Sex differences in language competence of three- to six-year old children

Benjamin P. Lange[1], Harald A. Euler[2], and Eugen Zaretsky[3]

[1] Department of Media Psychology, Institute for Human-Computer-Media, Faculty of Human Sciences, Julius Maximilian University of Wuerzburg, Germany

[2] Department of Developmental Psychology, University of Vienna, Austria

[3] Department of Phoniatrics and Pediatric Audiology, Goethe University of Frankfurt/Main, Germany

Corresponding Author: Dr. Benjamin P. Lange, Department of Media Psychology, Institute for Human-Computer-Media, Faculty of Human Sciences, Julius Maximilian University of Wuerzburg, Oswald-Kuelpe-Weg 82, D-97074 Wuerzburg, Germany. Email: benjamin.lange@uni-wuerzburg.de. Phone: +49 / (0) 931/31-86689. Fax: +49 / (0) 931/31-87364

Abstract

For decades, developmental research has involved the study of sex differences in language acquisition. Many studies of these differences have found a slight advantage in competence for females early in life that seems to wane with age. Yet, since most of these studies have focused on sex differences in mean values, they have mostly neglected sex differences in variance with males being more variable. In the current study, we examined sex differences in language competence in terms of mean values and variance in large samples (total $N >$ 10,000) of German children aged 3–6 years. We administered several tests to assess the childrens' vocabulary, grammar, speech comprehension, pronunciation, and the processing of sentences and nonce words. Girls performed better than boys in all domains, most often to a statistically significant degree, although the effect sizes were small. Differences decreased with age. Boys varied significantly more than girls in their language competence. In response, we discuss explanations for these findings, as well as recommend directions for future research.

Keywords: gender differences; language acquisition; language competence; language development; sex differences; variability

Introduction

Differences between the sexes have been extensively studied for decades (Kimura, 2000; Mealy, 2000; Zell, Krizan, & Teeter, 2015), and among them, language-related sex differences have attracted considerable scientific attention (Dindia & Canary, 2006). Research on the topic has revealed that, to gain and maintain status, men's verbal behavior is oriented more than women's toward competition, dominance, self-assertion, and hierarchy, especially in all-male groups (Eckert & McConnell–Ginet, 2003; Guiller & Durndell, 2007; Lange, 2011; Locke & Bogin, 2006). By contrast, women show more sociability in their communication (e.g., Haas, 1979).

Another branch of such research has revealed that women slightly outperform men on several language tasks, though the difference is usually small and robust for only some verbal fluency tasks (Hyde & Linn, 1988; Kimura, 2000; Maccoby & Jacklin, 1974; Wallentin, 2009). In terms of language acquisition, girls also seem to have a slight advantage over boys (Ullman, Miranda, & Travers, 2008), as Stern and Stern (1907/1987) proposed in the early twentieth century.

Indeed, the younger the child and the more basic the linguistic domain, the stronger this language-related sex difference seems (Wallentin, 2009). At the age of 8–16 months, girls outperform boys in word comprehension and, aged 8–30 months, in word production as well (Fenson et al., 1994). Feldman and colleagues (2000) have shown a female advantage in vocabulary production at the age of 1–2 years, which Galsworthy, Dionne, Dale, and Plomin (2000) and Szagun, Stumper, and Schramm (2009) have supported with similar findings. In a Swedish sample, Berglund and colleagues (2005) revealed a female advantage in both vocabulary production and vocabulary comprehension at the age of 18 months, a result that Bleses and colleagues (2008) similarly found in a Danish sample. The findings of another study involving French-speaking Canadian children aged 8–30 months were also comparable

(Bouchard, Trudeau, Sutton, Boudreault, & Deneault, 2009). Furthermore, as Bornstein and colleagues (2004) have demonstrated, at the age of 2–5 years females have an advantage over males in several linguistic domains, including vocabulary production, verbal comprehension, sentence complexity, and use of expressive language. Girls also reach the milestone of possessing a productive vocabulary of at least 50 words sooner than boys (Nelson, 1973). Importantly, this last threshold is used to distinguish so-called "late talkers" and children of non-delayed linguistic development (Rescorla, Hadicke–Wiley, & Escarce, 1993; Rescorla & Schwartz, 1990).

By contrast, concerning sex differences in the acquisition of morphosyntactic aspects of language and grammar, results either show far less consistency or small effect sizes (Bornstein et al., 2004; Fenson et al., 1994). The general heterogeneity of such findings suggests that, regarding grammar and in later childhood, sex differences are especially minor, if not non-existent (Bornstein et al., 2004; Hayiou–Thomas, Dale, & Plomin, 2012; Wallentin, 2009). Altogether, the various findings arguably support that girls' language acquisition is quicker than boys', though this advantage fades around school age—about 6 years (Beltz, Blakemore, & Berenbaum, 2013; Bornstein et al., 2004; Wallentin, 2009).

Importantly, however, all of the above sex differences refer to mean values, despite the fact that the sexes might also differ in variability (Feingold, 1992, 1995), as the male sex generally shows greater phenotypic variance than the female sex for several traits. Greater variance among males than females has been found, for instance, in terms of verbal proficiency (e.g., Hedges & Nowell, 1995; Strand, Deary, & Smith, 2006). Indeed, men seem to be overrepresented at both ends of the distribution of language competence. On the one hand, nearly all areas of life in which verbal proficiency facilitates social success and status are somewhat dominated by men. Most writers are men (Lange & Euler, 2014; Miller, 1999), for example, an overrepresentation that also characterizes other verbally oriented

occupations, such as those in law, journalism, politics, higher education, and public speaking (Burling, 1986, 2005; Halpern, 2000).

On the other hand, regarding poor verbal abilities at the other end of language competence, evidence of men's overrepresentation—even during childhood—is more convincing (Zubrick, Taylor, Rice, & Slegers, 2007). It is well known that language impairments, disorders, and pathologies, including dyslexia, stuttering, and developmental language disorders, are more prevalent in males than females (Andrews, 1964; Arden & Plomin, 2006; Ardila, Rosselli, Bateman, & Guzmán, 2000; Bloodstein, 1995; Canning & Lyon, 1989; Felsenfeld et al., 2000; Halpern, 2000; Thomson & Polnay, 2002). Such findings indicating that more males than females are affected by language problems are furthermore quite robust, as research has long shown (Eme, 1979; Fairweather, 1976; Ingram, 1975; Jespersen, 1922).

These examples of greater variance among males than females with respect to language are moreover supported by statistical analyses. For instance, Strand et al. (2006) demonstrated greater male than female variance regarding verbal classification, sentence completion, and verbal analogies. In addition, Hedges and Nowell (1995) found greater variance among males in terms of vocabulary, reading, and writing, similar to what Reynolds, Hewitt, Erickson, and Silberg (1996) found in terms of oral reading ability.

To our knowledge, however, only Eriksson et al. (2012) have studied sex differences according to variance in the language acquisition of children (aged 1–2 years). Moreover, the study did not find any evidence to suggest greater variance among males than females, at least not at such a young age. It is therefore possible that sex differences in variance do not emerge until around the age of 3 years (Arden & Plomin, 2006; Eriksson et al., 2012). At the same time, it could simply be too difficult to measure such young childrens' language competence to sufficiently detect sex differences in variance before the age of 3 years.

Since language-related sex differences in variance have been widely ignored, especially regarding early language development, the objective of this study was to assess sex differences—in terms of both means and variance—in language acquisition in several linguistic domains (e.g., vocabulary and grammar) among a representative sample of German preschool children aged 3–6 years. In so doing, we sought to discover whether not only sex differences in mean values but also in variance emerged in different domains of language development. The age range in our research (3–6 years) seemed more appropriate based on previous research on sex differences in language acquisition, as reported above.

Methods

Participants

The present study is a retrospectively conducted analysis of all available data originally collected during the development of a language screening test (Euler et al., 2010; Neumann et al., 2011a). Data were collected in several towns and rural areas in the states of Hesse and North Rhine–Westphalia in Germany during 2007–2012. As Table 1 shows, there were three parts with different test batteries and samples: Sample 1, Sample 2, and Sample 3. The study thus consisted of three samples because data were collected on several occasions during the construction of the language screening test.

--- Table 1 about here ---

For Sample 1, speech and language tests were administered in day-care centers in Hesse as part of a universal language screening. Participants were therefore recruited regardless of variables such as socioeconomic or ethnic background and the native language of parents. Since nearly all children in Germany attend day-care centers—2010 data shows

percentages ranging from 89% for 3-year-old to 96% for 4- and 5-year-old children (Statistisches Bundesamt, 2012)—Sample 1 can be considered unselected.

Contrary to Sample 1, participation in Samples 2 and 3 was not obligatory and required the informed consent of parents or caregivers. For Sample 2, letters were written to more than 1,000 day-care centers. Both the day-care centers and families were contacted without any preselection, though some facilities (e.g., Catholic ones) and parents (e.g., those with a very limited command of the German language) tended to abstain from participation in the study. The number of participants per day-care center varied greatly, from one to 60, depending on its size and enrollment of children of the required age group. For Samples 1 and 3, dropout from the study was rare since children were tested and, if necessary, retested until they completed the entire battery.

Participation in the study for Sample 2 was predominantly organized by the University Hospital of Frankfurt am Main, Germany, whereas that for Sample 1 was mostly organized by various public health departments. Therefore, although most children were tested in Hesse, the geographical overlap between the test locations of Samples 1 (predominantly Frankfurt am Main) and 2 (predominantly beyond Frankfurt am Main) was minimal.

For Sample 3, the parents of all children who had been invited to three public health departments for the nationally mandatory school enrollment examination were contacted. The tests administered in Sample 3 were selected to improve the typically brief and thus unreliable screening tests of language skills in the school enrollment examination. Accordingly, parents who gave their informed consent to participate in the study for Sample 3 might have been more concerned with the language development of their children than those in Samples 1 and 2. Some indirect evidence supports this assumption—namely, that the rate of children who stuttered was, at least on a descriptive level, greater in Sample 3

(5%) than in Samples 1 (1%) and 2 (2%). All samples consisted of both monolingually German and bi- or multilingual children.

Though children in Sample 1 did not reappear in either Sample 2 or Sample 3, 158 children from Sample 2 had also been invited to the public health departments for their school enrollment examination as tests for Sample 3 were being conducted there. As such, these children (94 males, 64 females) were retested for Sample 3. The timespan between participation in the tests in Samples 2 and 3 ranged from 0–33 months, with a median of 15 months.

Language Tests and Questionnaires

All participating children were tested with several language tests. The diversity of tests employed stemmed from the fact that data used in the current study originated during the development of the abovementioned language screening test, for which several different language tests were used. All available tests were reused for the statistical analysis of sex-related differences in language competence because it was possible that two tests addressing the same linguistic domain would yield contradictory results. The same reasoning motivated the use of not one, but all three available samples for this retrospective study.

Depending on their respective sample, children were tested with a validated and modified version of the language test *Marburger Sprachscreening* (MSS; Euler et al., 2010; Neumann et al., 2011a) or other validated speech and language tests or both. Pronunciation (i.e., phonetic and phonological skills) was tested with the common German-language tests MSS, the *Psycholinguistische Analyse kindlicher Sprechstörungen* (PLAKSS; Fox, 2005), and the *Screening des Entwicklungsstandes bei Einschulungsuntersuchungen* (S-ENS; Döpfner et al., 2005). The MSS includes, among others, a spontaneous speech task with two items: First, children were asked to describe what they see on the MSS picture and, some items later, they were allowed to ask the language tester anything they wanted in regard to the

same picture. Grammar was also examined with the MSS, as well as the *Sprachentwicklungstest für drei- bis fünfjährige Kinder* (SETK 3–5; Grimm, 2001), additional validated S-ENS items (S-ENS add.; Neumann et al., 2011b), and the *Entwicklungstest Sprache für Kinder von 4 bis 8 Jahren* (ETS 4–8; Angermeier, 2007). Meanwhile, tasks concerning vocabulary were part of the MSS, the *Aktiver Wortschatztest für 3- bis 5-jährige Kinder–Revision* (AWST-R; Kiese–Himmel, 2005), and the S-ENS add. Subtests addressing speech comprehension were part of the MSS, ETS 4–8, and SETK 3–5. The MSS, S-ENS, S-ENS add., SETK 3–5, ETS 4–8, and AWST-R use color illustrations of many child-relevant activities and objects to initiate utterance and test various abilities. The MSS was excluded from the test battery for Sample 3 because it was designed for younger children.

The MSS and S-ENS also examine phonological short-term memory by way of the repetition of sentences and nonce words. For Sample 1, however, the repetition of sentences and nonce words was not part of the MSS. At the same time, the Mottier test (Mottier, 1951), which examines phonological short-term memory and speech-motoric coordination and consists of nonce words that children have to repeat, was used in its updated version (Kiese–Himmel & Risse, 2009).

For the AWST-R, two different short versions with partly different items were designed and validated: one for Sample 2 (i.e., younger children) and one for Sample 3 (i.e., older children). The original German version of the AWST-R contains 75 items and is thus too extensive to be part of a test battery for preschool children. The most appropriate items were selected for inclusion on the basis of data collected in previous studies. The internal consistency of the abbreviated test AWST-R showed a Cronbach's alpha of .94 (cf. alpha of .95 for a full test version in our sample of 208 children). For further details on the included samples, see Table 1.

Questionnaires for parents (Samples 2 and 3) and day-care center teachers (Samples 1 and 2) were also developed during the construction of the abovementioned language screening test (see Appendix). Apart from addressing the demographic and sociolinguistic characteristics of children, these questionnaires sought the subjective judgments of day-care center teachers and parents regarding children's language competence (e.g., whether the child uses articles correctly).

Children's compliance was tested using a 3-point ordinal scale (*good*, *average,* and *poor*) with the MSS in Samples 1 and 2 and with the S-ENS add. in Sample 3. Poor compliance indicated that the child did not cooperate with the tester and that it was almost impossible to complete all tasks with the child.

Expert panels were also held for all samples to assess the children's language development based on questionnaires and language tests, including audio records. Experts such as speech and language pathologists, researchers in linguistics, and professors classified participants as (a) either needing or not needing additional education support (e.g., language courses) in acquiring and learning German; and (b) as needing or not needing additional medical help in acquiring and learning German.

Demographic and Sociolinguistic Characteristics of the Samples

The age difference between boys and girls in Samples 2 and 3 was not significant according to the Mann–Whitney *U* test ($ps \geq .263$). For Sample 1, no exact information regarding the age of participants was available, though children were normally tested within 2–3 months after reaching 4 years of age.

To control whether the sociolinguistic background of boys and girls was comparable in Samples 2 and 3—no such information was available for Sample 1—several variables from the questionnaires for parents and day-care center teachers were compared using cross tables. There were no significant differences in the chi-square values of variables that might

have influenced the language development of participants (ps ≥ .13). This group contained several variables: whether the child attends the day-care center for a half or whole day, attended nursery school during the first 2 years of life, receives educational support (e.g., language courses) in learning German, plays after day-care with peers who speak German—or, if not a native German speaker, plays with peers who speak the same other language—attends the day-care center regularly, enjoys coming to the day-care center, likes to play with other children, and is involved in extracurricular activities (e.g., clubs). The proportion of immigrant children was not significantly different between boys and girls (ps ≥ .480), and there were no significant differences in the length of contact with the German language between boys and girls in this group according to the Mann–Whitney U test (ps ≥ .309). Similarly, no significant differences were identified between boys and girls in Samples 2 and 3 in terms of length of day-care center attendance (ps ≥ .467).

To control whether medical variables were comparable for both sexes, further information provided by the questionnaires from Samples 2 and 3 was analyzed. No significant differences emerged in the following variables describing children as examined in cross tables using chi-square tests (ps ≥ .06): has a mental handicap, has a permanent hearing disorder, has a permanent sight disorder, has frequent inflammations of the middle ear, has an auditory processing disorder, takes medication, receives therapy, has considerable or severe illness, was a premature baby or had a difficult birth, and has a disorder or illness influencing language development. However, in Sample 3, boys showed significantly (5%) more disorders and illnesses that can influence language development than did girls ($\chi^2_{(1)} = 6.06$, $p = .014$, Ns = 396 males, 365 females), while in Sample 2, boys more often showed problems with hearing than did girls; the proportion of boys with such problems was often or always 2.3% greater ($\chi^2_{(4)} = 47.82$, $p < .001$, $N = 958$ males, 793 females). Yet, this result is unsurprisingly, since boys are generally overrepresented among children with language

deficits, as explained earlier, and hearing disorders (Commission of the European Communities, 1979; Finck–Krämer, Spormann–Lagodzinski, & Gross, 2000). Furthermore, there were no significant differences in the following family-related variables as cross-tables using chi-square tests showed ($ps \geq .082$): parents with problems with reading and writing, family member with a language disorder, and family member who stutters.

For Sample 2, total scores of correct answers on the language screening MSS were compared for three different age groups: 3-year olds ($N = 115$ males, 114 females; median age: 3;8 years), 4-year olds ($N = 1{,}288$ males, 1,062 females; median age: 4;3 years), and 5-year olds ($N = 148$ males, 127 females; median age: 5;4 years). Again, sociolinguistic and demographic characteristics of children were compared on the basis of questionnaires for parents and daycare center teachers. For the subsample of 3-year-old children, no statistically significant differences between boys and girls were identified regarding any demographic or sociolinguistic variables. Meanwhile, 4-year-old boys more often than girls did not hear well according to our data ($\chi^2_{(4)} = 48.62$, $p < .001$, $N = 860$ males, 691 females), though there were no further differences between the sexes. By greater contrast, according to our data, 5-year-old boys suffered more often than girls from language-related disorders or illness ($\chi^2_{(1)} = 5.63$, $p = .018$, $N = 18$ males, 23 females) and from hearing disorders ($\chi^2_{(4)} = 11.32$, $p = .023$, $N = 21$ males, 23 females), as well as significantly more often received some kind of therapy ($\chi^2_{(1)} = 5.54$, $p = .019$, $N = 104$ males, 80 females). After the application of either Bonferroni or Bonferroni–Holm adjustment and considering the low size of Sample 3, only results with $p < .001$ are considered statistically significant. In Sample 1, boys and girls were not significantly different according to the proportion of mental handicaps, hearing disorders, and family language disorders, as cross-tables using chi-square tests showed ($ps \geq .065$). Since no further demographic or sociolinguistic variables were available, the sexes in Samples 1–3

were assumed to be widely comparable and medical or socioeconomic confounding factors were assumed to not exist.

A small subgroup in Sample 2 also participated in the Colored Progressive Matrices (CPM) intelligence test (Raven, 2009). The difference between boys (n = 105, M = 16.07, SD = 4.11) and girls (n = 70, M = 16.76, SD = 3.39) was not significant (p = .138), though girls scored numerically higher and boys showed numerically greater variance, as could be expected (Arden & Plomin, 2006). Both for the CPM test and all language tests, only the total scores of correct answers were used in calculations, without any transformation of data or weighting of test items. The CPM test had a total of 29 items; the total numbers of items on the language tests and subtests appear in Tables 2–4.

Statistical Analyses

According to Kolmogorov–Smirnov tests, data were not normally distributed (p < .05), which recommended the use of non-parametrical tests. In most cases, differences between boys and girls were assessed with Mann–Whitney U tests, for which a non-parametric effect size value $\hat{p}$ (probability of superiority, PS) was provided (Grissom & Kim, 2012). If both groups did not differ, then $\hat{p}$ = .50, and if all values of one group were less or greater than all values of the other, then $\hat{p}$ = 1 ($\hat{p}$ = 0). A result of $\hat{p}$ = .50 is comparable to a Cohen's d = 0.00, whereas a result of $\hat{p}$ = .05 or $\hat{p}$ = .95 corresponds more closely to an exceptionally high negative or positive d effect size. Since non-parametric effect sizes are not widely known, Cohen's d (Cohen, 1988) was also calculated, though it applies strictly to parametric comparison. Negative d values indicate higher scores among females. For dichotomous or categorized variables (e.g., compliance and stuttering), cross-tables with chi-square values were calculated. To analyze whether boys demonstrated greater variance than girls in their language test scores, F-tests were used. For all tests, sample sizes were reported

as number of boys versus number of girls. All calculations were performed using the Statistical Package for the Social Sciences version 20.

Results

Sex Differences in Mean Values Among the Samples

Table 2 presents the differences between boys and girls in Sample 1.

--- Table 2 about here ---

Despite significant results in all linguistic domains, the effect sizes were small, as all values for Cohen's *d* were roughly 0.1–0.2. Exclusion of immigrant children, who generally scored lower than German children on all *MSS* subtests (*ps* < .001), did not change the results substantially shown in Table 2. Differences in vocabulary, grammar, articulation, and total score remained significant on the MSS (Zs ≤ -2.55, ds ≤ -0.09, $\hat{p}$s ≤ 0.49, ps ≤ .011), while differences in speech comprehension on the MSS were barely statistically significant (d = -0.06, $\hat{p}$ = 0.49, p = .051, N = 2,146 males, 2,133 females). Effect sizes remained almost the same (Table 2).

As shown in Tables 3 and 4, the same linguistic domains and the two additional repetition tasks (i.e., sentences and nonce words) that examined short-term phonological memory were assessed for Samples 2 and 3.

--- Table 3 about here ---

--- Table 4 about here ---

In most tests, girls showed numerically better performance than boys. Four of these differences—namely, speech comprehension, pronunciation, repetition of sentences, and total MSS score—were statistically significant in Sample 2 and two—namely, speech comprehension and repetition of nonce words—were significant in Sample 3, though again with small effect sizes.

Sex Differences in Mean Values Among Three Age Groups in Sample 2

In Sample 2, for children aged 3 years, girls scored significantly higher than boys on the MSS subtest for articulation (Z = -2.33, p = .02, d = -0.30, $\hat{p}$ = 0.41, N = 113 males,111 females). Concerning the MSS subtest for vocabulary, the difference was statistically significant by a slim margin (Z = -1.93, p = .053, d = -0.26, $\hat{p}$ = 0.42, N = 114 males, 109 females). Among children aged 4 years, girls significantly outperformed boys on the MSS subtests for articulation, repetition of sentences, and nonce words, as well as in terms of total MSS score (Zs ≤ -2.02, ps ≤ .044, $\hat{p}$s ≤ 0.48, ds ≤ -0.09, Ns ≥ 1,004 males, 835 females). However, the sample size was considerably larger, and the effect sizes were minimal. For 5-year-old children, no significant differences were found (ps ≥ .270). The median of the effect size values (Cohen's d) of differences between boys and girls in terms of *MSS* scores was -0.3 for 3-year olds—girls scored higher—as well as -0.1 for 4-year olds and 0.0 for 5-year olds, which indicates that sex differences decreased with age.

Sex Differences According to Language Experts' Judgments and Questionnaires

Poorer results for male participants are reflected in the greater proportion of boys classified as requiring educational support (e.g., language courses) according to the judgments of expert panels based on results of two of the three samples (Samples 1 and 2: χ^2s$_{(1)}$ ≥ 5.26, ps ≤ .022, Ns ≥ 1,407 males, 1,181 females; Sample 3: p = .114). Furthermore, in all three samples, the proportion of boys who stuttered according to the university language experts was significantly greater than that of girls (1.6– 6.7% for boys, 0.6–2.8% for girls; χ^2s$_{(1)}$ ≥ 3.87, ps ≤ .049, Ns ≥ 338 males, 306 females). The judgment on stuttering was based on the audio records of the test situation (Samples 1 and 3), judgments of language testers (MSS item "stuttering: yes/no"; Samples 1 and 2), questionnaires for parents (Samples 2 and 3) and for daycare center teachers (Samples 1 and 2). In addition, boys in all age groups were judged to show language impairments, pathologies, and disorders significantly

more often than girls, according to parental report (see section "Demographic and Sociolinguistic Characteristics of the Samples").

Though very few significant differences emerged in this respect according to the questionnaires, in all three samples boys were classified by university language experts as requiring medical examination significantly more often than girls (χ^2s$_{(1)} \geq 6.57$, ps $\leq .010$, Ns $\geq$ 363 males, 319 females). This result indicates that expert panels assumed, based on language tests including audio records and questionnaires, that children could have an illness, impairment, or disorder that harmed their language development. This does not mean, however, that these children had a diagnosed language-related medical issue. Furthermore, boys were more often in speech and language therapy according to the questionnaires from all three samples (χ^2s$_{(1)} \geq 6.94$, ps $\leq .008$, Ns $\geq$ 397 males, 362 females). However, according to both testers and questionnaires, in no sample did boys have a voice disorder significantly more often than girls (ps $\geq .084$). Voice disorder was defined as a frequently or permanently hoarse voice. Dichotomous items on the hoarseness (yes/no) were included both in MSS (Samples 1 and 2) and questionnaires for parents (Samples 2 and 3).

Some clues regarding sex-specific differences can be extracted from the questionnaires, in which parents and day-care center teachers provided additional information concerning the language development of the children tested. No sex differences emerged for three variables—namely, that everybody could understand what the child says, that the child can speak in complete sentences, and that, if not German speaking, then the child speaks his or her native language normally. However, boys were described significantly more often than girls as being unable to speak German appropriately according to their age (Sample 2: $\chi^2_{(1)} = 6.35$, $p = .012$, $N = 316$ males, 287 females; Sample 3: $p = .157$). Boys were also more often reported to have communication or speech comprehension problems ($\chi^2_{(1)} = 9.70$, $p = .002$, $N = 149$ males, 114 females), to not understand questions and requests (Sample 2: $\chi^2_{(4)}$

= 12.43, p = .014, N = 818 males, 689 females; Sample 3: p = .546) or sophisticated tasks ($\chi^2_{(4)}$ = 17.18, p = .002, N = 617 males, 481 females), to not have an age-appropriate vocabulary ($\chi^2_{(4)}$ = 18.58, p = .001, N = 476 males, 359 females) or pronunciation ($\chi^2_{(4)}$ = 19.98, p = .001, N = 478 males, 363 females), and to be unable to use articles properly ($\chi^2_{(4)}$ = 9.94, p = .041, N = 610 males, 476 females).

Only one finding supported the assumption that girls speak out more often than boys. Boys were less talkative in groups than girls according to a questionnaire from Sample 2 completed by day-care center teachers (item "The child speaks out during group activities" from the questionnaire for daycare center teachers; $\chi^2_{(4)}$ = 12.04, p = .017, N = 484 males, 374 females). Neither in Sample 1 nor in Sample 2 were significant sex differences on MSS tasks for spontaneous speech found (ps ≥ .606).

It is possible that the differences found between boys and girls in the subjective (questionnaires) and objective (language tests) assessment of language skills could have been influenced by the lower compliance shown by boys than girls in the test situation. The compliance values (*good*, *average*, and *poor*) of boys were indeed significantly lower than those of girls in all three samples (χ^2s$_{(2)}$ ≥ 6.82, ps ≤ .033, Ns ≥ 421 males, 384 females).

Sex Differences in Variance

The result that boys were described by parents, daycare center teachers, and university language experts as suffering from language abnormalities more often compared to girls, might be linked to the expected greater variance among males than among females in verbal proficiency.

Higher male than female variance in verbal proficiency was measured objectively based on total number of correct answers in the language tests. Only total scores from the MSS and S-ENS, including the S-ENS add., the tests which examined all important linguistic domains (i.e., speech comprehension, vocabulary, grammar, phonological short-term

memory, and pronunciation) were used for examining differences in variance between boys and girls.

Our data show that boys indeed demonstrated significantly greater variance than girls regarding total MSS performance in both Sample 1 (SD = 8.97 versus 8.49, $F_{(3,115,\ 3,026)}$ = 33.30, $p < .001$) and Sample 2 (SD = 12.85 versus 12.14, $F_{(1,100,\ 923)} = 9.63$, $p = .002$), but not for the S-ENS including the additional tasks in Sample 3 (SD = 8.89 versus 8.68, $F_{(355,\ 313)}$ = 2.67, $p = .103$). We also examined the proportion of the sexes in the lower and upper tails of the distribution by sorting participants into three different groups according to their total MSS score—namely, the upper 10% and lower 10%, with 80% in the middle. The sex differences within these groups were significant ($\chi^2_{(2)} = 37.82$, $p < .001$), which results from the overrepresentation of boys in the lower tail (1.41 boys : 1 girl). Furthermore, girls were overrepresented in the upper tail (1.42 girls : 1 boy) of the distribution. The overrepresentation of boys in the lower tail and of girls in the upper tail also surfaced in terms of the lower and upper 5% and the lower and upper 3%.

Discussion

In the current study, we examined sex differences in the language development of German preschool children, primarily aged 4–6 years. In accordance with previous research (Bornstein et al., 2004; Hayiou–Thomas et al., 2012; Wallentin, 2009), we found a slight female advantage over boys in several linguistic domains. As previous research also suggests, this sex difference seemed to vanish around school age (i.e., 6 years).

The relatively small developmental sex differences in verbal abilities found could be explained, for instance, by sex-different maturational rates. Females seem to mature faster than males, especially at young ages, and early maturation is correlated with better verbal abilities (Galsworthy et al., 2000; Waber, 1976). Accordingly, language-related sex differences seem to be affected by biological factors, as can be observed by referring to

genetic aspects (see below), but also due to hormonal effects. Transsexuals, for instance, experience a large decrease in performance in verbal fluency tests after androgen therapy (Kimura, 2000). Women seem to have greater verbal abilities (e.g., verbal fluency) mid-menstrual cycle, when progesterone and estradiol levels are high (Halpern, 2000; Hampson & Kimura, 1988; Kimura, 2000).

More importantly for our purposes, sex differences in language acquisition can also be partially caused by hormones. Albores–Gallo et al. (2009) found that prenatal testosterone, measured via second-to-fourth digit ratios (2D:4D), was negatively correlated with vocabulary development in preschoolers. In a similar vein, Lange (in press) reports correlations between 2D:4D and performance in language comprehension and verbal interaction in German kindergarteners, in that high 2D:4D ratios (i.e., low prenatal testosterone) were positively correlated with language performance. As such, fetal testosterone could explain boys' disadvantages—and thus girls' advantages—in lexical acquisition, as findings by Lutchmaya et al. (2002) have also shown. These authors found that girls have larger vocabularies than boys when aged 18–24 months, an advantage that is negatively correlated with fetal testosterone. Consequently, Lutchmaya et al. (2002) assume that fetal testosterone partially shapes neural mechanisms. More precisely, testosterone is likely responsible for advanced development of the right hemisphere in males, which could facilitate deficits and pathologies in the left hemisphere, where language functions are concentrated (Geschwind & Galaburda, 1987; Lust, Geuze, van de Beek, Cohen–Kettenis, Groothuis, & Bouma, 2010). However, it is also possible that a high level of estradiol, and not necessarily a low level of testosterone, is associated with better language development (Wermke, Hain, Oehler, Wermke, & Hesse, 2014). The question that remains is why sex differences in mean values vanish around the age of 6 years. To claim that this vanishing contradicts biological explanations would neglect the very nature of biological factors, which

are well known to work differently depending on age (e.g., puberty). So, the vanishing sex difference might still be explained by biological factors (e.g., maturational factors). Still, it might also be that around 6 years of age, which is strikingly the age when the children in our sample enter school, specific social factors increasingly take the reins with the result of a similarity of boys and girls in language skills. Future research needs to focus on identifying the biological and social factors (and their interaction) that cause this approximation of the sexes regarding language skills.

Along with sex differences according to means, we also assessed sex differences according to variance. Based on both theoretical and empirical work, we assumed that boys would be more variable than girls in their language competence. Ultimately, this assumption was mostly supported by our data. Our results suggest that boys could simply be overrepresented among children with poor verbal abilities. This does, of course, not mean that low test scores are necessarily equivalent with language impairment.

The fact that boys show greater variance than girls but are overrepresented only at the lower tail of the distribution could be explained by slightly greater female means. Taking former research (e.g., Eriksson et al., 2012) into consideration while interpreting our data, it is possible to conclude that sex differences in variance emerge around the age of 3 years.

As to why variance among males is greater than among females, reference to only social factors falls short of a sufficient explanation. Social factors cannot solely and consistently cause greater phenotypic variance among males than among females as this pattern can be found at young ages, which complicates explaining it only by environmental variables (Arden & Plomin, 2006; Ardila et al., 2000). Furthermore, if social factors could explain sex differences in variance, then for boys a broader variety of social factors than for girls—that is, more or more diverse environmental factors—would have to be assumed. This seems unlikely. However, it cannot be ruled out that boys and girls differ on multiple social

factors or that the same factors influence both sexes differently. Especially the latter argument, however, does not contradict biological explanations. Furthermore, the biological causation of language-related pathologies, which constitute one side of sex differences in variance, is evident (Alarcón et al., 2008; Dworzynski et al., 2007; Hayiou–Thomas, 2008; Jenkins, 2000; Stromswold, 2001, 2005). Moreover, such language pathologies are highly heritable in early life (Dale et al., 1998; Hayiou–Thomas et al., 2012), and sex-specific genetic factors responsible for language and language deficits have been reported. Galsworthy et al. (2000) conducted a study with children aged 3 years in order to identify sex-specific genetic factors in language development. They concluded that heritability of verbal abilities is greater in males than in females and that the etiology of vocabulary development might also differ between the sexes, thereby indicating sex-specific genetic factors regarding verbal abilities. This conclusion was based on a difference in correlations between dizygotic same-sex and dizygotic opposite-sex twins.

More notably still, results showing greater variance among males corroborate findings regarding the evolutionary psychology of sex differences in general and the evolution of language in particular. From an evolutionary perspective, it can be assumed that a greater phenotypic variance among males than among females mirrors the larger reproductive variance in males compared to females (Archer & Mehdikhani, 2003; Lange, 2012; Lange, Zaretsky, Schwarz, & Euler, 2014). Sex differences in variance have been reported for somatic traits as well (Lehre, Lehre, Laake, & Danbolt, 2008), which also undermines explanations citing environmental factors. Sex differences in variance, in which males show greater variance, are reflected in sex differences in reproductive strategies (Archer & Mehdikhani, 2003), in reproductive variance (Brown, Laland, & Borgerhoff Mulder, 2009), and, as reported above, in heritability (Galsworthy et al., 2000). Nevertheless, further

research is needed to explain why linguistic sex differences in variance exist and why they seem to emerge only around the age of 3 years (Eriksson et al., 2012).

A greater variance in language development among males than among females and a greater overrepresentation of males among children with a rather poor language acquisition respectively pose practical implications. It suggests that more boys may need to be offered educational support and, if necessary, special intervention regarding language. Though several biological factors influencing verbal abilities and high heritability estimations for verbal abilities have been reported, at least one third of variance in several verbal abilities is explained by environmental factors (Stromswold, 2001, 2005), thus suggesting that intervention can have an effect.

Since our database included numerous demographic variables, we could confirm that no major sex differences regarding extra- or sociolinguistic variables appeared that could have confounded our results. Our samples—at least Sample 1—could furthermore be deemed unselected. We therefore conclude that the findings regarding sex differences are valid. However, our effect sizes were relatively small, and statistical significance was thus found only due to relatively large sample sizes. Simply put, the sexes are more alike than they are different (Zell et al., 2015).

Future studies should aim to replicate our findings by focusing on sex differences in variance, which seems to be a neglected field of research compared to research on sex differences according to mean values. Further research should also seek to identify the causes of sex differences in language development. For sex differences in mean values, both social (e.g., sex-different rearing) and biological factors (e.g., prenatal testosterone and sex-different evolutionary selection pressures) seem to be possible explanations. Regarding sex differences in variance, genetic or chromosomal factors on a proximate level and evolutionary selection pressures linked to reproduction on an ultimate level (Archer &

Mehdikhani, 2003; Lange, 2012) might be worth further examination, but social/sociolinguistic factors should also be taken into consideration.

References

Alarcón, M., Abrahams, B. S., Stone, J. L., Duvall, J. A., Perederiy, J. V., Bomar, J. M., et al. (2008). Linkage, association, and gene-expression analyses identify CNTNAP2 as an autism-susceptibility gene. *The American Journal of Human Genetics, 82,* 150–159.

Albores-Gallo, L., Fernández-Guasti, A., Hernández-Guzmán, L., & List-Hilton, C. (2009). 2D:4D finger ratio and language development. *Revista de Neurología, 48*(11), 577–581.

Andrews, J. G. (1964). The nature of stuttering. *The Medical Journal of Australia, 191,* 919–924.

Angermeier, M. J. W. (2007). *ETS 4-8. Entwicklungstest Sprache für Kinder von 4 bis 8 Jahren* [ETS 4-8. A language development test for children aged between 4 and 8]. Frankfurt/Main, Germany: Harcourt Test Services GmbH.

Archer, J., & Mehdikhani, M. (2003). Variability among males in sexually-selected attributes. *Review of General Psychology, 7*, 219–236. Doi: 10.1037/1089-2680.7.3.219

Arden, R., & Plomin, R. (2006). Sex differences in variance of intelligence across childhood. *Personality and Individual Differences, 41*, 39–48. Doi: 10.1016/j.paid.2005.11.027

Ardila, A., Rosselli, M., Bateman, J. R., & Guzmán, M. (2000). Neuropsychological profile of stuttering children. *Journal of Developmental and Physical Disabilities, 12*(2), 121–130.

Beltz, A. M., Blakemore, J. E. O., & Berenbaum, S. A. (2013). Sex differences in brain and behavioural development. In P. Rakic & J. Rubenstein (Series Eds.), *Comprehensive developmental neuroscience*; H. Tager-Flusberg (Vol. Ed.), *Neural circuit development*

and function in the healthy and diseased brain (Vol. 3, pp. 467–499). Oxford, England: Elsevier.

Berglund, E., Eriksson, M., & Westerlund, M. (2005). Communicative skills in relation to gender, birth order, childcare and socioeconomic status in 18-month-old children. *Scandinavian Journal of Psychology, 46*, 485–491. Doi: 10.1111/j.1467-9450.2005.00480.x

Bleses, D., Vach, W., Slott, M., Wehberg, S., Thomsen, P., Madsen, T. O., & Basbøll, H. (2008). The Danish Communicative Developmental Inventories: Validity and main developmental trends. *Journal of Child Language, 35*, 651–669. Doi: 10.1017/S0305000907008574

Bloodstein, O. (1995). *A handbook on stuttering* (5th ed.). San Diego, CA: Singular Publishing Group.

Bornstein, M. H., Hahn, C.-S., & Haynes, O. M. (2004). Specific and general language performance across early childhood: Stability and gender considerations. *First Language, 24*(3), 267–304. Doi: 10.1177/0142723704045681

Bouchard, C., Trudeau, N., Sutton, A., Boudreault, M.-C., & Deneault, J. (2009). Gender differences in language development in French Canadian children between 8 and 30 months of age. *Applied Psycholinguistics, 30*, 685–707. Doi: 10.1017/S0142716409990075

Brown, G. R., Laland, K. N., & Borgerhoff Mulder, M. (2009). Bateman's principles and human sex roles. *Trends in Ecology and Evolution, 24*, 297–304. Doi: 10.1016/j.tree.2009.02.005

Burling, R. (1986). The selective advantage of complex language. *Ethology and Sociobiology, 7*, 1–16. Doi: 10.1016/0162-3095(86)90011-7

Burling, R. (2005). *The talking ape. How language evolved.* New York, NY: Oxford University Press.

Canning, P. M., & Lyon, M. E. (1989). Young children with special needs: Prevalence and implications in Nova Scotia. *Canadian Journal of Education, 14*, 368–380. Doi: 10.2307/1495365

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Erlbaum.

Commission of the European Communities (1979). *Childhood Deafness in the European Community. CEC Report EUR 6143 Medicine Luxembourg*. Luxembourg: Office for Official Publications of the European Communities.

Dale, P. S., Simonoff, E., Bishop, D. V. M., Eley, T. C. Oliver, B., Price, T. S., Purcell, S., Stevenson, J., & Plomin, R. (1998). Genetic influence on language delay in 2-year-olds. *Nature Neuroscience, 1*, 324–328.

Dindia, K., & Canary, D. J. (2006). *Sex differences and similarities in communication* (2nd edition). Mahwah, NJ: Lawrence Erlbaum Associates.

Döpfner, M., Dietmair, I., Mersmann, H., Simon, K., & Trost-Brinkhues, G. (2005). *S-ENS. Screening des Entwicklungsstandes bei Einschulungsuntersuchungen* [S-ENS. A screening of the developmental level for the school enrolment test]. Göttingen, Germany: Hogrefe.

Dworzynski, K., Remington, A., Rijsdijk, F., Howell, P., & Plomin, R. (2007). Genetic etiology in cases of recovered and persistent stuttering in an unselected, longitudinal sample of young twins. *American Journal of Speech Language Pathology, 16*, 169–178. Doi: 10.1044/1058-0360(2007/021)

Eckert, P., & McConnell-Ginet, S. (2003). *Language and gender*. Cambridge, MA: Cambridge University Press.

Eme, R. F. (1979). Sex differences in childhood psychopathology. *Psychological Bulletin, 86*, 574–595.

Eriksson, M., Marschik, P. B., Tulviste, T., Almgren, M., Pérez Pereira, M., Sonja Wehberg, S., et al. (2012). Differences between girls and boys in emerging language skills: Evidence from 10 language communities. *British Journal of Developmental Psychology, 30*, 326–343. Doi: 10.1111/j.2044-835X.2011.02042.x

Euler, H. A., Holler-Zittlau, I., van Minnen, S., Sick, U., Dux, W., Zaretsky, Y., & Neumann, K. (2010). Psychometrische Gütekriterien eines Kurztests zur Erfassung des Sprachstandes vierjähriger Kinder [Validity criteria of a short test to assess speech and language competence in 4-year-olds]. *HNO,* 58, 1116−1123.

Fairweather, H. (1976). Sex differences in cognition. *Cognition, 4*, 231–280. Doi: 10.1016/0010-0277(76)90019-6

Feingold, A. (1992). Gender differences in mate selection preferences: A test of the parental investment model. *Psychological Bulletin, 112*, 128–139. Doi: 10.1037/0033-2909.112.1.125

Feingold, A. (1995). The additive effects of differences in central tendency and variability are important in comparisons between groups. *American Psychologist, 50*(1), 5–13. Doi: 10.1037/0003-066X.50.1.5

Feldman, H. M., Dollaghan, C. A., Campbell, T. F., Kurs-Lasky, M., Janosky, J. E., & Paradies, J. L. (2000). Measurement properties of the MacArthur communicative development inventories at ages one and two years. *Child Development, 71*(2), 310–322.

Felsenfeld, S., Kirk, K. M., Zhu, G., Statham, D. J., Neale, M. C., & Martin, N. G. (2000). A study of the genetic and environmental etiology of stuttering in a selected twin sample. *Behavior Genetics, 30*(5), 359–366.

644 Fenson, L., Dale, P. S., Reznick, J. S., Bates, E., Thal, D. J., Pethick, S. J., Tomasello, M.,
645 Mervis, C. B., & Stiles, J. (1994). Variability in early communicative development.
646 *Monographs of the Society for Research in Child Development, 59*(5), i+iii-v+1–185.
647 Doi: 10.2307/1166093
648 Finck-Krämer, U., Spormann-Lagodzinski, M. E., & Gross, M. (2000). German registry for
649 hearing loss in children: results after 4 years. *International Journal of Pediatric*
650 *Otorhinolaryngology, 56*(2), 113–127.
651 Fox, A. (2005). *PLAKSS – Psycholinguistische Analyse kindlicher Sprechstörungen (2. Aufl.)*
652 [PLAKSS – A psycholinguistic analysis of speech impediments in children (2nd ed.)].
653 Frankfurt/Main, Germany: Harcourt Test Services.
654 Galsworthy, M. J., Dionne, G., Dale, P. S., & Plomin, R. (2000). Sex differences in early
655 verbal and non-verbal cognitive development. *Developmental Science, 3*(2), 206–215.
656 Doi: 10.1111/1467-7687.00114
657 Geschwind, N., & Galaburda, A. M. (1987). *Cerebral Lateralization: Biological mechanisms,*
658 *associations and pathology*. Cambridge, MA: MIT Press.
659 Grimm, H. (2001). *SETK 3-5. Sprachentwicklungstest für drei- bis fünfjährige Kinder* [SETK
660 3-5. Test of language development for three to five year old children]. Göttingen,
661 Germany: Hogrefe.
662 Grissom, R. J., & Kim, J. J. (2012). *Effect sizes for research: Univariate and multivariate*
663 *applications*. (2nd ed.). New York, NY: Taylor & Francis.
664 Guiller, J., & Durndell, A. (2007). Students' linguistic behaviour in online discussion groups:
665 Does gender matter? *Computers in Human Behavior, 23*(5), 2240–2255. Doi:
666 10.1016/j.chb.2006.03.004
667 Haas, A. (1979). Male and female spoken language differences: Stereotypes and evidence.
668 *Psychological Bulletin, 86*(3), 616–626. Doi: 10.1037/0033-2909.86.3.616

669 Halpern, D. F. (2000). *Sex differences in cognitive abilities* (3rd ed.). Mahwah, NJ: Erlbaum.
670 Hampson, E., & Kimura, D. (1988). Reciprocal effects of hormonal fluctuations on human
671 motor and perceptual-spatial skills. *Behavioral Neuroscience, 102*(3), 456–459. Doi:
672 10.1037/0735-7044.102.3.456
673 Hayiou-Thomas, M. E. (2008). Genetic and environmental influences on early speech,
674 language and literacy development. *Journal of Communication Disorders, 41*, 397–
675 408.
676 Hayiou-Thomas, M. E., Dale, P. S., & Plomin, R. (2012). The etiology of variation in
677 language skills changes with development: A longitudinal twin study of language from
678 2 to 12 years. *Developmental Science, 15*(2), 233–249. Doi: 10.1111/j.1467-
679 7687.2011.01119.x
680 Hedges, L. V., & Nowell, A. (1995). Sex differences in mental scores, variability, and
681 numbers of high-scoring individuals. *Science, 269*, 41–45.
682 Hyde, J. S., & Linn, M. C. (1988). Gender differences in verbal ability: A meta-analysis.
683 *Psychological Bulletin, 104,* 53–69. Doi: 10.1037/0033-2909.104.1.53
684 Ingram, T. T. S. (1975). Speech disorders in childhood. In E. H. Lenneberg & E. Lenneberg
685 (Eds.), *Foundations of language development* (pp. 195–261). New York, NY:
686 Academic Press.
687 Jenkins, L. (2000). *Biolinguistics. Exploring the biology of language*. Cambridge, MA:
688 Cambridge University Press.
689 Jespersen, O. (1922). *Language: Its nature, development, and origin*. London, England:
690 George Allen & Unwin.

691 Kiese-Himmel, C. (2005). *AWST-R. Aktiver Wortschatztest für 3- bis 5-jährige Kinder –*
692 *Revision* [AWST-R. An active vocabulary test for children aged three to five – a
693 revision]. Göttingen, Germany: Hogrefe.
694 Kiese-Himmel, C., & Risse, T. (2009). Normen für den Mottier-Test bei 4- bis 6-jährigen
695 Kindern [Norms for the Mottier Test for 4- to 6-year-old children]. *HNO, 57*, 943–948.
696 Kimura, D. (2000). *Sex and cognition*. Cambridge, MA: The MIT Press.
697 Lange, B. P. (2011). Male proneness to verbal display production. *Acta Linguistica, 5*(2), 97–
698 104.
699 Lange, B. P. (2012). *Verbal proficiency as fitness indicator. Experimental and comparative*
700 *research on the evolutionary psychology of language and verbal displays.* Saarbrücken,
701 Germany: Südwestdeutscher Verlag für Hochschulschriften.
702 Lange, B. P. (in press). Digit ratio as a predictor of language development and media
703 preferences in kindergarten children. *Acta Linguistica, 9*(2).
704 Lange, B. P., & Euler, H. A. (2014). Writers have groupies, too: High-quality literature
705 production and mating success. *Evolutionary Behavioral Sciences, 8*(1), 20–30. Doi:
706 10.1037/h0097246
707 Lange, B. P., Zaretsky, E., Schwarz, S., & Euler, H. A. (2014). Words won't fail:
708 Experimental evidence on the role of verbal proficiency in mate choice. *Journal of*
709 *Language and Social Psychology, 33*(5), 482–499. Doi: 10.1177/0261927x13515886
710 Lehre, A. C., Lehre, K. P., Laake, P., & Danbolt, N. C. (2008). Greater intrasex phenotype
711 variability in males than in females is a fundamental aspect of the gender differences in
712 humans. *Developmental Psychobiology, 51*, 198–206.
713 Locke, J. L., & Bogin, B. (2006). Language and life history: A new perspective on the
714 development and evolution of human language. *Behavioral and Brain Sciences, 29*,
715 259–280. Doi: 10.1017/S0140525X0600906X

716 Lust, J. M., Geuze, R. H., Van de Beek, C., Cohen-Kettenis, P. T., Groothuis, A. G. G., &
717 Bouma, A. (2010). Sex specific effect of prenatal testosterone on language
718 lateralization in children. *Neuropsychologia, 48*, 536–540. Doi:
719 10.1016/j.neuropsychologia.2009.10.014
720 Lutchmaya, S., Baron-Cohen, S., & Raggatt, P. (2002). Foetal testosterone and vocabulary
721 size in 18- and 24-month-old infants. *Infant Behaviour and Development, 24*, 418–424.
722 Doi: 10.1016/S0163-6383(02)00087-5
723 Maccoby, E. E., & Jacklin, C. N. (1974). *The psychology of sex differences*. Stanford, CA:
724 Stanford University Press.
725 Mealy, L. (2000). *Sex differences: Development and evolutionary strategies*. San Diego, CA:
726 Academic Press.
727 Miller, G. F. (1999). Sexual selection for cultural displays. In R. Dunbar, C. Knight, & C.
728 Power (Eds.). *The evolution of culture. An interdisciplinary view* (pp. 71–91).
729 Edinburgh, England: Edinburgh University Press.
730 Mottier, G. (1951). Über Untersuchungen der Sprache lesegestörter Kinder [On language
731 tests with dyslexic children]. *Folia Phoniatrica, 3,* 170–177.
732 Nelson, K. (1973). Structure and strategy in learning to talk. *Monographs of the Society of*
733 *Research in Child Development, 149*(38), 1–138.
734 Neumann, K., Holler-Zittlau, I., van Minnen, S., Sick, U., Zaretsky, Y., & Euler, H. A.
735 (2011a). Katzengoldstandards in der Sprachstandserfassung: Sensitivität und Spezifität
736 des Kindersprachscreenings (*KiSS*) [Fool's gold standards in language screening.
737 Sensitivity and specificity of the Hessian child language screening test
738 (Kindersprachscreening, KiSS)]. *HNO (eHNO), 59*, 97–109.
739 Neumann, K., Zaretsky, Y., & Euler, H. A. (2011b). Einführung einer flächendeckenden
740 Sprachstandserfassung in Hessen. Forschungsbericht 2011 [Introduction of the

741 language screening program in the state of Hesse. A research report]. Frankfurt/Main,
742 Germany: Universität Frankfurt/Main.
743 Raven, J. C. (2009). *Coloured Progressive Matrices*. Frankfurt/Main, Germany: Pearson.
744 Rescorla, L., Hadicke-Wiley, M., & Escarce, E. (1993). Epidemiological investigation of
745 expressive language delay at age two. *First Language, 13*, 5-22.
746 Rescorla, L., & Schwartz, E. (1990). Outcomes of toddlers with expressive language delay.
747 *Applied Psycholinguistics*, *11*, 393–407.
748 Reynolds, C. A., Hewitt, J. K., Erickson, M. T., & Silberg, J. L. (1996). The genetics of
749 children's oral reading performance. *Journal of Child Psychology and Psychiatry,*
750 *37*(4), 425–434. Doi: 10.1111/j.1469-7610.1996.tb01423.x
751 Statistisches Bundesamt. (2012). *89 % der 3-Jährigen besuchen Kindergarten – Deutschland*
752 *weit über OECD-Durchschnitt [89 % of three-year-olds attend kindergarten –*
753 *Germany way above OECD averaged]*. Retrieved from
754 https://www.destatis.de/DE/PresseService/Presse/Pressemitteilungen/2012/09/PD12_31
755 4_217.html
756 Stern, C., & Stern, W. (1907 / 1987). *Die Kindersprache. Eine psychologische und*
757 *sprachtheoretische Untersuchung* [Child language. A psychological and language
758 theoretic investigation]. Darmstadt, Germany: Wissenschaftliche Buchgesellschaft.
759 Strand, S., Deary, I. J., & Smith, P. (2006). Sex differences in cognitive ability test scores: A
760 UK national picture. *British Journal of Educational Psychology, 76*, 463–480.
761 Stromswold, K. (2001). The heritability of language: A review and metaanalysis of twin,
762 adoption, and linkage studies. *Language, 77*(4), 647–723. Doi: 10.1353/lan.2001.0247

763 Stromswold, K. (2005). Genetic specificity of linguistic heritability. In A. Cutler (Ed.),
764 *Twenty-first century psycholinguistics: Four cornerstones* (pp. 121–140). Mahwah, NJ:
765 Erlbaum.

766 Szagun, G., Stumper, B. & Schramm, S.A. (2009). *A normative study of early language*
767 *development in German using an adaptation of the long and short version toddler CDI.*
768 Retrieved April 2, 2014, from
769 www.giselaszagun.com/en/Szagun_et_al_FRAKIS_2009.pdf

770 Thomson, C., & Polnay, L. (Eds.) (2002). *Community paediatrics* (3rd ed.). Edinburgh,
771 England: Elsevier.

772 Ullman, M. T., Miranda, R. A., & Travers, M. L. (2008). Sex differences in the
773 neurocognition of language. In J. B. Becker, K. J. Berkley, & N. Gearyet al. (Eds.), *Sex*
774 *on the brain: From genes to behavior* (pp. 291-309). New York, NY: Oxford
775 University Press.

776 Waber, D. P. (1976). Sex differences in cognition: A function of maturation rate? *Science*
777 *192*(4239), 572–574. Doi: 10.1126/science.1257795

778 Wallentin, M. (2009). Putative sex differences in verbal abilities and language cortex: A
779 critical review. *Brain & Language, 108*, 175–183. Doi: 10.1016/j.bandl.2008.07.001

780 Wermke, K., Hain, J., Oehler, K., Wermke, P., & Hesse, V. (2014). Sex hormone influence
781 on human infants' sound characteristics: melody in spontaneous crying. *Biology Letters,*
782 *10*(5): 20140095. Doi: 10.1098/rsbl.2014.0095

783 Zell, E., Krizan, Z., & Teeter, S. R. (2015). Evaluating gender similarities and differences
784 using metasynthesis. *American Psychologist, 70*, 10–20. Doi: 10.1037/a0038208

Zubrick, S. R., Taylor, C. L., Rice, M. L., & Slegers, D. W. (2007). Late language emergence at 24 months: An epidemiological study of prevalence, predictors, and covariates. *Journal of Speech, Language, and Hearing Research, 50*, 1562–1592.

Appendix

Available demographic and sociolinguistic characteristics of children (questionnaires for parents and daycare center teachers):

Sample 1: (1) The child is mentally handicapped (yes/no), (2) The child undergoes/underwent a language therapy (yes/no), (3) Immigration background (yes/no), (4) The child has a permanent hearing disorder (yes/no), (5) The parents of the child have problems with reading and writing (yes/no), (6) Somebody in the family has a language disorder (yes/no).

Sample 2: (1) The child is mentally handicapped (yes/no), (2) Only for immigrants: How often the child plays with peers speaking the same (non-German) language (never-seldom-sometimes-often-always), (3) The child undergoes/underwent a language therapy (yes/no), (4) The child speaks out during group activities (never-seldom-sometimes-often-always), (5) Immigration background (yes/no), (6) The child attends daycare center for a half a day/whole day, (7) …attended day nursery in the first two years of life before attending the daycare center (yes/no), (8) …receives educational support in learning German (yes/no), (9) …plays after daycare center time with German children (yes/no), (10) …attends daycare center regularly (yes/no), (11) …likes to go to the daycare center (never-seldom-sometimes-often-always), (12) …likes to play with other children (never-seldom-sometimes-often-always), (13) …is involved in extracurricular activities (clubs etc.) (yes/no), (14) Length of contact with the German language (since birth-1st year of life-2nd year-3rd year-4th year-5th year), (15) Length of daycare center attendance in months, (16) The child has a permanent sight disorder (yes/no), (17) …has frequent inflammations of the middle ear (yes/no), (18) …has auditory processing

disorder (yes/no), (19) …has a permanent hearing disorder (yes/no), (20) …takes medication (yes/no), (21) …receives therapy (yes/no), (22) …has some considerable/severe illness(es) (yes/no), (23) …was a premature baby/had a difficult birth (yes/no), (24) …does not hear well (never-seldom-sometimes-often-always), (25) …has some disorder or illness influencing language development (yes/no), (26) The parents of the child have problems with reading and writing (yes/no), (27) Somebody in the family stutters (yes/no), (28) The parents of the child have problems with reading and writing (yes/no), (29) Somebody in the family has a language disorder (yes/no), (30) Somebody in the family has a hoarse voice (yes/no).

Sample 3: (1) Only for immigrants: How often the child plays with peers speaking the same (non-German) language (never-seldom-sometimes-often-always), (2) Immigration background (yes/no), (3) The child attends daycare center for a half a day/whole day, (4) …attended day nursery in the first two years of life before attending the daycare center (yes/no), (5) Length of daycare center attendance in months, (6) The child does not hear well (never-seldom-sometimes-often-always), (7) Somebody in the family has a language disorder (yes/no).

Subjective judgments of daycare center teachers and parents on children's German skills:

Sample 2: (1) The child speaks his/her mother tongue, if not German, age-appropriately (yes/no), (2) The child speaks German age-appropriately (yes/no), (3) The child stutters (never-seldom-sometimes-often-always), (4) The child has a hoarse voice often or always (yes/no), (5) The child can use articles correctly (never-seldom-sometimes-often-always), (6) The child understands sophisticated tasks (never-seldom-sometimes-often-always), (7) The child uses an age-appropriate vocabulary (never-seldom-sometimes-often-always), (8) Child's speech can be understood by everyone (never-seldom-sometimes-often-always), (9) Child's pronunciation is age-appropriate (never-seldom-sometimes-often-always), (10) The child can speak in full

sentences (never-seldom-sometimes-often-always), (11) The child can understand questions or requests correctly (never-seldom-sometimes-often-always), (12) The child has communication or speech comprehension problems (yes/no).

Sample 3: (1) The child speaks German age-appropriately (yes/no), (2) The child stutters (never-seldom-sometimes-often-always), (3) The child has a hoarse voice often or always (yes/no).