Contents lists available at ScienceDirect

Brain & Language

Putative sex differences in verbal abilities and language cortex: A critical review

Mikkel Wallentin *

Center for Semiotics, University of Aarhus, Denmark
Center of Functionally Integrative Neuroscience, Aarhus University Hospital, Aarhus Sygehus, byg. 30, Noerrebrogade, 8000 Aarhus, Denmark

ARTICLE INFO

ABSTRACT

This review brings together evidence from a diverse field of methods for investigating sex differences in language processing. Differences are found in certain language-related deficits, such as stuttering, dyslexia, autism and schizophrenia. Common to these is that language problems may follow from, rather than cause the deficit. Large studies have been conducted on sex differences in verbal abilities within the normal population, and a careful reading of the results suggests that differences in language proficiency do not exist. Early differences in language acquisition show a slight advantage for girls, but this gradually disappears. A difference in language lateralization of brain structure and function in adults has also been suggested, perhaps following size differences in the corpus callosum. Neither of these claims is substantiated by evidence. In addition, overall results from studies on regional grey matter distribution using voxel-based morphometry, indicate no consistent differences between males and females in language-related cortical regions. Language function in Wada tests, aphasia, and in normal ageing also fails to show sex differentiation.

© 2008 Elsevier Inc. All rights reserved.

1. Introduction

Sex is interesting, and the mass media and the general public are captivated by findings of differences between men and women. The book *Men Are From Mars, Women Are From Venus*, by John Gray (1992), which argued for large psychological differences between the sexes, has been translated into more than 40 languages and has sold over 30 million copies (http://www.marsvenus.com).

When cognitive neuroscientists are approached by the mass media, it is often with questions related to sex differences.

Textbooks routinely cite sex differences in language competence, usually regarding verbal fluency, as established fact (e.g. Kolb & Whishaw, 2001, p. 558; Mildner, 2008, p. 41; Pinker, 2007, pp. 85–86). Unfortunately, it is often difficult to see where these findings come from; this is not only the case in textbooks, but can also be found in research papers on the topic. For instance, in a study on sex differences (Weiss, Kemmler, Deisenhammer, Fleischhacker, & Delazer, 2003), one can find the following statement, without any data or references to back the claim:

> Women tend to be better than men in rapidly identifying matching items, a skill called perceptual speed. Common linguistic skills, in which females have been found to be superior, are verbal fluency, speech articulation, grammatical skills, and use of more complex and longer sentences.

Similar statements are found in all of the above citations.

A large number of studies reporting sex differences in brain structure and function underlying language processes have also been published (e.g. Shaywitz et al., 1995). The article by Shaywitz and colleagues, suggesting differences in language lateralization, has been cited more than 500 times, indicating the impact that these types of results have on the scientific community, not to mention the broader public opinion.

Most neuroimaging studies, however, do not distinguish between males and females. But if sex differences are real, they may possibly confound results from language studies obtained using neuroimaging techniques such as PET and fMRI. If so, measures must be taken to address the problem.

This review presents a broad overview of the multiple existing approaches to the investigation of sex differences in language performance and in the underlying brain structure and processing as well as in language-related disorders. An enormous number of such papers exist (e.g. Burman, Bitan, & Booth, 2008; Clements et al., 2006; Frost et al., 1999; Grabowski, Damasio, Eichhorn, & Tranel, 2003; Haut & Barch, 2006; Jaeger et al., 1998; Kaiser, Kuenzli, Zappatore, & Nitsch, 2007; Kansaku, Yamaura, & Kitazawa, 2000; Knecht et al., 2000; Plante, Schmithorst, Holland, & Byars, 2006; Schirmer, Zysset, Kotz, & Yves von Cramon, 2004; Shaywitz et al., 1995; Springer et al., 1999; van der Kallen et al., 1998; Weiss et al., 2003 to name but a few). It is not possible to cover all here and at

* Address: Center of Functionally Integrative Neuroscience, Aarhus University Hospital, Aarhus Sygehus, byg. 30, Noerrebrogade, 8000 Aarhus, Denmark. Fax: +4589494400.
E-mail addresses: mikkel@pet.auh.dk, mikkel.w@llentin.dk

the same time present studies from many fields. This review therefore relies primarily on results either from meta-analyses or from studies with a large number of participants ($n > 100$) when they exist.

1.1. Possible evolutionary origins

Before we look at the evidence it is interesting to ask why anyone thinks that there could be differences between the language systems for males and females at all. Two theories exist that include this hypothesis.

The "hunter-gatherer"-hypothesis (e.g. see Kolb & Whishaw, 2001; Mildner, 2008) states that putative sex differences in cognitive abilities arose from a division of labour between the sexes in prehistoric humans. Men were predominantly hunters, whereas women were predominantly gatherers. According to this line of thinking, males explored larger territories during hunting than females would do during gathering, which in turn made males develop better skills for navigation, whereas the females left behind in social groups would have benefited more from developing fine-tuned tools for social interaction, one of which is language. This theory therefore predicts that women are better at language than men.

Another possible evolutionary origin of cognitive sex differences has been suggested by Miller (2000). According to this theory, language may have evolved at least partly for purposes of sexual display. This means that language ability is used in sexual selection as a fitness-indicator. Individuals displaying a large capacity for language are considered fitter by potential partners compared to less eloquent individuals. According to this theory, language is the human equivalent of peacock feathers. Due to the special dynamics of sexual selection there are great differences between the feathers of a male peacock and those of a female. Sexual selection dynamics require that males, at great costs, try to display fitness, e.g. by walking around dressed in dangerously visible colours, while females have the power to decide which male to mate with and must therefore be good detectors of fitness. Evoking such a theory about human language therefore predicts great differences between male and female language capabilities. Only fit males will have the means and the time to develop and produce inciting language. The prediction is thus that males will be better at producing language and will try at great costs to exhibit this, while females, on the other hand, are better at understanding language in order to distinguish good talkers from bad talkers (Miller, 2000). Anecdotal evidence for this theory includes the observation that acclaimed writers are still predominantly male, whereas women are known to buy and read more books than men (e.g. Halpern, 2000).

When applied to language, however, both of these theories suffer from fact that this unique part of human behaviour per definition leaves very little room for cross-species comparison. This makes the theories very difficult to falsify or even investigate scientifically (Gannon, 2002). They are, nevertheless, part of the background for most of the research conducted on sex differences in language and therefore deserve to be mentioned.

2. Sex differences in the normal population

2.1. Verbal fluency

The task perhaps most cited as yielding sex differences is the verbal fluency task (e.g. Kolb & Whishaw, 2001; Mildner, 2008; Pinker, 2007; Sommer, Aleman, Bouma, & Kahn, 2004). The verbal fluency task is usually conducted in two parts, a "lexical" and a "categorical". In both conditions, subjects are asked to come up with as many words in a minute as possible. In the lexical conditions the words have to start with a particular letter, like F, A or S, and in the categorical condition subjects have to name as many

objects from a specific category as possible in one minute. Usually, the category is animals.

As an example, Weiss et al. (2003) examined 97 college students (51 women, 46 men) and found women to have a significant advantage in the lexical condition ($p < .02$), but no significant difference was seen in the categorical task ($p < .08$). Unfortunately these groups were not balanced on other important demographic measures, such as age ($p < .002$), and men were found on another test to have a significantly higher verbal IQ ($p < .004$). The results on the verbal fluency task, therefore, are not only weak, but also very difficult to interpret given these confounds in the data. Nonetheless, these findings are summarized in the following way in the abstract of the article: "In general, we found, that women tend to perform at a higher level than men on most verbal tests".

Tombaugh and colleagues, on the other hand, measured word fluency in 1300 individuals (Tombaugh, Kozak, & Rees, 1999). They set up booths at shopping centres, social organizations and places of employment. In this very large sample they found absolutely no differences between men and women. In the lexical condition, men on average came up with 37 words whereas women came up with 37.8 words. In the categorical condition, men named 17.4 animals whereas women named 16.5 animals. Neither of these differences was statistically significant. Performance on both tests, on the other hand, showed great, and greatly significant, dependence on age and educational level.

In summary, Sex is not a significant predictor of performance on the verbal fluency task, when using a large sample size, appropriately controlled for confounds such as age and education.

2.2. Overall performance

Hyde and Linn (1988) conducted a meta-analysis on 165 American language studies of both children and adults, including tests of vocabulary, analogies, anagrams, reading comprehension, speaking or other verbal communication, essay writing, Scholastic Aptitude Test (SAT)-Verbal, and general verbal ability tests such as the American College Testing Program Examination-Verbal. Forty-four (27%) of the studies had found females to perform significantly better than males, 109 (66%) found no significant sex difference, and 12 (7%) found males performing significantly better than females. Overall, this led to a slight female advantage, but when effect-size from the different studies was weighted by number of subjects, the effect was reversed, due to the largest study (including more than 900,000 subjects) finding a small male advantage. The authors conclude that: "Our meta-analysis provides strong evidence that the magnitude of the sex difference in verbal ability is currently so small that it can effectively be considered to be zero" (Hyde & Linn, 1988, p. 64). Interestingly, a significant effect of first-author gender was found in the 165 studies. Studies where the first author was a woman were significantly more likely to report a female language advantage than studies where the first author was a man (though both effects were close to zero).

2.3. Language acquisition

2.3.1. First language acquisition

The most frequently used tests for early language development are the MacArthur–Bates Communicative Development Inventories (Fenson et al., 1994). Fenson et al. (1994) studied 1803 and Feldman et al. (2000) 2156 sociodemographically diverse 1- and 2-year-old American children. The children were assessed by their mothers using a checklist questionnaire paradigm. Significant effects of gender were found in both 1- and 2-year olds on both vocabulary comprehension and vocabulary production. Girls scored significantly higher than boys, however, differences were very small, accounting for 1–2% of the variance. These findings, however, have been repli-