

PREDICTORS OF SEVERE LANGUAGE DEFICITS WITH CLINICAL IMPLICATIONS IN A REPRESENTATIVE SAMPLE OF GERMAN PRE-SCHOOL CHILDREN

Eugen Zaretsky^{1*} and Benjamin P. Lange²

¹Dr. phil., Department of Phoniatics and Pediatric Audiology, University Hospital of Frankfurt/Main, GERMANY, yevgen.zaretsky@kgu.de

²Dr. phil., Department of Media Psychology, Institute for Human-Computer-Media, Faculty of Human Sciences, Julius Maximilian University of Wuerzburg, GERMANY, benjamin.lange@uni-wuerzburg.de

*Corresponding author

Abstract

The language acquisition process is associated not only with social and demographic factors but also with the physiological condition of the child, that is, his or her physical ability to perceive and process enough language input. However, even in this case some sociolinguistic variables such as participation in therapies and high quality language input can help the child to cope with his/her problems. This study aimed at an examination of such variables. A sample of 741 German children acquiring German as their first or second language (393 boys, 348 girls, age median 70 months) was recruited for a series of language tests conducted during the school enrolment examination. On the basis of tests on articulation, phonological short-term memory, vocabulary, speech comprehension, and grammar (S-ENS, AWST-R, ETS 4-8), as well as questionnaires for parents, an expert panel classified all test participants as needing or not needing medical help in acquiring German. A subgroup of children ($N = 163$) was tested twice: at the age of four and a year or two later. In this subgroup, also data from questionnaires for daycare center teachers were available. Predictors for the classification of children as needing or not needing medical help were examined by classification trees, various association measures, and regressions. The most important sociolinguistic variables were related both to genetic factors and the quality and quantity of the language input. The factor "language disorders in the family" played the major role in the cross-sectional part of the study, the factor "need for additional educational help in acquiring German" in the follow-up part.

Keywords: language acquisition, language disorder, German language, language assessment, language impairment.

1. INTRODUCTION

As was shown in our previous studies, sociolinguistic factors are associated not only with the language competence of German preschoolers, measured by total scores of correct answers in the validated language tests (Zaretsky et al., 2013a), but even with certain grammatical error patterns (Zaretsky et al., 2013a, b; Zaretsky and Lange, 2014a). For the sample of study participants presented here, factors associated with the classification of children as needing or not needing additional educational support in acquiring/learning German were examined in one of our previous papers (Zaretsky et al., 2014). As the most important factor, father's educational level was identified, probably as an indirect link to the family income, followed by immigration background and language-related illnesses, diseases or impairments of the child. Furthermore, as was shown in Zaretsky et al. (2013) for some children from the sample used here, poor German language skills of Turkish preschoolers were associated with very unfavorable sociolinguistic conditions. When compared with Russian children who scored significantly higher in the same language tests, Turkish children attended significantly less often associations or study groups, they had more often a possibility to speak their L1 (Turkish) in the daycare centers, and they spoke less often German at home. In comparison with all other

immigrant children, Turks attended less often nursery schools in the first two years of life and they began later to learn/acquire German.

The study consists of two parts: first, all children were tested cross-sectionally, and, second, a subgroup was tested twice with a year or two in between. In both cases the association between the classification of children as needing or not needing medical help in acquiring/learning German and sociolinguistic factors was analyzed. In the cross-sectional part of the study, sociolinguistic factors were taken from the questionnaires for parents, which were filled out during the test session. In the follow-up part of the study, questionnaires for parents and daycare center teachers were utilized, which were filled out a year or two before the second test session. Consequently, in the follow-up part of the study, the predictive power of questionnaire items was analyzed.

It was hypothesized that in both study parts the quality and quantity of the language input would play an important role in the language acquisition because even children with medical issues can often overcome their language problems in course of time due to sufficient language input. Additionally, some links to the genetic predisposition of children classified as needing medical help were to be expected, for instance, in the form of language disorders in the family. Also, in most cases medical issues of children must have been identified and diagnosed in the first years of life, before the beginning of the presented study. Therefore, one could expect that most of the parents were aware of the medical issues of their children and mentioned them in the questionnaires.

2. METHODS

2.1. Test subjects and tests

In the cross-sectional part of the study, a sample of 741 German children acquiring German as their first or second language (393 boys, 348 girls, age range 60-99 months, median 70.0, 248 MO, 452 BM, 41 unknown) was recruited for a series of language tests conducted during the school enrolment examination. All children were tested by linguistics students and researchers in several public health departments in the German state of Hesse. Children were not pre-selected for the participation in the study, there were no exclusion criteria. Parents were asked to sign an informed consent prior to the language tests and then to fill out a questionnaire with some sociolinguistic variables, predominately related to the language contact and health status of the children.

The language competence of the study participants was tested by three language tests: AWST-R (vocabulary; Kiese-Himmel, 2005), ETS 4-8 (grammar and speech comprehension; Angermeier, 2007), and the official school enrolment test S-ENS with subtests on articulation, repetition of words and sentences (phonological short-term memory), and an insertion of missing sounds in words (Döpfner et al., 2005). Both ETS 4-8 and S-ENS have further subtests which are of no relevance here. AWST-R was used in a short, validated form with 25 items (Neumann and Euler, 2010). The analysis of the data was carried out retrospectively, the original study design aimed at the development of two language tests.

On the basis of these language tests, questionnaires for parents, and audio recordings of the tests, an expert panel (therapists, researchers) classified all test participants as needing (CLIN) or not needing (NCLIN) medical help in acquiring/learning German. For the first group, severe language deficits (mostly < 6th percentile of the reference group in at least one subtest) and additional information on sociolinguistic and medical issues made language-related impairments/illnesses highly probable according to the language experts. All in all, 67 children (9%) were classified as CLIN and 674 (91%) as NCLIN in the cross-sectional part of the study.

A subgroup of children was tested twice: at the age of four in the daycare centers and at the age of five or six in the public health departments. The sample consisted of 163 children (98 boys, 65 girls, age range 37-59 months in the first test session, median 50.0, and 60-81 months in the second test session, median 66.0). Less than a half of these children ($N = 59$ (36%)) were monolingual Germans, and 104 (64%) were bi-/multilingual children. In the first test session, 17 children were classified as CLIN (10%) and 146 as NCLIN (90%). In the second test session, 23 children were classified as CLIN (14%) and 139 as NCLIN (86%). In the first test session, children were tested with a short, validated version of the language test Marburger Sprachscreening (Euler et al., 2010, Neumann et al. 2011a, b) with subtests on speech comprehension, vocabulary, grammar, articulation, and repetition of sentences and nonce words (phonological short-term memory). Sociolinguistic data were documented in the questionnaires for parents and daycare center teachers. Both questionnaires were part of the short form of the language test Marburger Sprachscreening. In the second test session, children were tested with S-ENS, AWST-R, and ETS 4-8, with the same questionnaire for parents as in the cross-sectional part of the study. The time span between the first and the second test sessions was between 7 and 33 months, median 15.0. The variation in the time span values can be explained by the fact that both five-

and six-year-old children are usually invited to the school enrolment examination. For the second test session, no exclusion criteria were applied. In the first test session, only three- to five-year-old children were tested, which means that age was the only exclusion criterion.

In case of immigration background, the nationality and land of origin did not matter: all monolingual children acquiring German were classified as Germans (MO), and all bi-/multilingual children as immigrants (BM).

2.2 Statistical analyses

First, the cross-sectional part of the study was analyzed. Differences between children who were classified as CLIN and NCLIN were calculated by means of Mann-Whitney *U*-tests with total scores of language tests and additionally visualized by boxplots.

Predictors for the classification of children as CLIN/NCLIN were examined by two classification trees with the growing method “Exhaustive CHAID”. Independent variables were taken from the questionnaire for parents. The classification trees were calculated in two different ways for a comparison. Because of the Bonferroni adjustment of the *p*-value, it was not possible to include all the available sociolinguistic factors into the classification trees. Therefore, a pre-selection of the most important factors was necessary. For the first classification tree, the most important sociolinguistic variables were chosen subjectively by the expert panel. For the second classification tree, only those variables were included which were significantly associated with the classification of children as CLIN/NCLIN (Zaretsky and Lange, 2014b). This association was examined, depending on the characteristics of the data (nominal, ordinal, metric), by cross-tables with Chi-square (nominal), linear-by-linear associations (ordinal) or by Mann-Whitney *U*-tests (metric).

Results of the classification trees were controlled by binary logistic regressions with the “enter” method. The same two sets of independent variables were used as in the first and second classification trees. The dependent variable was again the CLIN/NCLIN classification.

For the subgroup of children who were tested twice, sociolinguistic factors documented in the questionnaires for parents and daycare center teachers in the first test session were analyzed in respect to their association with the classification of children as CLIN/NCLIN in the second test session. The association measures were the same as mentioned above for the first test session: Chi-square or linear-by-linear (l-b-l) association for the nominal and ordinal data respectively and Mann-Whitney *U*-tests for the metric values.

For the factors significantly associated with the CLIN/NCLIN classification, the data were additionally analyzed by a classification tree with the growing method “Exhaustive CHAID”. This analysis aimed at the identification of the most important factors, which constitute “the top of the tree”.

For the school marks given by daycare center teachers to the German language competence of the children in the first test session and at the beginning of daycare center attendance, Spearman-correlations with the total scores of the language tests S-ENS, AWST-R, and ETS 4-8 in the second test session were calculated in order to quantify the predictive power of these subjective estimations given by teachers.

Non-parametric tests were chosen because the metric data were not normally distributed according to Kolmogorov-Smirnov tests ($p < .05$). Sample sizes varied in different calculations depending on how many parents and daycare center teachers answered the questions and how many children completed the tests. Total scores were calculated only for those children who answered all questions.

3. RESULTS

First, differences between CLIN and NCLIN children in the cross-sectional part of the study are presented, see Table 1.

Table 1. Differences between children needing (CLIN) and not needing (NCLIN) medical help in acquiring German regarding total scores of the language tests; Mann-Whitney *U*-tests

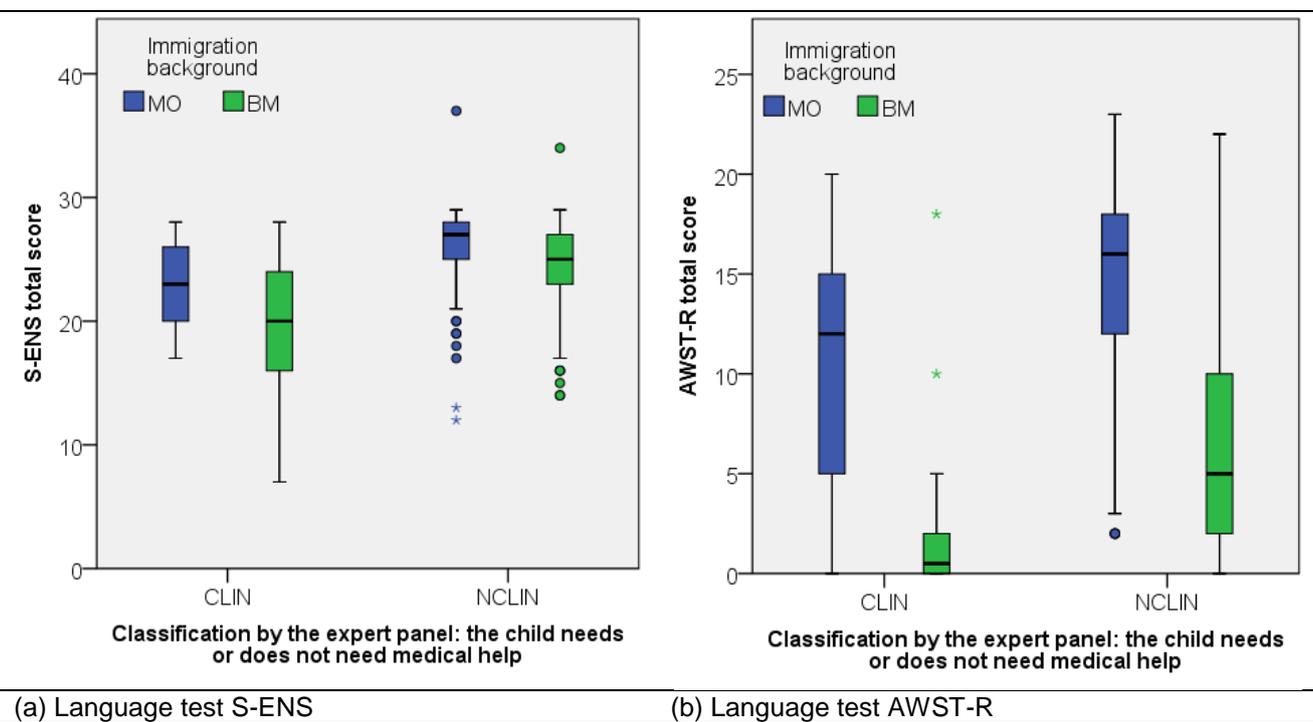
	ETS 4-8 grammar	ETS 4-8 speech comprehension	AWST-R vocabulary	S-ENS repetition of nonce words	S-ENS repetition of sentences	S-ENS articulation	S-ENS insertion of sounds in words
<i>U</i>	7016	6407	7549	9875	7693	9729	8687
<i>Z</i>	-6.81***	-8.44***	-7.16***	-4.28***	-6.23***	-5.39***	-5.22***
<i>N</i>	659	684	675	605	605	604	606
Mean	20.02	16.18	3.63 (5.58)	4.15 (1.73)	2.43 (2.21)	8.85 (1.52)	4.98 (2.41)

CLIN (SD)	(11.63)	(5.53)					
Mean	31.48	23.38	9.36 (6.58)	5.01 (1.11)	3.95 (1.42)	9.69 (0.72)	6.57 (1.40)
NCLIN (SD)	(8.29)	(5.02)					

*** $p < .001$

In Figure 1, the differences between CLIN and NCLIN children are exemplified by means of the total scores of the language tests S-ENS and AWST-R. Because immigrant children usually score lower than monolingual Germans (e.g., Zaretsky et al., 2014), the results were presented separately for these two groups. The boxes demonstrate the median, the first and third quartiles. Fifty percent of all cases (tested children) are represented within the box and 95% of all cases are located within the whiskers.

Figure 1. Total scores of the language tests S-ENS and AWST-R analyzed separately for children needing (CLIN) or not needing (NCLIN) medical help in acquiring German, subdivided into children with (BM) and without (MO) immigration background



Next, two classification trees were calculated. First, an expert panel chose subjectively most important factors which might be associated with the classification of children as CLIN/NCLIN. These factors were “sex of the child”, “age when the mother began to learn/acquire German” (for immigrants), the same for the father, “language disorders in the family” (yes/no), “problems with reading and writing in the family” (yes/no), “whether the child has some language-related illness, or impairment, or disease” (yes/no), “length of daycare center attendance in months”, “immigration background” (BM/MO), “whether the child hears well” (never – seldom – sometimes – often – always), “age of the child in months”, and “whether the child attended a nursery school in the first two years of life” (yes/no).

The chosen model predicted correctly the CLIN/NCLIN classification in 91% of the cases. The most important factors, that is, the top of the classification tree, can be summarized as follows:

- The highest priority for the classification was attributed to the “language disorders in the family” ($\chi^2_{(1)} = 27.63, p < .001$). Children from such families were classified as CLIN in 29% of the cases, whereas children from other families only in 8%.
- For children with language disorders in the family, no further factors were identified. For children without language disorders in the family, “immigration background” was the most important factor ($\chi^2_{(1)} = 10.73, p = .004$). Among the BM children, 11% were classified as needing medical help, whereas for MO children this value was only 4%.

- For MO children, no further important factors were found. For BM, “length of daycare center attendance” was identified as the most important factor ($\chi^2_{(1)} = 14.64, p = .007$). BM children who attended daycare centers for less than 23 months were more often classified as needing medical help (19%) in comparison with children in the subgroup with 23+ months of attendance (6%).

In the next classification tree, the classification of children as CLIN/NCLIN functioned again as the dependent variable. The selection method for independent variables, however, changed. Only those variables which were statistically significantly associated with the classification CLIN/NCLIN entered the analysis. Because the pre-selection is already described in another article (Zaretsky and Lange, 2014b), only the results of the classification tree are presented here. The following variables were excluded in the pre-selection: “age of the child in months”, “how often the child plays with children who speak the same non-German language” (never – seldom – sometimes – often – always; only for BM children), “hearing disorders of the child” (yes/no and “hearing well” never – seldom – sometimes – often – always), “whether the child attended the daycare center for half a day or a full day”, “attendance of a nursery school in the first two years of life” (yes/no), “father’s reading and writing skills in German” (not so good – good – very good), and “length of daycare center attendance in months”. The variables chosen for further analysis were “sex of the child”, “language disorders in the family” (yes/no), “problems with reading and writing in the family” (yes/no), “whether the child has some language-related illness, impairment, or disease (yes/no)”, “immigration background” (BM/MO), “educational level of the mother” (no school certificate – secondary school certificate – intermediate school leaving certificate – matura – high school; an ordinal variable), the same for the father, “which languages are spoken at home” (only German – German and other language(s) – only other language(s)), “mother’s reading and writing skills in German” (not so good – good – very good), “age when the mother began to learn German”, the same for the father, and “classification of the child as needing or not needing additional educational support in acquiring/learning German”. The classification of all study participants as needing or not needing educational support was carried out by the same expert panel as described above. Children who were classified as needing such support usually scored at least one standard deviation below the average value of the reference group in at least one subtest.

The second classification tree also predicted 91% of the CLIN/NCLIN classification correctly. However, the structure of the tree was somewhat different:

- “Language disorders in the family” was identified, again, as the most important variable ($\chi^2_{(1)} = 27.63, p < .001$). Among children with such disorders in the family, 29% were categorized as CLIN. In the group of children without such disorders, only 8% were categorized as CLIN.
- For children whose relatives have language disorders, no further factors were identified. For children without language disorders in the family, the next important factors was “mother’s reading and writing skills in German” ($\chi^2_{(2)} = 23.22, p < .001$). In case of “not so good” German skills, 19% of children were classified as CLIN. In case of “good” German skills, these were 10%. If the mother could read and write German very well, only 5% were classified as CLIN.
- One more important factor was found for the subgroup of children whose mothers could read and write German “very well”: “language-related disorders of the child” ($\chi^2_{(1)} = 9.85, p = .007$). Children with such disorders were classified as CLIN in 9% of the cases, whereas children without such disorders only in 3% of the cases.

Because both classification trees explained 91% of the variance, additional calculations were necessary to find out which classification tree model delivered more reliable results. Therefore, two binary logistic regressions were calculated with the same two sets of independent variables. In the first regression (Nagelkerke $R^2 = .259, N = 191$, constant: $B = 2.203$, Wald = 83.042, $p < .001$, $\text{Exp}(B) = 9.053$), the same variables were used as in the first classification tree. Two variables yielded significant results: “language disorders in the family” ($B = 2.730$, Wald = 7.285, $p = .007$, $\text{Exp}(B) = 15.328$) and “length of daycare center attendance in months” ($B = .052$, Wald = 3.924, $p = .048$, $\text{Exp}(B) = 1.053$). In the second regression (Nagelkerke $R^2 = .452, N = 199$, constant: $B = 1.940$, Wald = 82.285, $p < .001$, $\text{Exp}(B) = 6.960$), independent variables from the second classification tree were utilized. Only “language disorders in the family” reached statistical significance: $B = 2.444$, Wald = 5.350, $p = .021$, $\text{Exp}(B) = 11.518$.

For the subgroup of children who were tested twice, an analysis was carried out to find out which sociolinguistic factors documented one or two years before the second test session by parents and daycare center teachers were associated with the classification of children as CLIN/NCLIN in the second test session. Again, statistical associations between sociolinguistic variables and the CLIN/NCLIN classification were examined by means of cross-tables and Mann-Whitney U -tests. Not significantly associated were the variables “which languages are spoken at home” (only German, German and other language(s), only other language(s)),

“whether the child attended a nursery school in the first two years of life” (yes/no), “whether the child attends the daycare center for half a day or a full day”, “whether the child attends the daycare center regularly” (yes/no), “the child is mentally retarded” (yes/no), “the child does not hear well” (never – seldom – sometimes – often – always), “in the daycare center group there is at least one more child who speaks the same non-German language as the study participant” (yes/no), “how often the study participant speaks with this child” (never – seldom – sometimes – often – always), “whether the child likes to attend the daycare center” (never – seldom – sometimes – often – always), “whether the child speaks his/her mother tongue (if not German) appropriately for his/her age” (yes/no), “whether the child is in language therapy” (yes/no), “whether there are problems with reading and writing in the family” (yes/no), “whether there are language disorders in the family” (yes/no), “whether there is stuttering in the family” (yes/no), “first language of the mother” (German, Turkish, Russian, Italian, Croatian, Serbian, Greek, English, Arab, Spanish, other languages), “first language of the father” with the same options, “mother’s educational level” (no school certificate – secondary school certificate – intermediate school leaving certificate – matura – high school), “father’s educational level” with the same options, “whether the child has some illness/disease/impairment which might negatively influence language acquisition” (yes/no), “difficult or premature birth” (yes/no), “motor disturbance” (yes/no), “head and face malformations such as cleft lip and palate” (yes/no), “frequent otitis media” (yes/no), “sight disorder” (yes/no), “length of daycare center attendance in months”, “time span between the first and the second test sessions in months”, “age when the mother began to learn German”, and “age when the father began to learn German”.

The following variables were significantly associated with the classification of children as CLIN/NCLIN: “whether the child likes to play with other children” ($I-b-I_{(1)} = 4.55, p = .033$) with the options “never – seldom – sometimes – often – always”, “whether the child speaks out when playing” ($I-b-I_{(1)} = 22.14, p < .001$) with the same options, “whether the child plays with German speaking children” ($I-b-I_{(1)} = 6.97, p = .008$) with the same options, “age when the child had enough language contact to learn German” (in years, $I-b-I_{(1)} = 14.38, p < .001$), “school mark given by a daycare center teacher to the language competence of the child when he/she began to attend the daycare center” ($I-b-I_{(1)} = 15.39, p < .001$) with the options from 1 “excellent” to 6 “very bad”, “school mark given by a daycare center teacher to the language competence of the child in the first test session” ($I-b-I_{(1)} = 19.23, p < .001$) with the same options, “whether the child attends an association or a study group” (yes/no; $\chi^2_{(1)} = 8.84, p = .003$), “language the mother prefers to speak at home” ($\chi^2_{(6)} = 22.67, p = .001$) with the options “German, Turkish, Russian, Italian, Croatian, Serbian, Greek, English, Arab, Spanish, other languages”, “language the father prefers to speak at home” ($\chi^2_{(6)} = 14.92, p = .037$) with the same options, and “language the child prefers to speak at home” ($\chi^2_{(6)} = 20.96, p = .002$) with the same options. Also, in the first test session, CLIN children were significantly more often classified by the language experts as needing additional educational support than NCLIN children (yes/no; $\chi^2_{(1)} = 10.09, p = .001$). Marginally significant was an association between the CLIN/NCLIN classification and the variable “whether the child plays with German speaking children after the daycare center time” (yes/no): $\chi^2_{(1)} = 3.07, p = .080$. Also, in a Mann-Whitney *U*-test there was a marginally significant difference between CLIN and NCLIN children regarding their age in months in the first test session: $U = 1233, Z = -1.77, p = .077, M_{CLIN} = 51.4, SD = 2.62$, versus $M_{NCLIN} = 49.8, SD = 4.40$.

All the variables with statistically significant results entered a classification tree to identify the most important factors for the prediction of the CLIN/NCLIN classification. Also, the variables “sex” and “immigration background” entered the classification tree because they had already been shown above to be associated with the classification CLIN/NCLIN and they did not change between the two test sessions. The classification tree predicted 86% of the cases correctly. Only one factor yielded a significant result: classification of children as needing or not needing additional educational support in acquiring/learning German: $\chi^2_{(1)} = 10.09, p = .001$. Among children who were classified as needing educational support in the first test session, 30% were classified as CLIN in the second test session. Among children who were classified as not needing educational support in the first test session, 8% were classified as CLIN in the second test session. Indeed, there was a significant phi-correlation between these two classifications of children: $\phi = .249, p = .001, N = 163$.

Noteworthy are also significant correlations between the school marks given by the respective daycare center teachers to the language competence of the children in the questionnaires of the first test session and the total scores of the language tests in the second test session, see Table 2.

Table 2. Spearman correlations (r_s) between school marks given by daycare center teachers in the first test session and the total scores of the language tests in the second test session

	ETS 4-8 grammar	ETS 4-8 speech comprehension	AWST-R vocabulary	S-ENS repetition of nonce words	S-ENS repetition of sentences	S-ENS articulation	S-ENS insertion of sounds in words
Language competence in the first test session	-.766***	-.695***	-.814***	-.336*	-.646***	-.398**	-.586***
<i>N</i>	60	60	60	57	57	57	57
Language competence at the beginning of daycare center attendance	-.724***	-.591***	-.696***	-.373**	-.593***	-.307*	-.552***
<i>N</i>	60	60	60	57	57	57	57

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

In spite of numerical differences between the correlation coefficients of school marks given to the language competence in the first test session and at the beginning of daycare center attendance, Fisher's r -to- z transformation demonstrated in pairwise comparisons of correlation coefficients that none of these differences reached a statistically significant level ($ps > .05$).

4. DISCUSSION

The study aimed at an identification of sociolinguistic factors associated with the classification of children as needing (CLIN) or not needing (NCLIN) medical help in acquiring/learning German. The results of the language tests of children who were classified as CLIN by language experts usually corresponded to the lowest 5% of the reference sample. Additionally, children in the CLIN group often had language-related illnesses, impairments or diseases such as specific language impairment or Down syndrome, as far as one could deduce from the questionnaires for parents and daycare center teachers, test batteries, and audio recordings of the test situation. The reliability of these sources was shown to be good for stuttering and hearing disorders for children from the sample used here in the follow-up part of the study (Neumann et al., 2011b).

First, statistically significant differences in the total scores of the language tests ETS 4-8, S-ENS, and AWST-R were analyzed for CLIN and NCLIN children in the cross-sectional part of the study. CLIN children scored highly significantly lower in the subtests on grammar, speech comprehension, vocabulary, repetition of sentences and nonce words (phonological short-term memory), articulation, and insertion of missing sounds in words. Figure 1 visualizes these differences in two boxplots for S-ENS and AWST-R, separately for children with and without immigration background. According to this figure, both in S-ENS and AWST-R CLIN children scored on the average lower than NCLIN children. Additionally, BM children scored lower than MO children. The boxes and whiskers of AWST-R boxplots were much longer than those of S-ENS boxplots because of a higher number of items in AWST-R.

Two classification trees were tried out for the identification of the most important factors associated with the classification of children as CLIN/NCLIN. In both cases independent variables were taken from the questionnaires for parents, but in the first classification tree the pre-selection of sociolinguistic variables was carried out subjectively by an expert panel. In the second one the pre-selection was derived from various statistical tests which measured an association between sociolinguistic variables and the CLIN/NCLIN classification: Chi-Squares, linear-by-linear associations, and Mann-Whitney U -tests. Consequently, the pre-selection in the second classification tree was more objective and was supposed to yield a better result, measured by a percentage of correctly predicted cases (classifications of children as CLIN/NCLIN). However, both classification trees predicted correctly 91% of the cases and resulted in somewhat different "tops of the trees", although in both cases "language disorders in family" turned out to be the most important factor.

An additional analysis of the same two sets of sociolinguistic factors by two binary logistic regressions made clear that the objective choice of independent variables resulted in a more reliable model. The first regression, which was based on subjectively chosen factors, explained 26% of the variance, whereas the

second regression, which was based on objectively chosen factors, accounted for 45%. Therefore, in the further analysis of the sociolinguistic factors in the follow-up part of the study, only factors statistically associated with the classification of children as CLIN/NCLIN entered the classification tree.

In the following, the results of association measures will be discussed in more detail. Children who were classified as CLIN by the language experts displayed the following characteristics in the cross-sectional part of the study:

- They were more often boys than girls,
- Their relatives had language disorders and/or “problems with reading and writing”,
- They had more often some language-related illnesses, diseases or impairments,
- They were more often raised bi-/multilingually,
- Educational level of their parents was comparatively low,
- They often spoke not only German, but also (an)other language(s) at home or they did not speak German at home at all,
- Their mothers could read and write German either “not so well” or “well”, but not “very well”,
- Their parents began late to learn German,
- They needed more often, in comparison with NCLIN children, additional educational support in acquiring/learning German according to language experts.

The second classification tree which is considered more reliable here than the first one demonstrated that the most important of these factors was “language disorders in the family”, followed by “mother’s reading and writing skills in German” for those children who did not have relatives with language issues. For the subgroup of children whose mothers had a very good German competence, “language-related disorders of the child” was identified as the most important factor. The regression with the same independent variables confirmed that “language disorders in the family” was the most important factor in the chosen set of sociolinguistic characteristics.

A subgroup of children was tested twice: first when they were four years old and then a year or two later during the school enrolment examination. Questionnaires for parents and daycare center teachers from the first test session were utilized to predict the classification of children as CLIN/NCLIN in the second test session. Most variables were excluded because no statistically significant association with the classification of children was found. This was, unexpectedly, also valid for factors like “language-related illnesses of the child” and “language disorders in the family”, that is, variables which demonstrate a link to a genetic predisposition of the test subjects to have some language-related medical issues. Variables which were significantly associated with the CLIN/NCLIN classification were rather related to the quality and quantity of the language input and to the language competence of the child in the first test session as well as at the beginning of the daycare center attendance. Children who were classified as CLIN in the second test session, had the following characteristics in the first one:

- They did not like to play with other children,
- They did not like to speak out when playing,
- They played with German-speaking children not so often as NCLIN children, both during and after the daycare center time,
- They began later to learn/acquire German,
- They received worse school marks for their language competence in German from the daycare center teachers; this was valid both for the time of the first test session and for the time when children began to attend the daycare center,
- They did not attend any associations or study groups,
- Both parents and children preferred to speak other languages than German at home,
- They were more often classified as needing additional educational help in acquiring/learning German, although they were marginally significantly older than NCLIN children and thus had more time to profit from the language input.

These factors are more related to language contact than to medical issues, which reveals that in the follow-up part of the study not the physiological (medical) factors which caused the language deficits were identified but rather the sociolinguistic factors which prevented some children from overcoming these problems.

The list of significant factors in the follow-up study is longer than in the cross-sectional study because the number of available sociolinguistic variables in the follow-up study was higher: not only a questionnaire for parents, but also a questionnaire for daycare center teachers. In the cross-sectional study, children were

tested in the public health departments, which made the distribution of questionnaires among daycare center teachers impossible.

Again, all significant and marginally significant factors from the questionnaires entered a classification tree to identify the most relevant sociolinguistic variables associated with the classification of children as CLIN/NCLIN. Only one significantly associated factor was identified, namely whether the child was classified as needing or not needing additional educational support in the first test session. Indeed, out of 163 children, 12 (7%) were classified as needing both educational and medical support in acquiring German. The result was also confirmed by a phi-correlation for binary variables. Also, as was mentioned in the Introduction, the analysis of the sociolinguistic factors associated with the classification of children as needing or not needing additional educational help in acquiring/learning German revealed among important factors, although not as the most important one, language-related illnesses, diseases or impairments of the child, which constitutes a clear link to the CLIN/NCLIN classification.

The next finding demonstrated the predictive power of the teachers' subjective estimations of children's language competence measured by language tests one or two years later. School marks given to the language competence of children correlated almost in all cases highly significantly with the total scores of the language tests S-ENS, AWST-R, and ETS 4-8. Negative correlation coefficients mean that low (= good) marks for the language competence of the child in the first test session or at the beginning of daycare center attendance correlated with good results in the language tests of the second test session. This means that children with a good command of German in the first test session and/or at the beginning of daycare center attendance delivered high total scores of correct answers in the second test session. Children with a bad command of German in the first test session still had a comparatively bad command of German in the second one. Obviously, the daycare center teachers are capable of estimating adequately the language skills of children they take care of, otherwise the school marks would have had much lower predictive power for the results of language tests a year or even two years later.

Subjectively, the correlation coefficients between the estimation of the language competence in the first test session and the results of the language tests in the second one were considerably higher than the correlation coefficients of the language competence at the beginning of daycare center attendance and the results of the language tests in the second test session. However, Fisher's *r*-to-*z* transformation found no statistically significant differences in pairwise comparisons, probably due to low sample sizes in all correlations. It is highly probable that higher sample sizes would have resulted in significant differences between these two sets of correlations, simply because the estimation of the language competence of the child at the beginning of daycare center attendance, that is, several months or even up to two years before the first test session, cannot be as precise as the estimation of the language competence at the moment of filling out the questionnaire, that is, during the first test session.

To sum up, in the cross-sectional part of the study, a link between the CLIN/NCLIN classification and the genetic predisposition of children to language-related disorders is clear. Those who had relatives with language disorders and relatives who could not read and write properly were in a risk group for specific or not specific language impairments. However, the quality and quantity of the language input also played an important role: the educational level of parents, how often German was spoken at home, age when the parents began to learn German, and especially the mothers' reading and writing skills in German. In the follow-up part of the study, the link to sociolinguistic factors was even more prominent. Children who were classified as CLIN in the school enrolment examination were the same who did not play with other children, who began later to acquire/learn German, and who did not attend any association or study group, among other things.

Unexpectedly, the information on hearing disorders was not associated with the classification of children as CLIN/NCLIN. This can be partly attributed to the fact that obviously some families misinterpreted the question "Does your child hear well?" as "Does your child behave?" because in German the same verb *hören* is used in both meanings. Also, it should be taken into account that for ordinal data like "never – seldom – sometimes – often – always" linear-by-linear associations were chosen as the adequate association measure. If the Chi-square had been chosen instead, then the hearing disorders would have been among significantly associated factors in the cross-sectional part of the study (but not in the follow-up part).

Also, the findings indicate that children of immigrants are still disadvantaged in comparison with monolingual German children. In the presented study, immigrants were classified more often than Germans as needing medical help. The role of an immigration background can be attributed to two factors: (a) many children had no or only limited access to medical services comparable to those available in Germany in their lands of origin, and hence one can assume higher rates of various curable language-related illnesses/diseases

in this subgroup; (b) because of poor language skills of some immigrant parents their children remained inadequately treated or untreated in German medical facilities.

REFERENCE LIST

- Angermeier, M. J. W. (2007). ETS 4-8. Entwicklungstest Sprache für Kinder von 4 bis 8 Jahren. Frankfurt am Main: Harcourt Test Services GmbH.
- Döpfner, M., Dietmar, I., Mersmann, H., Simon, K., & Trost-Brinkhues, G. (2005). S-ENS. Screening des Entwicklungsstandes bei Einschulungsuntersuchungen. Göttingen: Hogrefe.
- 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. *HNO*, vol. 58.
- Kiese-Himmel, C. (2005). AWST-R. Aktiver Wortschatztest für 3- bis 5-jährige Kinder – Revision. Göttingen: Hogrefe.
- Neumann, K. & Euler, H. A. (2010). Einführung einer flächendeckenden Sprachstandserfassung in Hessen. Forschungsbericht 2010. Frankfurt am Main: Universität Frankfurt am Main.
- Neumann, K., Holler-Zittlau, I., Van Minnen, S., Sick, U., Zaretsky, Y., & Euler, H. A. (2011a). Katzensgoldstandards in der Sprachstandserfassung. Sensitivität und Spezifität des Kindersprachscreenings (KiSS). *HNO*, vol. 59.
- Neumann, K., Zaretsky, Y., & Euler, H. A. (2011b). Einführung einer flächendeckenden Sprachstandserfassung in Hessen. Forschungsbericht 2011. Frankfurt am Main: Universität Frankfurt am Main.
- Zaretsky, E. & Lange, B. P. (2014a). Influence of intra- and extralinguistic factors on the distribution of plural allomorphs in German. *California Linguistic Notes*, vol. 39.
- Zaretsky, E. & Lange, B. P. (2014b). Sociolinguistic factors associated with the classification of German preschoolers as needing or not needing educational or medical help in acquiring German. Conference "Language Acquisition & Learning Research: Tasks, Techniques, Outlook", Magnitogorsk, Russia, 27-29.11.2014.
- Zaretsky, E., Euler, H. A., Neumann, K., & Lange, B. P. (2014). Sociolinguistic predictors of language deficits in pre-school children with and without immigrant background. Proceedings of the Asian Conference on Language Learning, 17.-20.04.2014, Osaka, Japan.
- Zaretsky, E., Lange, B. P., Euler, H. A., & Neumann, K. (2013b). Acquisition of German pluralization rules in monolingual and multilingual children. *Studies in Second Language Learning and Teaching*, vol. 3.
- Zaretsky, E., Neumann, K., Euler, H. A., & Lange, B. P. (2013a). Pluralerwerb im Deutschen bei russisch- und türkischsprachigen Kindern im Vergleich mit anderen Migranten und monolingualen Muttersprachlern. *Zeitschrift für Slawistik*, vol. 58.