The effects of phonological skills and vocabulary on morphophonological processing

Boersma, T.; Baker, A.; Rispens, J.; Weerman, F.

DOI
10.1177/0142723717725430

Publication date
2018

Document Version
Final published version

Published in
First Language

Article

FIRST LANGUAGE

The effects of phonological skills and vocabulary on morphophonological processing

First Language
2018, Vol. 38(2) 147–174

DOI: 10.1177/0142723717725430
journals.sagepub.com/home/fla
SAGE

Tiffany Boersma
University of Amsterdam, The Netherlands

Anne Baker
University of Amsterdam, The Netherlands; Stellenbosch University, South Africa

Judith Rispens
University of Amsterdam, The Netherlands

Fred Weerman
University of Amsterdam, The Netherlands

Abstract

Morphophonological processing involves the phonological analysis of morphemes. Item-specific phonological characteristics have been shown to influence morphophonological skills in children. This study investigates the relative contributions of broad phonological skills and vocabulary to production and judgement accuracies of the Dutch past tense and diminutive, two morphophonological processes. Typically developing children (age 5;0–10;0, *N* = 114) were asked to produce and judge real and nonce diminutives and regular past tenses. Phonological processing skills were measured using a phonological awareness, digit span and nonword repetition task; vocabulary using the PPVT. Phonological skills and vocabulary contributed significantly to the production and judgement of the past tense and diminutive. The results underline the relation between phonological skills and the lexicon and the processing of morphophonology. These findings go further than showing the importance of the item-specific phonological context of the stem and suffix: they indicate that more general skills in the domain of phonology and vocabulary are involved.

Corresponding author:
Tiffany Boersma, Amsterdam Centre for Language and Cognition (ACLC), University of Amsterdam, Spuistraat 134, 1012 VB Amsterdam, The Netherlands.
Email: t.a.boersma@uva.nl

Keywords
Language acquisition, morphophonology, phonological processing skills, speech perception and production, vocabulary

Introduction

The interaction between morphosyntax and phonology (morphophonology) gives rise to many alternating forms. The acquisition of morphophonology has been investigated in different morphological paradigms and across different languages (Demuth & Tomas, 2016; Kerkhoff, 2007; Kernan & Blount, 1966; Royle & Stine, 2013; Tomas, Demuth, Smith-Lock, & Petocz, 2015; Tomas, van de Vijver, Demuth, & Petocz, 2017). Morphophonology follows the phonotactic restrictions of the language by changes in the surface phonological form of the stem or suffix. It is acknowledged as being one of the most complex aspects of grammar to acquire since it reflects not only higher levels of phonological structures but also the interaction between phonology and morphology (Buckler & Fikkert, 2016; Pierrehumbert, 2003; Zamuner, Kerkhoff, & Fikkert, 2012). The challenge for a child is to learn which variant forms are related and which phonological representations must be stored in the lexicon (Buckler & Fikkert, 2016; Fikkert & Freitas, 2006). The current study investigates morphophonological production and processing (Dutch past tense and diminutive) across 5- to 10-year-old typically developing (TD) children. The overall goal of this study is to form a better understanding of which factors contribute to the processing of morphophonology. Specifically, this study assessed whether phonological processing skills and vocabulary size are important contributing factors in morphophonological processing. The next section expands on the idea as to why these skills might be important.

Morpho(phono)logical development: Influences of phonology and vocabulary

Previous literature has shown that both phonological processing skills and vocabulary size contribute to the rate and manner of the acquisition and processing of morphophonological patterns (Marchman & Bates, 1994; Marshall & Van der Lely, 2006; Matthews & Theakston, 2006; Song, Sundara, & Demuth, 2009). A growing body of literature has established the importance of phonology in morphological development (Marshall & Van der Lely, 2006; Song et al., 2009; Tomas et al., 2015). Item-specific phonological characteristics significantly affect early production of inflections, with some phonological contexts being more challenging than others (Demuth & Tomas, 2016; Marshall & Van der Lely, 2006; Tomas et al., 2015). For example, Song et al. (2009) showed in their study of spontaneous speech of six English-speaking children between the ages of 1;3 and 3;6 that the item-specific variability in the production of third person singular *–s* could be accounted for by the phonological complexity of the coda. In another study, Jarmulowicz (2006) found that English-speaking children (age 6;6–10;6) learned and produced derived words with neutral suffixes that do not change stem stress (*–ment*, *–ful* and *–ness*) better than derived words with rhythmic suffixes that alter stem stress (*–tion*, *–ity* and *–ic*). Stress-changing suffixes that were phonologically consistent with their stems were thus easier to produce than words that exhibited phonological changes (Jarmulowicz, Hay, Taran, & Ethington, 2008; Jarmulowicz & Hay, 2009; Windsor, 1994). Studies on the English past tense, which has three allomorphs /Id/, /d/ and /t/, have demonstrated that the phonological composition of the verb stem influences past tense production (Blom, Paradis, & Duncan, 2012; Marchman, 1997). English-speaking children used the past tense inflection less when verb stems ended in an alveolar consonant, which requires the /Id/ allomorph, but used it more when stems ended in a vowel or a liquid. Past tense production thus seems to be dependent on an accurate phonological analysis of the individual verb stems (Blom & Paradis, 2013; Marchman, 1997; Oetting & Horohov, 1997; Rispens & De Bree, 2014). These studies thus show that allomorph production has a phonological component. The locus of some of the variability in allomorph production might therefore be due to limitations in children's phonological representations or phonological competence (Song et al., 2009).

The present study will expand on these findings. Specifically, we address the idea that children need good phonological processing skills to accurately process the morphophonological properties of the Dutch past tense and diminutive. As such, we moved outside the scope of the morpheme level and investigated whether broader phonological processing skills, i.e. nonword repetition (NWR), digit span and phonological awareness, are associated with both the production and perception (judgement) of these morphophonological phenomena. There is limited evidence on the relationship between phonological short-term memory and processing, and morphophonology. A study by Archibald, Joanisse, and Shepherd (2008) found that performance on an NWR task significantly correlated with regular past tense production (both real and nonce verbs) in a sample of monolingual TD children (age 6;9–11;1). Empirical evidence on associations between other phonological skills such as phonological awareness (potentially implied in the phonological analysis of morphemes) in TD children is limited. The present study fills this gap.

In addition to the influence of phonology, several studies have demonstrated strong relationships between lexical acquisition and morphosyntactic development (Blom & Paradis, 2013; Kidd & Kirjavainen, 2011; Marchman & Bates, 1994; Marchman, Saccuman, & Wulfeck, 2004). It has been suggested that an increase in vocabulary size allows children to abstract general patterns necessary for the productive usage of morphological rules (Blom & Paradis, 2013; Marchman & Bates, 1994). Song et al. (2009) found strong correlations between children's overall morphological/lexical development and their production of the third person singular morpheme *–s*. Older children with small vocabularies showed poorer grammatical performance, while children with larger vocabularies produced the third person singular morpheme more accurately (Song et al., 2009). Blom and Paradis (2013) found a relation between the past tense use of children with language impairment and their vocabulary size, indicating that deficits in building a lexicon affected past tense use. These and other previous studies show that vocabulary size is related to the production of morphosyntax. This study will expand on these findings by examining to what extent receptive vocabulary size is associated with both the production and perception of morphophonological processes.

What should be kept in mind is that almost all of the studies mentioned above tested production skills only, but not sensitivity to morphophonology using judgement tasks. No systematic research has investigated the question of whether children are able to detect inappropriate occurrences of allomorphs. Production data reflect children's ability to apply their knowledge, but not necessarily their processing abilities. In contrast, perception data reflect passive knowledge taking into account that children might know more than they produce (Buckler & Fikkert, 2016; Swingley & Aslin, 2000; Tomas et al., 2015). It may be the case that production is less accurate than perception due to phonological output complexities. Examining both the production and perception of allomorphs will provide a more complete perspective on morphophonological processing. It will allow us to examine whether phonological and vocabulary skills impact differently on morphophonological processing. Since perception skills generally precede production, both younger and older children were tested in this study. It was expected that the younger children would perform worse than the older children especially on the production task and to a lesser extent on perception.

In the previously mentioned studies relatively little attention has been paid to morphosyntactic processes with more than two or three allomorphs. In the current study we investigated the Dutch diminutive which has five allomorphs based on different phonological properties of the noun stem. The results provide insight into whether the previous findings on, for example, the English past tense, third person singular *–s* and Dutch plural and past tense are generalizable over other morphophonological phenomena (Kerkhoff, 2007; Matthews & Theakston, 2006; Rispens & De Bree, 2014; Tomas et al., 2015).

Dutch past tense and diminutive

The Dutch past tense consists of two allomorphs: The /də/ allomorph is attached to stems ending in a voiced consonant or vowel, e.g. *ren-de* (to run), *aai-de* (to pet), while the /tə/ allomorph is attached to stems ending in a voiceless consonant, e.g. *bak-te* (to bake), *knip-te* (to cut).

The Dutch diminutive is somewhat more complex than the past tense as it consists of five allomorphs: /jə/, /tjə/, /pjə/, /kjə/, /ətjə/, which are attached to noun stems based on their phonological properties. The final consonant of the stem determines the allomorph based on place assimilation between this consonant and the first consonant of the allomorph: stem ending in a labial *+pjə, raam-pje* (window), stem ending in an alveolar *+tjə, boon-tje* (bean), stem ending in an obstruent *+jə, huis-je* (house) and stem ending in a velar *+kjə, koning-kje* (king). For stems ending in a velar consonant, however, the number of syllables (one vs more than one) is also relevant; compare *koning-kje* (king) to *ring-etje* (ring). For stems ending in a labial or alveolar consonant, on the other hand, the length of the vowel is also relevant; compare *boom-pje* (tree) with a long vowel to *bom-etje* (bomb) with a short vowel, similarly *zoon-tje* (son) versus *zon-etje* (sun).

The current study

The purpose of the current study was to investigate the relationship between phonological processing skills, vocabulary size and morphophonological processing in Dutch TD children between the ages of 5;0 and 10;0 in both a production and judgement task with both real and nonce items. Because we were also interested in these relationships during development, we looked at developmental growth of accurately judging and producing the Dutch past tense and diminutive. It was important to test the use of the diminutive and past tense with real nouns/verbs as well as nonce nouns/verbs. The nonce forms typically follow the same structural patterns of real regular verbs and nouns but by definition have no lexical representation stored in long-term memory that can be drawn upon. The children were thus required to analyse the stem and make use of their knowledge of the morphophonological patterns to apply the appropriate allomorph.

We expected that in general the older the children, the better their performance. However, a study with adults using the same tasks found that for the diminutive allomorph /ətjə/ they were not 100% accurate when they had to use it with a nonce noun (Boersma, Baker, Rispens, & Weerman, submitted). That is, when asked to inflect and rate nonce nouns that should receive this allomorph according to the established linguistic description, they scored around 87% correct on the judgement task and only 43% on the production task. These findings indicate that /ətjə/ has a different status compared to the other allomorphs. As such, developmental growth in children can be expected for the four other allomorphs, but not for /ətjə/. For this allomorph it is not clear what can be expected. Also, as mentioned in the previous section, greater developmental growth was expected for the production task as the youngest children were expected to perform better on the judgement task due to it testing passive knowledge.

Based on the literature review above, it was predicted that both phonological processing skills and vocabulary size would be related to children's past tense and diminutive production and judgement. Since nonce verbs and nouns have no representation in the lexicon, it was hypothesized that vocabulary size would contribute more to the real verbs and nouns than to the nonce verbs and nouns. Furthermore, we expected that phonological processing skills would be associated with the production and judgement of both the real and nonce nouns and verbs.

Methodology

Participants

The TD children (*N* = 125) were between the ages of 5;0 and 10;0, native speakers of Dutch and raised in monolingual families. This broad age range was chosen to obtain insight into development. Eleven children had to be excluded. Two children had very low scores on all tasks and/or missing data on too many tasks. Two other children scored extremely low on the NWR task (administered as part of our test battery, see below) indicating a possibility of language impairment. Four children were raised in bilingual families and therefore had to be excluded (this was reported by the teacher after completion of the tasks). Finally, three children (aged between 7;5 and 8;4) scored 1 SD below the mean on the reading tasks (see Table 2 for an overview of background measures). For four additional children birth dates were missing. To compensate for this, the average age in months of the age group they belonged to was calculated and used in the analyses. After exclusion a total of 114 children (50 female) remained; they were divided into four age groups. See Table 1 for mean ages per age group. The children were recruited at five primary schools situated in the northern part of the Netherlands and were tested at their schools during school time. The schools reported no problems such as hearing, sight or language difficulties that could affect the outcome of the study in any of the children.

Table 1. Mean age in months for each age group.

Age group	Age in months	
	Mean	SD
5;0–5;11 (n = 34)	66.41	3.02
6;0–6;11 (n = 28)	78.14	3.74
7;0–7;11 (n = 24)	90.04	3.99
8;0–10;0 (n = 25)	105.64	6.49

Table 2. Measures of non-verbal intelligence, receptive vocabulary, phonological processing and reading skills.

Raven (standard progressive matrices) (Raven, Raven, & Court, 2003)	Measure of non-verbal IQ Children were asked to complete visual patterns by selecting the correct missing piece from an array of six or eight possible pieces
Phonological awareness, CELF-IV (Semel, Wiig, & Secord, 2010)	Measure of phonological awareness Children were asked to rhyme and manipulate sounds and syllables
Digit span, CELF-IV (Semel et al., 2010)	Measure of phonological working memory Children were asked to repeat a series of numbers forwards and backwards. A composite score of the forward and backward task was used in this study
NWR (Rispens & Baker, 2012)	Measure of phonological processing skills Children had to repeat 40 nonsense words with varying difficulties (two to five syllables and phonotactically frequent and infrequent)
Peabody Picture Vocabulary Test (PPVT) (Schlichting, 2005)	Measure of receptive vocabulary Children were presented with four pictures and were asked to point at the picture that matched the word spoken to them by the experimenter
Real word reading test (RWT) (Brus & Voeten, 1979)	Measure of real word reading Children were asked to read as many real words as they could in one minute
Pseudo word reading test (PWT) (Van Den Bos, Spelberg, Scheepstra, & De Vries, 1994)	Measure of nonce word reading Children were asked to read as many nonce words as they could in two minutes
Letter recognition task	Measure of letter knowledge (only used with children from grade 2) Children were asked to name letters presented to them and point to the correct letter on a sheet (16 of the most common letters of the Dutch alphabet were used)

Test material

Children were tested on a battery of four standardized tasks testing their receptive vocabulary and phonological processing skills (predictor variables in the multilevel analyses).

The Raven's progressive matrices task (Raven, Raven, & Court, 2003) was used to control for non-verbal IQ performance at time of testing. The reading measures and letter recognition tasks for the younger children were used to exclude children with lower than average reading scores but who had not been diagnosed with developmental dyslexia (3 children). An experimental production task and a grammaticality judgement task were conducted to test morphophonological performance. Table 2 presents the tasks used to measure non-verbal intelligence, receptive vocabulary, phonological processing and reading skills of the children.

Test items. The same items were used in the judgement and production task (see Appendix B for an overview of the items used). To make sure the real items were as homogeneous as possible and known to the children, they were selected from a list of words which had previously been established as known to children aged 4–6 years (Damhuis, 1992). The lexical frequencies of the nouns lay between 3 and 15 instances per million. The verbs were divided into items with high (45–960 instances per million) and low (1–40 instances per million) lexical frequencies.[1] All items were controlled for phonotactic frequency, which was calculated by averaging the log value of the biphone transitional probabilities (range real items −0.87 to −1.55; range nonce items −0.97 to −1.55).[2] The nonce nouns were constructed with a pseudo word generator called Wuggy (Wuggy: Keuleers & Brysbaert, 2010). Wuggy constructs nonce words based on given input by the researcher, which for this study were the real words. However, the items were adapted so that they consisted of one or two syllables and did not contain any consonant clusters to make sure that even the youngest children could produce them. They also had to conform to the phonological and phonotactic rules of Dutch.[3]

The diminutive and past tense stimuli were presented in the same task semi-randomly to ensure the children did not focus too much on either one of these aspects. Analysis of the diminutive nonce and real nouns was done without the allomorph /ətjə/ for the production task. As mentioned earlier, a study of adult production and judgement of the diminutive using the same tasks indicated that this allomorph is not as productive as expected from earlier literature (Boersma et al., submitted). This was especially the case for the production task in which the participants often used one of the other four allomorphs, based on the final consonant of the noun stem, instead of /ətjə/. Consequently, for this study we decided to exclude /ətjə/ from the analyses for the production task. Performance on the judgement task with this allomorph approached ceiling levels in adults and it was therefore decided to retain this allomorph in the judgement task. In addition, due to the nature of the stimuli set up it would not be possible to calculate A' scores for this task (see analysis section) if /ətjə/ were omitted from the analysis.

Due to restrictions in time, the stimulus set was divided into two. The stimuli were semi-randomly assigned to one of the two sets. Each child was randomly assigned to one of the two sets. No significant differences between the two sets were found so that analysis of the data was done on the two sets taken together.

Judgement task. A grammaticality judgement task was used in which the children had to judge a diminutive or past tense form presented auditorily via the computer. The items were embedded in a sentence and accompanied by pictures to facilitate comprehension

Figure 1. Example diminutive judgement task. Example is with a nonce noun.

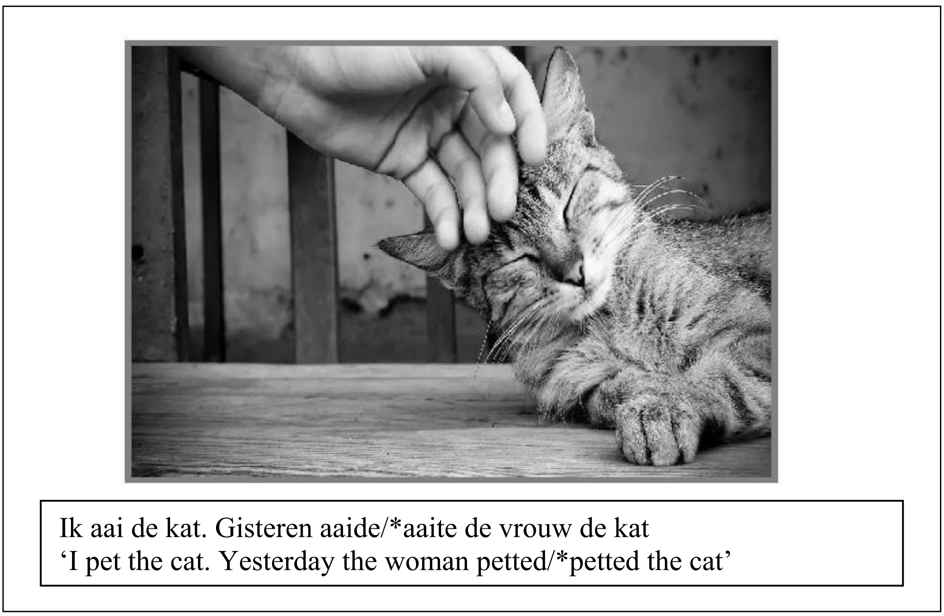

Figure 2. Example past tense judgement task. Example is with a real verb.

of the nonce nouns and verbs. Figure 1 shows a diminutive test item (using a nonce item) and Figure 2 a past tense test item (using a real item).

All nouns and verbs were presented twice: once with the target allomorph (expected according to the linguistic descriptions of the Dutch diminutive and past tense) and once with the non-target allomorph (unexpected according to the linguistic description). The non-target allomorphs violated one morphophonological pattern. To illustrate for the diminutive: in the case of the noun *koek* (cookie) the target allomorph is /jə/ due to the final consonant being an obstruent. The non-target allomorph chosen was /tjə/, **koektje*, which violates the obstruent pattern, but not for example the vowel length pattern. Other non-target stem+allomorph combinations could be **koeketje*, and **koekpje* but these violate two patterns or are phonotactically unlikely in Dutch.

The Dutch past tense has only two allomorphs and for the non-target stem+allomorph combinations they were interchanged with each other. For example, the stem of the verb *pakken* (to grab) is *pak*, which ends with a voiceless consonant and therefore receives the allomorph /tə/. The non-target form is then formed by attaching the voiced allomorph /də/ which makes **pakde*.

If the children decided the diminutive or past tense form was incorrect, they were asked to give their version of the correct answer. Each response was scored as correct (expected) or incorrect (unexpected) according to the existing linguistic descriptive rules.

Production task. A wug type task was used in which children had to provide the diminutive form of a noun and the past tense form of a verb. The stimuli were presented auditorily: first via the computer and a second time by the experimenter. Figure 3 shows a test item to elicit the diminutive. To elicit the past tense, the verb was introduced by a sentence in which the infinitive form of the verb was presented to the children twice. The child was then asked to tell the experimenter what happened yesterday by giving the past tense form of the verb. Figure 4 shows a test item used to elicit the past tense. The answer was scored as correct if the target stem+allomorph combination was produced. As the stimuli set was divided in two, each child was assigned to the same set as they had been assigned to for the judgement task.

Procedure

The study was reviewed and approved by the ethics review board of the Faculty of Humanities, University of Amsterdam.[4] Schools were approached and asked whether they wanted to participate in the study. Parents were then contacted and received an information letter with an active consent form to be signed if the parents and child were willing to participate.

Children were tested individually at their school during school time. Testing took three to five sessions of approximately 30 minutes. In the first session the children did the production task, in the second the first part of the judgement task and in the third the second part. In each session, the experimental tasks were always followed by two or three of the standardized tasks. In case not all tasks could be conducted in these three sessions, the children were asked to come back for a fourth or even a fifth session. A Sony Vaio and Dell laptop were used for the production, judgement and NWR tasks. An Olympus digital voice recorder was used to record all the sessions.

Scoring was done by two native speakers of Dutch who were trained linguists. Approximately 12% of the data from the production and NWR tasks were compared to calculate the interrater reliability. This was high for both the production (Cronbach's α = .82) and NWR (Cronbach's α = .96) tasks. Any discrepancies were solved by discussion until 100% consensus was reached.

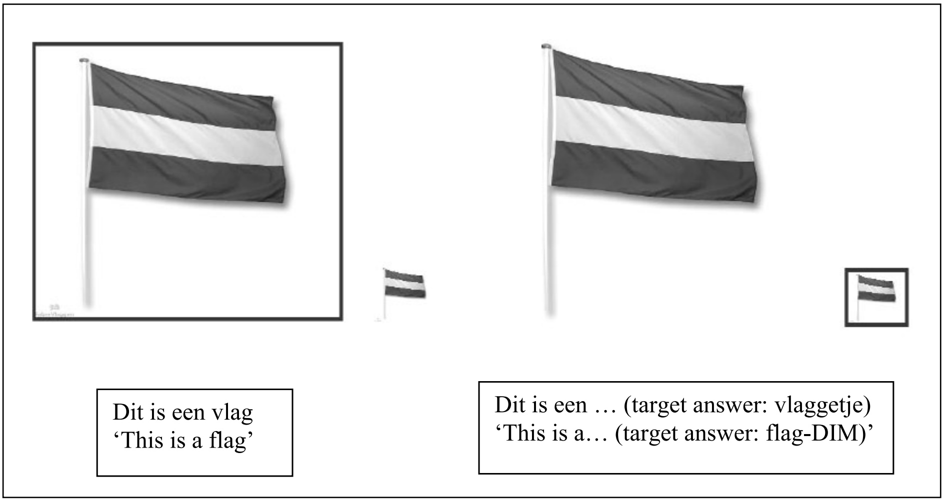

Figure 3. Example diminutive production task. The example is with a real noun.

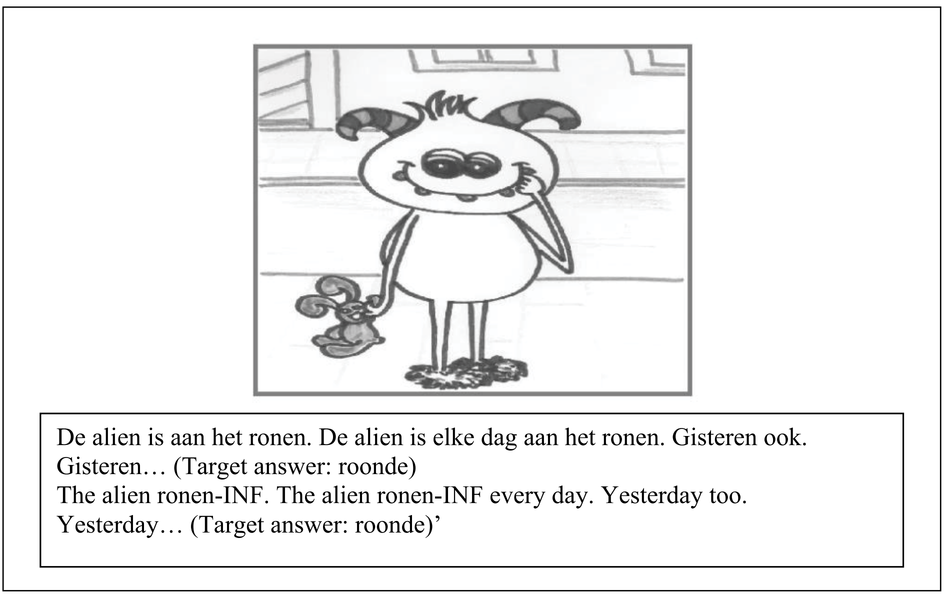

Figure 4. Example past tense production task. The example is with a nonce verb.

Analysis

ANOVAs were conducted to test for age-effects. Generalized linear mixed effects models were used to analyse the data further. These models are a type of regression model that control for participant and item variability by taking these as the random effect

Table 3. Formula to calculate A' scores.

A' = ½ + (hits – false alarms) * (1 + hits – false alarms) / (4*hits)*(1 – false alarms)

Note: A tendency to reject forms results in an A' score around 0, a tendency to accept in an A' score around .5 and good discrimination between grammatical and ungrammatical diminutives in an A' score around 1 (Rice, Wexler, & Redmond, 1999; Rispens, Roeleven, & Koster, 2004). Hits = correct judgements of target stem+allomorph combinations. False alarms = incorrect judgements of non-target stem+allomorph combinations.

factors (random intercept). Data were modelled in R (RStudio Team, 2015) using the lme4 package (Bates, Maechler, Bolker, & Walker, 2015). Separate models were built to predict production and judgement of the diminutive and past tense with real or nonce nouns and verbs from the four predictor variables taking age and IQ (Raven) into account, i.e. the dependent variables were diminutive and past tense real and nonce production and judgement, and the independent variables (fixed effect factors in the models) were performance on the phonological processing and PPVT tasks.[5] Random intercepts for participant, item and school the child attended were included to account for random by-participant, by-item and by-school variation in one model for the production task. Random intercept for item could not be included for the judgement task models as composite A' scores were calculated (see below). In the final mixed effects analysis, three children had to be excluded due to missing data in one of the tasks (2 NWR and 1 Raven. These children were included in the analyses testing for age-effects).

The dependent variable for the production task was whether the child gave a correct (target/expected) or incorrect (non-target/unexpected) response (binary data). In the judgement task the dependent variable was the A' score. An A' score gives an overall accuracy score in terms of accepting the target form of the items (hits) and rejecting the non-target items (false alarm: accepting where rejection is expected) (see Table 3) (Linebarger, Schwartz, & Saffran, 1983). The values can be interpreted as an answer to a two-alternative forced choice task. An A' value of .6 can thus be interpreted as a score of 60% correct if a participant had been asked to select which one of two forms is grammatical.

The raw scores of the predictor variables were zero centred to make sure the eigenvalue ratios of the models were not too large. Age was centred around 90 months (7;6) as this is the exact middle between the different ages tested in this study (5;0–10;0).

Forward selection was used to enter the predictor variables in the model as performance on the tasks testing IQ, receptive vocabulary size, phonological processing skills and also age in months were intercorrelated. To ensure that the predictor variables added to the prediction over and above age and IQ, both were always included in the models independently of whether they significantly added to the null and age models (Tabachnick & Fidell, 2013). As is common when using forward selection, predictor variables were added to the model based on the strength of their correlation with the dependent variable and excluded if they did not improve the model significantly. However, because the predictor variables were highly intercorrelated, it was also reported whether other predictor variables added to the age + Raven model significantly. Models were compared to each

Table 4. Mean raw and standardized scores for predictor variables for each age group.

Predictor variable		Age group			
		5;0–5;11	6;0–6;11	7;0–7;11	8;0–10;0
PPVT					
Raw	Mean	85.00	94.00	104.17	110.72
	SD	8.13	8.00	8.14	11.09
Standardized	Mean	111.50	108.64	109.43	105.82
	SD	9.64	10.70	8.73	13.48
Phonological awareness					
Raw	Mean	29.47	32.86	39.00	40.64
	SD	7.07	7.64	3.23	2.77
Standardized	Mean	11.59	11.07	11.26	9.68
	SD	1.59	1.71	2.05	2.72
NWR					
Raw	Mean	0.85	0.86	0.89	0.91
	SD	0.06	0.07	0.05	0.04
Digit span					
Raw	Mean	9.32	10.25	12.00	11.80
	SD	1.92	2.23	2.33	2.37
Standardized	Mean	12.38	11.32	11.13	9.59
	SD	2.30	2.59	2.21	2.59
Raven (IQ)					
Raw	Mean	22.35	23.50	28.42	32.88
	SD	4.33	6.93	5.75	5.96
Standardized	Mean	NA	59.04	54.47	49.74
	SD	NA	22.79	20.82	21.26

Note: The standardized scores indicate that all children scored within the norms of normal language development.

other using log-likelihood comparisons (using analysis of variance function in R). The model comparisons are presented in the results section.

Conditional R^2 values were calculated with the MuMIn package based on Nakagawa and Schielzeth (2013). The conditional R^2 was interpreted as variance explained by both the fixed and random effect factors.

Results

Descriptive statistics and age effects

The descriptive statistics for the standardized tasks (predictor variables) can be found in Table 4. It shows the mean raw and standardized scores on the non-verbal intelligence measure (Raven), vocabulary measure (PVVT) and phonological processing measures (phonological awareness, NWR and digit span composite score forward and backward). All children scored within the normal range on all tasks. In addition, as expected, the older the children were, the better they scored on the Raven (raw scores).

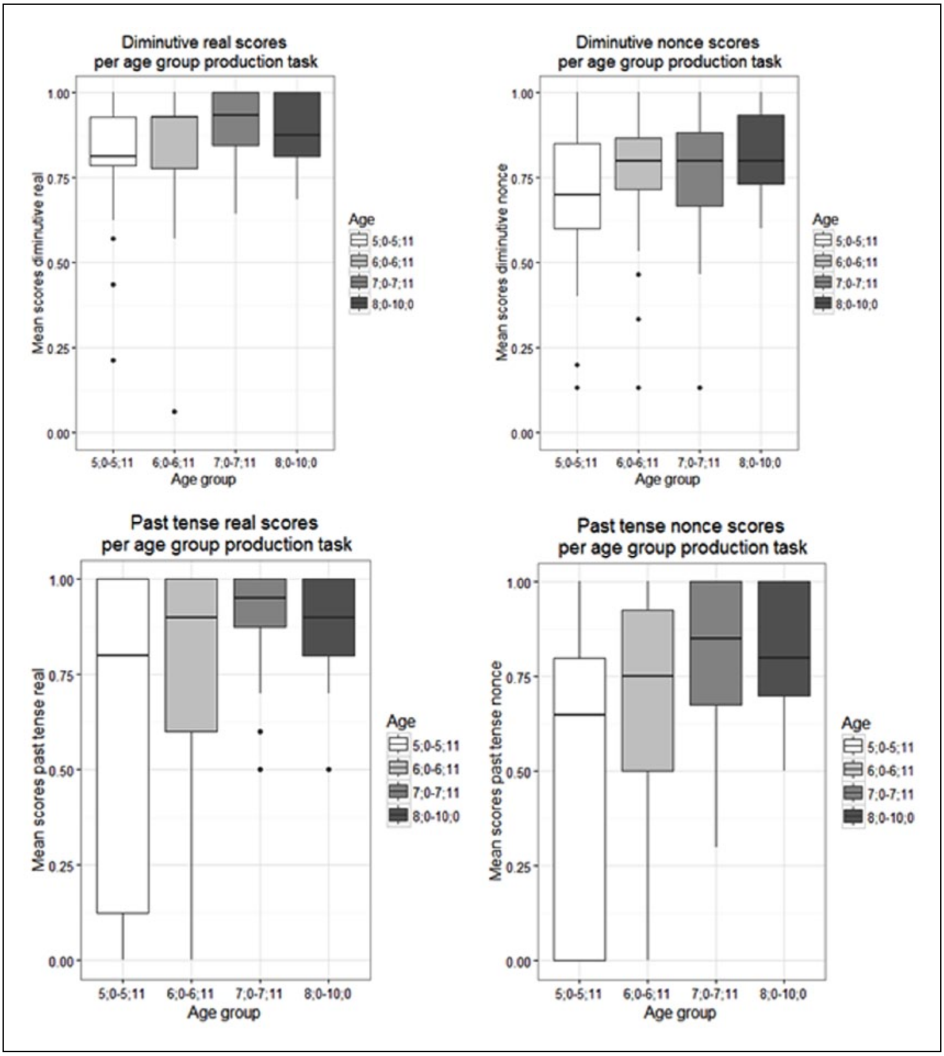

Figure 5. Boxplots production task diminutive and past tense real and nonce items.
Note: Significant differences were found for production of nonce diminutive items: 5- year-olds differed significantly from the 8-year-olds, $p < .05$; real and nonce past tense: 5-year-olds differed significantly from the 8-year-olds, $p < .001$.

The mean accuracy scores on the production task and the mean A' scores on the judgement task are presented in Figures 5 (production) and 6 (judgement). ANOVAs with Bonferroni post hoc tests were conducted to test whether the age groups were significantly different from each other. The significant results are presented below the figures. Overall, the results indicate a jump in performance between the 5- and 6-year-olds and the 7- and 8-year-olds on the production and judgement task with both the diminutive

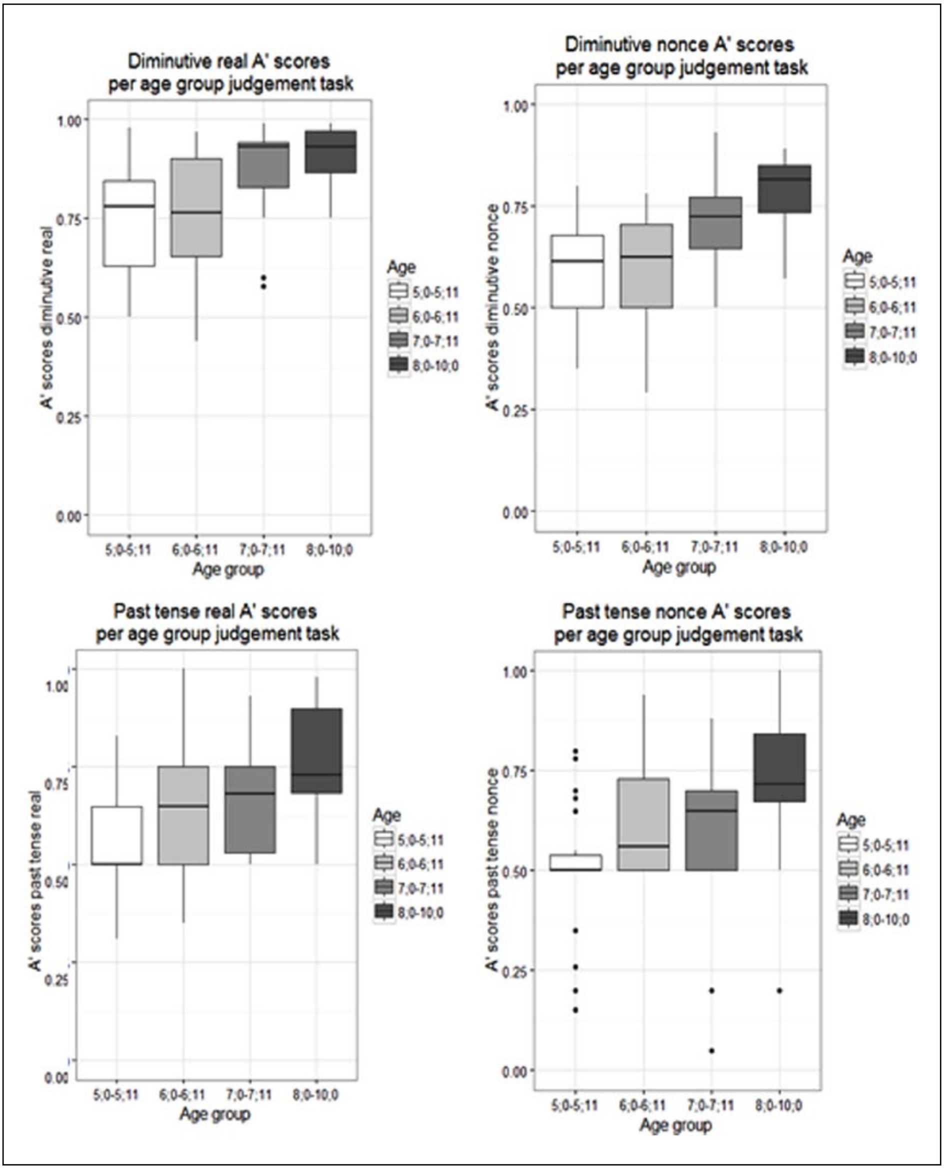

Figure 6. Boxplots judgement task diminutive and past tense real and nonce items.
Note: Significant differences were found for judgement of the real diminutive items: 5-year-olds differed significantly from the 7- and 8-year-olds, $p < .001$, 6-year-olds differed significantly from the 7-year-olds, $p < .05$, and 8-year-olds, $p < .001$; nonce diminutive items: 5- and 6-year-olds differed significantly from the 7-year-olds, $p < .05$ and 8-year-olds, $p < .001$; real past tense items: 5-year-olds differed significantly from the 6-year-olds, 7-year-olds, $p < .05$, and 8-year-olds, $p < .001$, and 6-year-olds differed significantly from the 8-year-olds, $p < .05$; nonce past tense items: 5-year-olds differed significantly from the 8-year-olds, $p < .001$.

Table 5. Pearson's correlations between the accuracy scores on the production task or A' scores on the judgement task and the predictor variables (raw scores).

	Diminutive real	Diminutive nonce	Past tense real	Past tense nonce
Production				
Age months	.233*	.271**	.344***	.392***
Raven	.236*	.259**	.351***	.295**
PPVT	.270**	.202*	.429***	.433***
Phonological awareness	.318***	.281**	.455***	.423***
NWR	.225*	.275**	.323**	.354***
Digit span total	.201*	.139	.379***	.397***
Judgement				
Age months	.486***	.541***	.582***	.449***
Raven	.477***	.485***	.451***	.351***
PPVT	.512***	.545***	.505***	.367***
Phonological awareness	.509***	.485***	.545***	.414***
NWR	.372***	.342***	.357***	.285**
Digit span total	.490***	.407***	.475***	.361***

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

and past tense. An interesting result is also that especially the 5-year-olds, and to a certain extent the 6-year-olds, had considerable difficulties with the past tense and the standard deviations are large, indicating a great deal of variation in past tense production.

The contribution of phonological processing skills and vocabulary

Table 5 shows the correlations between the production or judgement task and the predictor variables. All predictor variables were significantly correlated with diminutive and past tense production and judgement. Only performance on the digit span task was not correlated with the production of diminutive nonce nouns. As mentioned in the analysis section, the predictor variable with the highest correlation with the dependent variable was put in the model first (after age and IQ). It should be kept in mind that the predictor variables were all significantly intercorrelated and also correlated with age and IQ. This means that if two or more predictor variables overlap in how they explain the dependent variable in the linear mixed effect models, that overlap will not be reflected in either regression coefficient. As such, we report which model worked best, i.e. had the lowest *p*-value compared to the model without the predictor variable(s), but also report which models worked as well. The model comparisons of the linear mixed effects models are presented. Table 6 shows the model comparisons for the production task and Table 7 for the judgement task. The details of the final models can be found in Appendix A.

Production past tense. The null model for performance on the production tasks with real verbs was significantly improved after adding age. IQ significantly improved the age model. Accuracy scores correlated strongest with performance on the phonological awareness task. This task improved the age + Raven model significantly, $p = .002$.

Table 6. Model comparisons for performance on the production task with the real and nonce diminutive and past tense.

Production			
Model	Log-likelihood	χ^2	*p*
Past tense real			
Null + Age (df = 5)	–408.89	5.78	.016
+ Raven (df = 6)	–406.61	4.58	.032
+ Phonological awareness (df = 7)	–401.78	9.65	.002
+ PPVT (df = 8)	–399.40	4.762	.030
Past tense nonce			
Null + Age (df = 5)	–517.48	10.36	.001
+ Raven (df = 6)	–517.05	.85	.36
+ Digit span (df = 7)	–511.49	11.12	<.001
Diminutive real			
Null + Age (df = 5)	–508.93	4.30	.038
+ Raven (df = 6)	–508.35	1.16	.282
+ Phonological awareness (df = 7)	–505.90	4.91	.027
Diminutive nonce			
Null + Age (df = 5)	–770.10	7.529	.006
+ Raven (df = 6)	–768.89	2.408	.12
+ NWR (df = 7)	–766.25	5.284	.02

Note: Real diminutive model: conditional R^2 = .591; nonce diminutive model: conditional R^2 = .455; real past tense model: conditional R^2 = .745; nonce past tense model: conditional R^2 = .666.

Performance on the PPVT, the NWR and the digit span tasks also improved the age and Raven model significantly: respectively $p = .009$, $p = .02$ and $p = .03$, but the significance levels were higher compared to the phonological awareness task. Performance on the NWR ($p = .05$) and PPVT ($p = .02$) tasks improved the model with phonological awareness significantly. The PPVT improved the model with the lowest *p*-value. The NWR and the digit span tasks did not further improve the model with phonological awareness and PPVT.

The null model for performance on the production task with nonce verbs significantly improved after adding age, but the age model did not improve significantly after adding IQ. Performance on the PPVT tasks correlated the strongest with the accuracy scores and improved the age + IQ model significantly, $p = .04$, but performance on the phonological awareness task ($p = .005$), NWR task ($p = .007$) and the digit span task ($p < .001$) also improved this model. The model with the digit span task added was the best model. The other variables did not further improve this model.

Production diminutive. The null model for performance on the production task with real nouns significantly improved after adding age but the age model did not significantly improve after adding IQ. The phonological awareness task had the highest correlation with the accuracy scores. This was also the only predictor variable that significantly

Table 7. Model comparisons for performance on the judgement task with the real and nonce diminutive and past tense.

Judgement			
Model	Log-likelihood	χ^2	*p*
Past tense real			
Null + Age (df = 5)	25.32	14.66	<.001
+ Raven (df = 6)	28.24	5.84	.02
+ Digit span total (df = 7)	31.15	5.81	.02
Past tense nonce			
Null + Age (df = 5)	28.51	9.67	.002
+ Raven (df = 6)	29.38	1.73	.19
+ Digit span (df = 7)	31.71	4.66	.03
Diminutive real			
Null + Age (df = 5)	132.59	1.93	.17
+ Raven (df = 6)	137.11	7.10	.008
+ Digit span (df = 7)	142.26	10.31	.001
Diminutive nonce			
Null + Age (df = 5)	38.90	18.15	<.001
+ Raven (df = 6)	41.61	5.41	.02
+ PPVT (df = 7)	43.15	3.09	.08
+ Phonological awareness (df = 7)	43.29	3.37	.07

Note: Real diminutive model: conditional R^2 = .364; nonce diminutive model: conditional R^2 = .115; real past tense model: conditional R^2 = .200; nonce past tense model: conditional R^2 = .136.

improved the age + IQ model ($p = .027$). None of the other predictor variables significantly improved the model with phonological awareness.

The null model for performance on the production task with nonce nouns significantly improved after adding age, but the age model did not significantly improve after adding IQ. Again the phonological awareness tasks correlated strongest with the accuracy scores. However, it did not significantly improve the age + IQ model. Performance on the NWR task did significantly improve the age + IQ model ($p = .02$). None of the other predictor variables improved the age + IQ model nor the model with the NWR task included.

Judgement past tense. The null model for performance on the judgement task with real verbs significantly improved after adding age. The age model significantly improved after adding IQ. The A' scores correlated strongest with performance on the phonological awareness task, but this task did not significantly improve the age + IQ model ($p = .07$). Performance on the digit span task significantly improved the age + IQ model ($p = .02$). None of the other predictor variables significantly improved the age + IQ model nor the model with digit span included.

The null model for performance on the judgement task with nonce verbs significantly improved after adding age, but IQ did not significantly improve the age model. Again, performance on the phonological awareness task correlated strongest with the A' scores and this task significantly improved the age + IQ model ($p = .04$). However, performance on the digit span task also improved the age + IQ model with an even slightly lower p-value ($p = .03$). None of the other predictor variables further improved the digit span or phonological awareness models. The digit span and phonological awareness tasks also did not significantly improve the phonological awareness and digit span models respectively further.

Judgement diminutive. The null model for performance on the judgement task with real nouns did not significantly improve after adding age. However, IQ did significantly improve the age model. The PPVT task had the highest correlation with the A' scores and significantly improved the age + IQ model, $p = .02$. However, performance on the phonological awareness, NWR and digit span tasks also improved the age + IQ model: respectively $p = .007$, $p = .04$ and $p = .001$. Performance on the digit span task improved the age + IQ model with the lowest p-value. Adding the PPVT task improved the model with digit span near significance. None of the other predictor variables significantly improved the model with the digit span task included.

The null model for performance on the judgement task with nonce nouns significantly improved after adding age, and the age model significantly improved after adding IQ. Again performance on the PPVT task correlated the strongest with the A' scores. However, it did not significantly improve the age + IQ model ($p = .08$). None of the other predictor variables significantly improved the age + IQ model, although performance on the phonological awareness task approached significance ($p = .07$).

Discussion and conclusion

The present study was designed to increase our understanding of morphophonological processing by testing the Dutch diminutive and past tense in both a production and judgement task. More specifically the study examined to what extent broad phonological processing skills and vocabulary size contribute to the processing of the past tense and the diminutive.

A difference in performance between the younger and older children was expected and obtained. A (significant) jump in performance between the 6- and the 7-year-olds in both the past tense and diminutive was apparent. For the diminutive this finding confirms two earlier studies by Den Os and Harder (1987) and Peelaerts (2008). Both studies found that diminutive allomorphs are only fully acquired from the age of 7 onwards with some allomorphs being acquired at an even later age. In the present study even the oldest children, at age 10, had difficulties with appropriately inflecting nonce nouns with a diminutive marker.

Surprisingly, the 5-year-olds in particular scored relatively low and showed a great deal of variation in production of the past tense. Further error analyses will have to clarify whether the children in the present study had difficulty using the two allomorphs accurately or whether they simply had not yet acquired the concept of past tense to the extent that is required for the task. In Dutch, the simple past tense is less frequent in the input and in child usage at this age than the present perfect, which might have had an

Table 8. Results of the multilevel analysis showing the relationships between the tasks, structure and type of stimuli.

	Production		Judgement	
	Phonological processing	Receptive vocabulary	Phonological processing	Receptive vocabulary
Diminutive				
Real	Yes	No	Yes	No
Nonce	Yes	No	No	No
Past tense				
Real	Yes	Yes	Yes	No
Nonce	Yes	No	Yes	No

Note: Where the phonological processing and vocabulary tasks contributed to the accuracy scores, a Yes is indicated. Where this was not the case, a No is indicated. The findings that were hypothesized in the introduction are marked in grey.

effect. Although it was expected that the gap between the younger and older children would be larger in the production than in the judgement task, this was not necessarily found. For the past tense, the variation seems to be less, but the younger children are close to scoring at chance level. Also, the gap between the younger and older children seems to be even larger in the judgement task. We will expand below on why this might be the case.

Table 8 gives an overview of the results from the generalized linear mixed effect models. Based on previous studies it was hypothesized that phonological processing skills for the real and nonce items and vocabulary size for the real items would significantly contribute to the production and judgement of the past tense and diminutive.

As expected, performance on the phonological processing tasks was significantly associated with both the production and judgement of the past tense and diminutive. Interestingly, it differed per condition which phonological processing task made the largest contribution. For example, for production of the nonce nouns this was the NWR task, while for the nonce verbs this was the digit span task. Both are measures of phonological working memory. However, the NWR task taps more language-specific knowledge while the digit span task involves sequencing highly familiar items and is purely a phonological short-term memory task (Baddeley, 2003; Gathercole, Willis, Baddeley, & Emslie, 1994; Gathercole, 2006; Rispens & Baker, 2012; Rispens, Baker, & Duinmeijer, 2015). The NWR task might be more predictive for the diminutive, as selecting the appropriate allomorph is a phonologically more complex process than for the past tense, which has only two allomorphs and where the only phonological process that needs to be taken into account is voicing.

Interestingly, the digit span task made a significant contribution to performance on the judgement task with both the diminutive and past tense. The digit span task is, as mentioned above, a measure of how well children can temporarily store and manipulate verbal information, which closely resembles what the children had to do in the judgement task (Baddeley, 2003). It is also a good measure of attention and concentration span

(Semel, Wiig, & Secord, 2010), which might have been taxed to a greater extent in the judgement task (Astheimer, Janus, Moreno, & Bialystok, 2014).

The findings of this study underline the results of previous research in showing that both phonology and vocabulary are important contributing factors in the development of morphophonological processes. The study by Archibald et al. (2008) already found correlations between the NWR task and English past tense production. This article goes beyond this finding by showing that phonological processing skills in general contribute to performance on both processing and producing morphophonological patterns. Moreover, it supports, to a certain extent, the studies of, for example, Blom and Paradis (2013) and Marchman and Bates (1994), as vocabulary size was associated with past tense production with real verbs. However, although performance of the PPVT task correlated strongly with both production of the diminutive and judgement of the diminutive and past tense, it did not occur in the final most explanatory models. A possible explanation for this finding could be that past tense formation is explicitly learned at primary school where children learn how to spell the past tense. They learn the difference between the voiced and voiceless allomorphs explicitly, which might then be explicitly linked to the verbs in their mental lexicon (De Bree, Van der Ven, & Van der Maas, 2016). This difference between explicit learning (past tense) and implicit learning (diminutive) might have had an influence on children's performance. Also, the past tense is a morphosyntactic inflection, while forming the diminutive is a derivational process more based on phonological derivations of the stem. Choosing the appropriate allomorph when forming the diminutive is phonologically complex. Children, therefore, might have to rely more heavily on their knowledge of Dutch phonotactics when producing the diminutive.[6]

Although it is not possible to directly compare the findings from the production and judgement tasks as they involve different statistical analyses, we would like to speculate tentatively about the differences found between the two tasks. These differences might to a large degree be attributed to the extent that general metalinguistic competence is implied in a grammaticality judgement task (Rice et al., 1999). The metalinguistic demands might have been too challenging for some children, which means that their performance did not reflect their actual knowledge, i.e. they might have known what the appropriate and inappropriate diminutive and past tense forms were, but accepting or rejecting these and having to come up with the (in their view) correct form was too demanding. Children in general have a bias to say 'yes' in their responses and this might have been what especially the younger children resorted to (Rice et al., 1999). For example, the finding that none of the predictor variables made a significant contribution over and above age and IQ for judgement of the nonce diminutive may be attributed to the fact that the children benefited more from their increased age, non-verbal intelligence and metalinguistic awareness.[7]

In sum, in accordance with the hypotheses formulated on the basis of previous literature, this study found that phonological processing skills and receptive vocabulary significantly contribute to the processing of morphophonology, even though there are differences between the past tense and diminutive and between the production and judgement tasks. This implies a relation between phonological processing skills, the lexicon and the processing of morphophonology for both production and perception. Moreover, the results extend the findings of previous studies as they go beyond the importance of the item-specific phonological context of the stem and suffix and show that more general skills in the domain of phonology and vocabulary are involved.

Acknowledgements

We thank all the schools, children and parents who have participated in this study. We would also like to thank the research assistants of this project for their help with testing the children and the reviewers for their valuable comments on earlier versions of this article.

Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) received no financial support for the research, authorship, and/or publication of this article.

Notes

1. Taken from the SUBTLEX-NL database (Keuleers, Brysbaert, & New, 2010).
2. Phonotactic frequencies were calculated based on the Dutch phonotactic frequency database (Adriaans, 2006) derived from the corpus of spoken Dutch (Oostdijk, 2000).
3. The nonce items were tested with a word likeness task to make sure the items were considered by adult speakers as possible in the Dutch language
4. Dossier 2014-27.
5. In total eight different models were tested. Real and nonce nouns and verbs in the production and judgement task were thus tested in separate models.
6. Note that when doing the analysis including the diminutive /ətjə/, the PPVT is the only significant predictor, which indicates that children seem to store the stem + /ətjə/ combinations in their lexicon. This corresponds with findings in the adult study where the participants scored at ceiling on the real nouns but below 50% on the nonce nouns, indicating that adults stored the real stem + /ətjə/ combinations in their lexicon.
7. Although note that the models with either the phonological awareness or PPVT task included almost reached significance compared to the model with only age and the Raven included.

References

Adriaans, F. (2006). *PhonotacTools* (test version ed.). Utrecht, The Netherlands: Utrecht Institute of Linguistics OTS.

Archibald, L. M., Joanisse, M. F., & Shepherd, M. (2008). Associations between key language-related measures in typically developing school-age children. *Zeitschrift Für Psychologie/Journal of Psychology*, *216*, 161–171.

Astheimer, L., Janus, M., Moreno, S., & Bialystok, E. (2014). Electrophysiological measures of attention during speech perception predict meta-linguistic skills in children. *Developmental Cognitive Neuroscience*, *7*, 1–12.

Baddeley, A. (2003). Working memory and language: An overview. *Journal of Communication Disorders*, *36*, 189–208.

Bates, D., Maechler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, *67*(1), 1–48.

Blom, E., & Paradis, J. (2013). Past tense by English second language learners with and without language impairment. *Journal of Speech, Language and Hearing Sciences*, *56*, 281–294.

Blom, E., Paradis, J., & Duncan, T. S. (2012). Effects of input properties, vocabulary size, and L1 on the development of third person singular -s in child L2 English. *Language Learning*, *62*, 965–994.

Boersma, T., Baker, A., Rispens, J. E., & Weerman, F. (submitted). The Dutch diminutive allomorph /ətjə/: Description versus productivity. *Dutch Journal of Applied Linguistics*.

Brus, B. T., & Voeten, M. J. M. (1979). *EMT (one minute real word reading test)*. Amsterdam, The Netherlands: Pearson.

Buckler, H., & Fikkert, P. (2016). Using distributional statistics to acquire morphophonological alternations: Evidence from production and perception. *Frontiers in Psychology*, *7*: 540.

Damhuis, R. (1992). *Woordenlijst voor 4- tot 6-jarigen. Een streeflijst voor kleuters*. Rotterdam, The Netherlands: Projectbureau OVB.

De Bree, E., Van der Ven, S., & Van der Maas, H. (2016). The voice of Holland: Allograph production in written Dutch past tense inflection. *Language Learning and Development*, *13*, 215–240.

Demuth, K., & Tomas, E. (2016). Understanding the contributions of prosodic phonology to morphological development: Implications for children with specific language impairment. *First Language*, *36*, 265–278.

Den Os, E., & Harder, R. (1987). De verwerving van de regels voor meervouds- en verkleinwoordsvorming in het Nederlands [The acquisition of plural and diminutive rules in Dutch]. *De Nieuwe Taalgids*, *80*, 240–250.

Fikkert, P., & Freitas, M. J. (2006). Allophony and allomorphy cue phonological acquisition: Evidence from the European Portuguese vowel system. *Catalan Journal of Linguistics*, *5*, 83–108.

Gathercole, S. E. (2006). Nonword repetition and word learning: The nature of the relationship. *Applied Psycholinguistics*, *27*, 513–543.

Gathercole, S. E., Willis, C. S., Baddeley, A., & Emslie, H. (1994). The children's test of nonword repetition: A test of phonological working memory. *Memory*, *2*, 103–127.

Jarmulowicz, L. (2006). School-aged children's phonological production of derived English words. *Journal of Speech, Language and Hearing Research*, *49*, 294–308.

Jarmulowicz, L., & Hay, S. E. (2009). Derivational morphophonology: Exploring errors in third graders' production. *Language, Speech, and Hearing Services in Schools*, *40*, 299–311.

Jarmulowicz, L., Hay, S. E., Taran, V. L., & Ethington, C. A. (2008). Fitting derivational morphophonology into a developmental model of reading. *Reading Writing*, *21*, 275–297.

Kerkhoff, A. (2007). *Acquisition of morpho-phonology: The Dutch voicing alternation* (PhD thesis). University of Utrecht, The Netherlands.

Kernan, K. T., & Blount, B. G. (1966). The acquisition of Spanish grammar by Mexican children. *Anthropological Linguistics*, *8*(9), 1–14.

Keuleers, E., & Brysbaert, M. (2010). Wuggy: A multilingual pseudoword generator. *Behavior Research Methods*, *42*, 627–633.

Keuleers, E., Brysbaert, M., & New, B. (2010). SUBTLEX-NL: A new frequency measure for Dutch words based on film subtitles. *Behavior Research Methods*, *42*, 643–650.

Kidd, E., & Kirjavainen, M. (2011). Investigating the contribution of procedural and declarative memory to the acquisition of past tense morphology: Evidence from Finnish. *Language and Cognitive Processes*, *26*, 794–829.

Linebarger, M. C., Schwartz, M. F., & Saffran, E. M. (1983). Sensitivity to grammatical structure in so-called agrammatic aphasics. *Cognition*, *13*, 361–392.

Marchman, V. A. (1997). Children's productivity in the English past tense: The role of frequency, phonology and neighbourhood structure. *Cognitive Science*, *21*, 283–304.

Marchman, V. A., & Bates, E. (1994). Continuity in lexical and morphological development: A test of the critical mass hypothesis. *Journal of Child Language*, *21*, 339–366.

Marchman, V. A., Saccuman, C., & Wulfeck, B. (2004). Productive use of the English past tense in children with focal brain injury and specific language impairment. *Brain & Language*, *88*, 202–214.
Marshall, C. R., & Van der Lely, H. K. J. (2006). A challenge to current models of past tense inflection: The impact of phonotactics. *Cognition*, *100*, 302–320.
Matthews, D. E., & Theakston, A. L. (2006). Errors of omission in English-speaking children's production of plurals and the past tense: The effects of frequency, phonology and competition. *Cognitive Science*, *30*, 1027–1052.
Nakagawa, S., & Schielzeth, H. (2013). A general and simple method for obtaining R2 from generalized linear mixed-effects models. *Methods in Ecology and Evolution*, *4*, 133–142.
Oetting, J. B., & Horohov, J. E. (1997). Past-tense marking by children with and without specific language impairment. *Journal of Speech, Language and Hearing Research*, *40*, 62–74.
Oostdijk, N. (2000). The spoken Dutch corpus: Outline and first evaluation. In M. Gravilidou, G. Caravannis, S. Markantonatou, S. Piperidis, & G. Stainhaouer (Eds.), *Proceedings of the Second International Conference on Language Resources and Evaluation (LREC)* (pp. 887–894). Genoa, Italy.
Peelaerts, C. (2008). Van kleine boot tot scheepje: Experimenteel onderzoek naar de diminutiefvorming bij Vlaamse kinderen ['Kleine boot tot scheepje': Experimental research on diminutive production in Flemish children]. Unpublished manuscript.
Pierrehumbert, J. B. (2003). Phonetic diversity, statistical learning and acquisition of phonology. *Language and Speech*, *46*, 115–154.
Raven, J., Raven, J. C., & Court, J. H. (2003). *Manual for Raven's progressive matrices and vocabulary scales*. San Antonio, TX: Harcourt Assessment.
Rice, M. L., Wexler, K., & Redmond, S. M. (1999). Grammaticality judgements of an extended optional infinitive grammar: Evidence from English-speaking children with specific language impairment. *Journal of Speech, Language and Hearing Research*, *42*, 943–961.
Rispens, J. E., & Baker, A. (2012). Nonword repetition: The relative contributions of phonological short-term memory and phonological representations in children with language and reading impairment. *Journal of Speech, Language, and Hearing Research*, *55*, 683–694.
Rispens, J. E., Baker, A., & Duinmeijer, I. (2015). Word recognition and nonword repetition in children with language disorders: The effects of neighborhood density, lexical frequency and phonotactic probability. *Journal of Speech, Language and Hearing Research*, *58*, 78–92.
Rispens, J. E., & De Bree, E. H. (2014). Past tense productivity in Dutch children with and without SLI: The role of morphophonology and frequency. *Journal of Child Language*, *41*, 200–225.
Rispens, J. E., Roeleven, S., & Koster, C. (2004). Sensitivity to subject-verb agreement in spoken language in children with developmental dyslexia. *Journal of Neurolinguistics*, *17*, 333–347.
Royle, P., & Stine, I. (2013). The French noun phrase in preschool children with SLI: Morphosyntactic error analyses. *Journal of Child Language*, *40*, 945–970.
RStudio Team. (2015). *RStudio: Integrated development for R. RStudio*. Boston, MA: RStudio, Inc. Retrieved from http://www.rstudio.com/
Schlichting, L. (2005). *Peabody Picture Vocabulary Test-III-NL*. Amsterdam, The Netherlands: Harcourt Test Publishers.
Semel, E., Wiig, E. H., & Secord, W. A. (2010). *Clinical evaluation of language fundamentals* (3rd ed., W. Kort, E. Compaan, M. Schittekatte & P. Dekker, Trans.). Amsterdam, The Netherlands: Pearson Assessment and Information B.V. (Dutch version)
Song, J. Y., Sundara, M., & Demuth, K. (2009). Phonological constraints on children's production of English third person singular -s. *Journal of Speech, Language and Hearing Research*, *52*, 623–642.
Swingley, D., & Aslin, R. N. (2000). Spoken word recognition and lexical representation in very young children. *Cognition*, *76*, 147–166.
Tabachnick, B. G., & Fidell, L. S. (2013). *Using multivariate statistics* (6th ed.). Harlow, UK: Pearson.

Tomas, E., Demuth, K., Smith-Lock, K. M., & Petocz, P. (2015). Phonological and morphophonological effects on grammatical development in children with specific language impairment. *International Journal of Language and Communication Disorders*, *50*, 516–528.

Tomas, E., van de Vijver, R., Demuth, K., & Petocz, P. (2017). Acquisition of nominal morphophonological alternations in Russian. *First Language*, *37*, 453–474.

Van Den Bos, K. P., Spelberg, L. H. C., Scheepstra, A. J. M., & De Vries, J. R. (1994). *KLEPEL (two minute pseudo word reading task)*. Amsterdam, The Netherlands: Pearson.

Windsor, J. (1994). Children's comprehension and production of derivational suffixes. *Journal of Speech and Hearing Research*, *37*, 408–417.

Zamuner, T. S., Kerkhoff, A., & Fikkert, P. (2012). Phonotactics and morphophonology in early child language: Evidence from Dutch. *Applied Psycholinguistics*, *33*, 481–499.

Appendix A

Final most explanatory models linear mixed effects analysis.

Table A1. Optimal models predicting performance on the production task for the real and nonce past tense and diminutive.

Production				
Predictor variable	Estimate (β)	SE (β)	*Z*	*p*
Diminutive real				
Intercept	2.97	0.38	7.725	<.001
Age	0.01	0.016	0.703	.48
Raven	0.15	0.40	0.364	.72
Phonological awareness	0.94	0.41	2.255	.02
Diminutive nonce				
Intercept	1.65	0.28	5.958	<.001
Age	0.02	0.01	1.442	.15
Raven	0.33	0.29	1.137	.26
NWR	0.58	0.25	2.335	.02
Past tense real				
Intercept	2.00	0.39	5.037	<.001
Age	–0.03	0.03	–1.109	.27
Raven	0.45	0.75	0.591	.55
Phonological awareness	2.09	0.78	2.695	.007
PPVT	1.93	0.89	2.165	.03
Past tense nonce				
Intercept	–2.49	1.25	–2.00	.05
Age	0.05	0.02	2.53	.01
Raven	–0.04	0.59	–0.06	.95
Digit span	0.37	0.11	3.25	.001

Table A2. Optimal models predicting performance on the judgement task for the real and nonce past tense and diminutive.

Judgement				
Predictor variable	Estimate (β)	SE (β)	*t*	*p*
Real diminutive				
Intercept	0.64	0.06	10.81	<.001
Age	0.001	0.001	1.16	.25
Raven	0.05	0.03	1.80	.07
Digit span	0.02	0.005	3.24	.001
Nonce diminutive				
Intercept	0.69	0.01	51.66	<.001
Age	0.004	0.001	4.00	<.001
Raven	0.06	0.03	2.29	.002
Real past tense				
Intercept	0.48	0.08	6.38	<.001
Age	0.003	0.001	2.57	.01
Raven	0.06	0.04	1.67	.09
Digit span total	0.02	0.007	2.60	.009
Nonce past tense				
Intercept	0.627	0.018	34.25	<.001
Age	0.002	0.0013	1.84	.065
Raven	0.016	0.039	0.42	.68
Phonological awareness	0.091	0.045	2.01	.04

Appendix B

Test items.

Table B1. Real diminutive items.

Item	Target (correct) diminutive	Non-target (incorrect) form (judgement task)
Bol	Boll-**etje**	Bol-**tje**
Snor	Snorr-**etje**	Snor-**tje**
Vlam	Vlamm-**etje**	Vlam-**pje**
Schoen	Schoen-**tje**	Schoen-**etje**
Schaal	Schaal-**tje**	Schaal-**etje**
Ballon	Ballonn-**etje**	Ballon-**tje**
Klem	Klemm-**etje**	Klem-**pje**
Zaal	Zaal-**tje**	Zaal-**etje**
Boer	Boer-**tje**	Boer-**etje**
Teen	Teen-**tje**	Teen-**etje**
Worm	Worm-**pje**	Worm-**tje**
Zalm	Zalm-**pje**	Zalm-**tje**
Riem	Riem-**pje**	Riem-**etje**
Bezem	Bezem-**pje**	Bezem-**tje**
Pruim	Pruim-**pje**	Pruim-**etje**
Helm	Helm-**pje**	Helm-**tje**
Scherm	Scherm-**pje**	Scherm-**tje**
Kraam	Kraam-**pje**	Kraam-**etje**
Raam	Raam-**pje**	Raam-**etje**
Duim	Duim-**pje**	Duim-**etje**
Kring	Kring-**etje**	Krin-**kje**
Verwarming	Verwarmin-**kje**	Verwarming-**etje**
Leerling	Leerling-**etje**	Leerlin-**kje**
Woning	Wonin-**kje**	Woning-**etje**
Wandeling	Wandeling-**etje**	Wandelin-**kje**
Wang	Wang-**etje**	Wan-**kje**
Tekening	Tekening-**etje**	Tekenin-**kje**
Botsing	Botsin-**kje**	Botsing-**etje**
Pudding	Pudding-**kje**	Pudding-**etje**
Ontploffing	Ontploffin-**kje**	Ontploffing-**etje**
Wieg	Wieg-**je**	Wieg-**tje**
Vos	Vos-**je**	Vos-**tje**
Koek	Koek-**je**	Koek-**tje**
Schep	Schep-**je**	Schep-**tje**
Knoop	Knoop-**je**	Knoop-**tje**
Dwerg	Dwerg-**je**	Dwerg-**tje**
Stok	Stok-**je**	Stok-**tje**
Zeep	Zeep-**je**	Zeep-**tje**
Schaap	Schaap-**je**	Schaap-**tje**
Vaas	Vaas-**je**	Vaas-**tje**

Table B2. Nonce diminutive items.

Item	Target (correct) diminutive	Non-target (incorrect) form (judgement task)
Keen	Keen-**tje**	Keen-**etje**
Loen	Loen-**tje**	Loen-**etje**
Vom	Vomm-**etje**	Vom-**pje**
Gol	Goll-**etje**	Gol-**tje**
Raal	Raal-**tje**	Raal-**etje**
Poer	Poer-**tje**	Poer-**etje**
Neel	Neel-**tje**	Neel-**etje**
Zam	Zamm-**etje**	Zam-pje
Hil	Hill-**etje**	Hil-**tje**
Nan	Nann-**etje**	Nan-**tje**
Zuim	Zuim-**pje**	Zuim-**tje**
Liem	Liem-**pje**	Liem-**tje**
Reem	Reem-**pje**	Reem-**tje**
Zaam	Zaam-**pje**	Zaam-**tje**
Doom	Doom-**pje**	Doom-**tje**
Leem	Leem-**pje**	Leem-**etje**
Koem	Koem-**pje**	Koem-**etje**
Tuim	Tuim-**pje**	Tuim-**etje**
Buim	Buim-**pje**	Buim-**etje**
Gaam	Gaam-**pje**	Gaam-**etje**
Ming	Ming-**etje**	Min-**kje**
Tukking	Tukkin-**kje**	Tukking-**etje**
Loving	Lovin-**kje**	Loving-**etje**
Ting	Ting-**etje**	Tink-**je**
Ning	Ning-**etje**	Nink-**je**
Moging	Mogin-**kje**	Moging-**etje**
Boning	Bonin-**kje**	Boning-**etje**
Witting	Wittin-**kje**	Witting-**etje**
Sing	Sing-**etje**	Sink-**je**
Jing	Jing-**etje**	Jink-**je**
Lop	Lop-**je**	Lop-**tje**
Voes	Voes-**je**	Voes-**tje**
Poek	Poek-**je**	Poek-**tje**
Mip	Mip-**je**	Mip-**tje**
Tiek	Tiek-**je**	Tiek-**tje**
Foop	Foop-**je**	Foop-**tje**
Koep	Koep-**je**	Koep-**tje**
Moek	Moek-**je**	Moek-**tje**
Woeg	Woeg-**je**	Woeg-**tje**
Ries	Ries-**je**	Ries-**tje**

Table B3. Real past tense items.

Item	Target (correct) past tense	Non-target (incorrect) form (judgement task)
Aaien	Aai-**de**	Aai-**te**
Boren	Boor-**de**	Boor-**te**
Tillen	Til-**de**	Til-**te**
Kammen	Kam-**de**	Kam-**te**
Kauwen	Kauw-**de**	Kauw-**te**
Bellen	Bel-**de**	Bel-**te**
Duwen	Duw-**de**	Duw-**te**
Gooien	Gooi-**de**	Gooi-**te**
Delen	Deel-**de**	Deel-**te**
Horen	Hoor-**de**	Hoor-**te**
Kussen	Kus-**te**	Kus-**de**
Maken	Maak-**te**	Maak-**de**
Pakken	Pak-**te**	Pak-**de**
Liken	Lik-**te**	Lik-**de**
Bakken	Bak-**te**	Bak-**de**
Vissen	Vis-**te**	Vis-**de**
Werken	Werk-**te**	Werk-**de**
Dansen	Dans-**te**	Dans-**de**
Piepen	Piep-**te**	Piep-**de**
Dekken	Dek-**te**	Dek-**de**

Table B4. Nonce past tense items.

Item	Target (correct) past tense	Non-target (incorrect) form (judgement task)
Pirren	Pir-**de**	Pir-**te**
Lumen	Lum-**de**	Lum-**te**
Ronen	Roon-**de**	Roon-**te**
Monen	Moon-**de**	Moon-**te**
Zaren	Zaar-**de**	Zaar-**te**
Homen	Hoom-**de**	Hoom-**te**
Danen	Daan-**de**	Daan-**te**
Women	Wom-**de**	Wom-**te**
Norren	Nor-**de**	Nor-**te**
Lumen	Lum-**de**	Lum-**te**
Nikken	Nik-**te**	Nik-**de**
Doepen	Doep-**te**	Doep-**de**
Dappen	Dap-**te**	Dap-**de**
Guipen	Guip-**te**	Guip-**de**
Rakken	Rak-**te**	Rak-**de**
Lapen	Laap-**te**	Laap-**de**
Wauken	Wauk-**te**	Wauk-**de**
Vuppen	Vup-**te**	Vup-**de**
Naupen	Naup-**te**	Naup-**de**
Gikken	Gik-**te**	Gik-**de**