Acquisition of Morpho-phonology in Children with Specific Language Impairment and Typically Developing Children

Annemarie Kerkhoff and Elise de Bree

This paper addresses the acquisition of morpho-phonological alternations in Dutch plurals. A group of children with Specific Language Impairment (5;2) and their age-matched typically developing controls (5;1) were tested on productive knowledge of the alternation. Dutch plurals may alternate in voicing due to a process of final devoicing (/bed/ → [bet] ~ [bedən] beds). In an experiment, children were asked to form plurals of both words (e.g. [bet]) and non-words (e.g. [slat]). Results show that even though children with SLI produce more singulars than their controls, they are able to inflect non-words and regularise existing words. Furthermore, both groups of children show worse performance on alternating than on non-alternating words, showing an equal effect of irregularity. Voicing alternations for non-words (e.g. [slat] ~ [sladən]) were also produced by both groups. Like younger children, SLI children seem less sensitive to the distribution of alternating words in Dutch. Rather than interpreting children’s behaviour as rule-based, we argue for word formation based on analogy for both groups. SLI children seem poorer at forming associations among stored words and extracting generalisations from them.

1 Introduction
The present study addresses the acquisition of morpho-phonology in 25 Dutch children with Specific Language Impairment (SLI) with a mean age of 5;2 and 27 age-matched typically developing (TD) controls with a mean age of 5;1. We investigated acquisition of the Dutch plural, which shows an alternation due to final neutralisation of the voice contrast in the singular. For example, the target /bed/ bed is said to contain an underlyingly voiced obstruent which is neutralised in its singular form [bet] but surfaces in the plural form [bedən]. In contrast, words that end in underlyingly voiceless obstruents do not alternate (e.g. [pet] ~ [petən] caps), showing that there is no rule of intervocalic voicing in Dutch. The voicing alternation has generally been described as a phonologically conditioned alternation, caused by the cross-linguistically common process of final devoicing (Hubers and Kooij 1973; Booij 1995).

Children acquiring Dutch have been observed to mostly make errors in which they match the singular’s voicing value in the plural (e.g. *[betən] beds), opting for a
non-alternation. However, overgeneralisation of voicing has also been attested in spontaneous speech (e.g. *[boːdən] boats, *[əlifəndən] elephants (Kager 1999:336).

Kerkhoff (2004) studied the productive knowledge of the alternation in 59 typically developing children aged 2:9-7:8. These children were asked to form plurals of both existing words and non-words with the final obstruent placed in different phonological environments. The most common error type for words (*[bətən]) was produced in 37% of cases, with error percentages still high (31%) for the oldest children. Errors where voicing was overgeneralised (*[pədən]) occurred in 3% of cases, and in 4% of cases alternations for non-words (e.g. *[kət] ~ *[kədən]) were produced. The youngest children (2;9-4;0) seemed least sensitive to the distribution of the alternation in Dutch, since they produced plurals like *[tə:p] ~ *[tə:bən], occurring in a ‘lexical gap’ environment: p ~ b alternations in Dutch do not occur after long vowels and nasals (Zonneveld 1978:49). Kerkhoff argued that these unexpected alternations could be due to an early rule of intervocalic voicing, suggesting that young children generalise from phonological knowledge. It was also concluded that effects of analogy become stronger in the course of acquisition.

The present study was conducted to further investigate the issue of distinguishing rules from generalisations based on analogy. It is generally agreed upon by supporters of rule-based models that SLI children have an impaired grammar, which does not allow them to use (morphological or phonological) rules. By testing their morphophonological abilities, we hope to shed more light on the issue of linguistic rules. Following a discussion of the relevant literature, we will present the results of a production experiment onvoicing alternations in children with SLI and a control group.

1.1 Acquisition of morpho-phonology

It is generally assumed that a rudimentary lexicon is acquired by the age of one year. Knowledge of morpho-phonological alternations begins subsequently, with a strong dependence on vocabulary development. Regular and productive alternations (such as the voicing assimilation in the English plural suffix) appear early, while the knowledge of unproductive, irregular, or opaque alternations continues to develop into adulthood (see Bernhardt and Stemberger 1998; Pierrehumbert 2003).

In rule- or constraint-based models of language, such as the dual mechanism model (Pinker and Prince 1998), a strong distinction is made between the lexicon and the grammar. Irregularly inflected words are stored in the lexicon and regularly inflected words are formed by a (symbolic) rule that operates on a category (e.g. noun). Importantly, in rule-based models, productivity or frequency does not play a role in determining the regularity of a process. Evidence for symbolic rules could be found if processes apply across-the-board, since rules have either been acquired or not.

With respect to the Dutch voicing alternation, most phonologists assume that there is an underlying form /bɛd/ containing a voiced (or even underspecified) segment, which is regularly inflected by adding the –en suffix. The underlying voice specification of the final segment can only be deduced on the basis of alternations. This poses a problem for acquisition, since children need to undo the neutralisation in the singular to set up an underlying form that is not identical to the surface form. It is

---

4 The main mechanism for plural formation in Dutch is the addition of an /s/ or /z(n)/ suffix. The choice between the two productive suffixes depends on the phonological characteristics of the preceding context (de Haas and Trommelen 1993; Booij 1995).
thought that although words may be stored as wholes in the first stage of learning, the child will notice the semantic overlap between alternants and arrive at a more abstract lexical representation at a later stage. However, a child may go through intermediate stages in which overgeneralisations are produced that reflect phonological generalisations, mirroring cross-linguistically common natural processes such as postnasal or intervocalic voicing (Kager 1999). These processes could be phonetically motivated but might also reflect innate knowledge. Productive use of voicing in non-words (e.g. after nasals) would be indicative of children’s ability to set up ‘early’ phonological rules, even though these rules might turn out to be false at a later stage. Importantly, overextension of voicing need not necessarily reflect the distribution of voicing alternations in the input.

Alternatively, in single mechanism, connectionist or usage-based models of language processing, it is assumed that all inflected words may be stored in the lexicon and word formation is based on analogy (cf. Skousen 1989, Bybee 1985; 1994; McClelland and Patterson 2002). Here, the productivity of a process or pattern directly reflects its type frequency. Importantly, irregularly inflected words are not stored as unanalysed wholes: they form connections with related words in the same paradigm and in other paradigms that show morphological similarities. Bybee proposes that schemas emerge from these connections with varying lexical strengths. Stems are connected to inflected forms, but inflected forms themselves form connections to other inflected and non-inflected forms. A regular form may either be accessed from the lexicon or derived by applying a schema, depending on the token frequency of the form. This type of account is similar to a ‘dual-route’ model (Baayen et al. 1997), where storage and computation proceed in parallel. Importantly, the notion ‘rule’ as a symbolic operation that applies uniformly is replaced by the notion of analogy, which is probabilistic in nature. According to Bybee, phonological rules begin as phonetically motivated processes, which gradually become more and more involved in the morphology and the lexicon. They tend to lose their phonetic motivation and remain only as fossilised alternations in assorted morphological environments.

Under this view, Dutch alternating plurals could be much like other irregular processes (such as the limited set of Dutch plurals that exhibit vowel lengthening), in that they are productive only to a very limited extent due to their low type frequency. Also, alternating plurals are expected to be stored alongside their singular. Crucially, however, they need not be stored as unanalysed wholes. Children are expected to be sensitive to distributional properties of voicing (cf. Cutler and Carter 1987), and to apply word-formation according to similarity- or exemplar-based analogy. This strategy could lead to alternations for non-words, mirroring existing Dutch words. Words with high token frequencies are expected to attract more overgeneralisations. As the specific domain of generalisations is unknown, expectations are not clear-cut.

Even though the above theories seem to take a very different approach, they could nevertheless lead to highly similar predictions, especially with regard to overgeneralisations of voicing in non-words. The distinction between rules and analogical generalisations may turn out to be illusionary, as both models incorporate a distinction between storage and processing. Crucially, in order to use analogy, children need to be able to analyse stored words and abstract generalisations from them. This strategy leads to outcomes that are similar to the ones generated by a rule-

---

5 See Ernestus and Baayen (2001) for such results on Dutch past tense formation.
based strategy. Still, we hope to contribute to this ongoing debate by showing that one theory is better able to describe the behaviour of the two groups of children.

1.2 Morpho-phonology and SLI
Children with SLI have been found to be delayed in both their phonological and morphological development. In order to explain the deficits in SLI, rule-based models of language have often been applied. Those who claim that morphology itself is affected (cf. Gopnik and Crago 1991; Gopnik and Goad, 1997; Van der Lely and Ullman 2001), have pointed towards poor performance of SLI children on regular versus irregular past tense marking in English. They argue that a grammatical deficit causes inability to decompose inflectionally complex words into their constituent parts. SLI children rely on analogy with existing words to extend patterns to new forms. Furthermore, they use lexically based strategies instead of rules to compensate for impaired inflectional morphology, such as learning plurals by rote as if they were lexical exceptions. Goad and Rebelatti (1994) claim that plural realisation of English SLI children was based on explicit rule-learning (add -s to create a plural). According to these authors, SLI children are unable to use the (phonological) rules component to master the suffix alternation. This claim was further supported by the fact that they found a difference in performance between words and non-words.

Others claim that phonological complexity affects regular inflection more than irregular inflection in English, since regulars have word-final clusters much more often than irregulars (walked vs. ran) (Joanisse and Seidenberg 1998; Marshall 2004). Moreover, it has consistently been found that SLI children show more and persistent instances of phonological simplification compared to both age-matched and MLU-matched controls (e.g. Aguilar-Medinilla et al. 2003; Beers 1995). Evidently, weaknesses in phonology may have adverse effects on morphological abilities. Final consonant (cluster) omission, reduction, or substitution, for example, have been found to interfere with production of the grammatical morphemes for plurals, possessives, regular past tense and regular third-person singular (e.g. Bortolini and Leonard 2000; Owen et al. 2001). Plural marking has generally been found to be delayed in SLI (e.g. van Alphen et al. 2004; Oetting and Rice 1993), but compared to other morphological processes, such as verb inflection, plural acquisition seems more intact (e.g. Bishop 1994; Rice and Oetting 1993).

1.3 Production of the voicing contrast
Studying mastery of the voicing alternation demands an acoustic analysis of the realisation of the voicing contrast. It is assumed that phonetic acquisition is closely associated with timing constraints, and that inter- and intra-subject variability decreases with age (e.g. Smith 1992). Two prominent acoustic features of the Dutch voicing contrast are closure duration (CD) and burst duration (BD). Since voicing can only be maintained for a certain amount of time, CD is relatively short for voiced stops. Burst duration is also shorter in the voiced stop due to the lower intra-oral pressure (Slis and Cohen 1969). Kuijpers (1993) found that durational differences of typically developing Dutch children (4;5, 6;4 and 12;2) were similar to those of adults. The voicing contrast in intervocalic stops was present for all age groups, but 4-year-olds showed more variability (in terms of standard deviations) compared to the other age groups. Furthermore, children applied longer closure durations for voiceless stops than adults, and durations decreased with age.

Children with SLI have been found to exhibit motoric difficulties, already manifest at the babbling stage (e.g. Whitehurst et al. 1991). These motoric difficulties
persist, as for example, Edwards et al. (1999) found that phonologically disordered children (4;4) used less controlled gestures from stops to vowels than their age peers. With respect to voicing, some studies have established that voicing develops fast compared to other features in SLI children, and that it is maintained more than place and manner of articulation (e.g. Forrest and Morrissette 1999). However, others claim that language-delayed children’s control of the acoustic-phonetic details of the initial and final voicing contrast is less mature than that of normally developing children (Bond and Wilson 1980).

The question thus arises how intervocalic voicing is realised by the two groups of children and how this might affect the acquisition of morpho-phonological alternations.

1.4 Summary and predictions

The current study was carried out to assess productive knowledge of the voicing alternation in TD and SLI children. The expectation is that children with SLI will have more difficulty with plural marking than TD children. Under a rule-based approach, the absence of a morphological rule should leave SLI children unable to inflect non-words or produce overregularisations (*[betan]). Moreover, if all inflected words are treated as unanalysed wholes by SLI children, as most rule-based models would predict, alternating plurals should not be more difficult than non-alternating plurals (given that SLI children have no difficulty producing voiced obstruents). SLI children might even be expected to perform better on alternating plurals than on non-alternating plurals, since no knowledge of the alternation is needed to memorise inflected forms, and alternating plurals are more frequent.

On the other hand, if children store all plurals but are still able to derive generalisations from them (as would be evidenced by their ability to inflect non-words), alternating plurals are expected to be more difficult for both groups of children. The group of unimpaired children may treat alternating plurals as regular, and derive all plurals by adding a suffix to a stem (which may end in a voiced or voiceless obstruent). This means that the difficulty of learning alternating plurals would lie in learning the underlying voicing value of the singular. There seems to be no way of telling whether unimpaired children actually do this or also rely on storage for alternating plurals. However, even if both groups of children store alternating plurals alongside their singular and rely on analogy to form new words, unimpaired children are expected to be better at extracting generalisations from these words, resulting in a better performance on both types of words.

For non-words, predictions differ. Under a rule-based account, SLI children are expected to perform worse on inflecting non-words (compared to words), since they do not have access to a morphological rule. Also, if SLI children rely on explicit rule learning (add –en to create a plural), no overgeneralisations of voicing (e.g. [slat] ~ [sladen]) are expected for non-words. If occurrences of voicing alternations in lexical gap environments ([de:p] ~ [de:bo:n]) are taken to reflect a phonological rule mechanism, they should not occur for this group of children, unless we want to say that they do have access to ‘early rules’. Moreover, for both groups of children, these rules should apply across-the-board (i.e. consistently within a phonological context).

Lastly, a rule-based account would predict that the performance of SLI children on words should show a greater effect of lexical frequency than that of the control group, since they store all existing words. On the other hand, usage-based models would predict that lexical frequency affects the performance of both groups of children. It is possible that TD children are even more sensitive to frequency.
2 Method
In section 2.1, the distribution and frequency of voicing alternations in Dutch will be presented, followed by the production experiment in section 2.2.

2.1 The distribution and frequency of voicing alternations.
To determine the distribution of voicing alternations in Dutch, we considered Dutch word forms in the CELEX lexical database (Baayen et al. 1995). For experimental purposes we restricted environments to words ending in either [p] or [t], and three different types of rhymes: a short vowel ([bet] ~ [beddən] beds); a long vowel ([potlɔt] ~ [potlɔdən] pencils) or a vowel followed by a nasal ([hʌnt] ~ [hʌndən] hands). Note that the alternation also occurs in the Dutch verbal and adjectival paradigm (e.g. [rɔt] ~ [rɔdə] red⁶), which is taken into account in the frequency counts shown in table 1. Here, the frequency of the alternating inflected form (e.g. ‘bedden’ beds) is given as a percentage of all inflected forms in a particular environment, both alternating and non-alternating (e.g. the sum of ‘bedden’ beds, ‘hadden’ had, ‘petten’ caps, ‘zitten’ sit etc.). Frequency counts based only on noun plurals are also shown.

<table>
<thead>
<tr>
<th>Type</th>
<th>Frequency</th>
<th>Token frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td>All (N, V, A)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Short vowel</td>
<td>6.2%</td>
<td>1.7%</td>
</tr>
<tr>
<td>Long vowel</td>
<td>0.0%</td>
<td>14.2%</td>
</tr>
<tr>
<td>Vowel+nasal</td>
<td>0.0%</td>
<td>6.4%</td>
</tr>
<tr>
<td>Nouns</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Short vowel</td>
<td>7.2%</td>
<td>1.5%</td>
</tr>
<tr>
<td>Long vowel</td>
<td>0.0%</td>
<td>8.8%</td>
</tr>
<tr>
<td>Vowel+nasal</td>
<td>0.0%</td>
<td>5.7%</td>
</tr>
</tbody>
</table>

Table 1: Voicing alternations in the six experimental environments.

Clearly, type frequency of alternating words is lower than that of non-alternating words (i.e. the pattern itself is not dominant). For the environments under consideration, the average type frequency is only 7%. Notably, the high token frequency of /b/ alternations after short vowels is caused by the plural of only one verb ([hɛp] ~ [heban] have). If only nouns are considered, the distribution is similar in terms of type frequency (average 6%) but not token frequency. As mentioned above, there are no /b/ alternations after long vowels and vowel-nasal sequences, creating a so-called (accidental) lexical gap environment. In these contexts, voicing only occurs in non-final position in mono-morphemic words (e.g. [ɔkto:br] October and [tɔmbə] tomb). For nouns, most /d/ alternations occur in the long vowel and vowel-nasal environment, although in the former many high frequent words are characterised by a vowel alternation or open syllable lengthening (e.g. [ʃat] ~ [stedən] cities, [pɑt] ~ [pa:dən] paths), where the mapping between the singular and the plural is even less transparent⁷.

The frequency data in table 1 will be used to investigate whether children are sensitive to the distribution of voicing alternations in the input.

---

⁶ After long vowels, /d/ can also alternate with a glide in spoken language, as in [roːt] ~ [roːdə] red (cf. Zonneveld, 1978), so frequency of occurrence might actually be lower.

⁷ There are only four nouns with a t–d alternation after a short vowel (bed). Historically, they had a closed first syllable (bedde) which is why they escaped open syllable lengthening (Zonneveld, p.c.).
2.2 Experiment

2.2.1 Subjects
Two groups of 5-year-olds participated in the experiment; a control group of 27 typically developing children with a mean age of 61 months (sd 3 months) and a group of 25 children diagnosed with SLI with a mean age of 62 months (sd 5 months). The groups were matched on chronological age. SLI children were recruited through schools that provide full-time specialised education programs from 3 to 12 years. These children had been classified as language-impaired after extensive multidisciplinary assessment of their verbal and non-verbal abilities. The control children were matched in terms of chronological age and were contacted via day-care centres in Utrecht. All children were monolingual native speakers of Dutch and had no known hearing deficits.

2.2.2 Items
Children were asked to form plurals of both words and non-words. To elicit plurals of non-words, a classic wug-test (Berko, 1958) was used. The stimuli set consisted of 12 non-words that were created to match the six environments described above, yielding two items in each environment (e.g. [ket] and [slat], [dap] and [xop]). In the same contexts, 16 high-frequency words were chosen, 8 of which were non-alternating (e.g. [pet] hat, [kip] chicken) and 8 of which were alternating (e.g. [bet] bed, [krap] crab). The most frequent words were selected in each category, to maximise the likelihood that children knew them. Frequency counts of the test words based on the CELEX lexical database are given in table 2, showing that alternating plurals are more frequent than the non-alternating plurals.

<table>
<thead>
<tr>
<th>Word type</th>
<th>Singular</th>
<th>Plural</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Non-alternating</td>
<td>189</td>
<td>181</td>
<td>370</td>
</tr>
<tr>
<td>Alternating</td>
<td>1086</td>
<td>451</td>
<td>1537</td>
</tr>
</tbody>
</table>

Table 2: Token frequency of all singular and plural test words (n =16).

Words and non-words were mixed and 5 fillers were added to the list. The task was administered in two different orders to control for effects of fatigue and task novelty.

2.2.3 Procedure
Children were tested in the language lab at Utrecht University or at their school. Testing took place in a secluded room. The task was fifth in a session that included other language tests as well as IQ measures. Plurals were elicited through presentation of pictures in a PowerPoint slide show. For existing words, the child saw a picture of the object and had to name it. A second picture of the same object then appeared and the child had to complete the sentence "Now there are two .... ". For non-words, the child saw a fantasy animal and its name was presented through a loudspeaker. The non-words had been pre-recorded to ensure adequate and consistent production (including an audible release). The child had to repeat the name (until it was repeated correctly) and was then presented with the second image and asked to form the plural in the same way. Data were recorded on DAT (Tascam DA-P1) through a sensitive microphone (Crown PZM-185).

---

8 Only two plurals were found to be more frequent than their singulars in the CELEX database: ‘voeten’ feet and ‘schapen’ sheep, both non-alternating.
2.2.4 Analysis

Children’s realisations of both the singulars and the plurals were converted into .wav files. Phonological transcription of all utterances and acoustic analysis of the children’s plosives were then made by the authors (independently), and a portion of the data was judged by an additional 3 raters, with agreement reaching 90%. The items that were classified as voiced or voiceless by 4 out of 5 raters were used for further acoustic analysis in the Praat programme (http://www.fon.hum.uva.nl/praat/). Closure duration (CD) and burst duration (BD) were taken as acoustic variables; the beginning of closure was measured by the change in formant structure and amplitude, the end was marked at the release burst or the characteristic periodicity and amplitude of the vocalic segment.

3 Results and discussion

The result section consists of the results of the acoustic analysis in 3.1 and results of the perceptual analysis (phonological transcription) in section 3.2.

3.1 Acoustic analysis

Acoustic measurements were performed to establish the realisation of the voicing contrast. In table 3, results are reported for a total of 142 stimuli, containing obstruents rated as voiced and voiceless for both words and non-words. There were 38 stimuli rated as voiceless for both groups and 34 rated as voiced, matched for items as closely as possible. Acoustic measurements revealed that overall, both groups make a reliable difference between voiced and voiceless obstruents, in both closure and burst duration.

<table>
<thead>
<tr>
<th></th>
<th>TD</th>
<th>SLI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Closure</td>
<td>voiceless</td>
<td>voiced</td>
</tr>
<tr>
<td>Duration</td>
<td>118 (35)</td>
<td>129 (37)</td>
</tr>
<tr>
<td>Burst</td>
<td>35 (16)</td>
<td>30 (12)</td>
</tr>
</tbody>
</table>

Table 3: Mean CD and BD in ms, standard deviations between brackets.

It can be observed from this table that SLI children make a slightly bigger difference in closure duration (129 vs. 118 ms), due to a higher duration for voiceless [t]. However, this result is not significant (one tailed p=0.107). The trend might suggest that SLI children resemble younger Dutch children (cf. Kuijpers 1993).

In conclusion, both the acoustic and the perceptual analysis support the claim that the two groups are able to produce a voicing contrast.

3.2 Perceptual analysis

All stimuli were transcribed by both authors to determine voicing and error type. If only onsets were changed or simplified, items were counted as correct. First, plural realisation and error types are considered, followed by an analysis of the results for words and non-words separately.

<table>
<thead>
<tr>
<th></th>
<th>TD</th>
<th>SLI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Words</td>
<td>19 (4.7%)</td>
<td>56 (15.8%)</td>
</tr>
<tr>
<td>Non-words</td>
<td>18 (5.6%)</td>
<td>60 (21.4%)</td>
</tr>
</tbody>
</table>

Table 4: Null marking for words and non-words.
Table 4 shows that children with SLI are poorer in plural marking, resulting in more singulars for both words and non-words. A one-way ANOVA with plural as dependent variable yields a significant effect of group (F(1)=70.122, p<0.000). There was no significant difference in plural marking between words and non-words for either group, even though a bigger difference was found for the SLI children (p=0.080). There was no effect of plosive (P,T) or rhyme (short vowel, long vowel or nasal+vowel) on realisation of the plural for either group.

SLI children did not only produce more errors, their errors were also qualitatively different. The SLI group constructed 72 cases of ‘other error types’, opposed to 15 of the control group. These errors consisted mainly of final cluster simplifications ([bemp] ~ [bemən] or [bepən]). Sometimes these could have been real word substitutions ([kimp] ~ [kipən] *chickens*). Also, place changes were more frequent for this group ([hant] *hand* ~ [hɑŋkə]). Phonological complexity thus affected SLI children more than controls.

3.2.1 Performance on words

Results for existing words show that the dominant error type is that of ‘paradigm uniformity’ or regularising the alternating plural (*[bətən] beds*). Both groups of children would also occasionally overgeneralise voicing in non-alternating plurals (*[pədən] caps*). In some cases –s plurals were given (see footnote 3). Other error types were diminutive formation, usually with correct –s plural ([bɛtʃəs] *little beds*) and vowel or consonant substitutions. The latter were only counted as errors if they changed the phonological category of the test item (i.e. the rhyme). Results are summarised in table 5 below.

<table>
<thead>
<tr>
<th></th>
<th>TD</th>
<th>SLI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Correct non-alternation</td>
<td>167 (41.4%)</td>
<td>116 (32.9%)</td>
</tr>
<tr>
<td>Correct alternation</td>
<td>107 (26.4%)</td>
<td>34 (9.5%)</td>
</tr>
<tr>
<td>Error (p,t): *be[t]en</td>
<td>95 (23.4%)</td>
<td>99 (27.9%)</td>
</tr>
<tr>
<td>Error (b,d): *pe[d]en</td>
<td>9 (2.2%)</td>
<td>6 (1.7%)</td>
</tr>
<tr>
<td>S-plural</td>
<td>3 (0.7%)</td>
<td>10 (2.8%)</td>
</tr>
<tr>
<td>Other errors</td>
<td>6 (1.5%)</td>
<td>34 (9.6%)</td>
</tr>
</tbody>
</table>

Table 5: Performance on words.

The data show that the SLI group inflects both words and non-words. They even produce regularisations (*[pɛtən]*) at the same rate as the control group. An ANOVA with number of correct responses to words as dependent variable establishes a main effect of group: the control group performed better on both alternating and non-alternating words (F(1)=32.7, p<0.000). There is also an effect of word type: alternating words were more difficult than non-alternating words (F(1)=74.4, p<0.000). No significant interaction was observed: alternating words were not more difficult for SLI children, although there seems to be a trend in this direction (p=0.060).

Even though alternating plurals used in the test were shown to be more frequent than non-alternating plurals (see table 2), they seem to be more difficult for children to acquire. The role of frequency becomes more apparent if we take into account performance on the different test words. We found correlations between performance on a word (% correct) and the frequency of its plural form in the CELEX database (correlation with the singular was less strong). This effect of frequency was due to the alternating plurals, as we found no correlation if only non-alternating plurals were
taken into consideration. This is not surprising if alternating plurals are stored in the lexicon. Interestingly, correlations with frequency of alternating plurals are much stronger for the control group (Spearman’s rho=0.874, p<0.005) than for SLI children (Spearman’s rho=0.667, p<0.071), suggesting that SLI children are less sensitive to lexical frequency.

3.2.2 Performance on non-words

Intervocalic obstruents in non-words were most often realised as voiceless by both groups (e.g. [slAt´n]), matching the value of the singular. However, voicing alternations (e.g. [slAd´n]) were also produced by both groups (see table 6).

<table>
<thead>
<tr>
<th>TD</th>
<th>SLI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Non-alternation (p,t)</td>
<td>271 (83.9%)</td>
</tr>
<tr>
<td>Alternation (b,d)</td>
<td>12 (3.7%)</td>
</tr>
<tr>
<td>S-plural</td>
<td>13 (4.0%)</td>
</tr>
<tr>
<td>Other errors</td>
<td>9 (2.8%)</td>
</tr>
</tbody>
</table>

Table 6: Performance on non-words.

The control group produced more non-alternations than the SLI group, as they inflected items more often (t(48)=4.108, p<0.000). Furthermore, they produced fewer voicing alternations (12 vs. 15), but this difference was not significant (p=0.467). Voiced responses were produced by 6 children from the control group (4% of all inflected responses) and 7 SLI children (9% of all inflected responses). The distribution of voicing alternations is reported in table 7 below. Shaded cells are responses in ‘lexical gap’ environments. Responses are also given as a percentage of all opportunities (i.e. including all other responses).

<table>
<thead>
<tr>
<th>TD</th>
<th>SLI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Short vowel: [dup], [slat]</td>
<td>0 (0.0%)</td>
</tr>
<tr>
<td>Long vowel: [dep], [klat]</td>
<td>0 (0.0%)</td>
</tr>
<tr>
<td>Vowel+nasal: [bemp], [flant]</td>
<td>1 (0.3%)</td>
</tr>
</tbody>
</table>

Table 7: Distribution of voicing alternations for non-words.

SLI children seem to produce more /b/ than /d/ alternations, but overall numbers are too low to draw any conclusions. It is striking however that SLI children seem less sensitive to the distribution of alternations in Dutch, producing more ‘lexical gap’ alternations, both after long vowels and after nasals. For both groups, the highest number of alternations is found for /d/ after nasals, which is the category with the highest token frequency (for alternating nouns) in Dutch.

There is also some evidence for a general strategy of paradigm uniformity, since the rate of alternations is even lower than that expected on the basis of type frequency. Also, some children never produce any voicing alternations. These test results suggest that the process itself is quite unproductive but may be extended in case of similarity to existing words.

3.2.3 Comparing words and non-words

Despite the absence of a difference in inflection rate between words and non-words, we were interested to see if any acoustic differences could be found between them,
and whether these are comparable across groups. A set of words and non-words was selected for each group (32 for TD children and 44 for SLI children) with an equal number of voiceless and voiced responses per category, matched for final obstruent, rhyme and child as closely as possible. The results (see table 8) show that for both groups of children, closure durations were longer for non-words than for words.

<table>
<thead>
<tr>
<th></th>
<th>TD</th>
<th>SLI</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>voiceless</td>
<td>voiced</td>
</tr>
<tr>
<td>Words</td>
<td>99 (20)</td>
<td>70 (20)</td>
</tr>
<tr>
<td>Non-words</td>
<td>97 (24)</td>
<td>86 (35)</td>
</tr>
</tbody>
</table>

Table 8: Closure duration in ms with standard deviations between brackets.

By assuming that (both regular and irregular) inflected words may be stored, this difference can be attributed to the added processing time connected with constructing plurals (for non-words) on-line. Such an effect was also reported by Walsh and Parker (1983), who found that English /s/ in word-final position is longer in acoustic duration if it is the plural (laps) than if it is mono-morphemic (lapse). Losiewicz (1992) discovered a similar result for morphemic /d/ or /t/ (rapped) versus non-morphemic /d/ or /t/ (rapt). He argues that this is due to the non-morphemic segment being part of a lexical representation, while the morphemic one is added to the stem in processing.

In this light it is interesting to note that the only effect of word vs. non-word that reached significance was that found for voiced responses in SLI children. When 11 voiced words (i.e. correct voicing alternations such as [hændən] hands) and 11 voiced non-words (e.g. [fländən]) were compared, closure durations in the plurals of non-words were on average 22 ms longer than closure durations in the plurals of words (F(1)=5.345, p<0.032). There is actually a similar but non-significant difference (16 ms) for the control group (this may be due to the fact that results were based on only 8 words per category). This pattern of results is perhaps not surprising if we assume that alternating plurals are stored, whereas non-alternating plurals are processed on-line. The effect of processing (e.g. to create [flándən] from [flant]) may be somewhat greater for SLI children.

Finally, TD children seem to process both non-alternating words (e.g. [o:lfænt] elephants) and voiceless non-words (e.g. [flænt]) in the same way (computing them on-line), whereas alternating words (e.g. [hændən] hands) are treated differently (showing an effect of storage). SLI children seem to rely on storage for regular, non-alternating words more than TD children, as evidenced by shorter CD's of non-alternating words versus voiceless non-words.

4 Summary and conclusions

We can conclude from these results that both groups of children are able to use morphological rules (or generalisations) to inflect existing and novel words. SLI children were found to be poorer at marking the plural, but no significant difference between words and non-words was found, contrary to what rule-based models would predict. Furthermore, the regularisation rate (*[betɛn] beds) of SLI children was similar to that of the control group.

For both groups of children, alternating words ('bedden', beds) were shown to be more difficult than non-alternating words ('petten', caps), even though they are more frequent. This means that both groups of children were able to analyse these plurals in
comparison to their singular. Thus, it does not seem to be the case that SLI children store inflected words as unanalysed wholes, since they did not perform equally well or better on the alternating (irregular) forms.

Worse performance on alternating words does not seem to be due to an articulatory difficulty of producing voiced stops, as acoustic measurements established that both groups are able to make a reliable voicing contrast, matching the perceptual analysis. Items that were left out of analysis due to interrater disagreement are potentially relevant, as it could be the case that SLI children are less consistent in producing the contrast. However, we are confident that most of the disagreement was actually due to the quality of the sound files used.

We also found a correlation between performance on words and lexical frequency, with typically developing children showing a higher sensitivity to the frequency of existing (alternating) plurals in Dutch. This result is not expected if TD children treat all words (both alternating and non-alternating) as regular. For SLI children, the effect of processing (vs. storage) might have been higher than for typically developing children, as evidenced by longer acoustic durations for non-words ([flændən]) than for words ([hændən] hands). Also, they may rely more on storage for both types of words.

The results further show that both groups of children produced voicing alternations for non-words, which is not expected if SLI children only rely on explicit rule learning (add –en). Most alternations were produced in the environment where token frequency is high (e.g. [flændən]). Again, SLI children were found to be less sensitive to the distribution of alternations in Dutch, producing more voicing alternations in 'lexical gap' environments (e.g. [de:bən]). Rather than explaining the behaviour of (younger) typically developing children and SLI children as qualitatively different, we propose that these unexpected alternations are not due to early rules of intervocalic or postnasal voicing. Instead, they are caused by the fact that knowledge of the distribution of alternating patterns is not fully learned yet. Also, it remains to be seen whether the domain of generalisation can be so strictly defined as to exclude these lexical gap alternations. Alternatively, one would have to claim that both typically developing children and SLI children are able to form these early rules. However, if rules are taken to be categorical (instead of probabilistic), children would be expected to behave more consistently within environments (i.e. it was not the case that children who produced [de:bən] also produced [bo:bən]).

Taken together, these results lead to the conclusion that alternating plurals in Dutch are likely to be stored, and not derived by rule. Their low productivity and high frequency lends further support to this claim. It is also possible that knowledge of the alternation may be partly based on orthography (Dinnsen and Charles-Luce 1984; Warner et al. 2004).

To conclude, it seems that we need to reformulate the notion of a linguistic rule to account for these data, taking the form of a frequency-based generalisation. Crucially, in order to use analogy in the first place, all children would have to be able to analyse (stored) inflected forms. SLI children are then poorer at forming associations among stored words and extracting generalisations from them. It is not necessary to assume a dual mechanism model of language relying on symbolic rules.

\[^9\text{In a follow-up experiment, we found that even adults occasionally produce 'lexical gap' alternations after long vowels (Kerkhoff, forthcoming). However, Zonneveld (1978) notes that this is not a true lexical gap context since adjectives like 'xenofobe' xenophobe exist in the adult lexicon.}\]

\[^{10}\text{If rules are allowed to be probabilistic, the distinction between rules and analogy may turn out to be meaningless (see Albright and Hayes 2003).}\]
to account for the data described here. Usage-based (Bybee 1985) or parallel dual route models (Baayen et al. 1997) seem more suitable for describing the behaviour of both typically developing and SLI children, although more work is needed to investigate the precise nature of analogical processes.

References


