PRE-SCHOOLERS WHO STUTTER SCORE LOWER IN VERBAL SKILLS THAN THEIR NON-STUTTERING PEERS

Eugen Zaretsky
Department of Phoniatrics and Pediatric Audiology, University Hospital of Frankfurt/Main Germany

Benjamin P. Lange Department of Media Psychology, Julius Maximilian University of Wuerzburg Germany

Harald A. Euler Department of Phoniatrics and Pediatric Audiology, University Hospital of Bochum, Germany

Fiona Robinson Department of Phoniatrics and Pediatric Audiology, University Hospital of Frankfurt/Main Germany

Katrin Neumann Department of Phoniatrics and Pediatric Audiology, University Hospital of Bochum Germany

ABSTRACT

Purpose: The study aimed at the examination of a link between stuttering and verbal skills (speech comprehension, articulation, grammar, vocabulary, and phonological short-term memory) in three- to five-year-old children.
Method: Two samples with a total of 7,217 unselected German children were tested with the validated speech and language test *Marburger Sprachscreening – revised version (*MSSrev*).* Linguistic domains were compared for pre-school children who stuttered (CWS; *n*=110) and those who did not (CWNS; *n*=7,107) by means of Mann-Whitney *U* tests, general linear models, Spearman correlations, and cross-tables.
Results: In both samples, CWS scored lower in grammar, articulation, and overall performance on the MSSrev. Statistically significant associations between stuttering and (a) sex of the child, and (b) language disorders in the family were identified.
Conclusions: Taking into account the effect sizes, there appears to be a weak, but statistically significant link between stuttering and verbal skills.

Keywords: Language disorders, language acquisition, stuttering, bilingualism, German language

1 INTRODUCTION

The chronological overlap between the appearance of stuttering and the acquisition of language skills in childhood has led researchers to query whether there is a connection between stuttering and a child's verbal skills. Numerous studies have endeavoured to determine this connection across a variety of linguistic domains, summarized in a systematic review by Hall, Wagovich, and Bernstein Ratner (2007) and in a meta-analysis by Ntourou, Conture, and Lipsey (2011). It is well documented that disfluencies in children have diverse language performance effects, such as a higher probability of stuttering in longer sentences, in assertions compared to relaxed narrative speech, and generally in situations of high communicative demand (Bloodstein & Bernstein Ratner, 2008). The central and more relevant questions, however, are whether children who stutter (CWS) have limited verbal skills compared to children who do not stutter (CWNS), and whether persistent stuttering is associated with later language development in children. The answers so far to these questions are not undisputed (Nippold, 2012).

The 22 studies about verbal skills in CWS which Hall et al. (2007) present are inconclusive as to the central questions just stated. The picture of potential deficits in CWS is very heterogeneous: some studies did find them (e.g. Silverman & Bernstein Ratner, 2002), others did not (e.g. Nippold, Schwarz, & Jescheniak, 1991; Weiss & Zebrowski, 1994), and some (e.g. Watkins, Yairi, & Ambrose, 1999) even found CWS to be slightly above normal expressive verbal skills. Moreover, the studies reported by Hall et al. (2007) have mostly small sample sizes, with a median of 15.5 participants and a range from 8 to 45.

A clearer picture about differences in language competences between CWS and CWNS is presented by Ntourou et al. (2011). In their meta-analysis, 22 studies passed the inclusion criteria of a comparison between CWS and CWNS and the provision of norm-referenced measures of various verbal skills. Thus, questionnaire-based studies, e.g. asking speech-language pathologists about the occurrence of concomitant speech and language disorders (e.g. Arndt & Healey, 2001) from their case loads of CWS, were not included. Whether the inclusion of master's theses and doctoral dissertations reduces reporting bias, as these authors suggest, might be debatable, because the authors of such works probably prefer to present significant data as much as journal authors do. More helpful would be access to data that remain in the drawer because of non-significant differences in language competences between CWS and CWNS. As to the available results, Ntourou et al. (2011) showed that in some linguistic domains, namely receptive and expressive vocabulary, overall language (all linguistic domains taken together), and mean length of utterances (MLUs), CWS scored significantly lower than CWNS. The differences were of moderate effect size, except for MLUs where the effect size was small. The eight studies which addressed the difference between both groups with respect to syntactic complexity lacked significance

in the meta-analytic synopsis. Phonological differences were not examined. The authors arrive at the conclusion 'that CWS exhibit relatively consistent but subtle differences in language abilities when compared to their normally fluent peers' (Ntourou et al. 2011: 173).

The most pronounced sceptic of the 'stuttering-language connection' is Nippold (1990; 2012) who, in her 2012 paper, remains unconvinced even by the Ntourou et al. (2011) study, mainly for methodological reasons. She considers expressive vocabulary tests which require participants to name pictured objects unfair and invalid to test language competence because persons who stutter tend to avoid saying words, which for them are likely to induce disfluencies, and instead say easier words. For the future, Nippold (2012) advocates studies which draw larger samples as well as longitudinal studies from the onset of stuttering into school age.

The current study fulfils the sampling requests by Nippold (2012) because it presents population-based results with a relatively large number of CWS, thus avoiding any recruitment bias. Furthermore, the present report is unbiased as to favour detection of verbal skill differences between CWS and CWNS because it is a retrospective analysis of data collected from a state-wide application of a screening instrument obligatory for all kindergarten children of a certain age range (Sample 1) and of data which were collected by the authors for the purpose of the validation of a language test (Sample 2). For completeness, all language data relevant for the hypothesis of a stuttering-language connection are reported, namely results on articulation, grammar, vocabulary, phonological short-term memory, and speech comprehension.

The present study might be a valuable addition to the ongoing controversy about the stuttering-language connection because it has been performed in German. German is a synthetic language whereas English is an analytic language. Analytic languages use more unbound morphemes (separate words), more function words, and less suffixes and endings than synthetic languages. Synthetic languages, in contrast, have a higher morpheme-per-word ratio, fewer restrictions in word order, and use more grammatical cases instead of function words (Hawkins, 1986). Because stuttering occurs relatively frequently in function words (Dworzynski, Howell, & Natke, 2003; Dworzynski, Howell, Au-Yeung, & Rommel, 2004), demonstration of the link between stuttering and language skills in a language with fewer function words would add to the validity of the existence of such a link across language types.

The current study is restricted to the presentation of available data on three- to five-year-old children acquiring German as their first or second language. The research question was whether there is a link between stuttering and the verbal skills of German pre-school children: speech comprehension, grammar, articulation, vocabulary, and phonological shot-term memory, the latter quantified by the repetition of sentences and nonce words. Both available samples were analysed retrospectively for this purpose to examine the replicability of results. Because most of the previous studies, in spite of

different study designs and limited sample sizes, did demonstrate less advanced language skills of CWS compared to those of CWNS in most cases, it was hypothesized that the same tendency would be identified in the current study as well.

2   MATERIAL AND METHODS

For an overview of the two samples used in the current study, see Table 1.

Sample 1

Participants. Participants were 6,144 children aged 4;0 to 4;5 (years; months) and attending Hessian kindergartens. All children were obliged to participate in the state-wide screening programme, both in urban and rural areas. Immigrant children, who spoke (a) language(s) other than German in their family, participated as well, as did children with known language deficits and medical impairments.

Assessment. The language screening, the *Marburger Sprachscreening – revised version* (MSSrev), consisted of a 44-item speech/language test and several questions for parents and kindergarten teachers. This speech/language test is a validated screening instrument (REMOVED) for three- to five-year-old children. It employs the presentation of a large coloured picture with many child-relevant activities and objects, which are used to initiate utterances and test various abilities. The test has four subtests: 4 items about speech/language comprehension (receptive language skills), 11 items about expressive vocabulary, 12 items about articulation, and 15 items about expressive grammar. In two additional items, spontaneous speech was elicited. Furthermore, the examiner was asked whether the child stuttered and had a hoarse voice in the testing situation, with yes-no response options.

CWS/CWNS classification. Children were classified as CWS if they stuttered according to the examiners. The examiners were mostly kindergarten teachers who participated in a special training on speech and language disorders: six hours of theory (sensitisation for age-appropriate and pathological language(-related) phenomena including fluency disorders, MSSrev implementation and evaluation) and two hours of practical exercises, including some on the identification of stuttering. Examples were given for stuttering-related secondary behaviours, affective and cognitive aspects of stuttering as well as for concomitant disorders. Judgments of kindergarten teachers were controlled and, if necessary, corrected by speech-language pathologists in the local public health departments on the basis of test batteries and audio records.

Table 1. Characteristics of the two study samples.

	Sample 1	Sample 2
Recruitment	Obligatory language screening of kindergarten children	Voluntary participation of kindergarten children with written parental informed consent
Purpose of the original study	First implementation of a newly developed language screening (MSSrev)	Validation of MSSrev
Test location in Germany	State of Hesse	States of Hesse and North Rhine-Westphalia
Sample size (CWS/CWNS)	79/6,065	31/1,042
bi- or multilingual (CWS/CWNS)	23/1,841	15/418
monolingual Germans (CWS/CWNS)	56/4,224	16/624
CWS m/f	60/19 (ratio 3.2:1)	26/5 (ratio 5.2:1)
CWNS m/f	3057/3008 (ratio 1.0:1)	576/466 (ratio 1.2:1)
Age range (years; months)	4;0-4;5	3;0-5;11; median 4;3
Inclusion criteria	Age 4;0-4;5	Age 3;0-5;11
Exclusion criteria	Inappropriate age	Inappropriate age
Test applied	MSSrev	MSSrev
Examiners	Kindergarten teachers, seldom SLPs and researchers of linguistics and medicine	Kindergarten teachers, SLPs, students and researchers of linguistics and medicine
Assessed language skills	Articulation; receptive & expressive grammar; expressive vocabulary; speech/language comprehension; occurrence of spontaneous speech	Articulation; receptive and expressive grammar; expressive vocabulary; speech/language comprehension; occurrence of spontaneous speech; phonological short-term memory

Note: CWS: children who stutter; CWNS: children who do not stutter; m/f: males/females; MSSrev: Marburger Sprachscreening – rev. version (REMOVED); SLP: speech-language pathologist.

Sample 2

Participants. Parents or caretakers of children recruited from kindergartens in Hesse and North-Rhine Westphalia were asked to provide informed written consent for their children to participate in the study designed to validate MSSrev. The participants were 1,073 children aged 3;0 to 5;11. The tests were administered in a quiet room in the kindergartens or, comparatively seldom, university hospitals.

Assessment. The children were given the same validated language test as in Sample 1, MSSrev, however, with two new subtests on the phonological short-term memory, which did not yet exist when Sample 1 was examined: repetition of sentences (2 items with a total of 15 words) and repetition of nonce words (4 items; nouns with a German phonotactic structure but without meaning). Also, two questionnaires were integrated into the test: a 25-item questionnaire for parents and a 26-item questionnaire for the kindergarten teachers (some items were not relevant for this study). Questionnaire items are listed in the Results section in the description of calculations related to the choice of most relevant factors for general linear model (GLM). Both in Sample 1 and Sample 2 study participants were tested in the German language only, and not in their mother tongues, if these were not German.

Out of 31 CWS, 22 children could come to the Department of Phoniatrics and Pediatric Audiology in REMOVED and were tested by speech-language pathologists with the MSSrev and the *Stuttering Severity Instrument for children and adults* (SSI-3; Riley, 1994; German version: Sandrieser & Schneider, 2008). Also, several audio recordings of the children's speech in different situations (reading, dialogue, spontaneous speech) were analysed. Developmental or chronic stuttering was diagnosed in 18 out of 22 children.

There was no overlap between the samples, that is, children from one sample did not re-appear in the other one. The language tests were carried out by trained personnel consisting of speech-language pathologists, students of and researchers in clinical linguistics, PhD students of medicine, sometimes also kindergarten teachers. The latter two groups were obliged to participate in special courses dedicated to speech and language deficits (cf. Sample 1).

CWS/CWNS classification. The questionnaires for the kindergarten teachers contained, among others, one question about how often the child stutters, with the response options "never", "rarely", "sometimes", "often", "always". Stuttering was diagnosed if the option "often" or "always" was chosen. Again, the classification as CWS/CWNS was controlled and, if necessary, corrected by speech-language pathologists (University Hospital of REMOVED) on the basis of test batteries and audio records.

The tests in Samples 1 and 2 assessed expressive verbal skills, except the tests for speech comprehension and repetition of sentences. In the latter, not only phonological short-term memory, but also receptive grammar was assessed (cf. Stokes, Wong, Fletcher, & Leonard, 2006). The samples were not qualitatively different (kindergarten children tested with the same language screening in the same region without any preselection), the largest difference being, except for sample sizes, a broader age group in Sample 2. Two samples, instead of one, were utilized to examine whether the results of one sample could be confirmed by the other one, that is, for replicability purpose.

Data analysis

All calculations were carried out using SPSS 22 (International Business Machines Corp., New York, USA). According to Kolmogorov-Smirnov tests, most distributions were not normal ($ps<.05$). Therefore, the differences between CWS and CWNS were examined non-parametrically by the Mann-Whitney U test (one-tailed α levels due to the hypothesis formulated in the Introduction). Age was not considered an important covariate in these calculations because in Sample 1 all children were tested within a narrow time window of six months after their fourth birthday (usually within three months), and in Sample 2 no significant difference between CWS and CWNS in respect to their age (in months) was found in the Mann-Whitney U test ($z=-1.02$, $p=.306$).

The effect sizes of the CWS-CWNS differences were estimated with the probability of the superiority (PS) index (Grissom & Kim, 2012). This index measures the probability with which a randomly selected score from one group is larger (or smaller) than a randomly selected score from another group. If both groups do not differ, the PS index is $\hat{p}=0.50$. If all values of one group are lower (or higher) than all values of the other group, the index is $\hat{p}=1$ ($\hat{p}=0$). The more $\hat{p}$ deviates from 0.50 towards 1 or 0, the larger the effect size. A $\hat{p}=0.50$ would compare to a Cohen's d effect size of 0.00 (Cohen, 1988), whereas $\hat{p}$ close to 0 or 1 would correspond to a very large negative or positive d effect size measure. In Table 2, the d effect size estimations are presented in addition to the PS effect size measures, because the PS effect size indicator is not widely known. To our knowledge, conventions for the classification of small, medium, and large PS effect sizes, comparable to Cohen's d, are not available.

If there is a linear association between stuttering and verbal skills, one would expect a correlation between the amount of stuttering and language test scores. In Sample 2, the questionnaires for kindergarten teachers contained a question as to the frequency of stuttering ("never"=1; "seldom"=2; "sometimes"=3; "often"=4; "always"=5). Spearman correlations between these variables from questionnaires and total scores of MSSrev subtests were calculated. It was expected that a more differentiated classification of

stuttering might deliver more precise information on the associations between stuttering and language skills in comparison with Mann-Whitney U tests.

The link between German language skills and stuttering was also examined using general linear models, under consideration of factors interacting with stuttering. These models (univariate analyses of variance, ANOVA) were calculated with the total scores of correct answers in MSSrev subtests as dependent variables and demographic and sociolinguistic factors from questionnaires for parents and kindergarten teachers as independent variables. Because dozens of potential factors were available, only those factors were included as independent variables which were significantly associated with the CWS/CWNS classification according to the cross-tables (phi-correlation=ϕ or linear-by-linear associations=lbl) or Mann-Whitney U tests (two-tailed α levels because no hypotheses were formulated for these factors). Bonferroni adjustment of p values was applied. GLM aimed not at an identification of the most important factors related to the distribution of total scores of correct answers in MSSrev subtests, but at an identification of associations between children's German skills and stuttering, under consideration of possible links between stuttering and other sociolinguistic and demographic variables such as sex of the child. These associations should not be interpreted in terms of causality but, probably, rather in terms of interactions.

As a retrospective reanalysis of the anonymized data, this study did not require an additional approval of the university ethics committee. The original studies on language test development have been approved by REMOVED (Germany).

3   RESULTS

Comparison of CWS and CWNS in two samples: Verbal skills

Differences between CWS and CWNS in the linguistic domains were compared in Sample 1 using Mann-Whitney U tests by means of the total numbers of correct answers (see Table 2). In two out of four subtests (expressive grammar and articulation), and also in the MSSrev total score, CWS scored significantly lower than CWNS, but with small effect sizes.

In Sample 2, the same tendencies were found (see Table 3). Unlike in Table 2, not a single z value was marked with an asterisk because if the Bonferroni adjustment of the p value is applied (.05/8=.006), no result will remain statistically significant. However, effect sizes (p) demonstrate comparable values for both tables: on average, 0.39 in Table 2 vs. 0.38 in Table 3 (for the same subtests speech comprehension, vocabulary, articulation, grammar, and total score, that is, without phonological short-term memory).

Table 2. Results of the Mann-Whitney *U* test for the speech and language screening (*Marburger Sprachscreening – revised version*) scores in Sample 1; *n*=6,065 CWNS vs. 79 CWS. Mean (*M*) and standard deviation (*SD*) values are given first for CWNS, then for CWS; asterisks indicate significance at Bonferroni-adjusted *α*-level of *p*=.05/5=.01.

	Expressive vocabulary	Expressive grammar	Speech comprehension	Articulation	Total score
M	7.9/7.7	10.0/9.3	3.3/3.1	8.9/7.3	31.9/28.9
SD	2.7/2.7	2.6/2.8	1.0/1.1	4.4/4.0	8.8/8.5
U	223,956	193,889	218,374	184,210	187,911
Z	-1.01	-3.04*	-1.53	-3.54*	-3.30*
p	.157	.001	.063	<.001	.001
p̂	0.47	0.41	0.46	0.39	0.39
d	0.10	0.27	0.17	0.38	0.34

Note: CWS: children who stutter; CWNS: children who do not stutter.

Correlations between language test results and the severity of stuttering

For the data from Sample 2 the answers of kindergarten teachers to the question on the frequency of stuttering ("never"=1; "seldom"=2; "sometimes"=3; "often"=4; "always"=5) were rank-correlated with the available scores from the language tests (see Table 4). Negative correlations indicate that greater stuttering severity was associated with lower language scores.

In Sample 2, only one out of eight correlations yielded a significant result after the application of the Bonferroni-adjusted *α* level of *p*=.05/8=.006, namely the one between the severity of stuttering and the MSSrev subtest on articulation. However, correlations with the grammar, repetition of sentences, and the total score of MSSrev (with or without repetition tasks) were also significant, the latter being marginally significant even after application of the Bonferroni adjustment. In regard to the co-occurrence of statistically significant results for grammar and repetition of sentences, a highly significant correlation between total scores of correct answers in these two MSSrev subtests must be taken into account: *ρ*=.709, *p*<.001, *n*=746.

Table 3. Results of the Mann-Whitney *U* test for the speech and language screening (*Marburger Sprachscreening – revised version*) scores in Sample 2. Mean (*M*) and standard deviation (*SD*) values are given first for CWNS, then for CWS.

	Expressive vocabulary	Expressive grammar	Speech comprehension	Articulation	Repetition of sentences	Repetition of nonce words	Total score (with repetition)	Total score (without repetition)
M	5.7/5.7	7.1/5.8	2.2/2.0	8.00/7.1	9.5/5.1	2.2/2.1	35.9/31.8	24.1/21.4
SD	2.6/2.0	4.0/3.8	0.9/0.9	2.3/2.3	8.9/4.7	1.3/1.2	13.0/10.5	7.9/6.5
U	14,869	12,188	13,777	12,036	6,172	6,315	2,187	11,346
Z	-0.70	-1.88	-1.49	-2.42	-0.86	-0.52	-1.60	-2.29
p	.242	.030	.068	.008	.194	.303	.055	.011
d	0.04	0.35	0.24	0.38	0.62	0.10	0.35	0.38
$\hat{p}$	0.46	0.40	0.43	0.38	0.44	0.47	0.37	0.38
N CWS	31	30	31	31	19	18	13	30
N CWNS	1,042	1,016	1,041	1,033	734	754	455	1,001

Note: CWS: children who stutter; CWNS: children who do not stutter.

Table 4. Spearman correlations (*ρ*) between severity of stuttering and test scores of the speech and language screening (*Marburger Sprachscreening – revised version*) scores in Sample 2; asterisk indicates significance at Bonferroni-adjusted *α*-level of *p*=.05/8=.006.

	Expressive vocabulary	Expressive grammar	Speech comprehension	Articulation	Repetition of sentences	Repetition of nonce words	Total score (with repetition)	Total score (without repetition)
ρ	-.007	-.065	-.048	-.100*	-.078	-.035	-.091	-.078
p	.413	.018	.060	.001	.016	.167	.007	.006
N	1,066	1,046	1,072	1,067	756	774	732	1,031

Note: CWS: children who stutter; CWNS: children who do not stutter.

Univariate general linear model

GLM aimed at identification of links between stuttering and total scores of correct answers in MSSrev under consideration of interactions between stuttering and other sociolinguistic and demographic factors. In Sample 1, information on sex of the children (association with the CWS/CWNS classification: φ=-.058, *p*<.001) and language disorders in the family (φ=.036, *p*=.005) was included in GLM as independent demographic/sociolinguistic variables after the application of the Bonferroni adjustment of the *p* value. Variables on immigration background, hearing disorders, and intellectual disabilities were excluded because they were not significantly associated with the CWS/CWNS classification. No other demographic/sociolinguistic

variables were available in this study. GLM results for Sample 1 are given in Table 5.

Table 5. General linear model: Univariate analysis of variance. Associations between stuttering, including its interactions with sex of the child and language disorders in its family, and total scores of correct answers in the speech and language test _Marburger Sprachscreening – revised version_: Results of tests of between-subject effects (_F_ tests, _N_s=6,144).

	Expressive vocabulary	Expressive grammar	Speech comprehen-sion	Articulation	Total score
Corrected model	3.19**	30.56***	4.44***	5.45***	11.82***
Intercept	52,627***	95,077***	66,800***	25,644***	81,820***
Stuttering	0.69	6.34*	2.68	10.75**	8.74**
Stuttering * Sex of the child	6.89**	34.28***	5.85**	6.73**	16.64***
Stuttering * Language disorders in the family	0.74	38.95***	3.92*	1.51	8.55***

Note. *** $p < .001$, ** $p < .01$, * $p < .05$

Results presented in Table 5 demonstrate that stuttering was significantly associated with the MSSrev subtests on grammar and articulation as well as the total score of correct answers, which confirms the results presented in Table 2.

In Sample 2, none of the following demographic/sociolinguistic variable yielded a significant result in analyses of their associations with the CWS/CWNS classification after the application of the Bonferroni adjustment: stuttering in the family, language disorders in the family, sex of the child, whether the child plays with German speaking children in the kindergarten, whether the child can hear well, whether the child speaks his/her mother tongue, if not German, appropriately for his/her age; length of kindergarten attendance in months, attendance of the kindergarten for half a day or a full day, regularity of kindergarten attendance, attendance of a nursery school in the first two years of life, sociability ('the child likes to play with other children', '…plays with German speaking children after kindergarten hours', '…speaks out when playing'), immigration background, rates of sight disorders or other disorders/illnesses which might influence language development (frequent otitis media, permanent hearing and motor disorders, intellectual disability), whether there is at least one more child in the kindergarten group who speaks the same language, if not German, as the study participant, and how often the study participant plays with this/these child/children, language(s) preferred at home (by the mother, father, child), length of contact of the child with the German language, participation in associations, early or difficult birth, age when parents began to learn/acquire German, parents' educational level and first languages, 'problems with reading and writing' in the family, and age of the child in months. Due to the

lack of significant associations with the CWS/CWNS classification, no GLM was calculated for Sample 2.

4   DISCUSSION

The present study investigated the association between disfluencies of pre-school children and their performance in the language test MSSrev in two samples including a relatively large population sample. In spite of the low numbers of CWS in both samples, we found statistically significant support for such an association in the results of three- to five-year-old children. Relatively robust associations between stuttering and language skills were found for grammar, articulation, and a total MSSrev score of correct answers. In all statistical analyses stuttering was associated with lower language test scores.

These findings can be explained either in terms of study design or in terms of (probably genetic) associations between stuttering and language impairments (Hall, Yamashita, & Aram, 1993). In respect to the study design, for example, the tests of grammar required the children to say a particular word. Persons who stutter tend to replace words which they fear to result in a disfluency by another word, or they might hesitate if an alternative word is not immediately available (Van Riper, 1971; Natke, Sandrieser, Pietrowsky, & Kalveram, 2006). Even if only some CWS resort to such avoidance tactics, this might influence the expressive language tests to their disfavour due to a low sample size of CWS. The subtest on speech comprehension, where the children did not have to utter a word but were required to do or show something (e.g. 'Show me the boy with the blue pants'), did not reveal significant differences between CWS and CWNS. Also, statistically significant associations between language impairments and stuttering (Blood, Ridenour, Qualis, & Hammer, 2003) as well as stuttering-like dysfluencies in children suffering from specific language impairment (Befi-Lopes, Cáceres-Assenço, Marques, & Vieira, 2014) have already been described in the literature. Therefore, the second explanation for the significant differences between CWS and CWNS in both samples might be formulated in terms of limited language skills of CWS, probably due to a certain association between stuttering and other language-related medical issues such as hearing disorders of children (cf. a slight association with hearing disorders identified by St. Louis (1992)).

In Sample 2, probability values regarding differences between CWS and CWNS were less convincing than in Sample 1, which, however, can be explained by a comparatively limited sample size with only 31 CWS. The effect sizes in both samples were almost identical. Without application of the Bonferroni adjustment in Table 3, Sample 2 would have completely confirmed the limited German skills of CWS demonstrated in Sample 1, that is, those in grammar, articulation, and a total score of correct answers in the speech and language screening MSSrev (without repetition tasks). Taking into

account the fact that the necessity to apply the Bonferroni adjustment in case of so few—less than ten—statistical tests is questionable (Lang & Secic, 2006), Table 3 results (Sample 2) can be considered to a certain extent a confirmation of the results presented in Table 2 (Sample 1).

As the most prominent finding in the Mann-Whitney *U* tests, GLM, and correlations in both samples, the stuttering tended to co-occur with articulation disorders, which seems to be the most prominent finding in the previous research as well (e.g. Dworzynski & Howell, 2004; Pelczarski & Yaruss, 2014; Sasisekaran, 2014; St. Louis & Hinzman, 1998), although in some studies no significant differences were found between CWS and CWNS groups (Clark, Conture, Walden, & Lambert, 2013; Coalson & Byrd, 2016). This is, however, not surprising taking into account variability in study designs and research questions.

According to the meta-analysis of Ntourou et al. (2011), overall language was on average worse in CWS than in CWNS. In our study, the MSSrev can be considered a screening which tests all linguistic domains. In Sample 1, the total scores of correct answers of CWS were indeed significantly lower than the total scores of CWNS (cf. Table 2). In the smaller Sample 2, no significant differences were found for the total scores including repetition tasks (although the *p* value was marginally significant), but the total scores excluding them (which corresponds to the test version utilized in Sample 1) did identify statistically significant differences with the almost identical effect size as in Sample 1 (cf. Table 3), which confirms the results of the meta-analysis of Ntourou et al. (2011).

We cannot confirm the finding by Ntourou et al. (2011), also supported by Salihovic, Junuzovic-Zunic, Duranovic, and Fatusic (2010), that CWNS have a better command of expressive vocabulary than CWS. No significant differences between CWS and CWNS were found in Samples 1 and 2 in this respect. The choice of the test subjects might have contributed to this discrepancy. In the meta-analysis of Ntourou et al. (2011), only studies with children without language disorders were considered, whereas in our study the only exclusion criterion was inappropriate age.

One possible factor contributing to the differences in vocabulary skills between CWS and CWNS is phonological short-term memory, which is believed to play an important role in word learning (e.g. Gathercole, Service, Hitch, Adams, & Martin, 1999; Gathercole, 2006). Previous findings suggested only subtle differences in nonce word repetition scores between CWS and CWNS that were only significant at certain nonce word lengths (Hakim & Bernstein Ratner, 2004; Anderson, Wagovich, & Hall, 2006). The present study showed no statistically significant differences between CWS and CWNS in the repetition of nonce words, although results of CWS were numerically lower. It cannot be excluded that a larger sample size would have resulted in findings consistent with those of Hakim and Ratner (2004) as well as Anderson et al. (2006) in terms of either significant results or considerable effect sizes.

Also, in Sample 2, CWS scored numerically, but not significantly, lower in the repetition of sentences, the second task for the phonological short-term memory. In a Spearman correlation between severity of stuttering and total scores of correct answers, this domain even yielded a slightly significant result that, however, can be dismissed if the Bonferroni adjustment is applied. A co-occurrence of significant results for grammar and repetition of sentences (see Table 4) might have resulted from the fact that both tasks assess receptive grammar. This close link is reflected in a high correlation between total scores of correct answers in these two MSSrev subtests in Sample 2. Children without German grammar skills were usually unable to reproduce sentences of more than four to five words. Furthermore, disfluencies usually become more frequent when children attempt to produce long and grammatically complex utterances (Bernstein Ratner & Sih, 1987; Logan & Conture, 1995; Melnick & Conture, 2000), and this might have been the case in the sentence repetition tasks because many children tended not simply to repeat sentences, but to modify them syntactically and semantically. For example, instead of the correct repetition of the syntactically more difficult sentence in passive voice, 'Before Lena is picked up, she has to clean up her room', they tended to respond with their own modified sentences in active voice, e.g. 'Lena must clean her house, and then one will pick her up'. Therefore, repetition of sentences should rather be considered a mixed task both on phonological short-term memory and syntax in our view. Unfortunately, this subtest did not yet exist when Sample 1 was tested.

The correlations between the severity of stuttering and language test scores made up the last part of our study. Some researchers have proposed that stuttering restricts language development (Byrd & Cooper, 1989; Arndt & Healey, 2001). Hence, a negative correlation between stuttering severity and verbal skills (scores of language measures) could be expected (Nippold, 2012). However, in the studies by Ryan (1992) and Conture (2000) no statistically significant evidence for the link between stuttering severity and language skills was identified. Our study employed larger sample sizes than those by Ryan (1992) and Conture (2000) and had the potential to make low correlations visible. Sample 2 showed that seven out of eight correlations between the frequency of stuttering events and language test scores were more or less negligible if the Bonferroni adjustment is applied. Since, however, the correlations with grammar and total scores of correct answers yielded significant p values and, results of the Mann-Whitney U tests with both samples can be considered confirmed to a certain extent.

Lim, Lincoln, Chan, and Onslow (2008) reported that bilingual children may show more disfluencies in their less dominant language. In the current study, the percentage of CWS among Germans and immigrants was almost identical in both cases. Evidence for a possible misinterpretation of word finding difficulties of children with immigration background as stuttering was found only in the subsample of 22 children (identified as CWS in kindergartens) who were invited to the REMOVED hospital for a detailed

assessment of their fluency and language skills. Out of 22 children, 18 were confirmed as CWS and the other four were immigrants who might have used word repetitions as a compensation strategy for their word finding difficulties. These children had neither developmental nor chronic stuttering at the time point of their test session in the university hospital so that their word repetitions might have been misinterpreted as stuttering symptoms by kindergarten teachers. However, due to a time span of several months between study sessions in the kindergarten and in the university hospital, it cannot be excluded that some of these four children indeed might have shown some symptoms of developmental stuttering at the time point when they were tested by kindergarten teachers or earlier.

According to the cross-tables, a percentage of male CWS was significantly higher than the percentage of female CWS in both samples. In Sample 1, no other factor was so closely associated with the CWS/CWNS classification as the sex of test subjects, followed by language disorders in the family. The association of stuttering as well as generally language-related problems with the male sex (Ardila, Rosselli, Bateman, & Guzmán, 2000; Bloodstein, & Bernstein Ratner, 2008; Lange, Euler, & Zaretsky, in press; Yairi & Ambrose, 1992) and the heritability of stuttering (Kraft & Yairi, 2012) have been reported in many other studies.

The present study has several advantages as well as drawbacks. One of the advantages is that the data can be considered almost unselected. The second advantage is that the analysis was done on samples which were collected for the purpose of constructing and testing a language screening test. Thus, all persons involved in the data collection and data management were blind as to the purpose of the present analysis. None of these persons were sensitised to detect language comorbidities of stuttering, quite contrary to the study by Arndt and Healey (2001) where speech-language pathologists, who may be considered sensitised to the detection of language deficits, were asked about the occurrence of language disorders from their case loads of CWS. Thirdly, several subtests were conducted with both samples, thus giving an indication about the robustness and replicability of the findings. Fourthly, the current study was done for the German language, which adds to the generality of the findings about the link between stuttering and verbal skills because German is typologically not as analytic as the English language. Finally, the sample sizes were larger than in the previous studies, which increases the probability of detecting small magnitudes of the link between stuttering and language skills.

A limitation of the present study is the method used to diagnose stuttering. Stuttering in pre-school ages is difficult to assess with satisfactory sensitivity and specificity (Yairi & Ambrose, 2005; Neumann, Euler, & Schneider, 2014), even if the focus is on the valid identification of stuttering alone, which was not the case in the current study where the occurrence of disfluencies was one aspect among many others reported by the kindergarten teachers. The severity of stuttering was subjectively judged by kindergarten teachers, and no special test for the identification of stuttering was administered. However,

first, all results (CWS-CWNS classifications) were controlled by speech-language pathologists on the basis of test batteries and audio records. Second, a clinical examination of a subsample of CWS from Sample 2 confirmed that 82% of them indeed had a chronic or developmental stuttering (18 out of 22 children invited to the university hospital were diagnosed with stuttering). In addition, the operationalisation of stuttering was based on the reports by kindergarten teachers who received a special training on speech and language development and disorders. Also, they had a lengthy daily contact of months or even years with the test subjects and with their parents, and thus had sufficient time to find out whether the child stuttered or had difficulties in word finding. It should be also noted that the skills regarding identification of language and fluency disorders belong to the obligatory educational programme of German kindergarten teachers and might have sufficed even without our special training. Hence, the diagnostic criteria in the current research seem sufficient after all.

A misbalance in group sizes between CWS and CWNS is problematic for many statistical tests including cross-tables, Mann-Whitney U tests and GLM, which might be considered a limitation of the study, too. However, to make the calculations more reliable, one would have to increase the sample size considerably, which was not feasible. Also, the CWS group is inevitably much smaller than the CWNS group in any unselected sample.

Some of the most important failures of previous studies described by Nippold (2012) were (a) the failure to match the groups on key factors such as gender and socioeconomic status, (b) the use of screening procedures that excluded CWNS from participating if they showed signs of a language disorder but included CWS regardless of their language competence, and (c) the use of timed speaking tasks to compare the verbal skills of CWS to those of CWNS. In the present study, argument (a) was overcome by having unselected (Sample 1) or almost unselected (Sample 2) samples; (b) was overcome by not employing any exclusion criteria except inappropriate age; (c) was overcome by not employing any timing restrictions and any special termination criteria, that is, children were tested with the whole test, even if several appointments were necessary.

To conclude, we found a weak, but statistically significant link between stuttering and certain expressive language skills. Subtests on grammar and articulation as well as total scores of correct answers in the language screening MSSrev yielded significant results in all utilized statistical tests: Mann-Whitney U tests, correlations, and general linear models. Although some of the significant results in Sample 2 can be eliminated if the Bonferroni adjustment is applied, results of Sample 1 remain significant and demonstrate the same tendencies.

5 FUNDING

6 DECLARATION OF CONFLICTING INTERESTS

The authors declare that there is no conflict of interest.

7 REFERENCES

Anderson, J. D., Wagovich, S. A., & Hall, N. E. (2006). Nonword repetition skills in young children who do and do not stutter. *Journal of Fluency Disorders, 31*, 177–199.

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.

Arndt, J., & Healy, E. C. (2001). Concomitant disorders in school-age children who stutter. *Language, Speech, and Hearing Services in Schools, 32*, 68–78.

Befi-Lopes, D. M., Cáceres-Assenço, A. M., Marques, S. F., & Vieira, M. (2014). School-age children with specific language impairment produce more speech disfluencies than their peers. *Codas, 26*, 439–443.

Bernstein Ratner, N., & Sih, C. (1987). Effects of gradual increases in sentence length and complexity on children's dysfluency. *Journal of Speech and Hearing Disorders, 52*, 278–287.

Bishop, D. V. M. (2003). *Test for the reception of grammar (TROG-2)*. London, UK: Pearson.

Blood, G. W., Ridenour, V. J., Qualis, C. D., & Hammer, C. S. (2003). Co-occurring disorders in children who stutter. *Journal of Communication Disorders, 36(6)*, 427–448.

Bloodstein, O., & Bernstein Ratner, N. (2008). *A handbook on stuttering* (6th ed.). Clifton Park, NY: Delmar.

Byrd, K., & Cooper, E. B. (1989). Expressive and receptive language skills in stuttering children. *Journal of Fluency Disorders, 14*, 121–126.

Clark, C. E., Conture, E. G., Walden, T. A., & Lambert, W. E. (2013). Speech sound articulation abilities of preschool-age children who stutter. *Journal of Fluency Disorders, 38*, 325–341.

Coalson, G. A., & Byrd, C. T. (2016). Phonetic complexity of words immediately following utterance-initial productions in children who stutter. *Journal of Fluency Disorders, 47*, 56–69.

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

Conture, E. G. (2000). Dreams of our theoretical nights meet the realities of our empirical days: Stuttering theory and research. In H. G. Bosshardt, J. S. Yaruss, & H. F. M. Peters (eds.), *Proceedings of the Third World Congress on Fluency Disorders: Theory, research, treatment, and self-help* (pp. 329–344). Nijmegen, The Netherlands: University of Nijmegen Press.

Dollaghan, C., & Campbell, T. (1998). Nonword repetition and child language impairment. *Journal of Speech, Language, and Hearing Research, 41*, 1136–1146.

Dworzynski, K., Howell, P. (2004). Predicting stuttering from phonetic complexity in German. *Journal of Fluency Disorders, 29*, 149–173.

Dworzynski, K., Howell, P., Au-Yeung, J., & Rommel, D. (2004). Stuttering on function and content words across age groups of German speakers who stutter. *Journal of Multilingual Communication Disorders, 2*, 81–101.

Dworzynski, K., Howell, P., Natke, U. (2003). Predicting stuttering from linguistic factors for German speakers in two age groups. *Journal of Fluency Disorders, 28*, 95–112.

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.

Gardner, M. F. (1990). *Expressive One-Word Picture Vocabulary Test—Revised*. Novato, CA: Academic Therapy Publications.

Gathercole, S. E. (2006). Nonword repetition and word learning: The nature of the relationship. *Applied Psycholinguistics, 27*, 513–543.

Gathercole, S. E., Service, E., Hitch, G. J., Adams, A. M., & Martin, A. J. (1999). Phonological short-term memory and vocabulary development: Further evidence on the nature of the relationship. *Applied Cognitive Psychology, 13*, 65–77.

Grissom, R. J., & Kim, J. J. (2012). *Effect sizes for research: Univariate and multivariate applications* (2nd ed.). New York, NY: Taylor & Francis.

Hakim, H. B., & Bernstein Ratner, N. (2004). Nonword repetition abilities of children who stutter: An exploratory study. *Journal of Fluency Disorders, 29*, 179–199.

Hall, N. E., Wagovich, S. A., & Bernstein Ratner, N. (2007). Language considerations in childhood stuttering. In E. Conture & R. Curlee (Eds.), *Stuttering and related fluency disorders* (pp. 153–167). New York, NY: Thieme Medical.

Hall, N. E., Yamashita, T. S., & Aram, D. M. (1993). Relationship between language and fluency in children with developmental language disorders. *Journal of Speech and Hearing Research*, *36*, 568–579.

Hawkins, J. A. (1986). *A comparative typology of English and German: Unifying the contrasts*. Austin, TX: University of Texas Press.

Kraft, S. J., & Yairi, E. (2012). Genetic bases of stuttering: The state of the art, 2011. *Folia Phoniatrica et Logopaedica, 64*, 34–47.

Lang, T. A., & Secic M. (2006). *How to report statistics in medicine* (2nd ed.). Philadelphia, PA: American College of Physicians.

Lange, B. P., Euler, H. A., & Zaretsky, E. (in press). Sex differences in language competence of three- to six-year old children. *Applied Psycholinguistics*. doi: 10.1017/S0142716415000624

Lim, V. M., Lincoln, M., Chan, Y. H., & Onslow, M. (2008). Stuttering in English-Mandarin bilingual speakers: The influence of language dominance on stuttering severity. *Journal of Speech, Language, and Hearing Research, 51*, 1522–1537.

Logan, K., & Conture, E. (1995). Length, grammatical complexity, and rate differences in stuttered and fluent conversational utterances of children who stutter. *Journal of Fluency Disorders, 20*, 35–61.

Melnick, K. S., & Conture, E. G. (2000). Relationship of length and grammatical complexity to the systematic and nonsystematic speech errors and stuttering of children who stutter. *Journal of Fluency Disorders, 25*, 21–45.

Natke, U., Sandrieser P., Pietrowsky, R., & Kalveram, K. T. (2006). Disfluency data of German preschool children who stutter and comparison children. *Journal of Fluency Disorders, 31*, 165−176.

Neumann, K., Euler, H. A., & Schneider, P. (2014). *Identifikation von Stottern im Vorschulalter* [Identification of stuttering at pre-school age]. Cologne, Germany: Demosthenes-Verlag der Bundesvereinigung Stottern & Selbsthilfe e.V.

Neumann, K., Holler-Zittlau, I., van Minnen, S., Sick, U., Zaretsky, Y., & Euler, H. A. (2011). Katzengoldstandards in der Sprachstandserfassung: Sensitivität und Spezifität des Kindersprachscreenings (KiSS) [Fool's gold standards in language screening. Sensitivity and specificity of the Hessian child language screening test (Kindersprachscreening, KiSS)]. *HNO (eHNO), 59*, 97–109.

Nippold, M. A. (1990). Concomitant speech and language disorders in stuttering children: A critique of the literature. *Journal of Speech and Hearing Disorders, 55*, 51–60.

Nippold, M. A. (2012). Stuttering and language ability in children: Questioning the connection. *American Journal of Speech-Language Pathology, 21*, 183–196.

Nippold, M. A., Schwarz, I. E., & Jescheniak, J. D. (1991). Narrative ability in school-age stuttering boys: A preliminary investigation. *Journal of Fluency Disorders, 16*, 289-308.

Ntourou, K., Conture, E. G., & Lipsey, M. W. (2011). Language abilities of children who stutter: A meta-analytical review. *American Journal of Speech-Language Pathology, 20*, 163–179.

Pelczarski, K. M., & Yaruss, J. S. (2014). Phonological encoding of young children who stutter. *Journal of Fluency Disorders, 39*, 12–24.

Riley, G. D. (1994). *A stuttering severity instrument for children and adults – Third edition (SSI-3).* Austin, TX: ProEd.

Ryan, B. P. (1992). Articulation, language, rate, and fluency characteristics of stuttering and nonstuttering children. *Journal of Speech and Hearing Research, 35*, 333–342.

Sandrieser, P., & Schneider, P. (2008). *Stottern im Kindesalter* [Stuttering in childhood] (3rd ed.). Stuttgart, Germany: Thieme.

Salihovic, N., Junuzovic-Zunic, L., Duranovic, M., & Fatusic, A. (2010). Characteristics of vocabulary in school-age stuttering children. *The Journal of International Social Research, 3*, 399–406.

Sasisekaran, J. (2014). Exploring the link between stuttering and phonology: a review and implications for treatment. *Seminars in Speech and Language, 35*, 95-113.

Silverman, S., & Bernstein Ratner, N. (1997). Syntactic complexity, fluency, and accuracy of sentence imitation in adolescents. *Journal of Speech, Language, and Hearing Research, 40*, 93–106.

St. Louis, K. O. (1992). *Coexistence of communication disorders in school children (ASHA Monographs, No. 27).* Rockville, MD: American Speech-Language-Hearing Association.

St. Louis, K. O., & Hinzman, A. R. (1988). A descriptive study of speech, language, and hearing characteristics in school-aged stutterers. *Journal of Fluency Disorders, 13*, 331–355.

Stokes, S. F., Wong, A. M-Y., Fletcher, P., & Leonard L. B. (2009). Nonword repetition and sentence repetition as clinical markers of specific language impairment: The case of Cantonese. *Journal of Speech, Language, and Hearing Research, 49*, 219–236.

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

Van Riper, C. (1971). *The nature of stuttering*. Englewood Cliffs, NJ: Prentice-Hall.

Watkins, R. V., Yairi, E., & Ambrose, N. G. (1999). Early childhood stuttering III: Initial status of expressive language abilities. *Journal of Speech, Language, and Hearing Research, 42*, 1125–1135.

Yairi, E., & Ambrose, N. G. (1992). Onset of stuttering in preschool children. *Journal of Speech, Language, and Hearing Research, 35*, 782–788.

Yairi, E., & Ambrose, N. G. (2005). *Early childhood stuttering. For clinicians by clinicians*. Austin, TX: Pro-Ed.