Past tense productivity: Comparing bilingual children and children with SLI

Judith Rispens¹ and Elise de Bree
University of Amsterdam

1. Introduction

Successful language acquisition depends on several factors. Child-internal capacities play a role (e.g. the ability to hear and process language) and also external variables (for example quantity and quality of language input). Children with specific language impairment (SLI) show language learning difficulties. They have problems with processing and learning linguistic patterns. Bilingual children are often exposed less to their languages than their monolingual typically developing (TD) peers, due to shared resources, which may interfere with their language development. In the present study we compared past tense productions of children with SLI with those of bilingual children. We investigated whether 1) overall accuracy of Dutch past tense productions differed between these groups and 2) whether the groups of SLI and bilingual children were equally sensitive to frequency information involved in past tense production.

1.1. The Dutch past tense

Dutch regular past tense inflection is formed by adding the one-syllable allomorph -te or -de to the verb stem. The selection of the suffix is guided by an interaction between morphology and phonology. Verb stems ending in an underlyingly voiceless obstruent take -te, all other verbs take -de (Ernestus & Baayen, 2001). For example, the infinitive bakken (to bake), with the stem ending in a voiceless obstruent (bak), becomes bak-te (baked) in the past tense, whereas noemen (to name) with a stem-final sonorant becomes noem-de (named).

For the acquisition of the Dutch regular past tense, sensitivity to the interaction between morphology and phonology is thus necessary, similar to, for example, English that has past tense allomorphs /t/, /d/, and /Id/. In addition, lexical frequency information plays a role in the production of the past tense (Derwing & Baker, 1980; Marchman, 1997; Matthews & Theakston, 2006). Frequency can be divided into token frequency, i.e., the frequency with which a

¹ Both authors are affiliated with the University of Amsterdam. The corresponding author is Judith Rispens, j.e.rispens@uva.nl
specific verb and the inflected form of the verb occur, and type frequency. Words with high token frequency are assumed to have stronger memory representations, improving accessibility of that word, compared to low frequency items (e.g. Luce & Pisoni, 1998). Experimental evidence shows that token frequency influences children’s past tense marking of regular verbs (e.g. Oetting & Horohov, 1997; Matthew & Theakston, 2006). In other words, a higher frequency of the inflected form facilitates its retrieval. Ernestus and Baayen (2001) found evidence for token frequency effects in the productions of past tense forms of Dutch adults and Rispens and de Bree (2014) found that token frequency influenced past tense production in 7-year-olds.

Type frequency refers to the number of instances that lexical items occur in a certain pattern (Bybee, 2007). For regular past tense production, this refers to the frequency of past tenses expressed by the allomorph –te versus –de. Type frequency is assumed to affect morphological productivity as highly frequent suffixes tend to be applied to newly learned items rather than low frequent suffixes (Bybee 2008). Highly frequent past tense markers (types) are thus expected to be most productive. The type frequency of the two allomorphs, -de and -te in Dutch is different. The combination of voiceless consonants + -te is phonotactically more frequent than voiced consonants + -de (Rispens & de Bree, 2014). In this latter study of past tense production evidence of token frequency effects, as well as type frequency based on the frequency of the verb stem + allomorph (-te > -de) were found (Rispens & de Bree, 2014). Both monolingual TD Dutch 5- and 7-year-olds showed an effect of allomorph type on past tense production. The 5-year-olds showed an effect of type frequency (-te > -de) when they inflected lexical and pseudo-verbs for the past tense, and the 7-year-olds only showed this effect of type frequency for the inflections of novel verbs. Furthermore, for the 7-year-olds, not the 5-year-olds, the past tense production of lexical regular verbs was influenced by token (lexical) frequency, with verbs high in frequency being inflected more accurately. The absence of a token frequency effect in the younger group suggests that vocabulary size influences the occurrence of token frequency effects in verb productions.

Next to frequency effects, phonological preferences can also impact on inflection (see Song, Sundara, & Demuth, 2009; Marshall & van der Lely, 2007). In Dutch, there is evidence that early development of voiced segments/phonemes is different from that of voiceless ones. Dutch children around the age of three have been found to show a clear preference for voiceless over voiced segments (Kager, van der Feest, Fikkert, Kerkhoff & Zamuner, 2007) and acquire voiceless phonemes before voiced ones (e.g. Beers, 1995). Dutch shows final devoicing, and has the tendency of voiceless word-internal clusters (Zonneveld, 1983). The advantage and preference for -te allomorphs in Dutch in the study by Rispens and de Bree (2014) can thus also be accounted for by this phonological preference.
1.2 Past tense production: children with SLI and bilingual children

Children with SLI show severe and persistent difficulties in language learning from the input, despite sufficient child-external language opportunities. Severe problems with past-tense morphology have been reported cross-linguistically in children with SLI (see Leonard, 1998). In a previous study, Rispens and de Bree (2014) found that children with SLI had severe difficulty with past tense production. Nevertheless, like their typically developing peers, they did show an effect of type frequency, with an advantage for \(-te\) allomorph past tenses.

Bilingual children typically receive less input in either language than monolingual children learning just one language, which impacts on their rate of language development. Slowed acquisition in the domain of the past tense has been reported for this group (Blom & Paradis, 2013; Chondrogianni & Marinis, 2012; Nicoladis et al., 2007; Nicoladis & Paradis, 2012; Paradis, 2010; Thordardottir et al., 2006). The amount of input is an important factor in the acquisition pattern of bilingual children (e.g. Pearson et al., 1997; Paradis et al., 2010; Thordardottir, 2011; Unsworth et al., in press) and the typology of the other language of the bilingual children (Yip & Matthews, 2007). In the current study we were particularly interested in examining the effect of bilingualism (Dutch-Hebrew) on factors that have been found to influence Dutch past tense production in monolingual children (type and token frequency). Hebrew has a rich morphological system for the past tense, as person, number and gender are visible in the inflection (Schwarzwald, 2001) but allomorphs do not exist. In Hebrew, tense is marked by vowels in the stem, by suffixation or by prefixation (Armon-Lotem, 2014).

1.3 This study

In the present study, we compared Dutch–Hebrew bilingual children and children with SLI on their production accuracy of regular past inflections of lexical verbs as well as novel verbs in Dutch. We compared their results to those of monolingual TD children of the same age and to monolingual TD children with the same receptive vocabulary score as the children with SLI and the bilingual children. This was done to investigate a possible difference in morphosyntactic productivity as well as to assess whether such a potential difference reflects a rate of development comparable to younger children.

Both children with SLI as well as bilingual children have been reported to show difficulties with past tense acquisition. The causes for these difficulties differ (child-internal versus child-external) but they may impact in the same way on regular past tense production. The research question of the current study focuses on whether the two groups of children use token and type frequency information to the same extent as TD monolingual children.
2. Method
2.1. Participants

Four groups of children participated in this study. A group of children with SLI; a group of bilingual children, and two groups of TD monolingual children. One group included seven- to eight-year-olds (CA MONO TD), these children were age-matched with the children with SLI and with the bilingual children. The second group of TD monolingual children were five-year-olds (younger MONO TD). They were matched on the receptive vocabulary score (PPVT) of the SLI and bilingual groups, see Table 1.

Table 1. Participant characteristics

<table>
<thead>
<tr>
<th></th>
<th>Bilingual</th>
<th>SLI</th>
<th>Younger MONO TD</th>
<th>CA MONO TD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nr</td>
<td>11</td>
<td>25</td>
<td>21</td>
<td>30</td>
</tr>
<tr>
<td>Age (months)</td>
<td>91.9</td>
<td>96.9</td>
<td>70.2</td>
<td>92.9</td>
</tr>
<tr>
<td>Raw PPVT</td>
<td>92.4 (9.1)</td>
<td>90.8 (9.7)</td>
<td>92.0 (9.1)</td>
<td>103.2 (6.4)</td>
</tr>
<tr>
<td>Range PPVT</td>
<td>71-102</td>
<td>76-108</td>
<td>82-106</td>
<td>90-125</td>
</tr>
</tbody>
</table>

The CA MONO TD children all attended second grade primary school. The younger MONO TD children all attended kindergarten. For both MONO TD groups, only children who made normal progress in school and did not have any cognitive or emotional disturbances were included. All children had normal hearing and normal, or corrected-to-normal, vision. Children were all raised with Dutch being the language spoken at home from birth onwards, see Rispens and de Bree (2014) for a more detailed discussion.

The children with SLI were selected from three special schools for children with developmental language disorders. SLI was diagnosed when a child performed at least 1.5 SD below the mean in at least two language domains measured with Dutch standardized language tests, or when a child performed more than 2 SD below the mean on a Dutch standardized general language test. The testing for the diagnosis was carried out by a multidisciplinary team of the special needs schools, and the diagnosis of SLI was a requirement to be accepted by the special needs schools. This group has also been discussed in Rispens and de Bree (2014).

The children in the bilingual group all attended a Jewish school with a bilingual policy as education is provided both in Dutch and Hebrew. Our group of children entered this school at four years of age and thus were comparable in their Dutch-Hebrew bilingual language environment during school hours. As the bilingual children were all at the end of second grade when they participated, they had all experienced almost two years of bilingual formal instruction in the Dutch and Hebrew language and two years of bilingual education in kindergarten. Reports about language use at home revealed that all children grew up in a bilingual setting in which they all spoke both Hebrew and Dutch at home.
Matching on receptive vocabulary score was done by administering the Dutch version of the Peabody Picture Vocabulary Test (Schlichting, 2005). Both groups of children with SLI and bilingual children had significantly lower receptive vocabulary (both raw and standardized scores) compared to the CA MONO TD group, just like the younger MONO TD children (SLI versus CA MONO TD: \( p < .001 \); bilingual versus CA MONO TD: \( p = .01 \); younger MNO TD versus CA MONO TD: \( p = .001 \) see Table 1). The groups of children with SLI and the bilingual children did not show significant differences in their raw receptive vocabulary scores relative to the younger MONO TD children (SLI versus younger MONO TD \( p = .97 \); bilingual versus younger MONO TD \( p = .98 \)).

2.2. Materials

Past tense production was investigated through elicitation of past tense forms of existing verbs and novel verbs, see Rispens and de Bree (2014).

Task 1: Past tense production of existing verbs. Twelve regular verbs were selected based on high and low token frequency counts of the CELEX (Baayen, Piepenbrock, & van Rijn, 1993) database. The verbs were furthermore classified according to the type of past allomorph with which the verbs are inflected. Half of the verb stems ended in an underlying voiceless obstruent (four times /k/ and twice /s/), the other half in a voiced consonant (twice /l/, twice /r/ and twice /w/). These stems were chosen as they are unambiguous with respect to devoicing and rendered unambiguous -te or -de expectations. All verbs in the infinitival form had a CVCVC structure.

Task 2: Past tense production of novel verbs. Fifteen novel verbs were presented. Seven of the verbs ended in an underlying voiceless obstruent (4 times ‘p’ and 3 times ‘k’) and eight in a voiced consonant (4 times ‘m’, twice ‘n’ and twice ‘r’). Only verb stems ending in consonants that demand unambiguously the -te or -de were chosen. The verbs were controlled for phonotactic frequency by using the database of Adriaans (2006).

2.3. Procedure

The children were shown a picture in a PowerPoint presentation on a screen of a laptop computer that depicted an action (see also Rispens & de Bree, 2014). Children heard a prerecorded female voice pronounce the target. Pictures were presented to the children, one at a time, and they were accompanied by a little story during which the picture remained on the screen. This three line story first described what was going on in the picture and then stated that characters performed this action often or every day. The children were then asked to finish the final sentence in which they were prompted to give a past tense form.

The responses of the children were recorded and transcribed on-line and scored afterwards. An item would be scored as correct if the choice of the past-
tense suffix was correct. Violations of agreement between the verb and the subject were not taken into account.

The order of the presentation of the two tasks was pseudo-randomized: half of the children were first presented with task 2 (novel verbs) followed by task 1 (lexical verbs), the other half the other way around.

Qualitative error analysis

In addition to calculating the number of correct responses, we also assigned all erroneous responses to a category, based on the most frequently produced error types. For the regular existing verbs and the novel verbs, five categories were created (see Table 2). Note that especially the error category ‘incorrect allomorph’ is important for our research question. It measures the instances of assigning the wrong allomorph, which may indicate reduced sensitivity to morphophonology in past tense inflection and it may indicate overgeneralization due to the effect of type frequency. The error categories also include ‘other’. Responses that could not be classified as the other error types were assigned to this category. This means that all responses were assigned to a category as it never occurred that a participant responded with silence.

Table 2. Error patterns

<table>
<thead>
<tr>
<th>Error categories</th>
<th>Regular verbs</th>
<th>Novel verbs</th>
</tr>
</thead>
<tbody>
<tr>
<td>-en suffix (infinitive)</td>
<td>bellen (phone)</td>
<td>vuppen</td>
</tr>
<tr>
<td>Zero marking</td>
<td>bel (phone)</td>
<td>vup</td>
</tr>
<tr>
<td>-t suffix (present tense)</td>
<td>belt (phones)</td>
<td>vupt</td>
</tr>
<tr>
<td>Incorrect allomorph</td>
<td>bel (instead of belde)</td>
<td>vupde (instead of vupt)</td>
</tr>
<tr>
<td>Other: Various types of errors</td>
<td>Examples:</td>
<td></td>
</tr>
<tr>
<td></td>
<td>a). verb ‘go’ + verb:</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(ging bellen)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>b). change within the</td>
<td></td>
</tr>
<tr>
<td></td>
<td>verb stem (e.g. bemde)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>c). infinitive + ‘-de’</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(bellende)</td>
<td></td>
</tr>
</tbody>
</table>

3. Results

3.1 Correct scores

Results on accuracy of both existing and novel past tense productions are presented in Table 3. With respect to the regular lexical verbs, a repeated measures ANOVA with Group as between subjects factor (four-levels) and Type (two-levels) and Token frequency (two-levels) as within-subject factors showed a main effect of Group, $F(1, 83) = 39.2, p < .001$, a main effect of type frequency, $F(1, 83) = 17.4, p < .001$, and a main effect of token frequency, $F(1, 83) = 6.04, p = .016$. There were no significant interactions between type
frequency, token frequency, and/or group. Posthoc follow-up comparisons (Games Howell due to unequal group sizes) of the significant group effect showed that the children with SLI performed significantly more poorly than the other three groups ($p < .001$ for all comparisons). There were no other significant differences between the groups.

**Table 3. Past tense inflection (mean proportion scores correct and standard deviation between parentheses) of the four groups on the lexical and novel verbs**

<table>
<thead>
<tr>
<th></th>
<th>Lexical verbs</th>
<th>Novel verbs</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>HF –te</td>
<td>LF –te</td>
</tr>
<tr>
<td>SLI</td>
<td>.29 (.41)</td>
<td>.29 (.32)</td>
</tr>
<tr>
<td>Bilingual</td>
<td>.97 (.10)</td>
<td>.97 (.10)</td>
</tr>
<tr>
<td>TD 5-year-olds</td>
<td>.97 (.25)</td>
<td>.89 (.19)</td>
</tr>
<tr>
<td>TD 8-year-olds</td>
<td>.92 (.17)</td>
<td>.82 (.27)</td>
</tr>
</tbody>
</table>

For the novel verbs, a repeated measures ANOVA with group as the between-subjects factor (four-levels) and allomorph type as the within-subjects factor (two-levels) showed an effect of Group, $F(1, 83) = 32.9$, $p < .001$, an effect of type frequency, $F(1, 83) = 58.2$, $p < .001$, as well as an interaction between the two, $F(3, 83) = 5.9$, $p = .001$. Follow-up tests per group were carried out to unpack this significant interaction. The bilingual group and the two TD groups showed significant effects of type frequency (Bilingual group + younger TD group $p < .001$; CA TD group $p = .016$). The difference between the two allomorphs did not reach significance in the children with SLI ($p = .11$), possibly caused by the floor performance of the group. Posthoc tests (Games-Howell) were carried out to investigate the significant group effect and showed that the SLI group performed significantly more poorly than the other three groups ($p < .001$ in all comparisons). There were no other significant differences between the groups.

### 3.2. Error patterns

The distribution of errors produced in the regular verbs is presented in Table 4 and the distribution of errors produced in the novel verbs in Table 5. We calculated the proportion of errors for each response category relative to the total number of errors: the sum of the percentages within each category thus add up to 100%. The error pattern of the bilingual group is straightforward: the dominant error pattern is the incorrect allomorph, specifically, the production of
a –te allomorph when it should be –de. This pattern is observed in both the lexical as well as the novel verbs. For the SLI group, in contrast, the error pattern resembles that of the younger MONO TD children in showing a relative high degree of infinitives and present tense realizations instead of the past tense. The error pattern of the CA MONO TD group, finally, is characterized by the production of ‘other’ errors in the regular verbs (but note that their correct scores were high), and by a distribution over all categories in the novel verb productions with most errors being ‘wrong suffixes’ and ‘other’. The production of wrong suffixes is less dominantly present compared with the bilingual group.

Table 4. Percentages of errors (SD) for regular verbs per group

<table>
<thead>
<tr>
<th>Sum of errors</th>
<th>-en</th>
<th>Zero</th>
<th>-t</th>
<th>Incorrect allomorph</th>
<th>Other</th>
</tr>
</thead>
<tbody>
<tr>
<td>=100%</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bilingual</td>
<td>0 (0)</td>
<td>0 (0)</td>
<td>16.7 (40)</td>
<td>83.3 (40)</td>
<td>0 (0)</td>
</tr>
<tr>
<td>SLI</td>
<td>25.2 (25)</td>
<td>15.8 (24)</td>
<td>34.6 (32)</td>
<td>7.0 (23)</td>
<td>19.2 (25)</td>
</tr>
<tr>
<td>5-yr TD</td>
<td>24.1 (35)</td>
<td>13.7 (22)</td>
<td>44.2 (36)</td>
<td>12.7 (26)</td>
<td>5.4 (9)</td>
</tr>
<tr>
<td>8-yr TD</td>
<td>5.6 (17)</td>
<td>4.6 (13)</td>
<td>10.6 (22)</td>
<td>8.3 (18)</td>
<td>69.8 (31)</td>
</tr>
</tbody>
</table>

Table 5. Percentages of error types (SD) for novel verbs per group

<table>
<thead>
<tr>
<th>Sum of errors</th>
<th>-en</th>
<th>Zero</th>
<th>-t</th>
<th>Incorrect allomorph</th>
<th>Other</th>
</tr>
</thead>
<tbody>
<tr>
<td>=100%</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bilingual</td>
<td>4.3 (10)</td>
<td>0 (0)</td>
<td>4.5 (15.1)</td>
<td>65.9 (44)</td>
<td>25.3 (31)</td>
</tr>
<tr>
<td>SLI</td>
<td>26.4 (25)</td>
<td>8.4 (15)</td>
<td>33.5 (28)</td>
<td>6.4 (18)</td>
<td>26.2 (17)</td>
</tr>
<tr>
<td>5-yr TD</td>
<td>17.8 (27)</td>
<td>0 (0)</td>
<td>38.8 (35)</td>
<td>23.5 (29)</td>
<td>20.0 (27)</td>
</tr>
<tr>
<td>8-yr TD</td>
<td>18.7 (38)</td>
<td>0 (0)</td>
<td>12.1 (25)</td>
<td>36.5 (41)</td>
<td>32.6 (41)</td>
</tr>
</tbody>
</table>

4. Discussion

The present study compared the production of the past tense of children with SLI and bilingual children to investigate whether these groups were equally sensitive to type (-te and –de) and token (lexical) frequency when producing the past tense inflections. Their performances were compared to those of monolingual TD children of the same age and to those of younger control children with the same receptive vocabulary to investigate whether vocabulary proficiency would explain potential past tense difficulties.
Looking at the overall accuracy scores, a clear difference between the bilingual group and the children with SLI was found: the children with SLI performed significantly lower than the bilingual group and than the two TD monolingual control groups. In contrast, the bilingual children’s accuracy of past tense productions did not differ from that of the two monolingual control groups. This finding was somewhat unexpected based on other findings in the literature that showed past tense delays for bilingual children (Thordardottir et al., 2006; Nicoladis et al., 2007). Interestingly, this difference between the bilingual children and those with SLI cannot be explained by an advantage in vocabulary for the bilingual group as the receptive vocabulary scores for these two groups were the same. The severe problems that the children with SLI showed thus indicate a specific problem with past tense morphology.

Turning to the frequency effects, for all groups a significant influence of token frequency was found, with more frequently occurring existing verbs being inflected more correctly. In addition to token frequency, the past tense productions were found to be influenced by type frequency. Verbs demanding the -te suffix were produced significantly more accurately than –de verbs in all groups and in both task conditions (lexical and novel verbs). The exception on this effect was the performance of the children with SLI on their novel past productions, but they showed severe difficulties with this verb task (with an average error rate of 82%).

The significant effects of frequency showed that the bilingual children and the children with SLI were sensitive to the distribution of the frequency with which lexical verbs occur (token frequency) and to the distribution of the phonotactic environment of the allomorphs (type frequency). This sensitivity to frequency patterns was thus found despite the fact that the children with SLI showed clear problems with the past tense and despite the fact that the bilingual children and the children with SLI showed language acquisition delays in the form of depressed vocabulary scores.

The error analyses of the past tense productions established differences between the bilingual children and the children with SLI. The bilingual group, unlike the children with SLI and monolingual control children, almost exclusively produced incorrect allomorph errors. More specifically, they produced –te allomorphs instead of -de allomorphs. The bilingual children thus seemed to be less sensitive to the phonological constraints of past tense morphophonology. The SLI group showed a high rate of present tense realizations and infinitives instead of the target past tense productions. This pattern was similar to that of the younger monolingual TD children. The error rate of the CA TD group was low and showed some errors in several categories. Thus, even though the bilingual children were as accurate in past tense productions as the CA TD monolingual children and the younger monolingual children, they differ in the types of errors made. The bilingual children showed overgeneralizations of the –te allomorph within a morphophonological environment demanding a –de allomorph. The SLI children, however, showed severe difficulties with past tense productions, and were outperformed by the
bilingual and monolingual control groups, but they did show the same type of errors like the monolingual TD children.

There are a number of ways to interpret the different error pattern of the bilingual children. First, if children acquire two languages, one with and one without allomorphs, the absence of allomorphs in one language (Hebrew) could render higher sensitivity to allomorphs in the other language (Dutch). The increased frequency of final consonant + -te relative to final consonant + de in Dutch could lead to faster acquisition of the former pattern, as well using the –te as default in existing regular verbs with low token frequency and in novel verbs.

A different account for the findings would be based on language transfer. Although Hebrew does not have allomorphs, its rich productive past tense system contains several past tense suffixes with a /t/ + vowel, but none with a /d/ + vowel. These morphophonological characteristics of Hebrew past tense morphology might have influenced the bilingual children’s production of –te in Dutch past tenses (see also Nicoladis, Song & Marentette, 2012; Goad, White & Steele, 2003 for prosodic transfer effects). Our results showed that the disadvantage for -de allomorphs was most prominent in the verbs with low token frequency, and in the novel verbs. Thus, phonology of the other language may guide past tense marking in instances where the lexical representation is not available or not activated consistently due to low token frequency.

As the sample size of the bilingual children was limited, this study should be replicated with a larger sample, and should include past tense inflection in both Dutch and Hebrew. In addition, comparing past tense production and comprehension (Rispens & Woensdregt, 2013) is needed and additionally, more information of the amount and quality of exposure to Dutch is warranted (e.g. Blom, 2010, Paradis et al., 2010; Unsworth et al., in press). Nevertheless, the current findings demonstrated that children with SLI and bilingual children are sensitive to the distribution of the type and token frequency of the regular past tense, despite a delay in receptive vocabulary score (both groups); despite a severe problem with production of grammatical morphology (SLI group) and despite less exposure to Dutch than monolingual children (the bilingual group). Error analyses pointed to differences between the productions of the bilingual children relative to monolingual children, reflecting different language learning influences.

References


Unsworth, Sharon, Argyri, Froso, Cornips, Leonie, Hulk, Aafke, Sorace, Antonella, & Tsimpi, Ianthi (in press). On the role of age of onset and input in early child bilingualism in Greek and Dutch. *Applied Psycholinguistics*
