org/10.1007/s00213-006-0404-5>

Friedman, N., & Miyake, A. (2017). Unity and diversity of executive functions: Individual differences as a window on cognitive structure. *Cortex, 86,* 186-204. Retrieved from http://dx.doi.org/10.1016/j.cortex.2016.04.023

Friedman, N., Miyake, A., Corley, R., Young, S., DeFries, J., & Hewitt, J. (2006). Not All Executive Functions Are Related to Intelligence. *Psychological Science, 17(2),* 172-179. Retrieved from http://dx.doi.org/10.1111/j.1467-9280.2006.01681.x

Froming, K., Levy, M., Schaffer, S., & Ekman, P. (2006). *The Comprehensive Affect Testing System.* Psychology Software, Inc.

Gassió, R., Artuch, R., Vilaseca, M., Fusté, E., Boix, C., Sans, A. and Campistol, J. (2005). Cognitive functions in classic phenylketonuria and mild hyperphenyl-alaninaemia: experience in a paediatric population. *Developmental Medicine and Child Neurology, 47(7),* 443-448.

Gejão, M., Ferreira, A., Silva, G., Anastácio-Pessan, F., & Lamônica, D. (2009). Communicative and psycholinguistic abilities in children with phenylketonuria and congenital hypothyroidism*. Journal of Applied Oral Science, 17,* 69-75.

Geurts, H., Verte, S., Oosterlaan, J., Roeyers, H., & Sergeant, J. (2004). How specific are executive functioning deficits in attention deficit hyperactivity disorder and autism? *Journal of Child Psychology and Psychiatry, 45(4),* 836-854. Retrieved from http://dx.doi.org/10.1111/j.1469-7610.2004.00276.x

Godefroy, O., Cabaret, M., Petit-Chenal, V., Pruvo, J., & Rousseaux, M. (1999). Control Functions of the Frontal Lobes. Modularity of the Central-Supervisory System? *Cortex, 35(1),* 1-20. Retrieved from http://dx.doi.org/10.1016/s0010-9452(08)70782-2

Goel, V., & Dolan, R. (2001). The functional anatomy of humour: segregating cognitive and affective components. *Nature Neuroscience, 4(3),* 237-238. Retrieved from http://dx.doi.org/10.1038/85076

Griffiths, C. (2007). Pragmatic abilities in adults with and without dyslexia: a pilot study. *Dyslexia, 13(4)*, 276-296. Retrieved from http://dx.doi.org/10.1002/dys.333

Gunter, H., Ghaziuddin, M., & Ellis, H. (2002). Asperger Syndrome: Tests of Right Hemisphere Functioning and Interhemispheric Communication. *Journal of Autism and Developmental Disorders, 32(4),* 263-281. Retrieved from http://dx.doi.org/10.1023/a:1016326701439

Hall, D., Ouyang, B., Lonnquist, E., & Newcombe, J. (2011). Pragmatic Communication is Impaired in Parkinson Disease. *International Journal of Neuroscience, 121(5),* 254-256. Retrieved from http://dx.doi.org/10.3109/00207454.2010.550389

Halper, A., & Cherney, L. (1998). Cognitive-Communication Problems after Right Hemisphere Stroke: A Review of Intervention Studies. *Topics in Stroke Rehabilitation, 5(1),* 1-10. Retrieved from http://dx.doi.org/10.1310/r9c8-23c6-34v7-mb2h

Hardelid, P., Cortina-Borja, M., Munro, A., Jones, H., Cleary, M., Champion, M., Foo, Y., Scriver, C. and Dezateux, C. (2007). The Birth Prevalence of PKU in Populations of European, South Asian and Sub-Saharan African Ancestry Living in South East England. *Annals of Human Genetics, 72(1),* 65-71.

Hornberger, M., & Bertoux, M. (2015). Right lateral prefrontal cortex--specificity for inhibition or strategy use? *Brain, 138(4),* 833-835. Retrieved from http://dx.doi.org/10.1093/brain/awv027.

Jenkins, L., Myerson, J., Joerding, J., & Hale, S. (2000). Converging evidence that visuospatial cognition is more age-sensitive than verbal cognition. Psychology and Aging, 15(1), 157-175. Retrieved from http://dx.doi.org/10.1037//0882-7974.15.1.157

Jurado, M., & Rosselli, M. (2007). The Elusive Nature of Executive Functions: A Review of our Current Understanding. *Neuropsychology Review, 17(3),* 213-233. Retrieved from http://dx.doi.org/10.1007/s11065-007-9040-z

Kavé, G., Leonard, C., Cupit, J., & Rochon, E. (2007). Structurally well-formed narrative production in the face of severe conceptual deterioration: A longitudinal case study of a woman with semantic dementia. *Journal of Neurolinguistics, 20(2),* 161–177. doi:10.1016/j.jneuroling.2006.06.003

Koch, R., Burton, B., Hoganson, G., Peterson, R., Rhead, W., Rouse, B., … Azen, C. (2002) Phenylketonuria in adulthood: a collaborative study. *Journal of Inherited Metabolic Diseases 25(5),* 333-46.

Kuperberg, G., McGuire, P., Bullmore, E., Brammer, M., Rabe-Hesketh, S., & Wright, I. … David, A. S. (2000). Common and Distinct Neural Substrates for Pragmatic, Semantic, and Syntactic Processing of Spoken Sentences: An fMRI Study. *Journal of Cognitive Neuroscience, 12(2),* 321-341. Retrieved from http://dx.doi.org/10.1162/089892900562138

Landa, R. (2000) Social Language Use in Asperger Syndrome and High-Functioning Autism in Klin, A., Volkmar, F., & Sparrow, S. (2000). *Asperger syndrome* (1st ed.). New York: Guilford Press

Lewis, F., Woodyatt, G., & Murdoch, B. (2008). Linguistic and pragmatic language skills in adults with autism spectrum disorder: A pilot study. *Research in Autism Spectrum Disorders, 2(1),* 176-187. Retrieved from http://dx.doi.org/10.1016/j.rasd.2007.05.002

Luria, A. (1978). L. S. Vygotsky and the problem of functional localization. In M. Cole, The selected writing of A. R. Luria (1st ed., pp. 273-281). White Plains, NY: Sharpe.

MacWhinney, B., Fromm, D., Holland, A., Forbes, M., & Wright, H. (2010). Automated analysis of the Cinderella story. *Aphasiology, 24(6-8),* 856–868. doi:10.1080/02687030903452632

Manti, F., Nardecchia, F., Paci, S., Chiarotti, F., Carducci, C., Carducci, C., ... & Leuzzi, V. (2017). Predictability and inconsistencies in the cognitive outcome of early treated PKU patients. *Journal of Inherited Metabolic Disease,* ttps://doi.org/10.1007/s10545-017-0082-y

Marini, A., Andreetta, S., del Tin, S., & Carlomagno, S. (2011). A multi-level approach to the analysis of narrative language in aphasia. *Aphasiology, 25(11),* 1372–1392. doi:10.1080/02687038.2011.584690

Marini, A., Boewe, A., Caltagirone, C., & Carlomagno, S. (2005). Age-related Differences in the Production of Textual Descriptions. *Journal of Psycholinguistic Research, 34(5),* 439-463. Retrieved from http://dx.doi.org/10.1007/s10936-005-6203-z

Marini, A., Martelli, S., Gagliardi, C., Fabbro, F., & Borgatti, R. (2010). Narrative language in Williams Syndrome and its neuropsychological correlates. *Journal of Neurolinguistics, 23(2),* 97-111. Retrieved from http://dx.doi.org/10.1016/j.jneuroling.2009.10.002

Marini, A., Spoletini, I., Rubino, I., Ciuffa, M., Bria, P., & Martinotti, G., ... Spalletta, G. (2008). The language of schizophrenia: An analysis of micro and macrolinguistic abilities and their neuropsychological correlates. *Schizophrenia Research, 105(1-3),* 144-155. Retrieved from http://dx.doi.org/10.1016/j.schres.2008.07.011

Martin, I., & McDonald, S. (2003). Weak coherence, no theory of mind, or executive dysfunction? Solving the puzzle of pragmatic language disorders. *Brain and Language, 85(3),* 451-466. Retrieved from <http://dx.doi.org/10.1016/s0093-934x(03)00070-1>

McDonald, S. (1993). Pragmatic language skills after closed head injury: Ability to meet the informational needs of the listener. *Brain and Language, 44,* 28–46. doi: 10.1006/brln.1993.1003

McDonald, S., & Pearce, S. (1996). Clinical Insights into Pragmatic Theory: Frontal Lobe Deficits and Sarcasm. *Brain and Language, 53(1),* 81-104. Retrieved from http://dx.doi.org/10.1006/brln.1996.0038

McNamara, P., & Durso, R. (2003). Pragmatic communication skills in patients with Parkinson’s disease. *Brain and Language, 84(3),* 414-423. Retrieved from http://dx.doi.org/10.1016/s0093-934x(02)00558-8

Melnick, C., Michals, K., & Matalon, R. (1981). Linguistic development of children with phenylketonuria and normal intelligence. *The Journal of Pediatrics, 98(2),* 269-272.

Meteyard, L., & Patterson, K. (2009). The relation between content and structure in language production: An analysis of speech errors in semantic dementia. *Brain and Language, 110(3),* 121-134. Retrieved from http://dx.doi.org/10.1016/j.bandl.2009.03.007

Meyer, M., Alter, K., Friederici, A., Lohmann, G., & von Cramon, D. (2002). FMRI reveals brain regions mediating slow prosodic modulations in spoken sentences. *Human Brain Mapping, 17(2),* 73-88. Retrieved from http://dx.doi.org/10.1002/hbm.10042

Minshew, N., Goldstein, G., & Siegel, D. (1995). Speech and language in high-functioning autistic individuals. *Neuropsychology, 9(2),* 255-261. Retrieved from http://dx.doi.org/10.1037//0894-4105.9.2.255

Mitchell, R. L. C., & Crow, T. J. (2005). Right hemisphere language functions and schizophrenia: The forgotten hemisphere? *Brain, 128(5),* 963–978. doi:10.1093/brain/awh466

Miyake, A., Friedman, N., Emerson, M., Witzki, A., Howerter, A., & Wager, T. (2000). The Unity and Diversity of Executive Functions and Their Contributions to Complex “Frontal Lobe” Tasks: A Latent Variable Analysis. *Cognitive Psychology, 41(1),* 49-100. Retrieved from http://dx.doi.org/10.1006/cogp.1999.0734

Moran, J., Wig, G., Adams, R., Janata, P., & Kelley, W. (2004). Neural correlates of humour detection and appreciation. *Neuroimage, 21(3),* 1055-1060. Retrieved from http://dx.doi.org/10.1016/j.neuroimage.2003.10.017

Moyle, J. J., Fox, A. M., Arthur, M., Bynevelt, M., & Burnett, J. R. (2007). Meta-analysis of neuropsychological symptoms of adolescents and adults with PKU. *Neuropsychology Review, 17,* 91–101. Retrieved from http://dx.doi.org/10.1007/s11065-007-9021-2

Moyle, J. J., Fox, A. M., Bynevelt, M., Arthur, M., & Burnett, J. R. (2007). A neuropsychological profile of off-diet adults with phenylketonuria. *Journal of Clinical and Experimental Neuropsychology, 29,* 436–441. Retrieved from http://dx.doi.org/10.1080/13803390600745829.

Mueser, K., Bellack, A., Douglas, M., & Morrison, R. (1991). Prevalence and stability of social skill deficits in schizophrenia. *Schizophrenia Research, 5(2),* 167-176. Retrieved from http://dx.doi.org/10.1016/0920-9964(91)90044-r

Nettelbeck, T., & Burns, N. (2010). Processing speed, working memory and reasoning ability from childhood to old age. *Personality and Individual Differences, 48(4),* 379-384. Retrieved from http://dx.doi.org/10.1016/j.paid.2009.10.032

Nicholas, L., & Brookshire, R. (1993). A System for Quantifying the Informativeness and Efficiency of the Connected Speech of Adults with Aphasia. *Journal of Speech Language and Hearing Research, 36(2),* 338. Retrieved from http://dx.doi.org/10.1044/jshr.3602.338

Ozanne, A. E., Krimmer, H., & Murdoch, B. E. (1990). Speech and language skills in children with early treated phenylketonuria. *American Journal on Mental Retardation, 94(6),* 625-632.

Ozonoff, S., & Miller, J. N. (1996). An exploration of right-hemisphere contributions to the pragmatic impairments of autism. *Brain and Language, 52(3),* 411–434.

Palermo, L., Geberhiwot, T., MacDonald, A., Limback, E., Hall, S., & Romani, C. (2017). Cognitive outcomes in early-treated adults with phenylketonuria (PKU): A comprehensive picture across domains. *Neuropsychology, 31(3),* 255-267. Retrieved from <http://dx.doi.org/10.1037/neu0000337>

Palermo, L. MacDonald, A., Limback, E., Robertson2, L., Boocock, S., Geberhiwot, T, & Romani, C. (in preparation). Emotional health in early treated adults with phenylketonuria (PKU): Relationship with cognitive abilities and blood Phenylalanine.

Pearce, S., McDonald, S., & Coltheart, M. (1998). Interpreting Ambiguous Advertisements: The Effect of Frontal Lobe Damage. *Brain and Cognition, 38(2),* 150-164. Retrieved from http://dx.doi.org/10.1006/brcg.1998.1018

Pennington, B. F., Van Doorninck, W. J., McCabe, L. L., & McCabe, E. R. (1985) Neuropsychological deficits in early treated phenylketonuric children. *American Journal of Mental Deficiency, 89(5),* 467-474.

Phan, K., Wager, T., Taylor, S., & Liberzon, I. (2002). Functional Neuroanatomy of Emotion: A Meta-Analysis of Emotion Activation Studies in PET and fMRI. *Neuroimage, 16(2),* 331-348. Retrieved from http://dx.doi.org/10.1006/nimg.2002.1087

Rad, D. S. (2014). A review on adult pragmatic assessments. *Iranian Journal of Neurology, 13(3),* 113-118. Retrieved from https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4240926/

Rapp, A., & Wild, B. (2011). Nonliteral Language in Alzheimer Dementia: A Review. *Journal of The International Neuropsychological Society, 17(02),* 207-218. Retrieved from http://dx.doi.org/10.1017/s1355617710001682

Rapp, A., Leube, D., Erb, M., Grodd, W., & Kircher, T. (2004). Neural correlates of metaphor processing. *Cognitive Brain Research, 20(3),* 395-402. Retrieved from http://dx.doi.org/10.1016/j.cogbrainres.2004.03.017

Robinson, G., Cipolotti, L., Walker, D., Biggs, V., Bozzali, M., & Shallice, T. (2015). Verbal suppression and strategy use: a role for the right lateral prefrontal cortex? *Brain, 138(4),* 1084-1096. Retrieved from <http://dx.doi.org/10.1093/brain/awv003>

Rohrer, J., Sauter, D., Scott, S., Rossor, M., & Warren, J. (2010). Receptive prosody in nonfluent primary progressive aphasias. *Cortex, 48(3),* 308-316. Retrieved from http://dx.doi.org/10.1016/j.cortex.2010.09.004

Romani, C., Palermo, L., MacDonald, A., Limback, E., Hall, S., & Geberhiwot, T. (2017). The impact of phenylalanine levels on cognitive outcomes in adults with phenylketonuria: Effects across tasks and developmental stages. *Neuropsychology, 31(3),* 242-254. Retrieved from <http://dx.doi.org/10.1037/neu0000336>.

Romani, C., MacDonald, A, De Felice & Palermo, L. (in press). Speed of processing and executive functions in adults with phenylketonuria (PKU): Quick in finding the word, but not the ladybird*. Cognitive Neuropsychology*.

Romine, C., & Reynolds, C. (2005). A Model of the Development of Frontal Lobe Functioning: Findings from a Meta-Analysis. *Applied Neuropsychology, 12(4),* 190-201. Retrieved from http://dx.doi.org/10.1207/s15324826an1204\_2

Rossell, S., Van Rheenen, T., Groot, C., Gogos, A., O’Regan, A., & Joshua, N. (2013). Investigating affective prosody in psychosis: A study using the Comprehensive Affective Testing System. *Psychiatry Research, 210(3),* 896-900. Retrieved from http://dx.doi.org/10.1016/j.psychres.2013.07.037

Rousseaux, M., Sève, A., Vallet, M., Pasquier, F., & Mackowiak-Cordoliani, M. (2010). An analysis of communication in conversation in patients with dementia. *Neuropsychologia, 48(13),* 3884-3890. Retrieved from http://dx.doi.org/10.1016/j.neuropsychologia.2010.09.026

Saldert, C., Fors, A., Ströberg, S., & Hartelius, L. (2010). Comprehension of complex discourse in different stages of Huntington's disease. *International Journal of Language and Communication Disorders, 45(6),* 656-669. Retrieved from http://dx.doi.org/10.3109/13682820903494742.

Schaffer, S., Wisniewski, A., Dahdah, M., & Froming, K. (2009). The Comprehensive Affect Testing System-Abbreviated: Effects of Age on Performance. *Archives of Clinical* *Neuropsychology, 24(1),* 89-104. Retrieved from http://dx.doi.org/10.1093/arclin/acp012

Schnur, T. T., Schwartz, M. F., Brecher, A., and Hodgson, C. (2006). Semantic interference during blocked-cyclic naming: evidence from aphasia. *Journal of Memory and Language, 54,* 199–227. doi: 10.1016/j.jml.2005.10.002

Scriver C. R., & Kaufman S. (2001). Hyperphenylalaninemias: phenylalanine hydroxylase deficiency. In Scriver C. R., Beaudet A. L., Sly W. S., Valle D., Editors, Childs B., Kinzler K. W., Vogelstein B., Assoc. Editors. *The Metabolic and Molecular Bases of Inherited Disease* (8th ed.). New York: McGraw-Hill, 1667–1724.

Shallice, T. (1982). Specific Impairments of Planning. *Philosophical Transactions of The Royal Society B: Biological Sciences, 298(1089),* 199-209. Retrieved from http://dx.doi.org/10.1098/rstb.1982.0082

Shammi, P. & Stuss, D. T. (1999). Humour appreciation: a role of the right frontal lobe. *Brain, 122(4),* 657-666. Retrieved from http://dx.doi.org/10.1093/brain/122.4.657.

Shapiro, B. E., & Danly, M. (1985). The role of the right hemisphere in the control of speech prosody in propositional and affective contexts. *Brain and Language, 25(1),* 19–36. doi:10.1016/0093-934x(85)90118-x

Soleymani, Z., Keramati, N., Rohani, F., & Jalaei, S. (2015). Factors influencing verbal intelligence and spoken language in children with phenylketonuria. *Indian Pediatrics, 52(5),* 397-401. Retrieved from http://dx.doi.org/10.1007/s13312-015-0644-8

Soloway, A., Soloway, P., & Warner, V. (2013). Possible chemical initiators of cognitive dysfunction in phenylketonuria, Parkinson’s disease and Alzheimer’s disease. *Medical Hypotheses, 81(4),* 690-694. Retrieved from http://dx.doi.org/10.1016/j.mehy.2013.07.028

Sponheim, S., Surerus-Johnson, C., Leskela, J., & Dieperink, M. (2003). Proverb interpretation in schizophrenia: the significance of symptomatology and cognitive processes. *Schizophrenia Research, 65(2-3),* 117-123. Retrieved from http://dx.doi.org/10.1016/s0920-9964(02)00525-x

Stark, J. A. (2010). Content analysis of the fairy tale Cinderella – A longitudinal single-case study of narrative production: “From rags to riches.” *Aphasiology, 24(6-8),* 709–724. doi:10.1080/02687030903524729

Stuss, D. (2011). Functions of the Frontal Lobes: Relation to Executive Functions. *Journal of the International Neuropsychological Society, 17(05),* 759-765. Retrieved from <http://dx.doi.org/10.1017/s1355617711000695>

Tekin, S. & Cummings, J. (2002). Frontal–subcortical neuronal circuits and clinical neuropsychiatry. *Journal of Psychosomatic Research, 53(2),* 647-654.

Thompson, C. K. (2003). Unaccusative verb production in agrammatic aphasia: The argument structure complexity hypothesis. *Journal of Neurolinguistics, 16(s 2–3),* 151–167. doi:10.1016/S0911-6044(02)00014-3

Tonks, J., Huw Williams, W., Frampton, I., Yates, P., Wall, S. E., & Slater, A. (2008). Reading emotions after childhood brain injury: Case series evidence of dissociation between cognitive abilities and emotional expression processing skills. *Brain Injury, 22(4),* 325-332. doi:10.1080/02699050801968303

Uekermann, J., Channon, S., Winkel, K., Schlebusch, P., & Daum, I. (2007). Theory of mind, humour processing and executive functioning in alcoholism. *Addiction, 102(2),* 232-240. Retrieved from http://dx.doi.org/10.1111/j.1360-0443.2006.01656.x

Ventura, M., Baynes, K., Sigvardt, K., Unruh, A., Acklin, S., Kirsch, H., & Disbrow, E. (2012). Hemispheric asymmetries and prosodic emotion recognition deficits in Parkinson's disease. *Neuropsychologia, 50(8),* 1936-1945. Retrieved from http://dx.doi.org/10.1016/j.neuropsychologia.2012.04.018

Wang, A., Lee, S., Sigman, M., & Dapretto, M. (2006). Neural basis of irony comprehension in children with autism: the role of prosody and context. *Brain, 129(4),* 932-943. Retrieved from http://dx.doi.org/10.1093/brain/awl032

Webster, J., Franklin, S., & Howard, D. (2007). An analysis of thematic and phrasal structure in people with aphasia: What more can we learn from the story of Cinderella*? Journal of Neurolinguistics, 20(5),* 363-394. Retrieved from <http://dx.doi.org/10.1016/j.jneuroling.2007.02.002>

Wildgruber, D., Ackermann, H., Kreifelts, B., & Ethofer, T. (2006). Cerebral processing of linguistic and emotional prosody: fMRI studies. *Understanding Emotions,* 249-268. Retrieved from http://dx.doi.org/10.1016/s0079-6123(06)56013-3

Young, E., Diehl, J., Morris, D., Hyman, S., & Bennetto, L. (2005). The Use of Two Language Tests to Identify Pragmatic Language Problems in Children with Autism Spectrum Disorders. *Language Speech and Hearing Services in Schools, 36(1),* 62. Retrieved from http://dx.doi.org/10.1044/0161-1461

Zanini, S., Bryan, K., De Luca, G., & Bava, A. (2005). Italian Right Hemisphere Language Battery: the normative study. *Neurological Sciences, 26(1),* 13-25. Retrieved from http://dx.doi.org/10.1007/s10072-005-0377-x

Zartler, A. & Sassaman, E. (1981). Linguistic development in PKU. *The Journal of Pediatrics, 99(3),* 501.

Table 1. Demographic information, metabolic control and IQ for the whole sample of control and AwPKU groups. Blood Phe measured in μmol/L. Phe Fluctuations = average of Phe SD per year, in year band. Different subgroups were tested for different tasks.

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | | **Controls**  (N=39) | |  | **AwPKU** (N=38) | |  | **Controls vs AwPKU** |
| Mean | SD |  | Mean | SD |  |  |
| **Age** | **27.8** | *8.4* |  | **27.3** | *8.1* |  | t(1,75)= -0.2; p=-.5 |
| **Education years** | **14.6** | *2* |  | **14.1** | *2.1* |  | t(1,75)= -1.1; p=-.5 |
| **Gender (M/F)** | **10/29** |  |  | **13/25** |  |  |  |
|  |  |  |  |  |  |  |  |
| **0-10 Years-old** |  |  |  |  |  |  |  |
| Average Phe | -------- | ------- |  | **428.7** | *189.5* |  |  |
| Phe Fluctuations | -------- | ------- |  | **187.9** | *67.4* |  |  |
| **11-16 Years-old** |  |  |  |  |  |  |  |
| Average Phe | -------- | ------- |  | **674.8** | *322.6* |  |  |
| Phe Fluctuations | -------- | ------- |  | **144.3** | *52.1* |  |  |
| **17 years old to present** |  |  |  |  |  |  |  |
| Average Phe | -------- | ------- |  | **751.2** | *319.2* |  |  |
| Phe Fluctuations | -------- | ------- |  | **133** | *54.1* |  |  |
| **Phe at the time of testing\*** |  |  |  |  |  |  |  |
| Verbal Narratives (N=33) | -------- | ------- |  | **689.2** | *357.6* |  |  |
| Pragmatic tasks (N=21) | -------- | ------- |  | **794** | *427.7* |  |  |
|  |  |  |  |  |  |  |  |
| **FIQ** | **113.9** | *9.7* |  | **103.9** | *13.9* |  | t(1,75)= -3.5; p <.01 |
| **VIQ** | **112.2** | *9.2* |  | **101.4** | *13.8* |  | t(1,75)= -3.9; p <.01 |
| **PIQ** | **112.4** | *10.9* |  | **105.3** | *14.4* |  | t(1,75)= -2.4; p <.01 |

Table 2. Verbal narrative measures in control and PKU participants. Measures where there is a significant difference with controls are highlighted. Number of poor scores with Z score ≤ 2 from the control mean is also reported. CIUs: Correct Information Units; MUL: Mean Utterance Length.

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Cognitive Measures** | **Controls** | |  | **AwPKU** | |  | **Controls vs. AwPKU** |  | **Number of poor scores (**≤ 2 SD) | |
| (N=32) | | (N=33) | |
|  | **Mean** | **SD** |  | **Mean** | **SD** |  | ***t* and *p*** |  | **Controls** | **PKU** |
| **Age** | **27.1** | *7.5* |  | **27.9** | *8.2* |  | t(1,63)= -.4.; p=.68 |  |  |  |
| **Education years** | **14.9** | *1.7* |  | **14.4** | *2.1* |  | t(1,63)= 1.1; p=.26 |  |  |  |
| **Gender (M/F)** | **10/22** |  |  | **11/22** |  |  |  |  |  |  |
| **Speech Errors** | **1.8** | *1.7* |  | **2.1** | *1.4* |  | t(1,63)= -.7; p=.46 |  | 2 | 1 |
| (% on total words) |
| **Speech rate** | **2.4** | *0.4* |  | **2.2** | *0.4* |  | t(1,63)= 1.7; p=.10 |  | 1 | 2 |
| (word/sec) |
| **MUL** | **14.2** | *3* |  | **13.2** | *2.5* |  | t(1,63)= 1.5; p=.13 |  | 0 | 0 |
| (mean N of words) |
| **Main concepts** | **86.7** | *10.9* |  | **81.1** | *21.2* |  | t(1,63)= 1.3 ; p=.21 |  | 2 | 3 |
| (% correct answer) |
| **CIUs** | **91.7** | *4.6* |  | **87.7** | *7.8* |  | **t(1,63)= 2.5; p = .01** |  | 2 | 6 |
| (% on total words) |

Table 3. Individual score of the PKU group for verbal narrative measures. Speech errors combine articulatory, phonological, morphological and semantic errors; they are calculated out of total words produced. Z scores are calculated from the control mean. Where nedeed, Z scores have been sign reversed, so that negative Z scores always indicate worse performance compared to controls. CIU: Content Information Units; MUL: Mean Utterance Length. Z scores ≥1.5 Z are highlighted. \* p < .05; \*\* p < .01 according to single case analysis implemented in the program Singlims\_ES.exe (Crawford, Garthwaite, & Porter, 2010).

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | **Speech Errors** | |  | **Speech rate (words/sec)** | |  | **MUL** | |  | **Main concepts** | |  | **CIUs** | |
|  |  | |  |  | |  |  | |  |  | |  |  | |
| **AwPKU**  (N=33) | **%** | **Z-score** |  | **N words** | **Z-score** |  | **N words** | **Z-score** |  | **%** | **Z-score** |  | **%** | **Z-score** |
| 1 | 2.0 | -0.1 |  | 2.1 | -0.6 |  | 11.7 | -0.8 |  | 87.5 | 0.1 |  | 79.6\* | -2.6 |
| 2 | 1.7 | 0.1 |  | 2.9 | 1.3 |  | 14.1 | -0.1 |  | 84.4 | -0.2 |  | 88.9 | -0.6 |
| 3 | 3.3 | -0.9 |  | 1.9 | -1.3 |  | 10.0 | -1.4 |  | 25.0\*\* | -5.7 |  | 89.1 | -0.6 |
| 4 | 0.4 | 0.8 |  | 1.5\* | -2.0 |  | 10.6 | -1.2 |  | 71.9 | -1.4 |  | 71.2\*\* | -4.5 |
| 5 | 1.2 | 0.3 |  | 2.5 | 0.4 |  | 17.6 | 1.1 |  | 93.8 | 0.6 |  | 85.1 | -1.4 |
| 6 | 0.3 | 0.9 |  | 2.0 | -0.9 |  | 12.8 | -0.5 |  | 96.8 | 0.9 |  | 97.3 | 1.2 |
| 7 | 2.2 | -0.3 |  | 2.4 | 0.1 |  | 16.1 | 0.6 |  | 96.9 | 0.9 |  | 93.6 | 0.4 |
| 8 | 0.8 | 0.6 |  | 2.1 | -0.8 |  | 14.5 | 0.1 |  | 87.5 | 0.1 |  | 94.8 | 0.7 |
| 9 | 2.9 | -0.6 |  | 2.6 | 0.5 |  | 15.7 | 0.5 |  | 90.6 | 0.4 |  | 91.9 | 0.0 |
| 10 | 0.9 | 0.6 |  | 2.2 | -0.5 |  | 17.7 | 1.1 |  | 93.8 | 0.6 |  | 93.7 | 0.4 |
| 11 | 2.9 | -0.6 |  | 2.1 | -0.7 |  | 12.9 | -0.5 |  | 46.9\*\* | -3.7 |  | 70.8\*\* | -4.6 |
| 12 | 1.2 | 0.4 |  | 1.3\*\* | -2.8 |  | 9.7 | -1.5 |  | 34.4\*\* | -4.8 |  | 63.9\*\* | -6.1 |
| 13 | 4.3 | -1.5 |  | 1.9 | -1.2 |  | 11.5 | -0.9 |  | 100 | 1.2 |  | 90.1 | -0.3 |
| 14 | 1.6 | 0.1 |  | 2.6 | 0.5 |  | 17.4 | 1.0 |  | 90.6 | 0.4 |  | 84.3 | -1.6 |
| 15 | 1.7 | 0.1 |  | 1.6\* | -1.9 |  | 13.5 | -0.3 |  | 78.1 | -0.8 |  | 94.9 | 0.7 |
| 16 | 0.5 | 0.8 |  | 2.1 | -0.7 |  | 14.9 | 0.2 |  | 93.8 | 0.6 |  | 86.5 | -1.1 |
| 17 | 3.8 | -1.2 |  | 2.8 | 1.0 |  | 14.7 | 0.2 |  | 96.9 | 0.9 |  | 88.9 | -0.6 |
| 18 | 3.9 | -1.3 |  | 1.8 | -1.5 |  | 12.4 | -0.6 |  | 90.6 | 0.4 |  | 88.8 | -0.6 |
| 19 | 1.5 | 0.2 |  | 2.3 | -0.3 |  | 11.4 | -0.9 |  | 65.6 | -1.9 |  | 86.8 | -1.1 |
| 20 | 2.7 | -0.5 |  | 2.3 | -0.2 |  | 8.8 | -1.8 |  | 71.9 | -1.4 |  | 92.7 | 0.2 |
| 21 | 0.4 | 0.8 |  | 2.0 | -1.0 |  | 9.8 | -1.5 |  | 75.0 | -1.1 |  | 87.6 | -0.9 |
| 22 | 1.3 | 0.3 |  | 2.7 | 0.7 |  | 13.8 | -0.2 |  | 96.9 | 0.9 |  | 91.4 | -0.1 |
| 23 | 0.6 | 0.7 |  | 2.0 | -0.8 |  | 11.5 | -0.9 |  | 90.6 | 0.4 |  | 90.7 | -0.2 |
| 24 | 2.2 | -0.3 |  | 2.4 | 0.2 |  | 12.6 | -0.5 |  | 71.9 | -1.4 |  | 92.5 | 0.2 |
| 25 | 3.5 | -1 |  | 2.1 | -0.6 |  | 10.1 | -1.4 |  | 78.0 | -0.8 |  | 80.4\* | -2.5 |
| 26 | 4.1 | -1.4 |  | 2.6 | 0.5 |  | 15.5 | 0.4 |  | 78.0 | -0.8 |  | 88.4 | -0.7 |
| 27 | 1.6 | 0.1 |  | 2.5 | 0.2 |  | 14.9 | 0.2 |  | 81.0 | -0.5 |  | 96.7 | 1.1 |
| 28 | 1.2 | 0.4 |  | 1.9 | -1.2 |  | 16.1 | 0.6 |  | 84.0 | -0.3 |  | 94.8 | 0.7 |
| 29 | 1.4 | 0.3 |  | 2.5 | 0.4 |  | 12.6 | -0.6 |  | 100 | 1.2 |  | 91.8 | 0.0 |
| 30 | 1.4 | 0.2 |  | 2.3 | -0.3 |  | 12.7 | -0.5 |  | 96.0 | 0.9 |  | 96.6 | 1.1 |
| 31 | 6.3\* | -2.7 |  | 2.7 | 0.8 |  | 9.1 | -1.7 |  | 96.9 | 0.9 |  | 83.5 | -1.8 |
| 32 | 3.5 | -1 |  | 2.2 | -0.3 |  | 13.1 | -0.4 |  | 68.8 | -1.7 |  | 78.5\*\* | -2.9 |
| 33 | 1.6 | 0.1 |  | 2.3 | -0.3 |  | 15.8 | 0.5 |  | 93.8 | 0.6 |  | 86.9 | -1.1 |
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| **Average** | 2.1 | **0.1** |  | 2.2 | **-0.4** |  | 13.2 | **-0.3** |  | 82.0 | **-0.4** |  | 87.6 | **-0.9** |
| **Negative Z scores N (%)** | 14 (**42%**) | |  | 21 (**63%**) | |  | 21 (**63%**) | |  | 15 (**45%**) | |  | 21 (**63%**) | |

Table 4. Blocked Cyclic Naming: RTs= reaction times in millsec. Semantic interference is measured as the difference between related and unrelated sets. Related sets are faster for the first set and then become progressively slower. Number of participants with Z score ≤ 2 SD from the control mean (indicating poor performance) are also reported.

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | **Controls**  **(=24)** | |  | **AwPKU**  **N=15)** | |  | **Controls vs AwPKU** |  | **N of poor scores (**≤ 2 SD) | |
|  | **Mean** | **SD** |  | **Mean** | **SD** |  | ***t* and *p*** |  | **Controls** | **PKU** |
|
| **Age** | **28.2** | *9.1* |  | **25.1** | *8.7* |  | t(1,37)= -1.0 p= .81 |  |  |  |
| **Education years** | **14.2** | *2.0* |  | **14** | *1.7* |  | t(1,37)= -.4 p= .57 |  |  |  |
| **Gender (M/F)** | **8/16** |  |  | **5/10** |  |  |  |  |  |  |
| **Average RTs** | **679.3** | *143.1* |  | **704.5** | *147.3* |  | t(1,37)= -.5; p= .60 |  | 1 | 1 |
| **Total Errors** | **10.7** | *7.1* |  | **12.8** | *9.5* |  | t(1,37)= -.8; p= .44 |  | 1 | 1 |
| **Semantic Interference** | | |  |  |  |  |  |  |  |  |
| **1st cycle** |  |  |  |  |  |  |  |  |  |  |
| RT | **-27.5** | *76.2* |  | **-42.6** | *83.1* |  | t(1,37)= -.5; p= .56 |  | 0 | 0 |
| Errors | **0.2** | *1.8* |  | **0.4** | *2* |  | t(1,37)= .3; p= .71 |  | 1 | 1 |
| **2nd cycle** |  |  |  |  |  |  |  |  |  |  |
| RTs | **18.2** | *70.9* |  | **31.5** | *51* |  | t(1,37)= .6; p= .53 |  | 1 | 0 |
| Errors | **1** | *1.8* |  | **1.5** | *2* |  | t(1,37)= .6; p= .49 |  | 1 | 2 |
| **3rd cycle** |  |  |  |  |  |  |  |  |  |  |
| RTs | **27.1** | *50.7* |  | **53.8** | *103.2* |  | t(1,37)= 1.08; p= .28 |  | 0 | 1 |
| Errors | **1.3** | *2.0* |  | **0.7** | *1.4* |  | t(1,37)= -.8; p= .38 |  | 1 | 0 |
| **4th cycle** |  |  |  |  |  |  |  |  |  |  |
| RTs | **48.2** | *55.2* |  | **43.3** | *45.6* |  | t(1,37)= -.3; p= .77 |  | 1 | 0 |

Table 5. Pragmatic skills, prosody and search strategy (Hayling test) in controls and PKU participants. Tasks where there is a significant difference with controls are highlighted. Number of participants with Z score ≤ 2 SD from control mean (indicating poor performance) is also reported. ⸸ Inferred meaning N=21;

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | **Controls N=25** | |  | **PKU group**  **N=22** | |  | **Controls vs. PKU** |  | **N of poor scores (**≤ 2 SD) | |
|  | **Mean** | **SD** |  | **Mean** | **SD** |  | ***t* and *p*** |  | **Controls** | **AwPKU** |
| Age | **27.8** | *9.1* |  | **26.1** | *9* |  | t(1,45)= .7; p=.51 |  |  |  |
| Education years | **14.2** | *2* |  | **14** | *1.9* |  | t(1,46)= .4; p=.72 |  |  |  |
| Gender (M/F) | **8/17** |  |  | **8/14** |  |  |  |  |  |  |
| **PROSODY** |  |  |  |  |  |  |  |  |  |  |
| Non-emotional Prosody | |  |  |  |  |  |  |  |  |  |
| Errors | **0.4** | *0.8* |  | **0.2** | *0.5* |  | t(1,45)= .7; p= .49 |  | 2 | 1 |
| Emotional Prosody |  |  |  |  |  |  |  |  |  |  |
| Errors | **0.3** | *0.5* |  | **0.6** | *1.1* |  | t(1,45)= -1.1; p= .28 |  | 1 | 2 |
| Conflicting Prosody |  |  |  |  |  |  |  |  |  |  |
| Errors | **1.4** | *1.9* |  | **1.1** | *1.2* |  | t(1,45)= .7; p= .51 |  | 1 | 0 |
| **COMPLEX LANGUAGE TASKS** | | | | |  |  |  |  |  |  |
| Metaphor Interpretation | |  |  |  |  |  |  |  |  |  |
| Errors | **2.3** | *3.1* |  | **1.1** | *1.7* |  | t(1,45)= 1.6; p= .12 |  | 2 | 0 |
| Total time (sec) | **41** | *13* |  | **58.9** | *34.3* |  | **t(1,45)= -2.4; p < .05** | | 0 | 4 |
| Appreciation of Humour | |  |  |  |  |  |  |  |  |  |
| Errors | **2.2** | *1.6* |  | **2** | *1.3* |  | t(1,45)= .6; p= .57 |  | 1 | 0 |
| Total time (sec) | **76.9** | *23* |  | **91.6** | *44* |  | t(1,45)= -1.5; p= .15 |  | 1 | 2 |
| Comprehension of Inferred Meaning | | | | |  |  |  |  |  |  |
| Errors | **1.2** | *1.2* |  | **1.3** | *1.2* |  | t(1,44)= -.3; p= .79 |  | 1 | 1 |
| Total time (sec) | **119.2** | *40* |  | **157.9** | *54.6* |  | **t(1,44)= -2.8; p < .01** | | 0 | 4 |
| Hayling Sentence Completion | | |  |  |  |  |  |  |  |  |
| Part 1 Errors | **1.3** | *0.5* |  | **1.4** | *0.6* |  | t(1,45)= -.6; p= .58 |  | 0 | 1 |
| Part 1 Time (sec.) | **16** | *7.8* |  | **24.9** | *21.9* |  | t(1,45)= -1.9; p = .07 |  | 2 | 5 |
| Part 2 Errors A | **0.4** | *0.8* |  | **1** | *1* |  | **t(1,45)= -2.1; p < .05** | | 2 | 4 |
| Part 2 Errors B | **1.2** | *1.6* |  | **2.4** | *1.8* |  | **t(1,45)= -2.5; p < .05** | | 0 | 4 |
| Part 2 Total Errors | **1.6** | *2.1* |  | **3.4** | *2.5* |  | **t(1,45)= 2.6; p < .01** | | 1 | 5 |
| Part 2 Time (sec.) | **29.6** | *17* |  | **48.0** | *37.1* |  | **t(1,45)= -2.2; p < .05** | | 2 | 3 |
|  |  |  |  |  |  |  |  |  |  |  |
| **Total for complex language tasks** ( Hayling test only includes part 2) | | | | | | | |  | 8 | 19 |

Table 6. Individual Z scores for the PKU group. Cells are left blank when participant did not undertake the task. Highlighted scores are ≤ 1.5 Z scores. Where needed, Z scores have been signed reversed, so that negative Z score always indicate worse performance compared to controls.

\* p < .05; \*\* p < .01 according to single case analysis implemented in the program Singlims\_ES.exe (Crawford, Garthwaite, & Porter, 2010).

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **AwPKU** | **Average complex language tasks** (N=38) | **CIU** (N=33) | **Hayling Test**  **Part 2** (N=22) | |  | | | **Humour** (N=22) | |  | **Metaphors** (N=22) | |  | **Inferred Meaning** (N=21) | |
|  |  |  | **Errors** | **Time** | |  | **Errors** | | **Time** |  | **Errors** | **Time** |  | **Errors** | **Time** |
| 1 | -2.60\* | -2.60\* |  |  | |  |  | |  |  |  |  |  |  |  |
| 2 | -0.27 | -0.60 | 0.76 | -1.11 | |  | 0.12 | | -0.51 |  | 0.74 | -0.99 |  | 1.06 | -1.88 |
| 3 | -0.60 | -0.60 |  |  | |  |  | |  |  |  |  |  |  |  |
| 4 | -2.44\* | -4.50\*\* | -2.57\* | -3.45\*\* | |  | 0.73 | | -3.66\*\* |  | 0.74 | -5.55\*\* |  | 0.21 | -3.88\*\* |
| 5 | -0.31 | -1.40 | -0.67 | -0.02 | |  | -0.49 | | 0.12 |  | 0.10 | -0.59 |  | -0.65 | 0.84 |
| 6 | 1.20 | 1.20 |  |  | |  |  | |  |  |  |  |  |  |  |
| 7 | 0.40 | 0.40 |  |  | |  |  | |  |  |  |  |  |  |  |
| 8 | 0.70 | 0.70 |  |  | |  |  | |  |  |  |  |  |  |  |
| 9 | 0.00 | 0.00 |  |  | |  |  | |  |  |  |  |  |  |  |
| 10 | 0.40 | 0.40 |  |  | |  |  | |  |  |  |  |  |  |  |
| 11 | -1.12 | -4.60\*\* | -2.09\* | -0.05 | |  | -1.10 | | -0.61 |  | 0.42 | -2.27\* |  | 0.21 | 0.05 |
| 12 | -0.79 | -6.10\*\* | -1.14 | -1.16 | |  | 0.73 | | 0.03 |  | 0.42 | -0.21 |  | 1.06 | -0.78 |
| 13 | -0.64 | -0.30 | -0.29 | -0.23 | |  | 0.12 | | -1.71 |  | 0.42 | -0.17 |  | -1.51 | -2.12\* |
| 14 | -1.60 | -1.60 |  |  | |  |  | |  |  |  |  |  |  |  |
| 15 | 0.70 | 0.70 |  |  | |  |  | |  |  |  |  |  |  |  |
| 16 | -1.10 | -1.10 |  |  | |  |  | |  |  |  |  |  |  |  |
| 17 | -0.46 | -0.60 | -0.67 | -0.47 | |  | -0.49 | | 0.11 |  | 0.74 | -1.86 |  |  |  |
| 18 | -0.60 | -0.60 |  |  | |  |  | |  |  |  |  |  |  |  |
| 19 | -1.10 | -1.10 |  |  | |  |  | |  |  |  |  |  |  |  |
| 20 | -0.23 | 0.20 | -1.14 | -0.64 | |  | 1.35 | | -1.17 |  | 0.42 | -1.10 |  | -0.65 | 0.68 |
| 21 | -0.90 | -0.90 |  |  | |  |  | |  |  |  |  |  |  |  |
| 22 | -0.10 | -0.10 |  |  | |  |  | |  |  |  |  |  |  |  |
| 23 | -0.20 | -0.20 |  |  | |  |  | |  |  |  |  |  |  |  |
| 24 | -0.25 | 0.20 | -1.14 | -0.45 | |  | -0.49 | | 1.09 |  | -1.82 | 1.61 |  | -0.65 | -0.57 |
| 25 | -2.06\* | -2.50\* | -3.52\*\* | -9.42\*\* | |  | -0.49 | | -0.94 |  | 0.74 | -1.92 |  | 1.06 | -1.53 |
| 26 | 0.28 | -0.70 | 0.29 | 0.45 | |  | 0.73 | | 1.20 |  | 0.74 | 0.19 |  | 0.21 | -0.59 |
| 27 | -0.03 | 1.10 | -0.19 | 0.31 | |  | 0.73 | | 0.10 |  | 0.74 | -2.43\* |  | -0.65 | 0.05 |
| 28 | 0.70 | 0.70 |  |  | |  |  | |  |  |  |  |  |  |  |
| 29 | -0.09 | 0.00 | -2.09\* | -1.17 | |  | 1.35 | | 0.11 |  | 0.74 | -0.68 |  | 1.06 | -0.13 |
| 30 | 0.10 | 1.10 | 0.29 | 0.78 | |  | 0.73 | | -0.92 |  | 0.42 | -1.71 |  | 0.21 | 0.03 |
| 31 | -0.48 |  | -2.09\* | -0.49 | |  | 0.73 | | -0.23 |  | 0.10 | -0.27 |  | 0.21 | -1.79 |
| 32 | -1.50 | -1.80 | -1.62 | -3.64\*\* | |  | 0.73 | | -1.25 |  | -0.22 | -0.90 |  | -0.65 | -4.13\*\* |
| 33 | -0.18 | -2.90\*\* | -0.19 | -0.3 | |  | 0.73 | | 1.17 |  | 0.42 | 1.05 |  | -1.51 | -0.07 |
| 34 | 0.33 | -1.10 | 0.76 | 0.7 | |  | 0.12 | | 1.18 |  | 0.42 | 1.64 |  | 0.21 | -0.98 |
| 35 | 0.32 |  | 0.29 | 0.5 | |  | -1.10 | | 0.41 |  | 0.42 | 0.34 |  | 1.06 | 0.62 |
| 36 | -0.94 |  | -1.62 | -1.3 | |  | -1.10 | | 0.27 |  | 0.74 | -1.78 |  | -2.37 | -0.35 |
| 37 | -2.60\* |  | 0.29 | -1.3 | |  | -0.49 | | -7.56\*\* |  | 0.42 | -10.77\*\* |  | 1.06 | -2.42\* |
| 38 | -0.81 |  | -0.67 | -1.2 | |  | 0.12 | | -1.34 |  | 0.42 | -1.59 |  | -0.65 | -1.59 |
| **Average** | **-0.50** | **-0.88** | **-0.86** | **-1.08** | |  | **0.15** | | **-0.64** |  | **0.38** | **-1.36** |  | **-0.08** | **-0.98** |
| **AwPKU vs Control** | **t(1,75)= -2.1; p= .03** | **t(1,63)= -2.6; p= .01** | **t(1,45)= -2.9; p= .006** | **t(1,45)= -2;  p= .05** | |  | t(1,45)= 1.2; p= .22 | | t(1,45)= -1.7; p= .09 |  | t(1,45)= 1.1; p= .3 | **t(1,45)= -2.4; p= .02** |  | t(1,44)= -.9;  p= .35 | **t(1,44)= -2.7; p= .008** |
| **N (%) of negative z scores** | 27 (**71**%) | 21 (**63**%) | 16 (**72**%) | 17 (**77**%) | |  | 8 (**36**%) | | 11 (**50**%) |  | 2 (**9**%) | 17 (**77**%) |  | 9 (**42**%) | 15 (**71**%) |

Figure 1. Blocked cyclic naming semantic interference (related-unrelated) expressed in terms of speed and accuracy (RTs) across cycles. As expected. related set are faster in the first cycle. but then become progressively slower in both groups.

Appendix 1. Summary of the literature investigating linguistic functions in individuals with PKU.

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **EXPRESSIVE LANGUAGE & NAMING** | | | | | | | | | | |
| **Authors** | **Dev stage** | **Age range** | **N -PKU sample** | **Control Sample** | **Current or recent Phe level** | **Tasks** | **Impairment?** | **Correlated with Blood Phe?** | **Time of Phe-level correlation** | **Note** |
| **Brumm et al. 2004** | Adults | 21-32 yrs | 24 | Normative data | Range current Phe level: 157 - 1713 μmol/L (mean=1038) | Boston Naming Test (BNT) | YES | YES (childhood and concurrent levels only) | Childhood Phe levels (5.5-6 years; 9.5-10 years); Concurrent Phe; Brain Phe levels |  |
| **Faust et al. 1987** | Children and adolescents | 7-19 yrs | 10 | Normative data | n/a | No details provided. Tasks modified from Luria et al.. 1980 | NO | n/a |  |  |
| **Gejão et al. 2009** | Children | 1-120 mths | 25 | Normative data + 43 Children with hypothyroidism | n/a  . | Early Language Milestone Scale (ELMS); Language Development Evaluation Scale (LDES); Child Language Test Phonology (ABFW); Illinois Test of Psycholinguistic Abilities (ITPA)- verbal expression | NO for Phonology; YES for language use | n/a  . |  |  |
| **Melnick et al. 1981** | Children | 4-6 mths | 12 | Normative data | n/a | Bzoch-League Receptive-Expressive Emergent Language scale; Northwestern Syntax Screening Test (expressive portion) | YES (in 6/12 children) | n/a |  |  |
| **Ozanne et al. 1990 \*a** | Children | 7-47 mths | 11 | Outpatients and orthopaedic cases |  | Sequenced Inventory of Communication Development  (SICD) | NO – *but* two patients (18%) showed performance <75% of their chronological age | n/a |  |  |
| **Ozanne et al. 1990\*a** | Children and adolescents | 5-16yrs | 18 | Outpatients and orthopaedic cases |  | Test of Language Development (TOLD) | NO – *but* three (16%) patients showed performance below 10th percentile | n/a |  |  |
| **Palermo et al.. 2017** | Adults | 18-41 yrs | 37 | 30 healthy participants | Concurrent Phe level: 787 μmol/L ± 301 | Picture naming (accuracy and RTs); Colour Naming (accuracy and RTs)  Phonological processing (accuracy): Phoneme deletion. Spoonerisms deletion. Spoonerisms | YES in colour naming (RTs);  NO in accuracy tasks | n/a |  |  |
| **Soleymani et al. 2015** | Children | 4-6.5 yrs | 8 early treated. 9 late treated < 1 yr. 13 late-treated >1yr | 42 healthy participants | Range recent Phe level (six months before): 2.50-14.5 mg% | Test of Language Development-Primary (TOLD-P) | YES in all groups (early treated better than late treated < 1 yr and >1yr) | Yes in late treated < 1 yr; NO in the other groups | Recent Phe (six months before) |  |

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **SPONTANEOUS SPEECH** | | | | | | | | | | |
| **Authors** | **Dev stage** | **Age range** | **N -PKU sample** | **Control Sample** | **Current or recent Phe level** | **Tasks** | **Impairment?** | **Correlated with Blood Phe?** | **Time of Phe-level correlation** | **Note** |
| **Faust et al.**  **1987** | Children | 7-19 yrs | 10 | Normative data | n/a | No details provided. Tasks modified from Luria et al.. 1980 | NO | n/a |  |  |
| **Gassio et al. 2005** | Children and adolescents | 7-19 yrs | 26 | 21 healthy participants | Average Phe in the last six months: 449 μmol/L ±220 in 13 PKU; 529 μmol/L ±142 in 13 PKU with school problems | Spontaneous language assessment (form. content. use). No specific details or reference provided | NO | n/a |  |  |
| **Melnick et al. 1981** | Children | 4-6 mths | 12 | Normative data | n/a | Developmental Sentence Analysis using a spontaneous language sample of 50 consecutive utterances (Lee et al.. 1974); Mean Length of Utterance | YES in 6/12 children | n/a |  | Results not provided for spontaneous speech alone. but assembled with other formal tests of expressive language |

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **COMPREHENSION AND RECEPTIVE LANGUAGE** | | | | | | | | | | |
| **Authors** | **Dev stage** | **Age range** | **N -PKU sample** | **Control Sample** | **Current or recent Phe level** | **Tasks** | **Impairment?** | **Correlated with Blood Phe?** | **Time of Phe-level correlation** | **Note** |
| **Brumm et al. 2004** | Adults | 21-32 yrs | 24 | Normative data | Range Current Phe: 157 - 1713 μmol/L (mean=1038) | Peabody Picture Vocabulary Test Revised (PPVT-R) | NO | YES (childhood and testing time Phe-levels) | Childhood Phe levels (5.5-6 years; 9.5-10 years); Concurrent Phe; Brain Phe levels |  |
| **Faust et al.**  **1987** | Children and adolescents | 7-19 yrs | 10 | Normative data | n/a | No details provided. Tasks modified from Luria et al.. 1980 | NO | n/a |  |  |
| **Gejão et al. 2009** | Children | 1-120 mths | 25 | Normative data + 43 Children with hypothyroidism | n/a | Peabody Picture Vocabulary Test (PPVT); Language Development Evaluation Scale (LDES) | YES | n/a |  |  |
| **Melnick et al. 1981** | Children | 4-6 mths | 12 | Normative data | n/a | Bzoch-League Receptive-Expressive Emergent Language scale; Peabody Picture Vocabulary Test (PPVT); North Western Syntax Screening Test (receptive portion) | YES in 4/12 children | n/a |  |  |
| **Ozanne et al. 1990\*a** | Children | 7-47 mths | 11 | Outpatients and orthopaedic cases |  | Sequenced Inventory of Communication Development  (SICD) | NO – *but* three patients (27%) showed performance <75% of their chronological age | n/a |  |  |
| **Ozanne et al. 1990\*a** | Children and adolescents | 5-16yrs | 18 | Outpatients and orthopaedic cases |  | Test of Language Development (TOLD) | NO – but three (16%) patients showed performance below 10th percentile | n/a |  |  |
| **Soleymani et al. 2015** | Children | 4-6.5 yrs | 8 early treated. 9 late treated < 1 yr. 13 late-treated >1yr | 42 healthy participants | Range recent Phe level (six months before): 2.50-14.5 mg% | Test of Language Development-Primary (TOLD-P) | YES in all groups (early treated better than late treated < 1 yr and >1yr) | YES in late treated < 1 yr; NO in the other groups | Recent Phe levels |  |
| **Zartler et al. 1981** | Children | n/a | n/a | normative data | n/a | Peabody Picture Vocabulary Test (PPVT) | NO | n/a |  | It is a comment on the paper by Melinck et al.. 1981. |

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **READING AND SPELLING** | | | | | | | | | |  |
| **Authors** | **Dev stage** | **Age range** | **N -PKU sample** | **Control Sample** | **Current or recent Phe level** | **Tasks** | **Impairment?** | **Correlated with Blood Phe?** | **Time of Phe-level correlation** | **Note** |
| **Azen et al.**  **1996\*b#** | Children | 8 yr -longitudinal study\*\* | 54 | 54 Siblings | n/a | Wide Range Achievement Test (WRAT; Reading and Spelling) | YES in Reading; NO in Spelling | YES with Sibling-PKU differences | Current Phe | Age at dietary discontinuation was a good predictor of poorer reading and spelling. |
| **Brumm et al.**  **2004** | **Adults** | 21-32 yrs | 24 | Only normative data | Current Phe: 157 - 1713 μmol/L (mean=1038) | Wide Range Achievement Test (WRAT; Reading and Spelling) | NO  (one participant showed borderline performance) | NO | Childhood Phe levels (5.5-6 years; 9.5-10 years); Concurrent Phe; Brain Phe levels. | No significant correlations between cognitive variables and brain Phe levels. |
| **Chang et al. 2000** | Children and adolescents | 7-18yrs | 32 | Only normative data | Recent Mean Phe-level 13.6 mg/dl | Woodcock-Johnson Revised Test of Achievement (Letter-word identification; passage comprehension; dictation; writing samples) | YES in dictation and writing samples (Mild impairment);  NO in passage comprehension | YES in passage comprehension;  NO with other tasks | Recent Phe-level |  |
| **Faust et al.**  **1987** | Children and adolescents | 7-19 yrs | 10 | Only normative data | n/a | Reading and Spelling; No details provided. Tasks modified from Luria et al.. 1980 | YES | n/a |  | Author noted that problems in this area were secondary to problem in other areas (visuo-motor control. spatial analysis) |
| **Fishler et al. 1987\*c#** | Children | 6 to 10 yrs -longitudinal study\* | 105 (33 on diet -72 off diet) | Normative data + 51 siblings | *On diet:* average Phe levels at 10 years 11.8 mg/dl ± 3.5. *Off-diet*: average Phe levels at 10 years 25.2 mg/dl ±5.5 | Wide Range Achievement Test (WRAT; Reading and Spelling) | NO compared to normative data;  YES compared to siblings | YES with Sibling-PKU differences;  On-diet groups showed better performance. | Current Phe;  Off diet vs On diet. |  |
| **Koch et al. 2002 #** | **Adults** | Longitudinal study over a period of 32 years | 63 | Not reported | Current Phe: from 1448 μmol/L ±465 in PKU who discontinued the diet at 4.8-6.5 yrs to 926 μmol/L ±268 in PKU on diet. | Wide Range Achievement Test (WRAT; Reading and Spelling) | Not reported | YES (Phe at 6 yrs).  Current Phe level: no differences between adults with Phe < 960 μmol/L and >960 | Phe at 6 yrs; Phe at 10 yrs; Phe at 10; Phe at 18; adult Phe.  Adults with current Phe < 960 μmol/L vs >960 μmol/L |  |
| **Palermo et al. 2017** | Adults | 18-41 yrs | 37 | 30 Controls | Concurrent Phe level: 787 μmol/L ± 301 | Words and Nonwords spelling (accuracy); Words and Nonwords reading (accuracy and RTs) | YES in reading RTs;  NO in reading/spelling accuracy | n/a | Average Phe levels/fluctuations in Childhood (0-10 years). Adolescence (11-16 years). Adulthood (17 year-present). and Concurrent Phe. |  |
| **Pennington et al. 1985** | Children | 9-14 yrs | 6 | 12 Brain-damaged children | Life Phe range: 4.8- 18.9 mgm% (while on diet); 17.6-29.2 mgm% (while off diet) | Peabody Individual Achievement Test (PIAT; includes: Reading Recognition. Reading Comprehension. Spelling) | Better than brain damaged children | n/a |  |  |

\*a Ozanne et al. 1990 here. each experiment is cited separately for clarity. while it is considered as one study when listed in the text. \*b Azen et al. 1996 described data from a longitudinal study. Here we considered only the data for the age band (8 years) in which there was a comparison with sibling on the WRAT scale. \*c Fishler et al. 1987 is a longitudinal study. The same participants evaluated 5 times from 6 to 10 years on the WRAT scale. The comparisons with siblings are reported in Koch et al.. 1984 and Holtzman et al., 1986. # Data from the longitudinal collaborative study of children treated for Phenylketonuria (PKUCS).

Appendix 2. Macrostructural analysis of Verbal Narratives (Cinderella story). The table reports the 16 main concepts used for the analysis of the verbal narratives. adapted from Bartels-Tobin and Hinkley (2005).

|  |  |
| --- | --- |
|  | **Narrative main concepts/units of information** |
| 1 | Cinderella has to go to live with two step-sisters and step-mother |
| 2 | They’re mean. Cinderella does all work |
| 3 | Announcement to dance arrives |
| 4 | Cinderella wants to go to the ball but stepmother does not allow |
| opt | Step-sisters/step-mother go to ball |
| 5 | Fairy Godmother appears. |
| 6 | She transforms Cinderella. pumpkin. etc. |
| 7 | She warns Cinderella that at midnight everything has back to normal |
| 8 | Cinderella goes to/at ball |
| 9 | Cinderella and prince dance |
| opt | step mother and sisters are jealous |
| 10 | Cinderella has to get home (it’s midnight) |
| 11 | Cinderella runs. loses slipper |
| 12 | Everything goes back to normal |
| 13 | Prince brings shoe around town |
| 14 | Step-sisters/step-mother try on slipper. doesn’t fit |
| 15 | Cinderella tries on slipper. it fits |
| 16 | Cinderella and prince marry |

1. Note, however, one of these participants arrived in the UK as teenager. However, he had very good English and attended University. Moreover, our results are in z-scores from controls which, for fair comparisons, always included an individual with a similar history. All other participants were always resident in the UK. [↑](#footnote-ref-1)