Both Are Worse: Nonword Repetition Tasks With a High vs. Low Linguistic Load as Clinical Markers for Language(-Related) Impairments

Eugen Zaretsky, Marburg University Hospital, Germany
Benjamin P. Lange, IU International University of Applied Sciences, Germany
Christiane Hey, Marburg University Hospital, Germany

The Asian Conference on Education 2025
Official Conference Proceedings

Abstract

Purpose: Nonword repetition (NWR) tasks are widely used for the detection of children with deficits in phonological short-term memory (PSTM). Such deficits can be indicative of language(-related) medical impairments with or without comorbidities. However, previous studies showed a considerable influence of children's language competence on NWR performance. Therefore, some language-unspecific NWR tasks were developed that should minimize linguistic interferences and provide a more direct access to PSTM.
Methods: In the current retrospective study, language-specific and language-unspecific NWR tasks were compared regarding their associations with children's language skills and language(-related) impairments. Ten samples of four- and five-year-old monolingual and bilingual German children that were tested either with language-specific or language-unspecific NWR tasks were compared regarding linguistic interferences, associations between NWR performance and language(-related) medical issues as well as possible influence factors on the NWR scores. Sample sizes varied from 161 to 2,750. Seven NWR tests were utilized.
Results: In most cases, NWR performance depended on German articulation/phonotactics skills and was hardly associated with language(-related) impairments. Monolinguals outperformed bilinguals in language-specific nonwords under condition of significantly better German articulation skills. Language-unspecific nonwords minimized the influence of language competence on the NWR performance. However, this "culture-fairness" did not result in higher associations with language(-related) medical issues in comparison with language-specific nonwords.
Discussion and Conclusion: Neither language-specific nor language-unspecific NWR tasks contribute to the diagnostics of language(-related) impairments. Independent cut-off criteria for such impairments based on the NWR performance would be misleading due to strong linguistic interferences even in case of comparatively language-unspecific nonwords.

Keywords: working memory, phonological short-term memory, nonword repetition, language assessment, German language

iafor
The International Academic Forum
www.iafor.org

Introduction

Deficits in phonological short-term memory (PSTM) or, more generally, working memory, are often utilized as a clinical marker for language(-related) impairments and disorders (e.g., Masoura et al., 2020; Penke, 2018). Numerous studies demonstrated significant associations between such deficits and Developmental Language Disorder (Schwob et al., 2021) as well as language impairments with comorbidities such as severe hearing loss (Ching et al., 2019) or Attention-Deficit/Hyperactivity Disorder (Redmond et al., 2011).

However, PSTM tests that are based on the repetition of nonwords, real words, sentences, and digit spans also assess language competence, because these tasks are usually language-specific (Grimm, 2016; Verhagen et al., 2017). For instance, sentence repetition requires not only PSTM, but also vocabulary, articulation, and grammar skills and is, therefore, more suitable for language assessment than for PSTM assessment (Polišenska et al., 2015). Among repetition tasks, nonwords have the lowest linguistic load, but many of them contain real morphemes, rhyme with existing words, and follow the phonotactic rules of the language they are based on. Therefore, they can put children with a weak competence in the language nonwords are based on (e.g., immigrants) at a disadvantage (Zaretsky & Lange, 2022). Bilinguals' weak performance in repetition tasks can be misinterpreted as deficits in PSTM (Engel de Abreu et al., 2013; Grimm, 2016).

Various tasks were developed for the reduction of the linguistic load in PSTM tests: meaningless sentences, low frequency words (that are unknown to most children and can be used as nonwords), nonwords without morphological elements (but with language-specific phonemes and consonant clusters), and "quasi-universal"/"cross-linguistic" nonwords. The latter avoid all more or less language-specific features (e.g., morphological elements, consonant clusters, stress patterns, rhymes with existing words) and can, therefore, be considered most "culture-fair" and language-unspecific. "Quasi-universal" nonwords (QUNW) developed by Chiat (2015) seem to represent the most promising new approach for PSTM assessment and were, therefore, in the focus of the present study.

Nonword repetition (NWR) tasks were integrated in many language tests as possible clinical markers for language impairments. For instance, the German language screening "Kindersprachscreening" ([KiSS], Holler-Zittlau et al., 2011) utilizes them, as independent cut-off criterion, for the detection of monolingual German children who need medical assistance in acquiring German. In spite of the wide-spread use of NWR tasks, hardly any statistical evidence can be found that their dichotomous results can be "translated" into acceptable sensitivity and specificity values for the detection of children with language(-related) impairments (LRI) (Schwob et al., 2021). Most studies only found evidence for statistically significant differences in NWR scores of correct answers between typically developed and language-impaired children (e.g., Boerma et al., 2015). However, such differences do not guarantee that NWR tasks can be used as a reliable clinical marker for LRI. Also, there is no evidence that a relative "culture-fairness" of language-unspecific nonwords can improve the quality criteria of such clinical markers in comparison with language-specific nonwords.

In the current study, seven NWR tasks with zero to high linguistic loads were scrutinized on the basis of ten samples of four- and five-year-old German children. The study aimed to quantify associations between total scores of correct answers in such tasks and children's LRI, under consideration of their German language competence. Also, possible linguistic interferences and other influence factors, including monolingualism vs. bilingualism, on the NWR performance were analysed. Two gold standards were used for the assessment of children's LRI: running or completed language therapy (LT) and LRI that is known to parents (documented in questionnaires).

Methods

Test Subjects

For this study, ten samples of German four- and five-year-old kindergarten children were re-analysed retrospectively. No inclusion or exclusion criteria were applied in the original prospective studies except for appropriate age (4;0–4;11 or 5;0–5;11 years) and a written consent signed by parents.

Information on LT and LRI was documented in questionnaires for parents in dichotomous items "Does/did the child undergo a therapy due to a language impairment?" and "Has the child any language(-related) impairment or disorder?" (yes/no). Phi-correlations between these two items were predominantly low. Therefore, these items were analysed separately. For sample characteristics, including phi-correlations between LT and LRI, see Table 1.

All children were classified as monolinguals or bi-/multilinguals (called "bilinguals" from now on) on the basis of questionnaires for parents. The largest subgroups of bilinguals spoke Turkish, Russian, and Arabic languages at home.

In all samples, monolinguals and bilinguals did not differ in age in months according to Mann-Whitney U-tests and in biological sex according to Chi-square tests ($p > .05$).

All children were tested in German kindergartens by PhD students, researchers, and interns of three university hospitals. Children's answers were audio-recorded and controlled by postdoc researchers (linguists).

Test Materials

All children were tested with language tests that included NWR tasks. Additionally, their German vocabulary and articulation skills were assessed. All tests except QUNW are widely used, standardized, and validated for German pre-schoolers. Characteristics of language tests can be found in Table 2, characteristics of nonwords in Table 3.

NWR tasks differ in terms of item length, number of items, and linguistic load. The latter is defined as frequency of morphological elements (suffixes, prefixes, and stems) and of phonological elements (language-specific phonemes and consonant clusters) of the German language. Only one NWR task, QUNW, did not contain any of such elements and can be considered language-unspecific. All other NWR tasks were classified as language-specific. Generally, articulation errors (e.g., rhotacism, sigmatism) were not considered as NWR (or PSTM) errors in NWR tasks.

Table 1a
Sample Characteristics: Four-Year-Old Children

	S1	S2	S3	S4	S5
n	2,750	1,735	161	194	1,589
Nonword repetition task	KiSS.2 subtest	SSV subtest	HASE subtest	Mottier	QUNW
Vocabulary and articulation tasks	KiSS.2 subtests	KiSS.2 subtests	KiSS.2 subtests	KiSS.2 subtests	KiSS.2 subtests
Age (in months): *M* ±*SD*	52.8 ±3.4	53.4 ±3.4	51.5 ±1.9	51.8 ±2.5	53.9 ±3.2
Age span	48–59	48–59	48–55	48–59	48–59
Monolinguals	1,364 (49.6%)	804 (46.3%)	104 (64.6%)	148 (76.3%)	742 (46.7%)
Bilinguals	1,389 (50.4%)	931 (53.7%)	57 (35.4%)	46 (23.7%)	847 (53.3%)
Male	1,479 (53.8%)	886 (51.1%)	87 (54.0%)	123 (63.4%)	809 (50.9%)
Female	1,271 (46.2%)	849 (48.9%)	74 (46.0%)	71 (36.6%)	780 (49.1%)
% LT	220 (8.0%)	119 (6.9%)	15 (9.3%)	52 (26.8%)	107 (6.7%)
% LRI	161 (5.9%)	108 (6.2%)	—	23 (11.9%)	121 (7.6%)
LT vs. LRI: phi-correlations	.338***	.326***	—	.354***	.311***

Note. LT language therapy, LRI language(-related) impairments, QUNW quasi-universal nonwords, *M* mean, *SD* standard deviation

Table 1b
Sample Characteristics: Five-Year-Old Children

	S6	S7	S8	S9	S10
n	727	656	338	423	716
Nonword repetition task	KiSS.2 subtest	SSV subtest	S-ENS subtest	SOPESS subtest	QUNW
Vocabulary and articulation tasks	KiSS.2 subtests	KiSS.2 subtests	S-ENS subtests (“Add-on”)	SOPESS subtest (articulation only)	KiSS.2 subtests
Age (in months): *M* ±*SD*	63.5 ±2.8	63.2 ±2.7	67.4 ±3.6	67.0 ±2.2	63.3 ±2.7
Age span	60–71	60–71	60–71	60–71	60–71
Monolinguals	253 (34.8%)	234 (35.7%)	139 (41.1%)	139 (32.9%)	244 (34.1%)
Bilinguals	474 (65.2%)	422 (64.3%)	199 (58.9%)	284 (67.1%)	472 (65.9%)
Male	378 (52.0%)	338 (51.5%)	166 (49.1%)	215 (50.8%)	368 (51.4%)
Female	349 (48.0%)	318 (48.5%)	172 (50.9%)	208 (49.2%)	348 (48.6%)
% LT	110 (15.1%)	83 (12.7%)	51 (15.1%)	37 (8.7%)	91 (12.7%)
% LRI	55 (7.6%)	46 (7.0%)	24 (7.1%)	24 (5.7%)	49 (6.8%)
LT vs. LRI: phi-correlations	.431***	.434***	.495***	.322***	.428***

Table 2a
Characteristics of Tests: KiSS.2, SSV, HASE, S-ENS

Short test names	**KiSS.2**	**SSV**	**HASE**	**S-ENS**
Full test names in English*	"Language Screening for Children", 2nd version	"Language Screening for the Preschool Age", 2nd version	"Heidelberg Auditive Screening for the School Enrolment Examination"	"Develop-mental Screening for the School Enrolment Examination"
Test authors	Holler-Zittlau et al. (2011)	Grimm (2017)	Brunner & Schöler (2002)	Döpfner et al. (2005)
Subtests	(1) Speech comprehension (2) vocabulary (3) articulation (4) grammar (5) repetition of nonwords (6) repetition of sentences (7) spontaneous speech	(1) repetition of nonwords (2) repetition of sentences	(1) repetition of sentences (2) repetition of digit spans (3) repetition of nonwords	(1) repetition of nonwords (2) filling out gaps in words (3) repetition of sentences (4) articulation (5) speech comprehension (6) vocabulary (7) grammar**

Table 2b
Characteristics of Tests: SOPESS, Mottier, QUNW

Short test names	**SOPESS**	**Mottier**	**QUNW**
Full test names in English*	"Social-Paediatric Developmental Screening for School Enrolment Examinations"	"Mottier Test"	"Quasi-Universal Non-words"
Test authors	Daseking et al. (2009)	Mottier (1951)	Chiat (2015)
Subtests	(1) grammar (2) repetition of nonwords (3) articulation (4) spontaneous speech	(1) repetition of nonwords	(1) repetition of nonwords

Note. * For full German test names, see References; ** Tasks 5–7 are included in the extended version of S-ENS known as S-ENS-Add-on (Neumann et al., 2011).

Table 3a
Characteristics of Nonwords: KiSS.2, SSV, HASE, S-ENS

	KiSS.2	**SSV**	**HASE**	**S-ENS**
n (nonwords)	4	18	10	6
German prefixes	2	1	0	0
German suffixes	2	8	0	1
German stems	0	2	0	3
Consonant clusters				
(a) 2 consonants	4	28	2	1
(b) 3 consonants	0	1	0	0
Phonemes /r, ʒ, ʃ, η, ʔ/*	4	12	6	5
Item length				
(a) 2 syllables	1	3	3	0
(b) 3	0	6	3	1
(c) 4	3	4	3	5
(d) 5	0	5	1	0
(e) 6	0	0	0	0
Linguistic load	high (morphology, phonology)	high (morphology, phonology)	low (phonology)	low (phonology)
Language-specific	yes	yes	yes	yes
Nonword examples	*Triser*, *Nabolira*	*Kalifeng*, *Gattwutz*	*Godepi*, *Wunore*	*Fangofänger*, *Maramula*

Table 3b
Characteristics of Nonwords: SOPESS, Mottier, QUNW

	SOPESS	**Mottier**	**QUNW**
n (nonwords)	6	30	16
German prefixes	0	0	0
German suffixes	2	0	0
German stems	0	0	0
Consonant clusters			
(a) 2 consonants	1	0	0
(b) 3 consonants	0	0	0
Phonemes /r, ʒ, ʃ, η, ʔ/*	4	16	0
Item length			
(a) 2 syllables	0	6	4
(b) 3	3	6	4
(c) 4	3	6	4
(d) 5	0	6	4
(e) 6	0	6	0
Linguistic load	low (phonology)	low (phonology)	zero
Language-specific	yes	yes	no
Nonword examples	Not public	*Mera*, *Dugabe*	*Lumika*, *Litisaku*

Note. QUNW quasi-universal nonwords; * German-specific phonemes

Statistical Analyses

First, monolinguals' and bilinguals' performance in NWR, articulation, and vocabulary tasks was compared by means of Mann-Whitney *U*-tests. It was assumed that (a) monolinguals would outperform bilinguals in language-specific NWR tasks under condition of better articulation and/or vocabulary skills, (b) bilinguals and monolinguals would perform equally in the language-unspecific NWR task (QUNW) irrespective of their German language competence. Effect sizes in Mann-Whitney *U*-tests were quantified by the probability of superiority index ($\hat{p}$; Grissom & Kim, 2012). Values close to .5 are considered low, those close to .0 high.

Second, possible transfers from bilinguals' first language in NWR performance were exemplified by simplifications/omissions in consonant clusters in the word final position. For this purpose, repetition of all six word final consonant clusters in SSV nonwords was compared for monolinguals and bilinguals by means of Chi-square tests, without differentiation between four- and five-year-old children (samples S2 and S7 were merged). Because word-final consonant clusters are highly language-specific and do not occur in some languages that are widely spoken in Germany (e.g., Turkish; Dirim, 2005), it was assumed that bilinguals would simplify them more often than monolinguals.

Third, associations between the NWR performance and questionnaire items on LT/LRI were quantified by point-biserial correlations. Due to linguistic interferences, it was assumed that the correlations would be low both for monolinguals and bilinguals, at least in language-specific NWR tasks.

Finally, possible influence factors on the NWR performance were assessed by linear regressions (method "enter"). The independent variables were (a) vocabulary test score, (b) articulation test score, (c) children's age in months, (d) LT, (e) LRI, and (f) classification monolinguals vs. bilinguals. It was assumed that, under consideration of children's German language skills, the associations between NWR performance and LT and/or LRI would be very weak, except in QUNW.

All calculations, except for those with consonant clusters in the word final position, were carried out separately for four- and five-year-old children as well as separately for samples S1–S10.

For retrospective studies based on anonymized data, no ethical approval is necessary.

Results

When comparing monolinguals and bilinguals regarding their NWR performance as well as articulation and vocabulary skills (see Table 4), considerable, that is, highly significant differences were found only for vocabulary: Monolinguals outperformed bilinguals in all cases. In the samples that were tested with NWR tasks with a high linguistic load (S1, S2, S6, S7), monolinguals outperformed bilinguals in NWR, but also in articulation and vocabulary in almost all cases. In the samples that were tested with NWR tasks with a low linguistic load (S3, S4, S8, S9), monolinguals and bilinguals did not differ significantly in NWR and articulation. However, monolinguals outperformed bilinguals in vocabulary tasks. In the samples that were tested with NWR tasks without any language-specific load (S5, S10), monolinguals showed weaker results than bilinguals in NWR (in one case significantly, in the other case numerically), although they outperformed bilinguals in articulation and vocabulary. Thus, monolinguals outperformed bilinguals in NWR tasks with a high linguistic load, did not differ from bilinguals in NWR tasks with a low linguistic load, and were weaker than bilinguals in NWR tasks with zero linguistic load.

Bilinguals simplified word final consonant clusters significantly more often than monolinguals in five SSV nonwords out of six (see Table 5).

Almost all correlations between NWR scores and children's LT and LRI were low to non-existent (see Table 6). Correlation coefficients of monolinguals were predominantly higher than those of bilinguals. Correlations between NWR scores and LT were higher than correlations between NWR and LRI in 17 calculations out of 18.

In the linear regressions, vocabulary and articulation skills yielded significant results in all cases, children's age in eight calculations out of ten, LT in nine, the classification monolinguals vs. bilinguals in seven, and LRI did not yield any significant results (see Table 7). Higher total scores of correct answers in the vocabulary and articulation tasks, higher age, being bilingual, and no LT were associated with better NWR results. In two regressions, the standardized beta of LT was higher than the betas of the vocabulary and articulation tasks (S3, S8), that is, LT constituted a more important influence factor. In most other regressions, including those with QUNW, vocabulary and articulation skills showed higher beta values than LT.

Table 4
Differences in the Language Test Scores of Monolinguals and Bilinguals: Means ±Standard Deviations and Mann-Whitney U-Tests

Sample		Nonword repetition	Articulation	Vocabulary
		Four-year-old children		
S1 (KiSS.2)	**MO**	2.4 ±1.2	8.6 ±2.0	6.8 ±1.8
	BI	2.1 ±1.2	8.2 ±2.0	3.8 ±2.7
	Z/p̂	-5.39***/.44	-6.88***/.43	-28.83***/.18
S2 (SSV)	**MO**	10.2 ±3.7	8.7 ±1.7	6.7 ±1.7
	BI	9.2 ±4.2	8.2 ±1.9	3.6 ±2.7
	Z/p̂	-4.39***/.44	-6.70***/.41	-23.12***/.18
S3 (HASE)	**MO**	6.7 ±2.3	9.6 ±1.9	7.1 ±1.7
	BI	6.3 ±2.3	9.4 ±1.9	4.81 ±2.7
	Z/p̂	-0.96/.45	-1.15/.45	-5.35***/.25
S4 (Mottier)	**MO**	15.3 ±5.8	7.7 ±2.6	6.9 ±1.9
	BI	15.5 ±6.8	7.3 ±2.3	4.3 ±2.5
	Z/p̂	-0.33/.48	-1.22/.44	-6.04***/.21
S5 (QUNW)	**MO**	10.9 ±2.7	10.7 ±1.9	6.7 ±1.7
	BI	11.1 ±3.0	10.0 ±2.2	3.4 ±2.7
	Z/p̂	-2.12*/.47	-6.63***/.41	-23.05***/.17
		Five-year-old children		
S6 (KiSS.2)	**MO**	2.5 ±1.2	9.2 ±1.3	7.6 ±1.4
	BI	2.4 ±1.2	8.8 ±1.6	4.5 ±2.8
	Z/p̂	-1.04/.48	-4.24***/.41	-14.58***/.17
S7 (SSV)	**MO**	11.4 ±3.6	9.3 ±1.2	7.7 ±1.4
	BI	10.7 ±3.7	8.9 ±1.6	4.5 ±2.8
	Z/p̂	-2.33*/.45	-4.39***/.40	-14.26***/.17
S8 (S-ENS)	**MO**	4.8 ±1.4	9.6 ±0.9	3.9 ±1.2
	BI	5.00 ±1.1	9.6 ±0.9	1.8 ±1.8
	Z/p̂	-1.10/.47	-0.48/.49	-10.11***/.18
S9 (SOPESS)	**MO**	5.2 ±1.1	9.4 ±1.0	—
	BI	5.0 ±1.2	9.4 ±1.1	—
	Z/p̂	-1.58/.46	-0.25/.49	—
S10 (QUNW)	**MO**	11.9 ±2.5	11.3 ±1.3	7.6 ±1.4
	BI	12.1 ±2.6	10.7 ±1.8	4.4 ±2.8
	Z/p̂	-1.83/.46	-5.02***/.39	-14.90***/.16

Note. *** $p < .001$, * $p < .05$; QUNW quasi-universal nonwords, MO monolinguals, BI bilinguals, − no data available, $\hat{p}$ probability of superiority index

Table 5
Monolinguals' and Bilinguals' Errors (Simplifications/Omissions) in Word-Final Consonant Clusters in SSV Nonwords: Chi-Square (Combined Samples S2, S7)

Consonant clusters	Monolinguals	Bilinguals	Chi-Square
/nt/ in *Entiergent*	43/1,027 (4.2%)	110/1,335 (8.2%)	15.74***
/ts/ in *Gattwutz*	80/1,028 (7.8%)	189/1,338 (14.1%)	23.22***
/st/ in *Seregropist*	179/976 (18.3%)	346/1,277 (27.1%)	23.72***
/bst/ in *Nebatsubst*	382/998 (38.3%)	534/1,308 (40.8%)	1.54
/nk/ in *Ippazeumerink*	213/950 (22.4%)	336/1,228 (27.4%)	6.94**
/st/ in *Vominlapertust*	208/945 (22.0%)	360/1,224 (29.4%)	15.11***

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

Table 6
Associations Between Nonword Repetition Performance and (a) Language Therapies (LT), (b) Language(-Related) Impairments (LRI): Point-Biserial Correlations

Sample		LT	LRI
		Four-year-old children	
S1 (KiSS.2)	**MO**	.209***	.097***
	BI	.142***	.089**
S2 (SSV)	**MO**	.136***	.104**
	BI	.209***	.092**
S3 (HASE)	**MO**	.455***	—
	BI	.108	—
S4 (Mottier)	**MO**	.380***	.123
	BI	.365*	.241
S5 (QUNW)	**MO**	.152***	.124**
	BI	.220***	.095**
		Five-year-old children	
S6 (KiSS.2)	**MO**	.267***	.136*
	BI	.220***	.081
S7 (SSV)	**MO**	.167*	.131*
	BI	.156**	.047
S8 (S-ENS)	**MO**	.391***	.187*
	BI	.290***	.065
S9 (SOPESS)	**MO**	.209*	.053
	BI	.111	.149*
S10 (QUNW)	**MO**	.199**	.148*
	BI	.159*	.084

Note. *** $p < .001$, ** $p < .01$, * $p < .05$; QUNW quasi-universal nonwords, MO monolinguals, BI bilinguals, — no data available

Table 7a
Influence Factors on the Nonword Repetition Performance: Linear Regressions

	S1 (KiSS.2)		S2 (SSV)	
	Stand. *β*	***T***	**Stand. *β***	***T***
Vocabulary	.310	14.00***	.347	12.93***
Articulation	.183	9.37***	.254	11.00***
Age (in months)	.046	2.61**	.067	3.12**
LT	.083	4.25***	.077	3.37**
LRI	.021	1.13	.020	0.89
MO vs. BI	.085	4.03***	.122	4.78***
Constant	—	-4.02***	—	-3.86***
Corr. R^2	.169		.236	
Model (*F/df*)	94.01***/6, 2,743		90.15***/6, 1,728	
Age group	Four-year-olds		Four-year-olds	
	S3 (HASE)		**S4 (Mottier)**	
Vocabulary	.269	3.31**	.364	4.82***
Articulation	.205	2.78**	.204	2.49*
Age (in months)	.026	0.37	-.007	-0.11
LT	.325	4.55***	.176	2.15*
LRI	—	—	.038	0.57
MO vs. BI	.006	0.07	.179	2.49*
Constant	—	-0.86	—	-0.23
Corr. R^2	.243		.273	
Model (*F/df*)	11.26***/5, 155		13.07***/6, 187	
Age group	Four-year-olds		Four-year-olds	
	S5 (QUNW)		**S6 (KiSS.2)**	
Vocabulary	.233	7.74***	.169	3.86***
Articulation	.219	8.64***	.221	5.82***
Age (in months)	.124	5.27***	.069	1.97*
LT	.108	4.33***	.173	4.39***
LRI	.031	1.29	.004	0.11
MO vs. BI	.206	7.17***	.077	1.89
Constant	—	-3.00**	—	-2.66**
Corr. R^2	.169		.146	
Model (*F/df*)	54.19***/6, 1,566		21.70***/6, 720	
Age group	Four-year-olds		Five-year-olds	

Note. *** $p < .001$, ** $p < .01$, * $p < .05$; LT language therapy, LRI language(-related) impairment, QUNW quasi-universal nonwords, MO monolinguals, BI bilinguals

Table 7b
Influence Factors on the Nonword Repetition Performance: Linear Regressions

	S7 (SSV)		**S8 (S-ENS)**	
	Stand. *β*	***T***	**Stand. *β***	***T***
Vocabulary	.335	7.38***	.234	3.88***
Articulation	.178	4.59***	.176	3.32**
Age (in months)	.096	2.70**	.087	1.76
LT	.089	2.24*	.270	4.54***
LRI	.028	0.71	-.045	-0.80
MO vs. BI	.114	2.71**	.189	3.15**
Constant	—	-2.34*	—	-1.36
Corr. R^2	.185		.192	
Model (*F/df*)	25.83***/6, 649		14.33***/6, 331	
Age group	Five-year-olds		Five-year-olds	
	S9 (SOPESS)		**S10 (QUNW)**	
Vocabulary	—	—	.135	2.99**
Articulation	.260	5.46***	.244	6.32***
Age (in months)	.061	1.30	.096	2.72**
LT	.082	1.66	.118	3.02**
LRI	.066	1.32	.048	1.24
MO vs. BI	-.066	-1.41	.172	4.11***
Constant	—	-0.43	—	-1.09
Corr. R^2	.091		.132	
Model (*F/df*)	9.40***/5, 413		19.05***/6, 706	
Age group	Five-year-olds		Five-year-olds	

Note. *** $p < .001$, ** $p < .01$, * $p < .05$; LT language therapy, LRI language(-related) impairment, QUNW quasi-universal nonwords, MO monolinguals, BI bilinguals

Discussion and Conclusion

Although nonword repetition tasks were developed to detect deficits in children's PSTM, as a possible clinical marker for LRI, the current study showed that NWR performance depends on children's articulation skills in most cases. Also, NWR performance is hardly associated with children's (running or completed) LT and LRI (documented in questionnaires for parents). Reduction of the linguistic load in NWR tasks resulted in a relative "culture-fairness" of nonwords, that is, monolinguals could not outperform bilinguals in NWR, but this did not make such tasks more appropriate for the detection of children with LRI.

In all samples, monolinguals outperformed bilinguals in the vocabulary tasks. Therefore, the condition of comparable vocabulary skills could not be tested. Also, monolinguals outperformed bilinguals in articulation tasks in six out of ten samples. In three of these samples (S1, S2, S7), namely those with language-specific NWR tasks, this resulted in a better NWR performance of monolinguals compared to bilinguals (also, the same finding did not achieve statistical significance in S6). On the contrary, comparable articulation skills of monolinguals and bilinguals (S3, S4, S8, S9) resulted in a comparable NWR performance, without exceptions. Thus, good performance in German-based nonwords presupposes sufficient competence in German articulation/phonotactics and puts those bilinguals at a disadvantage who lag behind in its acquisition.

Language-unspecific nonwords (QUNW) showed a very different NWR performance pattern. In S10, bilinguals outperformed monolinguals in NWR in spite of weaker German language skills (vocabulary, articulation). However, the “culture-fairness” of QUNW did not result in higher correlations between NWR scores and LT/LRI, compared to German-based nonwords. Irrespective of the linguistic load, NWR tasks demonstrated far too low correlations with LT and LRI.

Generally, the importance of the linguistic load and “culture-fairness” of NWR tasks seems to be overestimated in the research on possible indicators of children’s LRI. In the current study, children’s German articulation skills were shown to be a more important predictor than LT/LRI for the NWR performance in both samples that were tested with QUNW (S5, S10, see Table 7). Additionally, children’s age and German vocabulary skills as well as the classification monolinguals vs. bilinguals showed a higher influence on NWR scores, compared to LT, in S5. For both monolinguals and bilinguals taken together, the best regression result regarding associations between NWR tasks and LT was achieved not by QUNW, but by HASE (four-year-olds) and S-ENS (five-year-olds), that is, German-based NWR tasks with a low linguistic load (German phonology, but not morphology). The same tasks yielded the highest point-biserial correlations with LT for monolinguals. The highest point-biserial correlations with LT for bilinguals were achieved by the Mottier test (four-year-olds) and, again, S-ENS (five-year-olds) that are both German-based in terms of phonology. Thus, NWR tasks with a low linguistic load appear to be more promising than language-unspecific tasks, but none of them can contribute considerably to the detection of LRI (quantified by LT). Also, when LRI was quantified by the questionnaire item on LRI, this variable did not show any significant results in the regressions and yielded very low to non-existent point-biserial correlations with NWR scores.

It should be noted that the better performance of bilinguals, compared to monolinguals, in QUNW, in spite of weak German language skills, as well as associations between bilingualism and higher NWR scores in the regressions, might be indicative of the so-called “bilingual advantage” that consists in better-developed cognitive skills, including PSTM, of bilinguals compared to monolinguals (Gunnerud et al., 2020). The reflection of the “bilingual advantage” in QUNW results may mean that language-unspecific nonwords indeed provide a more direct access to PSTM, which, however, does not make them a more reliable clinical marker for LRI, probably because of a weak association between deficits in PSTM and (some) LRI.

Quality criteria of NWR tasks in terms of sensitivity and specificity were not determined in the present study because correlations below .5 (and mostly below .3) between NWR scores and LT/LRI cannot be “translated” into acceptable sensitivity (at least 90%) and specificity (at least 70%). Therefore, no Receiver Operating Curves (ROC) were calculated. In other studies, such calculations resulted in insufficient quality criteria of both language-specific and language-unspecific NWR tasks for LRI (e.g., Zaretsky & Hey, 2022).

Because only 9–27% of the variance of the NWR performance were explained in the regressions (see Table 7), it can be assumed that a number of other factors that were not considered in the present study influenced NWR results (e.g., attention, long-term memory, German grammar skills, year and region of the language test, typology of bilinguals’ first language, oromotor skills, intelligence quotient). Also, due to the low percentage of explained variance in the regressions, it cannot be excluded that vocabulary and articulation skills do not constitute the most important influence factors on the NWR performance.

Since children in the current study did not undergo any medical examinations, it remains open whether parents' information on LT and LRI can be considered reliable. Information on LRI may be comparatively subjective, in comparison with children's participation in LT (that can hardly escape parents' attention because they are (co-)responsible for its organisation).

To sum up, no evidence was found that language-specific or language-unspecific NWR tasks can contribute considerably to the LRI detection. Hence, none of them can be recommended for this purpose. Although NWR tasks with a low linguistic load appear more promising, their results can be misinterpreted in case of weak German articulation skills. Relative "culture-fairness" of language-unspecific nonwords (QUNW) did not result in a higher predictive power for LRI. Unexpectedly, children's performance in QUNW depended, first and foremost, on their German articulation skills, followed by vocabulary skills. It cannot be excluded that some NWR tasks may provide sufficient predictive power for one or two LRI that are associated with considerable deficits in PSTM but not for all LRI taken together.

References

Boerma, T., Chiat, S., Leseman, P., Timmermeister, M., Wijnen, F., & Blom, E. (2015). A quasi-universal nonword repetition task as a diagnostic tool for bilingual children learning Dutch as a second language. *Journal of Speech, Language, and Hearing Research*, *58*(6), 1747–1760. https://doi.org/10.1044/2015_JSLHR-L-15-0058

Brunner, M., & Schöler, H. (2002). *HASE − Heidelberger Auditives Screening in der Einschulungsuntersuchung [HASE – Heidelberg Auditive Screening for the School Enrolment Examination].* WESTRA.

Chiat, S. (2015). Non-word repetition. In S. Armon-Lotem, J. de Long, & N. Meir (Eds.), *Assessing Multilingual Children. Disentangling Bilingualism from Language Impairment* (pp. 125–150). Multilingual Matters.

Ching, T. Y. C., Cupples, L., & Marnane, V. (2019). Early cognitive predictors of 9-year-old spoken language in children with mild to severe hearing loss using hearing aids. *Frontiers in Psychology*, *10*, 2180. https://doi.org/10.3389/fpsyg.2019.02180

Daseking, M., Petermann, F., Röske, D., Trost-Brinkhues, G., Simon, K., & Oldenhage, M. (2009). Entwicklung und Normierung des Einschulungsscreenings SOPESS [Development and standardization of the school enrolment test SOPESS]. *Das Gesundheitswesen*, *71*(10), 648–655. https://doi.org/10.1055/s-0029-1239511

Dirim, I. (2005). Deutsch lernen auf der Grundlage der Erstsprache Türkisch [Learning German with Turkish as the first language]. In H. Bartnitzky, & A. Speck-Hamdan (Eds.), *Deutsch als Zweitsprache lernen* (pp. 53–57). Grundschulverband.

Döpfner, M., Dietmair, I., Mersmann, H., Simon, K., & Trost-Brinkhues, G. (2005). *S-ENS. Screening des Entwicklungsstandes bei Einschulungsuntersuchungen* [S-ENS. Developmental Screening for the School Enrolment Examination]. Hogrefe.

Engel de Abreu, P. M., Baldassi, M., Puglisi, M. L., & Befi-Lopes, D. M. (2013). Crosslinguistic and cross-cultural effects on verbal working memory and vocabulary: Testing language minority children with an immigrant background. *Journal of Speech, Language and Hearing Research*, *56*(2), 630–642. https://doi.org/10.1044/1092-4388(2012/12-0079)

Grimm, A. (2016). Quatschwörter nachsprechen – gleiche Anforderungen für alle Kinder? [Repetition of nonwords: The same requirements for all children?] *Diskurs Kindheits- und Jungendforschung*, *11*(1), 113–117. https://doi.org/10.3224/diskurs.v11i1.22252

Grimm, H. (2017). *SSV. Sprachscreening für das Vorschulalter. Kurzform des SETK 3-5* [SSV. Language Screening for the Preschool Age. Short form of SETK 3-5] (2nd ed.). Hogrefe.

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

Gunnerud, H. L., Ten Braak, D., Reikerås, E. K. L., Donolato, E., & Melby-Lervåg, M. (2020). Is bilingualism related to a cognitive advantage in children? A systematic review and meta-analysis. *Psychological Bulletin*, *146*(12), 1059–1083. https://doi.org/10.1037/bul0000301

Holler-Zittlau, I., Euler, H. A., & Neumann, K. (2011). Kindersprachscreening (KiSS) – das hessische Verfahren zur Sprachstandserfassung [Language Screening for Children (KiSS) – a Hessian language screening tool]. *Sprachheilarbeit*, *5*(6), 263–268.

Masoura, E., Gogou, A., & Gaterhcole, S. E. (2020). Working memory profiles of children with reading difficulties who are learning to read in Greek. *Dyslexia*, *27*(3), 312–324. https://doi.org/10.1002/dys.1671

Mottier, G. (1951). Über Untersuchungen der Sprache lesegestörter Kinder [On the analysis of the language of children with dyslexia]. *Folia Phoniatrica*, *3*, 170–177. https://doi.org/10.1159/000262507

Neumann, K., Zaretsky, Y., & Euler, H. A. (2011). *Einführung einer flächendeckenden Sprachstandserfassung in Hessen. Forschungsbericht 2011* [Introduction of a state-wide language screening program in Hesse. Research report 2011]. Goethe-University, Frankfurt/Main.

Penke, M. (2018). Verbal agreement inflection in German children with Down syndrome. *Journal of Speech, Language, and Hearing Research*, *61*(9), 2217–2234. https://doi.org/10.1044/2018_JSLHR-L-17-0241

Polišenska, K., Chiat, S., & Roy, P. (2015). Sentence repetition: What does the task measure? *International Journal of Language and Communication Disorders*, *50*(1), 106–118. https://doi.org/10.1111/1460-6984.12126

Redmond, S. M., Thompson, H. L., & Goldstein, S. (2011). Psycholinguistic profiling differentiates specific language impairment from typical development and from attention-deficit/hyperactivity disorder. *Journal of Speech, Language, and Hearing Research*, *54*(1), 99–117. https://doi.org/10.1044/1092-4388(2010/10-0010

Schwob, S., Eddé, L., Jacquin, L., Leboulanger, M., Picard, M., Oliveira, P. R., & Skoruppa, K. (2021). Using nonword repetition to identify developmental language disorder in monolingual and bilingual children: A systematic review and meta-analysis. *Journal of Speech, Language, and Hearing Research*, *64*(9), 3578–3593. https://doi.org/10.1044/2021_JSLHR-20-00552

Verhagen, J., de Bree, E., Mulder, H., & Leseman, P. (2017). Effects of vocabulary and phonotactic probability on 2-year-olds' nonword repetition. *The Journal of Psycholinguistic Research*, *46*(3), 507–524. https://doi.org/10.1007/s10936-016-9448-9

Zaretsky, E., & Hey, C. (2022). Deutsch-basierte und (quasi-)universelle Kunstwörter als Prädiktoren für sprachbezogene medizinische Störungsbilder bei Vorschulkindern [German-based and (quasi-)universal non-words as predictors for language(-related) medical impairmens in preschoolers]. In M. Spreer, M. Wahl, & H. Beek (Eds.), *Sprachentwicklung im Dialog. Digitalität – Kommunikation – Partizipation* (pp. 198–204). Schulz-Kirchner Verlag.

Zaretsky, E., & Lange, B. P. (2022). Language-specific non-words for the assessment of working memory: Dealing with bilingual children. *Communication Disorders Quarterly, 44*(4). https://doi.org/10.1177/15257401221113064

Contact email: zaretsky@staff.uni-marburg.de