Bilingual Arabic-German and Turkish-German Children with and without Specific Language Impairment: Comparing Performance in Sentence and Nonword Repetition Tasks

Lina Abed Ibrahim and Cornelia Hamann

1. Introduction

In recent research on language development attention has been focused on bilingualism and on the question of how to identify language impairments in bilingual situations, see for instance COST Action IS0804. In this context, specific language impairment (SLI) has been particularly interesting because it is well established that areas of difficulty and error patterns in the language performance of monolingual children with SLI (Mo-SLI) and bilingual typically developing children (Bi-TD) can overlap (Håkansson & Nettelbladt 1996; de Jong 2010, Paradis 2010, Hamann 2012 and many others). Since specific language impairment is a developmental language disorder which manifests itself in low language skills, usually in morpho-syntax and phonology, in the absence of primary deficits such as hearing impairment, neuro-cognitive anomalies, or sensorimotor and socio-emotional impairments (Leonard 2014, Stark & Tallal 1981), the overlap situation is theoretically relevant as well as important from a practical point of view. From the theoretical perspective, it can be asked whether these similar error patterns should be derived from a similar underlying deficit in processing or language representation, or whether (not so) subtle differences can be discovered that can demonstrate that in contrast to the children with SLI, bilingual typical children possess essentially intact mechanisms for language acquisition and language processing. From the
practical point of view, the problem of identifying language impairment is more and more acute: as Tomblin et al. (1997) established, 7% of the 5-year-olds have SLI, and this should also be the case in the growing bilingual population worldwide. It is therefore necessary to identify individuals with SLI as early as possible in monolingual and bilingual settings in order to provide not only language support but therapy if required. Inadequate assessment tools, however, lead to over- and underdiagnosis (Genesee et al. 2004) which, in selective educational systems, may lead to early educational dead-ends. Another question which has theoretical and practical relevance is whether the difficulties could be cumulative, i.e. whether a bilingual child with SLI (Bi-SLI) will show more severe language deficits than a monolingual child with SLI. In this paper we address the practical side, the identification of SLI in bilingual settings, by demonstrating that Mo-SLI and Bi-SLI children perform on the same level in the German sentence and non-word repetition task developed in the BiLaD-project (see also *), whereas their performance is significantly below that of their peers without an impairment. An analysis of error profiles will also allow us to speculate about the theoretical questions raised here: We do not find a cumulative effect of bilingualism and SLI and the profiles are different enough to stipulate different underlying deficits.

1.1. Theoretical background on assessment tools: sentence repetition and nonword repetition tasks

Sentence repetition and nonword repetition have been shown to reliably identify monolingual SLI (Conti-Ramsden et al. 2001) and are integrated into many normed assessment tools. These tasks can, depending on their construction, assess phonological working memory and/or phonological representation/derivation and/or syntactic representation/derivation as we discuss in the next paragraphs. They can target quite specific linguistic areas and can even be used to empirically settle intricate theoretical problems (Hamann et al. 2017, Stzerman and Friedmann 2015). Because of this versatility, they are ideal for targeting cross-linguistically difficult phonological/syntactic structures as well as language specific structures.

We first focus on the assessment of morpho-syntax where difficulties usually manifest themselves in monolingual and bilingual children with SLI (Leonard 2014, Armon-Lotem et al. 2015). In addition to tense and agreement marking, which constitute an area of overlap, syntactically complex constructions, especially Wh-questions, finite complement clauses, passives and (object) relative clauses are reported to be particularly challenging for language impaired children across languages. Because avoidance of complexity has been observed in many elicitation tasks, sentence repetition has been suggested as a method for eliciting such structures minimizing the option of avoidance. Sentence repetition tasks (SRT) have been cross-linguistically implemented as a
reliable assessment tool for these linguistic domains, since it has been shown that they can be constructed so as not measuring (phonological) working memory, but as drawing on morpho-syntactic knowledge - if their design is complex enough or controlled enough (Polišenská et al. 2014; Vinther 2002; Szterman & Friedmann 2015). SRTs including a broad range of cross-linguistically complex structures (Wh-movement, embedding and intervention) as well as language-specific structures known to pose difficulties for children with atypical language development, are found to have good specificity and sensitivity in identifying SLI not only in monolingual but also in bilingual children (Marinis & Armon-Lotem 2015; Tuller et al. 2015). A further advantage of SRTs is that they are fast and easy to administer. For these reasons we constructed a German SRT conform to the principles developed in COST Action IS0804 (Marinis & Armon-Lotem 2015) and explore its possibilities in this study.

It is also well established that children with SLI can be identified by their poor performance on nonword repetition tasks which usually measure (phonological) working memory (WM) capacities (Archibald & Gathercole 2006). However, further research has shown that NWRTs are not pure measures of phonological working memory: they draw on storage and tap into underlying phonological representations (Marshall & van der Lely 2009). Therefore it has been argued by Gallon et al. (2007) and Ferré et al. (2015) that it is phonological complexity rather than WM (measured by increasing number of syllables) which is vulnerable in children with SLI.

Unlike other language assessment measures, NWR tasks can be constructed so that they do not primarily rely on prior knowledge of vocabulary and morpho-syntax, and thus have a good potential for identifying SLI in bilingual children with limited exposure to the L2 (Chiat 2015). Moreover, NWR counts as a culturally fair measure which could be used to assess children with diverse socio-economic backgrounds (Chiat & Polišenská 2016). Therefore, NWR seems to be a promising clinical assessment tool for the identification of SLI in bilingual contexts (e.g., dos Santos and Ferré 2016).

Construction of an NWRT that does not disadvantage bilingual children is not straightforward, however, and although NWR, unlike other language assessment measures, is less reliant on language specific knowledge, it might still be affected by language experience and levels of exposure to the target language depending on its design (Chiat 2015). Several studies showed that children’s performance on NWR is affected by the characteristics of the test items such as word-likeness, articulatory complexity and length in terms of number of syllables (see Graf Estes et al. 2007 for a meta-analysis). For example, children were found to perform significantly better on nonwords that are more word-like, carry language-specific (LS) stress patterns, contain LS-root morphemes and affixes or are characterized by higher phonotactic probability (Marshall & van der Lely 2009; Jones et al. 2010; Leclercq et al. 2013. a.o.). Furthermore, Gathercole (2006) found a correlation between NWR and vocabulary size in monolingual children. Studies on bilingual children showed
that the accuracy on NWR was affected by vocabulary size, phonotactic knowledge, length of exposure to the L2 and L1 background of the children (Thordardottir & Brandeker 2013; Sorensen Duncan & Paradis 2016). Since our primary goal is disentangling bilingualism from SLI and given the diverse linguistic profiles of bilingual children, a NWR task should crucially test phonological processing abilities/competence of bilingual children rather than working memory or lexical abilities. To achieve this goal, new NWRTs focusing on phonological complexity with less word-like nonwords were constructed within the framework of the COST Action IS0804 (Chiat 2015). Both these LITMUS tools, SRT and NWRT, are included in the test battery used by our Franco-German research project Bilingual Language Development (BiLaD) for the collection of comparative data. The following paper reports on German data only.

2. Method

2.1. LITMUS Sentence Repetition (SR) and Nonword Repetition (NWR) tasks

As we pointed out in 1.1., both tasks investigated in this study, namely the German SRT and NWRT were developed within COST Action IS0804 as part of the LITMUS (Language Impairment Testing in Multilingual Settings) toolbox. Both are linguistic tasks and are proposed to identify SLI in bilingual settings by targeting language phenomena known to pose difficulties for children with atypical language development.

2.1.1. The German LITMUS Sentence Repetition Task

The German SRT (Hamann et al. 2013) was designed according to the LITMUS principles (Marinis & Armon-Lotem 2015) and in close parallel to the task used by our French partners in the BiLaD project (Fleckstein et al. 2016). It thus includes object questions, subject and object relative clauses, complement clauses and passives, as well as structures found to be challenging for SLI children in German e.g., topicalization and the sentence bracket, which is a reflex of the V2 and OV properties of German (see Hamann et al. 2017 & Lein et al. 2016 for details).

The current version\(^1\) of the German LITMUS-SRT (45 sentences) is constructed in three levels of increasing complexity controlled for syllable number (five conditions per level, three items per condition). The test stimuli are incorporated into a child friendly PPT and are presented in a pseudorandomized order. Level 1 focuses on SVA, tense and sentence bracket (see (1)). Level 2 tests bare object Wh-questions (bare WH), object Wh-questions with an intervening lexical DP (Wh+NP), finite (see (2)) and non-finite complement

\(^1\) The original long version of the German LITMUS-SRT was shortened to meet the needs of the age range investigated in the BiLaD project.
clauses as well as coordinate structures. Level 3 comprises the most complex structures and targets topicalization, long passives, subject and object relatives with (see (3)) and without an intervener. It should be noted that the task does not explicitly test case marking.

(1) Sentence bracket:
Die Köchin hat den Cowboy geweckt
The/nom. cook has the/acc. cowboy woken up
‘the cook woke the cowboy up’

(2) Finite complement clause:
Der Prinz will, dass der Ritter die Affen jagt.
The/nom. prince wants, that the/nom. knight the/acc. monkeys hunts
‘the prince wants that the knight hunts the monkeys’

(3) Object relative with intervention:
Ich sehe den Clown, den der Wikinger umarmt.
I see the/acc. clown who/acc. the/nom. viking hugs
‘I see the clown who(m) the wiking hug’

The task takes about 10 minutes to administer and can be scored either by ‘identical repetition’ (only disregarding phonological errors) or by ‘target structure met’. The latter scoring measure compensates for L2-errors such as lexical substitutions and certain recurrent case errors which do not affect the realization of the targeted structure.

2.1.2. The German LITMUS Nonword Repetition Task

As previously mentioned, in order to assess the phonological abilities of bilingual children without penalizing them, NWRTs should be designed to be less language dependent and include complex phonological structures known to pose difficulties for children with (phonological) SLI, e.g. consonant clusters (Ferré et al. 2015). In order to achieve this goal, the German NWRT (Grimm et al. 2014) was constructed in accordance with the LITMUS principles (Chiat 2015). It contains a (quasi-) Language Independent (LI) and a Language Dependent (LD) part. In both parts, the maximal nonword length is limited to three syllables to minimize the effect of working memory. The LI-part contains 30 items built using typologically well-attested phonemes and phonotactic properties (Maddieson et al. 2011) and systematically varied
in segmental (only consonantal), syllabic and sequential\textsuperscript{2} complexity (see dos Santos & Ferré 2016; Grimm & Hübner to appear for details). It further contains syllables with branching onsets, which are cross-linguistically widely attested despite their relative complexity. The latter should be problematic for children with SLI but not for bilingual children with typical language development. The LD-part consists of 36 items. It was constructed using the same phonemes and syllable types as the LI-part in addition to the extrametrical /s, ʃ/, which are restricted to word-initial and word-final positions. Such sequences violate the Sonority Sequencing Principle, are considered to be more complex and thus form a locus of difficulty for language impaired children (dos Santos & Ferré 2016). Despite their cross-linguistic scarcity, sC# and #Cs clusters are not unique to German and are not expected to be a particular source of difficulty for Bi-TD children. Table 1 gives an overview.

Table 1: Description of LITMUS-NWRT-German

<table>
<thead>
<tr>
<th></th>
<th>vowels</th>
<th>consonants</th>
<th>syllable types</th>
<th>examples</th>
</tr>
</thead>
<tbody>
<tr>
<td>Language-Independent</td>
<td>/a, i, u/</td>
<td>/p, k, f, l/</td>
<td>CV</td>
<td>faku</td>
</tr>
<tr>
<td>part (LI)</td>
<td></td>
<td></td>
<td>CCV, CVC#</td>
<td>klipafu</td>
</tr>
<tr>
<td>30 items</td>
<td></td>
<td></td>
<td></td>
<td>kafip</td>
</tr>
<tr>
<td>23 test items, 7 controls</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Language-Dependent part</td>
<td>/a, i, u/</td>
<td>/p, k, f, l/</td>
<td>same syllable</td>
<td>skifapu</td>
</tr>
<tr>
<td>(LD)</td>
<td></td>
<td></td>
<td>types plus</td>
<td></td>
</tr>
<tr>
<td>36 Items</td>
<td></td>
<td></td>
<td>#sCV, #sCCV,</td>
<td></td>
</tr>
<tr>
<td>32 test items, 4 control items</td>
<td></td>
<td></td>
<td>Cs#, internal</td>
<td>fikapuks</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>/s/</td>
<td>kufiski</td>
</tr>
</tbody>
</table>

The administration of the task takes about 10 minutes. In order to make the test appealing to the children, the nonwords are incorporated into a PowerPoint presentation. The child is usually told that she has to repeat words said by a friendly alien who would like to teach its language to the child. The test stimuli are presented in a pseudo-randomized order and are scored by whole item accuracy. Systematic substitutions, e. g. /t/ for /k/ as well as errors concerning vowels or voicing of consonants are not counted as errors if their ordering corresponded to the targeted form. Furthermore, substitution of extrametrical /ʃ/ by [s] is disregarded since there is no phonemic contrast between the two in syllable initial position in German (see Grimm & Hübner to appear for details).

\textsuperscript{2} Both consonant sequence and syllable sequence affect the relative complexity of the nonwords. For example, a consonant sequence with alternating place and/or manner of articulation is considered to be more complex. In trisyllabic items, the medial syllable was found to be particularly vulnerable especially if it showed aspects of complexity such as branching onsets (see dos Santos and Ferré 2016).
2.2. Participants

In this paper, we are reporting on the results of 54 children recruited for the BiLaD project. The subjects of the study are aged 5;6 to 9;3 and are divided into four groups of monolingual and bilingual children (both simultaneous and early successive) with and without SLI. All of the bilingual Arabic-German and Turkish-German children started acquiring their L2 German before the age of three and had a length of exposure of at least 18 months. The SLI children were recruited from special speech language therapy centers and kindergartens and met the exclusionary criteria for SLI (Leonard 2014). Due to the risk of misdiagnosis (Grimm and Schulz 2014), all of the participating children underwent extensive language screening in their L1 and L2 to verify their initial clinical status as ± language impaired. In addition to that, the Questionnaire for Parents of Bilingual Children (PaBiQ, Tuller 2015) was used to gather background information about the children and determine the type of child bilingualism. Relevant aspects such as age of onset of exposure to the L2 (AoO), length and richness of exposure to the L2 (LoE), current language usage and skills as well as risk factors for SLI were documented using the PaBiQ. Only those children who scored below the norms in two language domains (applying the adjusted cut-offs proposed by Thordardottir 2012) in both of their L1 and L2 were assigned to the SLI group (refined clinical status). Table 2 provides an overview of the participants based on their refined clinical status (see also Abed Ibrahim et al. to appear). In this table, the bilingual children are grouped together as Bi-TD or Bi-SLI regardless of their L1, as statistical analysis on the relevant comparisons showed no significant differences within the two bilingual groups as distinguished by their home language.

Table 2: Participant groups, age-range, L1 and AoO

<table>
<thead>
<tr>
<th>Group</th>
<th>Age range</th>
<th>Mean Age (SD)</th>
<th>N</th>
<th>L1</th>
<th>AoO</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mo-TD</td>
<td>5;6-7;8</td>
<td>6;6 (0;11)</td>
<td>11</td>
<td>German</td>
<td>0</td>
</tr>
<tr>
<td>Mo-SLI</td>
<td>5;8-7;5</td>
<td>6;4 (0;7)</td>
<td>12</td>
<td>German</td>
<td>0</td>
</tr>
<tr>
<td>Bi-TD</td>
<td>5;8-8;9</td>
<td>7;1 (1;0)</td>
<td>22</td>
<td>Arabic (n=10)</td>
<td>&lt;3;0</td>
</tr>
<tr>
<td>Bi-SLI</td>
<td>5;7-9;3</td>
<td>6;6 (1;4)</td>
<td>9</td>
<td>Arabic (n=4)</td>
<td>&lt;3;0</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Turkish (n=12)</td>
<td>&lt;3;0</td>
</tr>
</tbody>
</table>

2.3. Procedure

Both of the LITMUS tasks were administered using a pseudo-randomized computerized version. Each of the tests started with two training items. To minimize the effect of background noise, the participants were equipped with noise cancelling headphones. They were instructed to repeat each test stimulus

---

3 A total of 92 children were recruited for the BiLaD project in Germany.
verbatim. There is no abort criteria for the tests, i.e. if an incorrect response is given, the subjects advance to the next item without receiving corrective feedback. The responses of the children were audio-recorded, transcribed, verified and coded for errors. Overall item accuracy “correct identical repetition” is the main measure used in this study for the NWR task. For the SRT, the measure of “target structure met” is applied as well.

Given the small sample sizes and violation to normality assumption (Shapiro-Wilk test), non-parametric statistical tests were used. For each of the tasks, we will first compare the overall performance of typically developing and language impaired children on the tasks as a whole to see whether they can distinguish TD from SLI. Next, we will look into the effects of syntactic and phonological complexity. This will allow us to verify whether structurally complex constructions are particularly vulnerable in child language impaired populations. Since this is still work in progress, the investigation of phonological complexity effects is limited to the LI-part of the NWRT.

3. Results and discussion

As follows from the introduction, the objective of this study is to determine whether the LITMUS SRT and NWRT (using overall accuracy scores) can identify language impairment in (bilingual) children. The study further aimed at investigating whether complex constructions pose more difficulties for monolingual and bilingual children with SLI. For each task, we first compare the global scores/overall results of the children with and without SLI (Mo-TD/Mo-SLI, Bi-TD/Bi-SLI) and compare the Bi-TD group to the Mo-SLI group to check whether the tests can avoid the frequently reported overlap problem. We then look at the performance of the groups on syntactically and phonologically complex structures to judge whether the SLI groups perform significantly worse on them than their TD peers. These comparisons will be presented and discussed in the subsequent sections.

3.1. Global scores on the German LITMUS SR and NWR tasks

Starting with the SRT, using the scoring measure “identical repetition” we arrive at the distribution depicted in figure 1. On the whole task, monolingual and bilingual children with SLI performed significantly worse than their typically developing peers. Statistical analysis of the percentages of identical repetition for each group yielded significant differences between Mo-TD and Mo-SLI (Z: -3.766, \(p < 0.001\)), and most importantly between Bi-TD and Bi-SLI (Z: -4.203, \(p < 0.001\)), but also between Bi-TD and Mo-SLI (Z: -4.464, \(p < 0.001\)). In contrast, no significant differences were found between Mo-TD and Bi-TD as well as Mo-SLI and Bi-SLI. Figure 1 also shows that Mo-SLIs/ Bi-SLIs as well as Mo-TDs/Bi-TDs pattern alike with no overlap between Bi-TD and Mo-SLI.
Turning now to the group results on the nonword repetition task as shown in figure 2, the statistical analysis of the overall accuracy scores reveals a similar pattern to the SRT (see figure 2). The Mo-TD group performed significantly better than the Mo-SLI group (U=2, Z= -3.610, p < 0.001). Likewise, the Bi-SLI children were outperformed by their Bi-TD peers (U=5.5, Z= -3.940, p < 0.001). There was no overlap between Mo-SLIs and Bi-TDs since typically developing children scored significantly better than monolingual children with SLI (U=13, Z= -3.940, p < 0.001). Furthermore, no significant differences emerged between Mo-SLI and Bi-SLI as well as between Mo-TD and Bi-TD.

Considering group results, both tasks appear to sharply distinguish between SLI and TD. For overall accuracy scores on the French LITMUS SRT and NWRT, a 60% cut-off was reported to have a good diagnostic accuracy (90%)\(^4\) in separating Bi-SLI from Bi-TD (Tuller et al. 2015; Fleckstein et al. 2016). Applying this 60% cut-off to our data set yields diagnostic accuracy of 97% for the SRT as measured by identical repetition: 9/9 bilingual children with SLI performed below 60% (sensitivity\(^5\) = 100%), while 21/22 typically developing children performed above 60% (specificity = 95%). As expected, scoring by “target structure met” on the SRT raised the specificity of the task to 100%, since it compensates for typical L2 errors. As for NWRT, the task identified all of the 9 Bi-SLI children as having SLI and 20/22 Bi-TD children as typically developing (sensitivity = 100%, specificity = 95%). Thus, it could be said, that

\(^4\) According to Vance and Plante (1994), a sensitivity or specificity level above 90% is considered to have good discriminatory power. High sensitivity helps avoid under-diagnosis, whereas high specificity ensures that typically developing children are not over-diagnosed with SLI.

\(^5\) Sensitivity measures the proportion of language impaired children who score below a particular cut-off on a task and are thus identified correctly as being language impaired by the task, while specificity measures the proportion of typically developing children who do not score below cut-off and are correctly identified as TD (Marinis & Armon-Lotem 2015).
both of these tasks are well-suited for the identification of SLI in bilingual
children.

### 3.2. Syntactic Complexity effects in Children with SLI

In order to investigate the effect of syntactic complexity we used Wh-
movement with and without intervention (Rizzi 2004; Friedmann et al. 2009;
Friedmann & Novogrodsy 2011; Hamann et al. 2017) and embedded structures
with and without Wh-movement and intervention. We therefore compare bare
object Wh-questions that do not involve embedding nor intervention (*Whom
does the penguin hug today?*), Which-NP questions, which have intervention but
no embedding (*Which clown does the magician visit*?), finite complement
clauses, which do not involve Wh-movement, see (2), subject relatives (SR),
which involve embedding and Wh-movement but do not have intervention (*I
see the king, who hugs the clown*) and object relatives, which can involve
embedding, Wh-movement and intervention (OR + intv) as shown in (3). We
will not report on passives here, which generally are well mastered (see Hamann
et al. 2017). In order to keep memory load as constant as possible in the object
and subject relatives, we only used two lexical DPs in these structures and the
verb *see* consistently in the matrix clause. Additionally, we had included
coordinate structures in the German version of the SRT, which allowed us to
compare the effect of the mere presence of two clauses expressing two
propositions (in coordinations) with the effect of embedding (in finite
complement clauses). Importantly, coordinations contained a higher number of
syllables (13/14) than finite complement clauses (11). In these comparisons we
used the measure of “target structure met” which ignores lexical substitution
errors and systematic case errors such as the use of “wem” – ‘who-
dative’ instead of “wen” – ‘who-accusative’, an error which seems to appear with
schooling.

Figure 3 and Figure 4 show that sentences with higher degree structural
complexity, Wh-object questions, finite complement and relative clauses are
particularly difficult for Mo-SLI and Bi-SLI children. Children with SLI seem to
have about equal difficulties with Wh-movement (object questions, WH-bare
and Wh-NP) and simple embedding (finite complement clauses, CompFin).
Note that in these structures, bilingual children with SLI perform slightly better
than their monolingual peers, though these differences are not significant. Note
also that in these particular groups we do not see an effect of intervention in
object questions. Because of the strong effect of either embedding or Wh-
movement in isolation, an interaction of embedding and Wh-movement could
not be measured in that performance on SR was not significantly worse than
performance in either object questions or finite complement clauses. However,
the most complex construction, object relatives with intervention (embedding,

---

6 Detailed statistical analysis on syntactic and phonological complex constructions will be provided in Abed Ibrahim et al. (in preparation).
Wh-movement and intervention), is difficult for all groups. Even monolingual TDs of this age group produce less than 50% correct repetition using the measure of “target structure met”. Children with SLI, however, perform at 0%.

The main error source for TD children is erroneous case marking of the relative head pronoun, e.g. double accusative, or canonization, i.e. the production of a subject relative instead of an object relative exchanging the thematic roles, whereas children with SLI tend to avoid embedding and produce monoclausal instead.

That children with SLI have extraordinary difficulties with embedding is evident in the comparison of coordination and finite complement clauses in Figure 4: Unlike TDs, both Mo-SLI and Bi-SLI children perform significantly worse on structurally more complex biclausal: Coord. is significantly easier than either CompFin or SR, and OR is not produced at all.

We also investigated L1 effects, but could only observe tendencies which did not reach significance: In the production of ORs and CompFin Bi-TD-Arabic children performed better than Turkish Bi-TD children, and in the production of SRs Bi-TD-Turkish children outperformed their Bi-TD Arabic speaking peers.

3.3. Phonological Complexity effects in Children with SLI

As mentioned in (2.1.2), the LITMUS-NWRT consists of two parts differing in their inherent complexity. The LI-part is considered to be structurally less complex as it includes typologically well-attested structural variables. The LD-part is phonologically more complex, since it contains three-member consonant clusters including clusters violating the Sonority Sequencing Principle due to the presence of the extrametrical /s, f/. Given its higher degree of complexity, the LD-part is expected to be particularly difficult for children with SLI.
A cross-group comparison (see fig.5) of the performance on the NWRT LI and LD parts showed that both SLI groups (monolingual and bilingual) performed significantly worse than the TD groups on both parts. Statistical analysis yielded the following significant comparisons: **LI:** (Mo-TD/Mo-SLI U=11, Z= -2.932, \( p = 0.002 \), Bi-TD/Bi-SLI U= 8.5, \( Z = -3.808, p < 0.001 \), Mo-SLI/Bi-TD U= 21, \( Z = -3.604, p < 0.001 \)), **LD:** (Mo-TD/Mo-SLI U=1, Z= -3.637, \( p < 0.001 \), Bi-TD/Bi-SLI U= 9, \( Z = -3.772, p < 0.001 \), Mo-SLI/Bi-TD U= 9, \( Z = -4.115, p < 0.001 \)). No significant differences were found between Mo-TD and Bi-TD as well as Mo-SLI and Bi-SLI on either of the parts. The fact that the SLI and TD groups significantly differed on both LI and LD parts implies that the phonological difficulties encountered by the Bi-SLI children cannot only be ascribed to the acquisition of L2 phonology/phonological properties (see also dos Santos & Ferré 2016).

Interestingly, comparing the means of accuracy scores on the LD-part shows that the Bi-SLI children perform slightly better than the Mo-SLIs on the LD-part (Bi-SLI: mean = 31.17%, Mo-SLI: mean = 29.36%). A possible explanation for this finding is that the Bi-SLI children profit from the enriched phonological input they receive through exposure to two languages or from bilingual advantages in executive function.

Within-group comparisons showed that only the performance of Mo-SLI and Bi-SLI children significantly drops on the structurally more complex LD part (Mo-SLI: \( Z = -2.847, p = 0.004 \), Bi-SLI: \( z = -2.660, p = 0.008 \)). Moreover, the performance gap between SLI and TD was greater on the LD part, not only for the bilingual but also for the monolingual groups. These findings corroborate that difficulties with phonological complexity are indicative of SLI not only in monolingual but also in bilingual populations (see also Grimm and Hübner to appear).

---

**Fig. 5:** NWRT: % correct identical repetition in the LI vs. LD parts split by group
In the LI part, phonological complexity effects were mainly observed in disyllabic and trisyllabic nonwords. As shown in figure 6, monolingual and bilingual children with SLI performed significantly worse than their TD peers on structurally more complex nonwords. Nonwords containing consonantal coda, e.g. CVCVCVC or tautosyllabic consonant clusters like CCVCCV, CVVCVCVCV and CVCCVCVCV led to a significant decrease in the performance of both SLI groups.

Furthermore, the position of the consonant cluster within the nonword seemed to affect the performance especially of the SLI children. Branching onsets were particularly challenging for SLI children if they occurred word-medially (see figure 7 & 8). The presence of word initial clusters did not affect the performance of the SLI groups on two-syllable nonwords. On the contrary, SLI children performed TD-like on CCVCV as opposed to CVCCV (see figure 7). Although the SLI groups performed significantly worse than their TD-peers on trisyllabic items with branching onsets, they performed even worse if the consonant cluster occurred word-medially (see figure 8).
The data further show that longer nonwords (in terms of syllable number) were not necessarily more difficult for the SLI groups. SLI children perform significantly worse on structurally more complex 2-syllable non-words as opposed to the structurally less complex trisyllabic CVCVCV (see fig. 9). The latter observation allows us to assume that it is phonological complexity and not mere working memory limitations which are responsible for the reduced phonological abilities of children with SLI.

![Fig. 9 NWRT (LI): % Mean rate of identical repetition per structure](image)

4. Conclusion

The results presented here clearly demonstrate that both, the German LITMUS SRT and the German LITMUS NWRT, identify children with SLI also in bilingual populations and can be applied in different L1-L2 combinations. The fact that both tests differentiate typical and impaired language development in monolingual children shows that these tasks are well constructed and target the relevant syntactic and phonological structures. The findings that a) neither test shows an overlap of monolingual SLI and bilingual typical development and b) both tests identify SLI in bilingual children identifies them as reliable diagnostic tools in multilingual settings and thus provides corroboration for the COST/LITMUS construction principles, especially the decision of targeting complex structures.

We find that linguistic complexity significantly influences the performance of children with SLI (Mo-SLI and Bi-SLI) and manifests itself in difficulties with embedding, Wh-movement and intervention in the SRT (fig. 3 & 4) and in particular difficulties with e.g. word-medial consonant clusters in both disyllabic and trisyllabic nonwords in the NWRT (fig. 6, 7 & 8). Additionally, we argue that working memory deficits alone cannot account for the poor performance of Mo-SLI and Bi-SLI in complex shorter items in SRT (Finite complements vs. Coordination) and NWRT (CCVCCV vs. CVCVCV). The tasks used here can therefore be assumed to target linguistic complexity and exploit the weak performance of children with SLI in linguistically complex structures.
Because the profiles of typically developing bilingual children differ significantly from those of children with SLI, particularly in complex constructions, it can be stated that typical development allows the acquisition of core properties of complex syntax (Wh-movement, embedding, intervention) and phonology, whereas children with SLI show deficits. Note also that a cumulative effect of SLI and bilingualism cannot be found in any of the comparisons. On the contrary, in some structures Bi-SLIs outperform their monolingual peers: in object questions (fig. 4), in the language specific part of the NWRT (fig. 5) and for example in CCVCCV nonwords (fig.6). We attribute the better performance to an input which can be called enriched in that it provides more variety. In further research we will explore whether bilingual advantages in executive function such as superior inhibition and/or switching contribute to the patterns we found in our repetition tasks (see also Zebib et al. 2016). We will also investigate the influence of L1 in more detail since the marginal L1-effects we found in certain conditions indicate that SRTs need to include a range of structures in order to retain cross-linguistic validity. Finally, we intend to systematically explore the effect of length of exposure by contrasting the performance of our groups with the performance of late successive child bilinguals with around 12 months of exposure.

References


Friedmann, Naama & Rama Novogrodsky. (2011). Which questions are most difficult to understand? The comprehension of Wh-questions in three subtypes of SLI. *Lingua* 121, 367-382.


