UNIVERSITEIT VAN AMSTERDAM

The role of phonology in the categorization of nouns and verbs

Thesis for the Research Master in Linguistics

Sophie ter Schure

Supervisor	**Dr. J. Don, Universiteit van Amsterdam**
Second assessor	**Dr. A. M. Kerkhoff, Universiteit Utrecht**
Date	**January 2010**
Institute	**University of Amsterdam**

Abstract

In the last decade, several studies have found that there are correlations between the forms of words and their grammatical category (Cassidy & Kelly, 1991; Durieux & Gillis, 2001; Monaghan, Christansen & Chater, 2007). Nouns and verbs, for example, can be distinguished by their number of syllables: in both English and Dutch, disyllabic words tend to be nouns, while the opposite holds for monosyllabic words. This thesis investigates whether Dutch adults and children are sensitive to such properties by presenting them with a set of novel words distinguished by their phonological characteristics on a category choice task. It was found that the number of syllables did influence participants' choice, primarily for nouns. This contrasts with findings in English, where sensitivity to phonological cues was found mainly for verbs. The asymmetry is explained by the different morphological characteristics of English and Dutch.

Table of contents

1. Introduction

In the process of acquiring their native language, children discover whether a novel word they encounter is a noun or a verb quite soon. Those newly encountered words do not come with a label, so how do they find out to which category each word belongs? On top of that, how do they discover that there are different categories of words in the first place? This thesis tries to investigate how children are able to learn linguistic categories by looking at phonological information available in the input.

After producing the first words around age one, children are thought to acquire up to ten words a day reaching a receptive vocabulary of around 14,000 words at age six (Clark, 2003). Each newly acquired word appears to be incorporated in some early linguistic system instantly. For example, Mintz (2006) has found that English 12-month-olds can already distinguish between correct noun and verb use in an *aXb* context that is either exclusively nominal or verbal, a finding that was replicated for Dutch 16-month-olds in an experiment by Kerkhoff, Erkelens, de Bree & Wijnen (in prep.). These studies tested novel verbs and nouns with neutral phonological cues, but where Mintz found significant results using a frequently occurring frame of two words (e.g. the verb frame *you X it*), Kerkhoff et al. used frames of noun- or verb differentiating morphemes (e.g. the noun frame *het X-je* 'the X-dim'). Their findings suggest that infants are already sensitive to differences between actual noun and verb contexts and are capable of generalizing these to novel nouns and verbs.

Gómez and Lakusta (2004) argue that apart from co-occurring words or morphemes, phonological distinctions between groups of words are beneficial in the acquisition of categories. They trained 12-month-old infants on *aX* and *bY* strings, i.e. infants had to learn that a-elements grouped with *X*s and not *Y*s, which themselves differed systematically in number of syllables (respectively di- and monosyllabic). During testing, new *X* and *Y* elements were introduced in a preferential-listening procedure with legal versus illegal strings. Results showed that 12-month-olds were capable of forming categories of *X*s and *Y*s on the basis of one shared phonological feature – the number of syllables – and distributional elements (*a* or *b*). Apart from the context of a word, as in Kerkhoff et al. (in preparation) and Mintz (2006), it seems that phonological features can support the acquisition of categories as well. However, this is still far from being an established fact: phonological cues have not been tested in isolation on infants yet, and it is unclear what the correlation is between adult syntactic categories and those of young children at least until they are able to use combinatorial speech.

The findings above suggest that young children are already able to perceive deviations from normal noun and verb use. On top of that, they also seem to use their linguistic categories

correctly in production. Erkelens (2009) for example finds in a corpus of four two- and three-year-olds that Dutch children make virtually no errors in using verbs, nouns and adjectives in their correct syntactic function in spontaneous speech. Concerning experimental evidence, Olguin and Tomasello (1993; Tomasello & Olguin, 1993) find a similar age of productive use of nouns and verbs in a study with English-speaking children. They used a structured test setting in which they exposed two-year-olds to new words of both categories. They found that after seven exposure sessions, the children could productively generalize inflections over groups of words, albeit more readily so for nouns than for verbs.

How is it possible that children acquire these abstract features of language so quickly? Linguistic categories cannot be innate, considering that they are language specific – for example, whereas Dutch has fixed lexical categories, which means that words from one group can only occur in restricted syntactic environments (Hengeveld et al., 2004), Samoan is said to show no clear distinction between nouns and verbs, the status of each word being derivable only through their actual use in a sentence (Don & Van Lier, fc.; Mosel & Hovdhaugen, 1992). Thus, each child has to acquire the appropriate categories of its own native language(s). More importantly, category is a feature that has to be learned for each lexical item separately. In what way do children discover these categories and how do they find out to which each new word belongs?

One potential discovery route is that the language-learning child pays attention to the environment of words: words that have the same distributional properties are grouped in the same category. For example, a word following *the* is likely to be a noun, whereas a word following a pronoun like *he* will probably be a verb (e.g., Maratsos & Chalkley, 1980; Höhle et al., 2004). Logical as this strategy may seem, by itself it is not sufficient, according to Cassidy and Kelly (1991). In an analysis of a large corpus of child-directed speech, they found that up to 15% of the caregivers' words were presented in isolation, and these words belonged to a varied set of syntactic categories. Also, it is unquestionable that children will use other information to solve the question of category as well. Pinker (1984, 1987) proposed that children use the semantic properties of words to reach a lexical distinction. Words that refer to objects are usually nouns, while verbs usually refer to actions. However, again, the extralinguistic context is not always sufficiently restrictive so that the child would be unable to discover to which entity or activity in the environment a word refers to if no other cue like distributional information is available (Gleitman, 1990; Tomasello, 2000).

A more recent hypothesis is that children could use phonological properties of speech to discover lexical distinctions between words (e.g., Morgan & Demuth, 1996; Echols, 2001). This hypothesis has at its base the idea that words that belong to the same functional category also group together phonologically, which could be at suprasegmental, phonetic, phonotactic and

prosodic levels. Although this conception is counterintuitive for anyone trained in a Saussurian frame of linguistics, which is grounded on the claim that the relation between form and meaning is arbitrary, there are numerous studies that show that the lexicon is seeped with iconic form-meaning relations (e.g., Gasser et al., 2009; Hinton et al., 1994; Nänny & Fischer, 1999; Nygaard et al., 2009), one of the most well-known being that in antonyms of for example size or place, the word for the small and close thing tends to be the one with the high vowel, whereas for the large and far thing a low vowel is used (e.g., Ramachandran & Hubbard, 2001). Since the advent of corpus linguistics, it has been found that these relations exist not only between form and meaning, but also between form and function: e.g., open-class words differ in shape from words belonging to the closed class (Monaghan et al., 2007).

A significant difference has also been found between the phonological forms of nouns and verbs. Kelly (1992), for example, observed that the number of syllables of a word in English is quite a reliable estimator of its grammatical category, polysyllabic words being less likely to refer to a verb. For Dutch, as well, Trommelen (1989) reported that Dutch underived verbs are generally monosyllabic, whereas only nouns can have polysyllabic stems. Kelly (1992) pinpointed a variety of other phonological cues that correlate with either noun or verb class in English, for example consonant quality, vowel length and word stress. The very same features were found to be even better differentiators between Dutch nouns and verbs by Durieux and Gillis (2001) in a computational learning algorithm that had to predict the grammatical category of words on the basis of phonological information. Combined, the cues could successfully categorize 68% of the words in a corpus of English, but even 75% of the words in a Dutch language corpus.

Thus, it been established that phonological cues are useful for computerized categorization of words. But are these cues actually used in language acquisition? Cassidy and Kelly (1991) found that indeed, the number of syllables of a new (pseudo)word did influence their 3- to 6-year-old participants' choice for a noun or verb interpretation of that word. Their participants seemed to think that monosyllabic words probably referred to verbs, whereas polysyllabic words were more likely to have an object meaning. In an empirical study with adults, Don and Erkelens (2008) demonstrated that this cue also influences speakers of Dutch: disyllabic stems trigger noun-assignment in 85% of the cases, while of the monosyllabic words only 42% were thought to be a noun. Concerning cues like consonant quality and vowel length (Durieux & Gillis, 2001; Monaghan et al., 2007), Fitneva, Christiansen and Monaghan (2009) showed that these cues also influence 7-year-old children's hypotheses for noun- and verb-assignment.

In this study, some incentives to test the phonological bootstrapping hypothesis for Dutch will be discussed. The following theoretical section will broaden the discussion on bootstrapping and the role of phonology in category formation, arriving at the research questions and hypotheses of the current study. Subsequently, in section 3, 4 and 5, several experiments will be reported. Section 6 will summarize the findings and put them in perspective.

2. Theoretical background

The language learning child is confronted with an enormously complex task. Even if we confine ourselves to comprehension only and disregard the many difficulties concerned with producing language, the new language learner needs to locate words in the continuous speech stream, determine the meaning of the words, distinguish among different types of words and, possibly the most complex of all, identify the syntactic relations between those words. Since syntactic rules apply to sets of words, the child's grammatical knowledge can be said to start with learning the different types of words, i.e. syntactic categories (Langacker, 1987, in Tomasello & Olguin, 1993).

The category of a word is something that has to be learned per lexical item and that is not directly available to the child. The paradox is that in order to get information about the category of the item, the syntactic relations between words also have to be known. So how do early language learners solve this problem? There must be something to help them discover that there are different categories in the first place, to *bootstrap* them into their language system. A clear definition of bootstrapping is given in Weissenborn and Höhle (2000: vii): "The notion of bootstrapping implies that the child (on the basis of already existing knowledge and information processing capacities) can make use of specific types of information in the linguistic and non-linguistic input in order to determine the language particular regularities which constitute the grammar and the lexicon of the native language."

One such proposed route to grammar is that the child makes use of the phonology of words (e.g., Morgan & Demuth, 1996). Because the bootstrapping hypothesis is based on the premise that the 'accessible' feature is actually learnt before the 'abstract' feature, and that there is indeed a more or less reliable link between the accessible and the abstract element, this means that phonology should in principle be available to the child before syntactic categories are in place, and that there is some relation between phonology and syntactic categories.

Concerning the first premise, empirical evidence shows that children have acquired the receptive phonological inventory of their first language when they are approximately 12 months old (Werker & Tees, 1984). For vowels, the emergence of language-specific categories starts

with 6 months (e.g. Kuhl et al., 1992; Polka & Werker, 1994). This means that if, for example, /i/ and /I/ represent one and the same phoneme in a language, the infant acquiring that language will not be able to distinguish /I/ from /i/ anymore when it is 6 months old. If they are meaningfully contrasted, however, the child will still be able to discriminate them after this age. For consonants, fine-tuning to the categories of one's own language seems to be a little later than for vowels: it has been established for infants aged 10-12 months (e.g., Kuhl et al., 2006; Rivera-Gaxiola et al., 2005; Werker & Tees, 1984).

Thus, indeed, the receptive phonological inventory is acquired very early. The next step is to find out whether syntactic categories are learned after this. The current methods available to test infants are unable to reveal the nature of infants' representations – there is no way of telling whether they are similar in nature to those of adults. Therefore, we can only be relatively sure about children's use of categories when they are able to use novel words correctly in a sentence. Combinatorial speech arises when children reach their second birthdays (Olguin & Tomasello, 1993; Schaerlaekens & Gillis, 1987). Krikhaar and Wijnen (1996) suggest that children around that age indeed have acquired syntactic categories or their preambles. They presented Dutch 2- and 3-year-olds with novel words in verbal (e.g., *die is aan het Xen* 'that one is Xing') or neutral (*kijk, X* 'look, X') contexts and let them choose between either nominal or verbal referents. They found that their participants tend to associate new words occurring in morphosyntactic context prototypical for the category 'verb' with action or change-of-state referents, not with objects. Krikhaar and Wijnen conclude that 'at a relatively early age, not very long after the onset of combinatorial speech, children have already abstracted some of the most typical structural phenomena associated with syntactic categories and that they are able to use this knowledge in predicting the category of a new word' (1996: 205).

The next question is whether children also use *phonological* information in predicting the category of novel words. The first study to investigate whether there actually is a link between phonology and grammatical category in the input children receive was Cassidy and Kelly (1991). These researchers observed that in a corpus of parental speech to 15-month-old children, there were almost no polysyllabic verbs, whereas polysyllabic nouns were frequent. Cassidy and Kelly systematically investigated this corpus (Landau & Gleitman, 1985) as well as a corpus of adult written language (Francis & Kucera, 1982) and found that in both cases, there was a highly reliable syllable number difference between nouns and verbs.

Kelly (1992) proceeded from this finding by giving an overview of which phonological characteristics are correlated with specific grammatical categories in the literature, covering a number of distinctive noun- or verb features. Beginning with Cassidy and Kelly (1991)'s finding on the distinction between the number of syllables in nouns and verbs, he reports that "in a

corpus of English parental speech, the observed probability that a monosyllable is a noun was 38%. For disyllabic words this figure went up to 76% and for trisyllabic words to 94%. All words of four syllables were nouns." Kelly also describes a stress difference in English nouns and verbs, concerning only disyllabic words: where nouns are more likely to receive first-syllable stress, disyllabic verbs are more likely to receive second-syllable stress (*re*cord vs. re*cord*). Thirdly, it has been reported that English nouns and verbs differ in duration, but this distinction has been found only in sentence context. Another phonological feature that is correlated to category is the voiced-voiceless distinction. In some specific noun-verb pairs, the only difference between the two is in the voicing of the final phoneme, with the last phoneme of the verb being voiced, the noun voiceless. Homographic examples are *house* and *use* (1992: 354), with the noun taking a voiceless /s/ and the verb /z/. Fifth, Kelly reports a study by Soreno and Jongman (1990) who found that the stressed syllables of nouns in the Francis and Kucera (1982) word frequency count more often contained back vowels than front vowels, while verbs demonstrated the opposite pattern.

Following Shi, Morgan and Allopenna (1998) who stated that phonological cues to categories seem to be generalizable to more than one language on the basis of a corpus study on Turkish and Mandarin, Durieux and Gillis (2001) checked Kelly's findings on both English and Dutch, using a sample of 5000 word forms and lemmata in the CELEX corpus (Baayen et al., 1993). They used a computational learning algorithm which had to assign word forms to their appropriate grammatical class by looking at various types of phonological and prosodic information. Durieux and Gillis conclude that "there is more than an arbitrary relationship between (...) nouns and verbs and their phonological form." (p. 203). Their findings suggest that the mapping between form and function is more transparent in Dutch than in English. This suggests that an experiment that finds a sensitivity to phonological cues in speakers of English might be even more successful when investigating speakers of Dutch.

Trommelen (1989) observed another set of phonological differences between Dutch nouns and verbs. She noticed that Dutch verbs generally conform to a set of phonological restrictions that nouns do not have to fulfill: underived verbs do not have disyllabic stems[1], verb stems do not end in a monophtongal vowel, do not have more than three elements as rhyme, and finally, do not end in final schwa followed by /m/. As opposed to these binary cues, the ones named in Kelly (1992) and Monaghan, Christiansen and Chater (2007) are only probabilistic. This last group of researchers extended the research on the phonology-syntax relation by

[1] Don and Erkelens (2008) discuss and refute counterexamples to this generalization: disyllabic verbs with a schwa as kernel vowel in the second syllable are phonologically speaking monosyllabic; plurisyllabic verbs ending in *–eer* are not underived; other exceptions seem to be converted verbs (e.g., *olie* 'to oil' is derived from the noun *olie* 'oil').

including the impact of distributional cues on category discrimination. They calculated the relative weight of distributional and phonological information per word using speech corpora of English, Dutch, Japanese and French and used this in a discriminant analysis to see which property contributed most to discrimination in nouns and verbs. Monaghan et al. (2007)'s results suggest that when phonological and distributional information is combined categorization success is highest, but phonological and distributional cues could separately also categorize words with a success rate above chance level. Interestingly, they also found that phonological and distributional cues seem to operate in different areas of the lexicon – i.e., that they classify a complementary set of words. Where contextual cues are less strong, phonological cues are found to be more effective.

The studies above point to a broad range of information in the input that can potentially be employed to categorize words into nouns and verbs. More relevant to our question is whether language users are actually sensitive to this information. Some of the studies above have examined this issue. For example, the noun-verb stress-difference that is reported in Kelly (1992) was tested on adult native speakers of English in Kelly (1988a). He presented participants with recorded disyllabic pseudowords that differed only in stress pattern, asking the participants to create a sentence containing the pseudoword. When the word had first-syllable stress, it was more likely to get the role of a noun in a sentence, whereas second-syllable stress increased the likelihood of verb use. Another experimental study assessing one of the cues in Kelly (1992) was performed by Soreno and Jongman (1990), who found that nouns and verbs were differentiated by vowel quality. In an experiment where participants had to classify existing words as being nouns or verbs, words were classified correctly more rapidly if they conformed to the typical pattern, e.g. nouns were classified faster if they contained back vowels than if they contained front vowels.

A sensitivity for the relationship between syllable number and grammatical class has also been found in both adults and children. Cassidy and Kelly (1991) created an experiment similar to that in Kelly (1988a): adult speakers of English were asked to construct sentences with mono-, di- and trisyllabic pseudowords. The number of syllables was found to affect grammatical category use in the expected direction: noun use increased with syllable number. They pursued this positive result through an experiment with 3- to 6-year old children. The influence of syllable number on noun and verb category assignment was tested by assessing children's choice for the potential referent of a pseudoword, providing two options: one action and one concrete object. Twelve videotaped scenes consisting of one common object performing a common action were used as experimental trials, e.g. a door opening. Participants were told that they were going to hear words from a puppet's alien language. With each scene, participants

were presented with a pseudoword that was either mono- or trisyllabic (e.g., *pell* or *delfinate*). Children had to guess which referent they thought was meant by the word. This resulted in a significant effect of stimulus length on referent choice, with actions being chosen more often for monosyllabic pseudowords and objects more often for trisyllabic words.

Cassidy and Kelly (2001) expanded this result by looking at the effect of phonological cues on learning new words. Here, participants were submitted to three blocks of trials: after each guess in the first block, 3- to 6-year-old children received feedback to teach them the 'correct' referent according to the condition the participant was in. In the consistent condition, the mapping between stimulus and referent followed the pattern in English, such that monosyllabic words referred to actions, and trisyllabic words to objects. In the inconsistent condition, the mapping was opposite to the typical pattern. In the independent condition, there was no systematic mapping between stimulus length and referent. In blocks 2 and 3, the influence of condition on learning was assessed. Cassidy and Kelly (2001)'s results show that phonological cues have an effect both on the initial guess in that children tend to assume that trisyllabic words are nouns, and on word learning: children in the consistent condition learned more words than participants in the other conditions.

Fitneva, Christiansen and Monaghan (2009), instead of testing the number of syllables of a word as a cue to category, tried to incorporate all probabilistic cues available in monosyllabic pseudowords only. Using the CELEX-database (Baayen, Piepenbrock & Van Rijn, 1993), they created eight stimuli that where phonologically typical for English nouns and eight that conformed to typical verbs. In three blocks of each sixteen trials that consisted of two pictures, one of an object and one of an action, and one auditory stimulus presented through a computer, their 7-year-old participants had to decide to which picture they thought the pseudoword referred to (e.g., 'popcorn' or 'chop'). After feedback, in a second and third block it was investigated whether phonological information could overrule another important constraint on word learning, namely learning by exclusion. Fitneva et al. found that phonology had an influence on initial categorization (children's guesses) but that this was mitigated as soon as the child had received feedback, i.e. on the second and third block they no longer chose on the basis of phonology but on the basis of what they had learned. However, when looking at verb-like stimuli only, children in the consistent condition performed better than those in the other two. Thus, there was only an effect of phonology when learning verb-like items. According to Fitneva et al., this is in accordance with findings in Monaghan et al.'s computational analysis of noun- and verb-categorization in English (2007): nouns benefit most from distributional information, while verbs are best categorized on the basis of phonological cues.

Seeing that the corpus studies have found that cues for Dutch might be even stronger than those for English, and experimental results on English look promising, it is surprising that there has been only one study on sensitivity for phonological cues for Dutch. Don and Erkelens (2008) assessed the phonological characteristics of nouns and verbs reported by Trommelen (1989) in a decision task with adults. Using stimuli that were either ambiguous to noun or verb category or conformed to phonological characteristics of nouns, unwitting participants had to decide whether they thought each stimulus was a verb- or a noun stem. Don and Erkelens found that Dutch adults are indeed sensitive to almost all of the cues listed in Trommelen (1989). The two strongest cues for nounhood were the following: when the pseudoword contained a final schwa, 88% of the participants classified it as a noun, and disyllabic stimuli elicited noun-classification in 85% of the subjects. Ambiguous stimuli elicited a noun response in only 21% of the cases, so this effect was is not due to a tendency to interpret any novel word as a noun.

These results suggest that, if children show the same sensitivity, phonology might indeed be adopted in learning syntactic categories, as Don and Erkelens conclude: "If adults can actively use the link between phonology and category membership, the idea that language-learning children may use this information to get at the proper categorization becomes more promising." (2008: 680). Thus, first we have to test whether Dutch children are sensitive to these correlations as well or whether possibly sensitivity to phonological differentiations between nouns and verbs only arises after extensive exposure to the language, which would imply that it would not be employed in the initial learning of syntactic categories. The goal of this study is to examine the sensitivity of Dutch children to phonological cues available in the input as testified by Don and Erkelens (2008), Durieux and Gillis (2001) and Monaghan et al. (2007), which are summarized in Table 1 below. To accomplish this, several experiments will be carried out along the lines of Cassidy and Kelly (2001)'s referent choice experiment.

The number of syllables of a word as tested by Cassidy and Kelly (1991/2001) seems to be the most promising cue: with English children, it was found to have a significant effect on initial categorization of novel words and on learning them. Furthermore, Dutch adults were found to employ this cue in a decision task (Don and Erkelens, 2008). To examine whether this cue is relevant for Dutch children as well, after a pilot study with adults, the first experiment in this thesis employs the method used by Cassidy and Kelly (2001), in an attempt to investigate both the influence of phonology on children's initial guess of the category of novel words and the effect of phonological typicality on learning the correct referent of the words.

Phonological cues for Dutch			
Category	**Cue**	**Strength**	**Source**
nouns can have	disyllabic stem	85% vs. 15%	Don & Erkelens, '08 (percentage of participants triggered by this cue to choose noun over verb answer)
	final long vowel	83% vs. 17%	
	final schwa	88% vs. 12%	
	final schwa + nasal	82% vs. 18%	
nouns are more likely to have	a high number of phonemes	66.80% success	Durieux & Gillis, '01 (success rate of algorithm in choosing correct category)
	nasal consonants	59.06% success	
	low vowels	66.46% success	
	bilabial in onset	0.10 vs. 0.05	Monaghan et al., '07 (distribution difference in nouns vs. verbs and verbs vs. nouns in corpus at $p < 0.001$)
	plosive in onset	0.20 vs. 0.15	
	palatal in word	0.06 vs. 0.01	
verbs are more likely to have	fricative in onset	0.30 vs. 0.21	
	labial in word	0.10 vs. 0.05	

Table 1: Phonological characteristics correlated with Dutch noun- or verb-category.

Apart from testing the influence of the number of syllables, it will also be investigated whether a combination of the more subtle phonological cues to category assignment are employed by adults in Experiment 2. Finally, Experiment 3 combines all cues available in an experiment with children in an adaption of the method used by Storkel (2001/2003). This method is more sensitive than that of Cassidy and Kelly (2001) because it does not only assess the learning of the correct referent, but also the influence of phonology on acquiring the word form in comprehension and in production. This experiment will look at the influence of phonology in the presence of distributional information, which is a more natural situation than when words are presented in isolation.

This study will thus be the first to assess the influence of phonological information on category assignment in Dutch children. Furthermore, until now, only sensitivity to the cues reported by Trommelen (1989) was tested in an experimental setting. This study will add to the experimental evidence by also assessing the cues in Durieux and Gillis (2001) and Monaghan et al. (2007).

3.1 EXPERIMENT 1A: SENSITIVITY TO NUMBER OF SYLLABLES IN ADULTS

Before testing the effect of the number of syllables on children's category assignment, this experiment intends to replicate the finding of Don and Erkelens (2008) that this phonological cue influenced initial categorization of pseudowords in Dutch adults. In Don and Erkelens, the number of syllables was only one of the cues tested; here, all items are intended to test sensitivity to this phonological characteristic. Twenty monosyllabic and twenty disyllabic pseudowords were presented to participants in the context of a decision task.

3.1.1 Method

Participants

Thirty-one adult native speakers of Dutch were tested in this study. Each had finished their university education or was currently studying at university level. The age range was 18-40 years. Participants were approached through email and pseudo-randomly assigned to one of four versions of the experiment varying in the order of the test items and the order of the answer columns. There were twelve participants in version 1, seven in version 2, five in version 3 and seven in version 4.

Materials

The test involved 40 pseudowords that were mostly created for earlier studies (Kerkhoff et al., in preparation; Don & Erkelens, 2008). All were phonologically plausible for Dutch but nevertheless non-existent words. For a list of the stimuli, see Table 2. The 40 words were administered in a decision task consisting of three columns on a single computer-presented page. In the first column, each pseudoword was listed in isolation, i.e. without a semantic, morphological or syntactic context. Half of the pseudowords were mono-, the other half disyllabic, so that twenty words were ambiguous to either the Dutch verb- or noun category, and twenty conformed to the phonology of typical Dutch nouns.

The second and third columns on the page were answering columns which were headed by either 'noun' or 'verb'. To help participants remember that they were deciding whether each pseudoword was a possible noun or verb *stem*, the noun column contained the indefinite determiner *een* 'a', which is a typical context for a singular noun (which is identical to the stem form in Dutch), and the verb column contained the first person singular pronoun *ik* 'I', which is also always followed by the stem of a verb in Dutch.

Design & Procedure

The test was introduced in an email which informed possible participants about what was expected from them if they would agree to take part in the experiment. They were told that the attachment of the email contained a list of non-existing word stems for each of which they had to decide whether they thought it was a noun stem, like *een <u>boek</u>* 'a book' or a verb stem, as the form that follows after *ik* 'I' as in *ik <u>zie</u>* 'I see'. They were instructed to look at each pseudoword briefly and then put a mark in the noun or verb box behind the word depending on which category they thought it would most probably belong to if it existed.

The set of phonological possibilities for verbs is a proper subset of the phonological options of nouns (Don and Erkelens, 2008) which means that disyllabic underived verb stems are exceptions in the language but there is no such restriction for the number of syllables of noun stems. It was thus expected that nonce stems with two syllables would incline participants to give a noun answer, whereas for monosyllabic words a noun or verb choice should be at chance level.

3.1.2 Results

All 31 participants showed a tendency of choosing the nominal category more often than the verbal category. As Table 2 below shows, only two of the stimuli received exclusively noun responses: *prellon* and *smoza*. Disyllabic items evoked a noun response in 88% of the participants, whereas monosyllabic items yielded a mean of only 42% of noun answers.

Monosyllabic items			
Items	Pct.	Items	Pct.
frep	6%	dieg	41%
blaap	18%	pla	41%
snoeg	26%	diel	44%
pluig	26%	book	50%
krei	29%	plif	50%
sook	29%	drauf	56%
taaf	29%	nort	59%
kloep	32%	pierst	76%
klui	35%	boogst	79%
snef	35%	kleest	85%
		Mean total	42%

Disyllabic items			
Items	Pct.	Items	Pct.
prellon	100%	boele	91%
smoza	100%	strempa	91%
donkam	94%	zierem	91%
fidong	94%	riele	88%
wadim	94%	ploren	88%
krile	94%	wirfem	88%
bolee	94%	meliens	79%
doelem	91%	nado	79%
kellem	91%	balter	74%
giveno	91%	slamper	50%
		Mean total	88%

Table 2: Percentage of noun responses for each stimulus across participants in Experiment 1a.

The table above shows the percentage of noun answers across participants for each item separately, ordered in such a way that the best items for each category are at the top. Thus, *frep* was categorized as a verb by 94% of the participants while the monosyllabic stimulus that was ranked lowest, *kleest*, was considered to be a verb stem by only 15%.

In accordance with our expectations, the response to the monosyllabic verb-like items did not differ significantly from chance (50%) on a two-sided t-test (t [19] = -1.67, p = 0.09). Disyllabic noun-like items did yield a difference from chance (t [19] = 15.41, $p < 0.001$). Thus, the number of syllables of a word does seem to influence participants in categorizing pseudowords as nouns as opposed to verbs.

It is interesting to note that the three items that were the least likely to be categorized as verbs (*pierst, boogst,* and *kleest*) conform to one of the cues for nounhood indicated by both Trommelen (1989) and Durieux and Gillis (2001), that of 'superheavy rhyme' or 'high number of phonemes' respectively. The response to these three items differed significantly from response to the other 17 on a two-sided t-test with $p < 0.001$. However, in the current experiment, the goal was to establish only the effect of the number of syllables on categorization.

3.1.3 Discussion

Conforming to results found by earlier studies (Cassidy & Kelly, 1991; Don & Erkelens, 2008), we found that although monosyllabic items did not yield a difference between noun and verb assignment, disyllabic pseudowords were interpreted as being nouns by the majority of the participants. This suggests that the number of syllables is used as a cue by adults when it comes to guessing the category of a novel word. [2]

Fitneva et al. (2009) suggest that it might obscure results if the stimuli have not been controlled for all possible cues to category: phonological cues could confound each other's effect. For example, *fidong* which was created to be noun-like by including two syllables also conforms to certain verb characteristics in that it starts with a fricative and contains a high vowel (Monaghan et al., 2007). Nevertheless, it seems safe to argue that if we did conform to all possible cues, the established effect would only increase.

[2] Further, it would be interesting to see whether this information is also used in grammatical processing, as suggested by Kelly (1992): although distributional rules are usually enough in determining grammatical class of words, there are domains of human perception in which probabilistic information is used even though sufficient other information is available, e.g. depth perception, letter identification (1992: 350). This might be the same in the case of grammatical processing. E.g. Kawamoto, Farrar and Overbeek (1990) find that on-line classification of words (in sentence contexts) into grammatical categories was affected by stress: disyllabic nouns were named more rapidly if they conformed to the usual stress pattern (i.e. trochaic) in English than if they were iambic, whereas for disyllabic verbs the opposite pattern emerged. Also, Monaghan, Christiansen and Chater (2003) find that the phonological properties of syntactic categories influence word-naming and lexical decision in adults which suggests that there is a relation with lexical access.

The results of this experiment indicate that native Dutch speaking adults are sensitive to the correlation between grammatical category and syllable number – albeit subconsciously – and also, that this information is used when they have to determine the grammatical category of a novel word in the situation that there is no other information available, just as was found by Cassidy and Kelly (1991) in their study on English speaking adults and by Don and Erkelens (2008) with Dutch adults. These findings suggest that it could be fruitful to investigate sensitivity to syllable number in Dutch children as well, using a procedure similar to that of Cassidy and Kelly (2001).

3.2 EXPERIMENT 1B: SENSITIVITY TO NUMBER OF SYLLABLES IN CHILDREN

In this experiment, the influence of disyllabicity is tested on a category assignment task with Dutch 7- and 8-year-olds. The effect of consistent vs. inconsistent matching of pseudowords and referents is investigated in a replication of Cassidy and Kelly (1991). Six mono- and six disyllabic pseudowords from the previous test are presented to the children in the context of twelve videotaped scenes containing one object and one action. An alien puppet that does not speak Dutch is pictured on the side of the screen 'watching' the scene with the child. Participants are then asked to guess to what the alien pseudoword is referring to: the object or the action. By repeating the procedure, it will also be tested whether phonological correlations facilitate the learning of novel words.

3.2.1 Method

Participants

The test group consisted of five 8-year-olds (mean age 8.6 [*SD* 3.8 months], 2 girls, 3 boys) from a primary school in West-Friesland. This age group was chosen because we wanted to make sure that the test was comprehendible and feasible for the children, since we hoped to extend the research to younger children eventually. Also, the positive findings on Dutch adults and on English 3- to 6-year-olds (Cassidy & Kelly, 1991, 2001) suggest that the chance of finding an effect in Dutch 8-year-olds is high.

Materials

Apart from the categorization experiment, children were presented with a phonological test battery consisting of the Een-Minuut-Test (EMT; Brus & Voeten, 1972) and the Klepel (Van den Bos, Lutje Spelberg, Scheepstra & De Vries, 1994), the digit span test from the Wechsler Intelligence Scale for Children (WISC-III; Wechsler, 2002) and a vocabulary measure, the

Peabody Picture Vocabulary Test (PPVT; Dunn & Dunn, 2005). These measures were taken in order to make sure that the children were normally developing and to be able to compare the result of the categorization experiment per child with their vocabulary estimates. It was thought that any effect of phonology would possibly increase with vocabulary size. The mean standardized PPVT score was 111 (*SD* 7.5), the raw mean score on the digit span 12.2 (*SD* 3.1), on the EMT 49 (*SD* 8.1) and on the Klepel 43.6 (*SD* 9.4).

The actual test consisted of a word learning task based on Cassidy and Kelly (2001). It included twelve of the phonotactically legal pseudowords that were used in Experiment 1a. Six disyllabic pseudowords with a relatively high noun score and six monosyllabic pseudowords with a relatively low noun score were chosen in order to maximize the possible result. Care was taken to include pseudowords that differed maximally in phonemes so as not to thwart word learning unnecessarily. Stimuli were recorded by one female speaker of Dutch to ensure consistent presentation.

Type	**Items**	**Object**		**Action**	
Monosyllabic	frep	*tijger*	'tigre'	*vallen*	'to fall'
	blaap	*potlood*	'pencil'	*kleuren*	'to colour'
	pluig	*egel*	'hedgehog'	*likken*	'to lick'
	krei	*theepot*	'teapot'	*zinken*	'to sink'
	sook	*kikker*	'frog'	*springen*	'to jump'
	taaf	*brommer*	'moped'	*rijden*	'to ride'
Disyllabic	prellon	*beker*	'cup'	*drijven*	'to float'
	smoza	*schildpad*	'turtle'	*glijden*	'to glide'
	donkam	*vogel*	'bird'	*buigen*	'to bow'
	fidong	*wortel*	'carrot'	*dansen*	'to dance'
	wadim	*varken*	'pig'	*rollen*	'to roll'
	strempa	*appel*	'apple'	*groeien*	'to grow'

Table 2: Pseudowords and target words used in Experiment 1b.

Pseudowords were matched to twelve short videotaped scenes containing one common object performing or undergoing one common action. Scenes were acted out with clay dolls, photographed and made into five-second stop-motion videos. Each pseudoword was played three times during the scene.

It was important for the current study that the movies triggered only a single specific noun and verb, because the intended target words should be matched for frequency and number of syllables. Five adults were asked to write down which noun and verb they would use to describe the scene. Their word choice matched fully with the intended design.

Both verbs and nouns are known by the large majority of Dutch 6-year-olds (Kohnstamm et al., 1981, 1999) and were tested for frequency in CELEX (Baayen et al., 1993). Noun and verb frequencies were as balanced as possible (see appendix). The nouns and verbs in each scene were matched on the number of syllables.

Pseudowords were matched to scenes in such a way that for each participant, six trials were consistent with the phonological pattern in Dutch and six inconsistent. Since in half of the scenes the intended referent was the action, this means that there were three monosyllabic pseudowords matched to action referents (consistent), three disyllabic pseudowords matched to action referents (inconsistent), three disyllabic pseudowords matched to object referents (consistent) and three to action referents (inconsistent). Different versions of the test were created varying in the order of the trials.

Design & Procedure

Participants were tested separately in a room of their primary school. The procedure was administered through a laptop with separate speakers. During the procedure, the experimenter was seated to the right of the child. The experiment started with an introduction scene in which the non-Dutch speaking puppet ('Blip') presented itself. After that, a training scene featuring one common object performing a common action was shown – a toy house turning – with Blip watching from one corner of the computer screen. While the scene was playing, a pseudoword *X* was offered three times. The experimenter then asked, 'What do you think Blip means with *X*? House or turn?' (*Wat denk je dat Blip bedoelt met X? Huisje of draaien?*).

When the child understood the procedure, the test was started. Twelve scenes similar to the training scene were shown in succession. Following each answer, the child received positive or negative feedback ('yes, it was *Y*' or 'no, actually, it was *Z*'). After this first block, the same twelve scenes were shown again. Now, the experimenter asked 'Do you remember what Blip meant with *X*? *Y* or *Z*?' (*Weet je nog wat Blip bedoelde met X? Y of Z?*). The third block followed the same procedure.

Answers in the first block of trials were scored for phonological guess, i.e. does the child choose the item that corresponds with the phonological cue in the pseudoword? It was expected that children would be inclined to give a noun answer if the pseudoword was disyllabic and a verb answer if it was monosyllabic. On the second and third block of testing, the effect of consistency could be measured. The hypothesis was that it would be easier to learn items that were consistent with the phonological tendency in Dutch, i.e. monosyllabic words that referred to nouns and disyllabic to verbs, than to learn the inconsistent items.

3.2.2 Results

Concerning the results on the first block, no effect of phonology could be found: participants chose the referent that was cued by phonology in only 51.7% of the cases. In order to look at the effect of phonology on learning, Block 1 answers were also scored for accuracy so that they could be compared to scores on Blocks 2 and 3. Table 3 shows the percentage correct for each participant per condition separately.

Participant	Condition	Block 1	Block 2	Block 3
1	consistent	66.7	100	100
1	inconsistent	50.0	83.3	100
2	consistent	16.7	100	100
2	inconsistent	33.3	100	100
3	consistent	33.3	66.7	50.0
3	inconsistent	33.3	50.0	50.0
4	consistent	50.0	50.0	66.7
4	inconsistent	66.7	66.7	66.7
5	consistent	50.0	83.3	83.3
5	inconsistent	50.0	83.3	83.3

Table 3: Accuracy percentages on each block per participant per condition in Experiment 1b.

There is an effect of block in that children learn the correct referent very quickly: Figure 1 shows a very steady rise in accuracy from Block 1 to 2 but not from 2 to 3. An effect of consistency could not be established. Note that these results are all established on the data of only five participants.

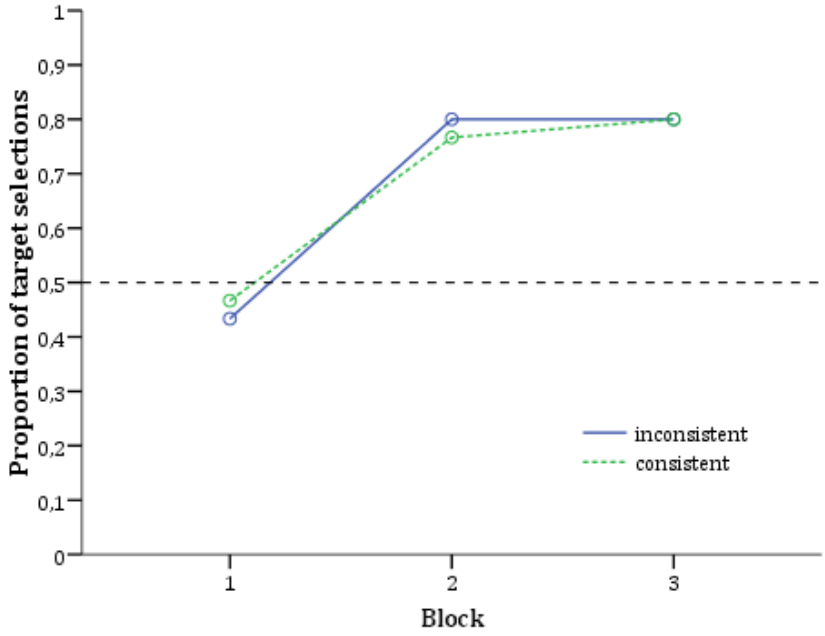


Figure 1: Answers over blocks Experiment 1b.

If we split the data on referent type (objects vs. actions), it becomes clear that children were more successful in learning action names than object names. This difference is significant on a t-test (t [8] = -2.630, p = 0.030).

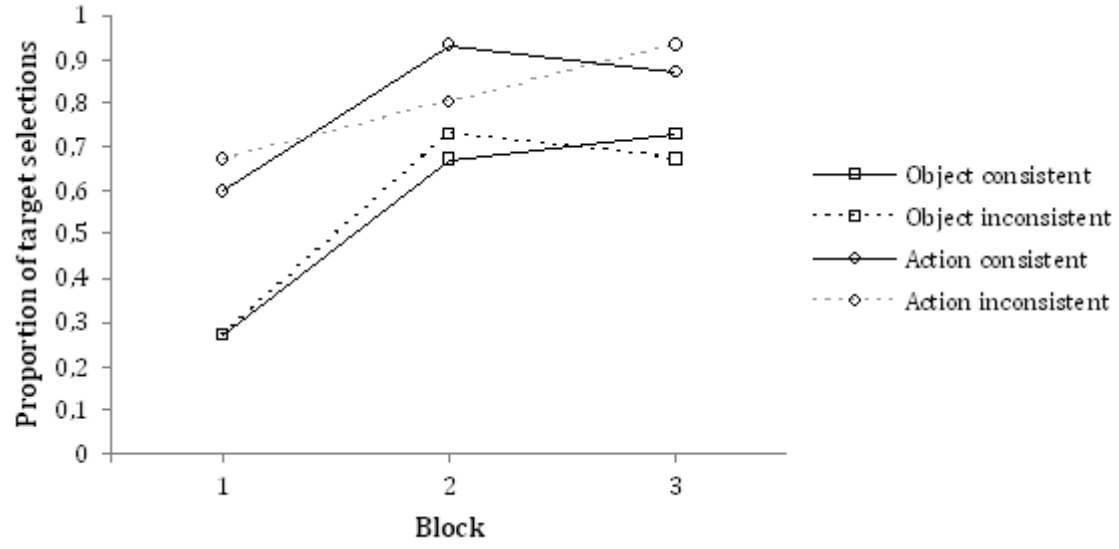


Figure 2: Answers over blocks per referent type on Experiment 1b.

On the basis of these data, all that can be concluded is that no clear picture emerges. It seems preliminary to say that there is no effect of phonology at all, since only five participants were tested.

3.2.3 Discussion

The results of this study do not support the phonological cue-hypothesis, contrary to what was found in Experiment 1a. Why does this experiment not do what we expected? One problem could be that the task does not trigger phonological cues at all. This is improbable since a very similar experiment did yield the expected results with English children (Cassidy and Kelly, 1991, 2001). At least we know that children in our experiment were paying attention, since their accuracy clearly increased, but on the first block they seemingly randomly chose one of the options presented to them by the experimenter. The scores of all participants on the first block were around chance level.

Concerning the learning measure, Figure 2 displays a tendency in the object referents that consistent items were learnt better, while for action referents inconsistent referent trials yielded higher scores. This means that in both cases children were more successful in learning disyllabic words than monosyllabic words. Since one would expect that in general shorter words should be easier to learn than longer ones, we have to look for an external reason for this finding. It is possible that the disyllabic advantage comes from the fact that the target words are disyllabic as well. If children simply learned the novel words by relating them to the correct target word in the scene, for nouns this would be easier if the novel word was consistent with phonological cues (i.e. disyllabic), since e.g. consistent *strempa* is more like *egel* 'hedgehog' than inconsistent *frep*, because both the consistent item and the target word have two syllables. For verbs, it would actually be easier to match novel to target word if the novel word was inconsistent with phonological cues (disyllabic in the case of verbs), because it seems that Dutch verbs are stored in the lexicon with their infinitival suffix *-en*, which makes them disyllabic also (e.g. Ernestus, 2005) although their stems are monosyllabic.

If indeed the participants learned the pseudowords by relating them to the appropriate label for that referent in their lexicon, this could be the cause for the disyllabic advantage and for the non-effect of consistency, because inconsistent verbs and consistent nouns would now be easier to learn then consistent verbs and inconsistent nouns, thereby cancelling out any consistency effect of phonological cues. Remember that in Experiment 1a, monosyllabic pseudowords were linked to verbal stems, not infinitives. Infinitives are always at least disyllabic in Dutch (except for atypical irregular verbs like *gaan*, *slaan*, *staan*, *zijn*), whereas in English, infinitives are in most cases exactly the same form as the stems (e.g., 'I turn', 'to turn'). In English, thus, the monosyllable cue for verbs holds both for stems as for infinitives, so it is not a problem if seeing an action triggers an infinitive, not a stem.

Not only is the fact that infinitives are disyllabic in Dutch a problem, but also the *-en*-affix itself: since seeing a scene like this triggers participants to think of an infinitive verb (as pilot

testing with 5 adults revealed), Dutch 8-year-olds might be confused to learn that a pseudoword without such an affix can be a verb as well, albeit in some alien language that they only learn for the current experiment. That would mean that morphological cues to category are much stronger here than phonological cues. In that case, however, we would expect children always to choose the object referent in the first block, which they did not: in 70.6% of the items they picked the action referent after hearing the pseudoword.

To solve this confound in the experiment, regardless of whether it lies in the *-en* suffix or the number of syllables of the target word, it seems necessary to find a way to elicit stems as opposed to infinitives, which would imply providing a stem context for the pseudowords, e.g. by putting the first singular pronoun before the pseudoword. In that way, however, an experiment like that of Cassidy and Kelly, for which it is necessary to have no context so that phonological cues can be tested in isolation, is impossible.

Another option would be no longer to test the disyllabicity cue but to proceed with the other possible categorization cues that are proposed in the literature. Durieux and Gillis (2001) investigated a set of 'less reliable cues' proposed by Kelly (1992) and found them to be good predictors of grammatical class in a machine learning experiment on a corpus of Dutch when these cues were combined. The following experiment tries to establish an effect of these minor cues in adults' initial categorization of pseudowords.

4. EXPERIMENT 2: SENSITIVITY TO MINOR CUES IN ADULTS

In this experiment, an attempt is made to investigate the influence of the non-exclusive cues to category that exist in Dutch. We were interested in finding out whether it was possible to find a sensitivity for phonological characteristics other than that of the number of syllables. It was decided to investigate the cues as listed in Durieux and Gillis (2001), see Table 1. If indeed the adult participants in this study reveal an effect of phonological cues here, we would be able to test phonological cues using Cassidy and Kelly (2001)'s method without the confounding relation between the number of syllables in the target and the pseudoword.

4.1 Method

Participants

Thirty-seven students at the University of Amsterdam took part in this experiment. None had participated in the preceding adult categorization task. All were native speakers of Dutch and were approached in university cafeterias by the experimenter. The test lasted a maximum of five minutes and subjects received a small bar of chocolate as an incitement to participation.

Materials

The test consisted of a list of 20 pseudowords and was equivalent to the list used in Experiment 1: three columns were printed on one sheet of paper, the left column listing each of the pseudowords, the other two being answer columns headed by either 'noun' and '*een*…' (singular noun distinctive context) or 'verb' and '*ik…*' (first person singular verb context), of which only one should be ticked. Twelve of the 20 pseudowords were test items, of which six were formed according to typical phonological characteristics of Dutch nouns according to Durieux and Gillis (2001) and six according to phonological characteristics typical for verbs. Briefly repeating those characteristics: nouns are more likely than verbs to have a high number of phonemes, which results in complex clusters in onset and rhyme; nouns are more likely to contain nasal consonants (/m/, /n/, /ŋ/); nouns are more likely to have low vowels (/A/, /a/). Finally, it was decided to include a factor stress into the items as well: although the dominant pattern in Dutch is for disyllabic words to have trochaic (initial) stress, Durieux and Gillis note that Dutch verbs are more likely to have secondary stress in the case that they do have a disyllabic stem, while Dutch nouns are more likely to have initial word stress. Stimuli are listed in Table 4 below.

Procedure

Participants were told that the experiment consisted of pseudowords for each of which they had to decide whether they thought it was a possible singular noun or a singular first person form of a verb. These last notions were explained by giving two examples of a verb stem and two of a noun stem, preceded by respectively the first person singular pronoun and the indefinite determiner. Participants were given a fill-in form on which they could mark their choice, while the experimenter read each item out loud so that word stress was unambiguous.

It was expected that those items that were formed according to phonological characteristics of Dutch nouns would yield a higher noun categorization score than items that did not conform to these characteristics.

4.2 Results

Table 4 shows the percentage of participants that gave a noun answer per stimulus. The left column shows the items that conformed to phonological characteristics of Dutch nouns, ordered in such a way that the items that gathered the highest percentage of noun responses are at the top. In the right column, presenting the non-noun-like pseudowords, the top items are the ones that gathered the highest percentage of verb responses.

Although the words constructed to be more verb-like than noun-like did yield a lower noun score than the other stimuli, this difference was not significant on a t-test (two-sided, p =

0.124). Looking at the percentages in Table 4, this is not surprising: the low average of the non-noun-like items is mostly due to the high verb-response of one item, *rideer*. Also, the disyllabic pseudowords that now also contained noun-like complex phoneme clusters, low vowels and nasal consonants did not yield a noun response as high as the items in the first experiment.

High noun cue items	
Items	**Noun pct.**
'bronkasp	89%
'smozarp	82%
fli'bolt	72%
'strempa	70%
prel'bans	68%
dra'nomp	65%
Mean total	74%

Low noun cue items	
Items	**Noun pct.**
ri'deer	11%
'rieluf	57%
'lebik	57%
se'gie	62%
'tikep	68%
ti'soel	84%
Mean total	57%

Table 4: Items and percentage of noun answers across participants on Experiment 2.

Nevertheless, the percentage of noun answers on noun-like items is still significantly different than chance (t [5] = 6.448, $p < 0.001$ (two-tailed)), while the percentage of verb answers on non-noun-like items is not ($p = 0.555$). The difference between the groups is non-significant (t [10] = 1.679, $p = 0.124$). Remember that apart from a difference in phonological category cues, there was a stress distinction included in the stimuli: half of the words in this experiment had trochaic, the other half iambic stress. This distinction did not result in a significant difference on the percentage of noun answers.

4.3 Discussion

The goal of this experiment was to replicate the results of Experiment 1a with phonological cues to category other than the number of syllables. These cues were found to be much less apparent than the mono- vs. disyllabic cue in an experiment with 37 adults. Although there was a significant effect of item group compared to chance in the expected direction, it cannot be concluded that there is a proper effect of phonological cues here: there is no significant difference between the two groups of items.

Subjects categorized pseudowords as nouns more often than as verbs. In Experiment 1a, a noun choice was made for 65% of the items, but it was clear that monosyllabic items yielded a lower score than disyllabic items. The current experiment yields exactly the same noun percentage over all items, 65%, but there is no clear difference between the two sets of stimuli. Perhaps speakers of Dutch are only sensitive to the disyllabicity cue to category. Another

possibility is that having more than one syllable is such a strong indicator of noun category that it obscures any other less reliable cues. Investigating these cues would then mean using monosyllabic pseudowords only.

The nine highest noun scoring items in Don and Erkelens (2008) were all disyllabic, and of the five cues to category they tested, only the four that concerned pseudowords with more than one syllable were successful (the anomalous cue was 'more than superheavy rhyme' which triggered nominal category choice in only 55% of the items). The high noun choice in this experiment overall could also be due to the tendency that novel words that we encounter are probably more often nouns than verbs. Several participants noted that they felt that all the items in this task were more noun- than verblike.

The only item that produced a high verb score was *rideer*. Although this item was created according to the phonological characteristics of verbs, seeing that the other items designed for this purpose did not yield a similar score means that there must be another reason for its success. Although there is no reported rule that dictates that 'stem + *-eer*' makes a verb from a noun or adjective, there is a whole set of derived verbs and loanwords ending in *-eer*: for example, *kampeer* 'to camp', *paneer* 'to coat with breadcrumbs', *halveer* 'to divide into halves', *parkeer* 'to park', *probeer* 'to try', *signeer* 'to autograph', *markeer* 'to mark', *fileer* 'to fillet', *arceer* 'to hatch', *studeer* 'to study', *geleer* 'to jellify', *bekeer* 'to convert'. However, I can think of only two nouns that are disyllabic and have –eer: *meneer* 'sir' and *verkeer* 'traffic'. The high verb score for *rideer* could thus well be a lexical neighbourhood effect.

Nevertheless, even if the cues under investigation did not trigger an effect in this experiment, they might still have an effect in concord with others. In Durieux and Gillis' categorization algorithm, it was found that even those cues that are not very informative in predicting category by themselves do contribute significantly to categorization when they are combined with other phonological cues (Durieux and Gillis, 2001: 209). Thus, although sensitivity to these cues by themselves could not be established, it is likely that if they were to be combined with the first cue that was tested, the number of syllables in a word, there would be an effect. Therefore, in Experiment 3, we are going to test the number of syllables-cue again in a test with children by reversing the set-up of Experiment 1b: no longer does the child have to choose to what referent the word they hear designates, but it is asked to decide which of two words it thinks more likely refers to the action or object they see. Because there is no longer a choice between a noun or a verb, the pseudoword does not have to be presented in isolation. Now, the same word is tested both in a noun context (*een*) and a verb context (*ik*). This means that the influence of phonology on choice can be tested in the presence of distributional info, which is very interesting in itself.

5. EXPERIMENT 3: SENSITIVITY TO PHONOLOGICAL CUES IN CONTEXT IN CHILDREN

In this experiment, the influence of phonological characteristics of words in the initial categorization of words is evaluated in the presence of distributional cues. Now, this is not tested by letting participants actively choose to what category novel words belong by picking one of two referents or checking a box; here, subjects are presented with only one picture and are asked to choose which of two words – one with noun, one with verb characteristics – most likely refers to it. Thus reversing the original procedure, we can present the words as stems by embedding them in a distinctive stem context while phonological characteristics of those words can be tested by evaluating the difference between presenting them as noun or as verb.

The effect of consistent or inconsistent mapping of phonological noun or verb characteristics and their referents will be explored by looking at participants' initial choice of word form and the speed and accuracy of learning the correct word-referent mapping. This is a combination of the procedure in Cassidy and Kelly (2001) and that in Storkel (2001/2003). Cassidy and Kelly (2001) assessed whether children's initial referent choice was influenced by phonological characteristics as well, and examined the influence of consistent versus inconsistent mapping on the accuracy of learning the intended referent of a word. Storkel (2001/2003) investigated the relation between phonology and word learning regardless of grammatical category. Hypothesizing that the phonological make-up of a word influences the ease with which children store a new word in their lexicon, she found that 3- to 6-year-olds had more difficulty in remembering words that contain phonotactically less plausible sequences compared to words that have a high phonotactic probability (a measure of the likelihood of occurrence of a sound sequence computed over the whole lexicon). She exposed children to pseudowords that had either common or rare sound sequences coupled with unfamiliar objects and tested them on the speed and accuracy of word learning. Children were tested after 1, 4, and 7 exposures on their passive learning of the word form in a task where they had to pick the right word from a set of three (target, distractor and filler) and on their active learning in a production task. It was found that words containing common sound sequences were easier to learn than words with a lower phonotactic probability, both for nouns (2001) and verbs (2003). The procedure used by Storkel is more sensitive than Cassidy and Kelly's concerning the learning of words since apart from learning the correct referent for a word, it includes two measures of learning the phonological word form itself: after acquiring the correct word-referent mapping, children are asked to identify the correct word from three, and in a subsequent block, children are required to produce the word themselves.

Although Cassidy and Kelly (2001) found an effect of consistent versus inconsistent mapping on both initial referent choice and on word learning, Fitneva et al. (2009), using a similar procedure, found the strongest effect of phonology on the initial guessing block. Thus, in order to maximize a possible effect, it is crucial to include such a guessing block into the current procedure. Therefore, before teaching participants the intended form-referent relation, they are asked to guess which of two words most likely refers to the picture they are shown. The hypothesis is that children will use the phonological make-up of the words to establish which word is the most probable candidate. After this guessing block, the children are taught the intended mappings according to the condition they were in. In the consistent group , children were taught word-referent pairs that conformed to the typical pattern in Dutch, which means that object pictures were coupled with disyllabic words, actions with monosyllabic items. The inconsistent group was taught the opposite pattern. Finally, an independent group received mixed picture-word combinations. It was hoped that in this way, a clearer picture would emerge than in Experiment 1b. Also, this group division allows us to compare results to those in Cassidy and Kelly (2001) and in Fitneva et al. (2009), who used the same design for English participants.

5.1 Method

Participants

Thirty-six monolingual Dutch-speaking primary school children were recruited from groups 3 and 4 (equivalent to US 1st and 2nd grade) of two primary schools in the West Friesland area of the Netherlands (20 girls, 16 boys). Each child was subjected to a phonological test battery identical to the one in Experiment 1b to control for atypical language development. Table 5 below shows the age in months and phonological (EMT, Klepel), digit span (WISC) and vocabulary (PPVT) test scores for all participants and for each test condition separately.

Because the independent group consisted only of participants recruited from the younger grade, who had only started their reading instruction three months before they were subjected to the test, this group has lower EMT and Klepel scores than the children in the other two conditions. This is not likely to corrupt the results because the experiment does not involve written language at all, and as can be seen in Table 5 the mean vocabulary score of this group is similar to that of the other groups. In the consistent condition, there were eight girls and four boys, the inconsistent and independent condition contained six girls and six boys each. Children were given a small toy for their participation.

	Total *M*	*SD*	Range	Consistent	Inconsistent	Independent
Age in months	82.6	6.8	70.5-95.3	83.3	83.7	80.7
EMT	24.36	17.8	5-61	29.17	27.58	16.33
Klepel	22.86	16.9	1-66	28.08	26.33	14.17
Digit span	9.92	10.7	8-12	10.67	9.27	9.73
PPVT WBQ	106.17	1.8	83-131	105.75	105.25	107.5

Table 5: Participant characteristics in Experiment 3.

Materials

Pseudowords

Sixteen test items were constructed of which eight conformed to phonological characteristics of Dutch verbs and eight to those of nouns. Verb-like items were monosyllabic, preferably contained a fricative in the onset, a labial in the word, and high vowels rather than low vowels. Noun-like items consisted of a disyllabic stem, preferably contained a bilabial or plosive in the onset, could have a final long vowel, schwa or schwa + nasal in the final syllable and contained low vowels rather than high (see Table 1 for an overview). An additional set of eight items was created to serve as fillers.

Item type	Test items 1	Test items 2	Fillers
verb-like	ruip	voek	wijg
	ries	goop	weep
	hook	hig	gol
	zil	vijs	guik
noun-like	gater	wamer	pineer
	banijn	mapier	safel
	wegel	tinger	fato
	bodee	nado	gappel

Table 6: Items Experiment 3.

Pseudowords were controlled for phonotactic probability (the likelihood of occurrence of sound sequences based on a corpus of spoken Dutch (Corpus Gesproken Nederlands: Nederlandse Taalunie, 2004) using the database of Adriaans, 2006) and lexical neighbourhood (the number of existing words differing only one or two phonemes from the test item according to a corpus of Dutch child-directed speech (Van Kampen, 1997). It was attempted to construct the pseudowords in such a way that there was as little overlap between the phonemes of the different words as possible, in order not to let similarity between items influence the difficulty of

the task. Nevertheless, not wanting to include any incorrect cues meant that some phonemes had to be used more than once within one group.

All stimuli were recorded by one female speaker. Items were preceded by the first person singular pronoun *ik* or the indefinite singular determiner *een*. The same token of a pseudoword was cut out of the speech stream once and attached to both introductory function words, so that there was no difference between the tokens of words in the different conditions.

Referents

Object pictures contained four unfamiliar objects taken from a series of storybooks (Busser & Schröder, 2005, 2006, 2007). It was important that pictures did not trigger a single lexical item because this could influence results, e.g. it could be easier to remember a referent-word pair if the referent is a well-known action than if the picture shows an unfamiliar action. In a trial session with six adults, it was tested whether the referents were sufficiently ambiguous. Two of the initial object pictures were found to elicit an identical noun response in several pilot participants. These pictures were consequently adapted. Action pictures were taken from the same storybooks and featured a girl performing an unfamiliar action. None of these elicited identical responses.

Pictures were presented through a slideshow on a laptop. Auditory stimuli were shown above each picture through small speaker icons and could be activated by clicking on them with a mouse. The order of the auditory stimuli was counterbalanced across conditions.

Design & Procedure

The aim of this experiment was to test both the effect of phonological cues on guessing (monosyllabic vs. disyllabic, controlled for other cues as in Table 1) and the effect of phonological cues on learning - identification and production (consistent vs. inconsistent). In the first block, participants were asked to guess which word most likely referred to the picture. In the learning phase, they were taught the 'correct' matching according to the condition they were in: consistent, inconsistent, or independent. The design is schematized in Table 6 below.

Condition	**Block 1: guess**	**Learning block**	**Block 2: identification**	**Block 3: production**
Consistent	ik zil ik bodee	*ik zil*	ik guik ik bodee *ik zil*	*ik zil*
Inconsistent	ik zil ik bodee	*ik bodee*	ik guik *ik bodee* ik zil	*ik bodee*
Independent	*half of the items are consistent, the other half inconsistent*			

Table 6: Design Experiment 3: example for an action picture (correct answer in italics).

Children were tested individually in a quiet room of their primary school. The procedure started with a training block, in which the child was familiarized with the test set-up. This phase consisted of four separately offered pictures, each shown with two auditory stimuli. The child was told that it had to decide which of the two computer-presented pseudowords referred to the picture. They could either repeat the stimulus they chose or point to the speaker icon that had produced the sound when it was clicked. After their choice, the experimenter always provided positive feedback. All children collaborated readily and understood what was asked of them.

The first test block consisted of eight pictures, four objects and four actions, each combined with two auditory stimuli. The experimenter introduced the test block by saying the equivalent of the following in Dutch: "Those where all pictures and words you already know. Now we are going to do the same with pictures and words you do not know. In a little while, we are going to learn what everything is called, but first you can just guess, ok?" For each of the eight pictures, the child had to choose which of two auditory stimuli it thought would be the most likely label.

After the first block, the child was told that it was going to learn the proper words for each picture. Thus, the same eight pictures were offered, but now with only the 'correct' auditory stimulus, according to condition. It was stressed that the child had to listen and watch carefully, because in the last block it would be asked to produce each pseudoword. Each pseudoword-picture pair was offered twice. This learning phase was followed by an identification block. Each picture was shown again with three auditory stimuli: the target, the alternative from the phonological guess block and a filler. The child was asked whether it remembered which word was the one they heard in the learning phase. The participant could either point to the correct speaker or repeat the word. Feedback followed after each choice.

The third block of trials tested the production of the target form. While once more going through the pictures the child was asked to produce the correct form for each one. Because of the difficulty of this task for the age group, the child was told that if it didn't remember anything of the word, it could click on the speaker above the picture, which would then produce the stimulus itself. However, the child was stimulated to produce any phoneme it remembered.

Only in the learning phase and the test blocks after that, the three conditions listed above became important. The hypothesis was that the consistent group would perform better than the other two groups. The independent group should score better than the inconsistent group, because four items do conform to the phonological pattern in Dutch.

Coding

In the first block, where participants have not received feedback yet, the 'phonological guess score' is measured. If the child picks the word that is consistent with phonological patterns in

Dutch, a score of 1 is rewarded for the answer, 0 otherwise. For the identification block, the accuracy of choice is coded. If the child chooses the word that was taught according to the group it was in, a score of 1 is given, 0 otherwise. For the production block, three scores are calculated. Strict scoring means that only when the child produces the correct target word fully it gets a score of 1, 0 otherwise. On the lax scoring criterion, a score of 1 is given if the child uses any correct phoneme of the target word. Finally, a percentage score in the production block was calculated by measuring the percentage of correct phonemes in the right place of the produced word compared to those in the target word.

5.2 Results

First, the results on the phonological guessing block will be discussed. Because there was no difference between conditions in this block, we interpret the data of all three groups at once. The mean percentage of choosing according to phonological cues was 52.4%. This does not differ from chance (50%) on a t-test.

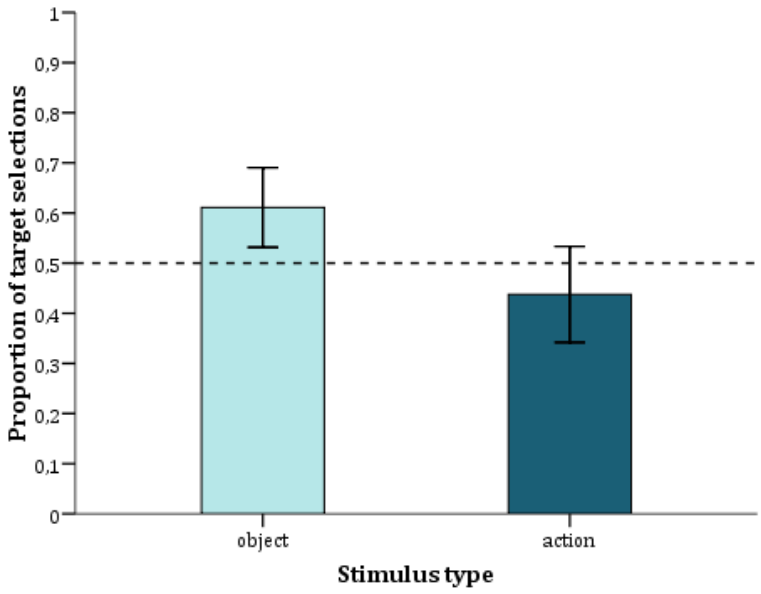


Figure 2: Proportion of target selections in Block 1 of Experiment 3 split out for picture type (T-bar represents the 95% confidence interval).

On objects, the score was 61.11% (*SD* 23.49) of correct choice according to phonology (i.e. they chose the disyllabic item if the picture showed an object). This is significantly different from chance (t [35] = 2.839, p = 0 .007). They chose the monosyllabic pseudoword if the picture contained an action in 43.75% (*SD* 28.27) of the cases (non-significant at $p < 0.05$). This means

participants did tend to choose according to phonology in the first block for objects, not for actions. The difference between the two groups is significant as well (t [70] = 2.834, p = 0.006).

The disyllabic item was chosen more often than the monosyllabic pseudoword in general, i.e. disregarding which was the target choice (58.7% disyllabic). However, this is not dependent of the picture type: on action trials, children chose the disyllabic item for 56.3% of the cases, for objects, 61.1%, which is a non-significant difference.

We now move on to the results on the measurements of learning. A one-way ANOVA on the three groups yields no significant result at $p < 0.05$ on any of the measures. Scores of each of the three groups are reported in Table 7.

		Identification		**Prod. (strict)**		**Prod. (pct)**		**Prod. (lax)**	
Condition	**N**	***M***	***SD***	***M***	***SD***	***M***	***SD***	***M***	***SD***
Consistent	12	76.04	20.9	7.29	6.4	15.38	13.94	21.88	23.3
Inconsistent	12	65.63	16.1	3.13	5.7	10.39	9.37	18.75	16.4
Independent	12	78.13	14.2	3.13	7.8	12.71	12.17	20.83	17.1
Total	36	73.26	17.7	4.51	6.8	12.82	11.81	20.49	18.7

Table 7: Mean percentages and standard deviations on learning measures Experiment 3 per group.

A plot of the identification block provides the following picture (Figure 3). The consistent group performed better than the inconsistent group on learning the correct word for each picture, but this difference was not significant at $p < 0.05$ on a two-sided t-test.

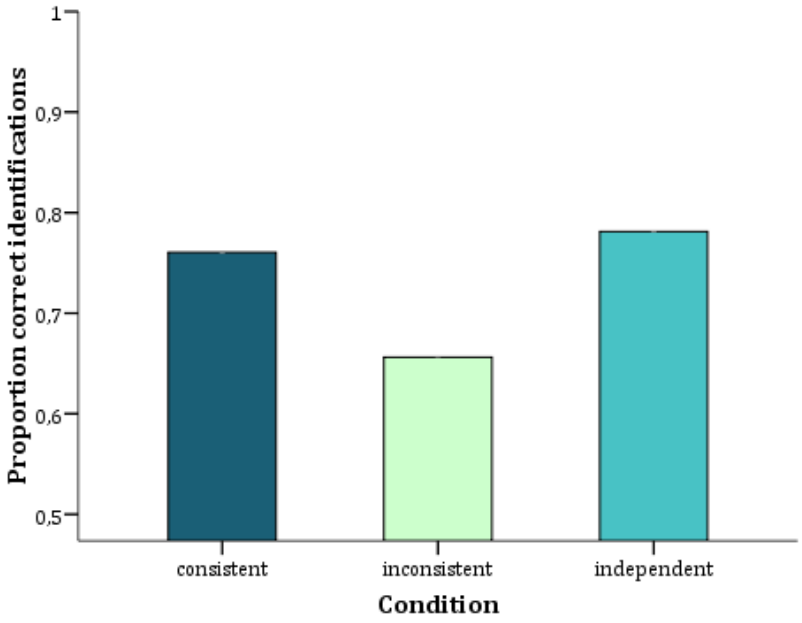


Figure 3: Proportion of correct identifications on Block 2 of Experiment 3 (please note that the y-axis does not start at zero).

Also, the independent group does not indicate the expected behaviour of performance between the scores of the consistent and inconsistent group: participants in this group actually seemed to score slightly better than those in the other two groups. This difference is not significant at $p < 0.05$. Thirteen out of 36 children made only 1 or 0 mistakes in this block, and these came from all three groups (consistent: N = 6, inconsistent: N = 2, independent: N = 5); the effect is thus probably not due to outliers in the independent group.

Figure 4 shows the performance at the third block. Results on this graph do render the expected behaviour: the consistent group scores best, and the independent group is between the consistent and the inconsistent. However, the difference between the three groups was not significant at $p < 0.05$.

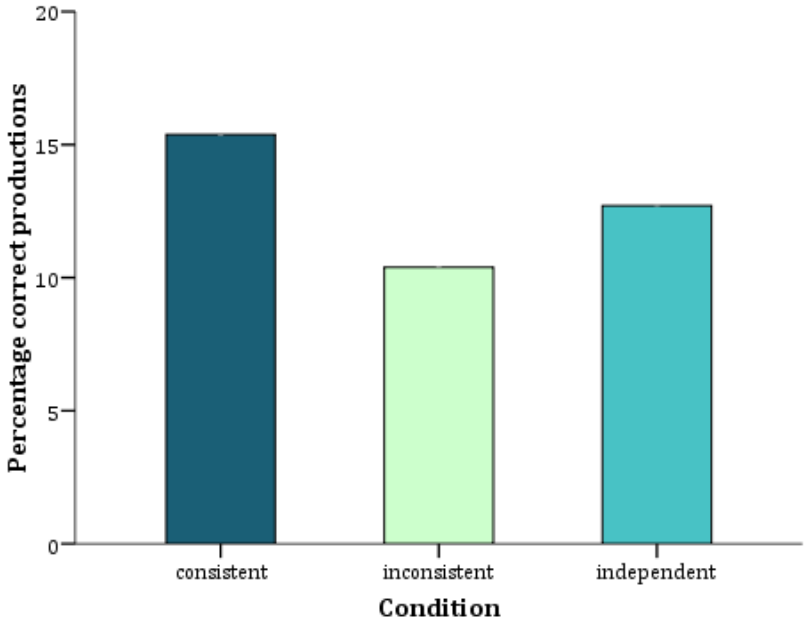


Figure 4: Percentage of correct productions on Block 3 of Experiment 3.

Again, when considering the data for just the object pictures, we find a significant difference between consistent and inconsistent in the identification block. The mean score of the consistent group on identification is here 85.42% (*SD* 19.82), of the inconsistent group 56.25 (*SD* 30.39), which is significant on a two-sided t-test with (t [22] = 2.785, p = 0.011). The same test on the strict production measurement yields 8.33% (*SD* 12.31) vs. 2.08% (*SD* 7.22), but this is not significant at $p < 0.05$. The percentage measurement is non-significant either with 14.27% (*SD* 14.41) vs. 7.27% (*SD* 10.93).

When looking only at the action pictures, no effect whatsoever can be found. On identification, the consistent group scores 66.67% (*SD* 28.57) versus inconsistent 75.00% (*SD* 23.84), not significant at $p < 0.05$. On production (strict), the consistent group reaches 6.25% (*SD* 11.31) where the inconsistent group scores 4.17% (*SD* 9.73), non-significant. An effect of group can thus only be established for object pictures.

In the two paragraphs above I looked at the consistent vs. the inconsistent group and the object vs. the action pictures (since the biggest effect was found for object pictures). When looking at both picture types together, however, a marginal effect of group on the strict production measure can be found as well, at t [22] = 1.685, p = 0.053 (one-tailed). None of the other measures yield a notable effect.

An ANOVA on only the objects for all three groups on block 2 and 3 yields the following results (means in Table 8). There is a significant effect only for the scores in the identification block, F [2, 33] = 5.360, p = 0.010. The strict production measure yields F [2, 33] = 1.222, p = 0.308; the lax production measure F [2, 33] = 1.081, p = 0.351; the percentage production measure F(2, 33) = 0.743, p = 0.483. Bonferroni post-hoc comparisons of the three groups revealed that on the identification block, the consistent group had significantly higher scores than the inconsistent group (M = 29.17, 95% CI [4.86, 53.47]) at p = 0.014. The independent group scored better than the inconsistent group at p = 0.042 (M = 25. 95% CI [0.07, 49.30]). No difference could be established between the consistent and the independent group.

		Identification		**Prod. (strict)**		**Prod. (pct)**		**Prod. (lax)**	
Condition	**N**	**Mean**	**SD**	**Mean**	**SD**	**Mean**	**SD**	**Mean**	**SD**
Consistent	12	85.42	19.82	8.33	12.31	14.27	14.41	22.92	27.09
Inconsistent	12	56.25	30.39	2.08	7.22	7.27	10.93	12.50	19.94
Independent	12	81.25	18.84	4.17	9.73	14.83	16.26	22.92	24.91
Total	36	74.31	26.38	4.86	10.04	12.13	14.07	19.44	23.99

Table 8: Mean percentages of all three groups on object pictures only, Blocks 2 and 3.

Finally, an item analysis was executed to reveal whether the effects we found could be due to certain stimuli. Table 9 shows the results for each of the items separately. Note that the first column shows the percentage that each pseudoword was chosen in accordance with the phonological characteristics, i.e. *voek* was chosen in the case of an action picture by 67% of the children; *banijn* was chosen in the case of an object picture by 44%.

	First guess		Identification		Prod (perc)	
Item	***M* (%)**	***SD***	***M* (%)**	***SD***	***M* (%)**	***SD***
banijn	44	51.1	78	42.8	10.94	29.24
bodee	56	51.1	100	0	17.78	33.53
gater	61	50.2	83	38.3	13.00	32.64
goop	39	50.2	61	50.2	4.78	15.99
hig	50	51.4	67	48.5	0	0
hook	44	51.1	83	38.3	5.94	17.91
mapier	61	50.2	94	23.6	39.83	38.90
nado	67	48.5	83	38.3	12.50	32.37
ries	56	51.1	44	51.1	7.39	24.38
ruip	56	51.1	78	42.8	11.11	28.73
tinger	61	50.2	78	42.8	16.72	29.72
vijs	39	50.2	67	48.5	26.39	38.80
voek	67	48.5	56	51.1	12.78	30.64
wamer	61	50.2	83	38.3	12.22	26.69
wegel	33	48.5	72	46.1	6.67	23.76
zil	44	51.1	44	51.1	7.11	13.72
Total	52	50.0	73	44.3	12.82	28.60

Table 9: Mean percentages per item on each test phase in Experiment 3.

Because the items that contain a schwa in the second syllable were thought to be possibly more verb-like than others, t-tests were run to reveal whether their inclusion could have corrupted results. Statistical comparisons demonstrate that the four items that contained a schwa-syllable did not differ significantly from the other stimuli on any of the measures on two-sided t-tests: in the first guess-block, the difference between schwa-items and others is non-significant with $p = 0.748$; on identification, $p = 0.420$; on production (strict), $p = 0.819$; on production (percentage), $p = 0.876$; on production (lax), $p = 0.874$.

5.3 Discussion

In this experiment, it was investigated whether an effect of phonological characteristics could be found on Dutch 7-year-old children's initial categorization and learning of pseudowords. It was found that only for object pictures, children seemed to rely on phonological characteristics of stimuli. This guided them in their first guesses towards which word was the most likely label for a picture, and influenced the accuracy with which they learned words.

Initially, one would think that a difference between learning nouns and verbs is not very surprising: numerable studies in the field of word learning in both children and adults have

pointed towards cognitive and perceptual biases for learning object names over action names (for an overview, see Woodward & Markman, 1998; a recent study is Piccin & Waxman, in press). However, though this may explain why there was a significant result on the identification block for object pictures, it cannot give account for the finding that the participants in this study showed a preference for choosing disyllabic items as noun labels over monosyllabic ones in the first guess-block of the test. Also, there can no longer be a possible relation to the target words as in Experiment 1b, because the objects and actions in this experiment were unfamiliar.

Furthermore, results in a similar experiment with English 7-year-olds established an advantage for verb learning over object learning (Fitneva et al., 2009). Their participants were found to relate words that were phonologically more typical for the verb class more easily to action pictures and than they matched typically noun-like stimuli to object pictures. On learning, there was only a significant effect for verb-like items, not for noun-like stimuli. Fitneva et al. ascribe this effect to statistical differences between the reliability of phonological versus distributional cues in the lexical classification of nouns and verbs that were found by Monaghan et al. (2005, 2007). According to these researchers, the same picture emerges in English, French, Japanese and Dutch: "phonological cues resulted in significantly more accurate and complete classification of verbs than nouns. Distributional cues provided more accurate and complete classification of nouns" (Fitneva et al., 2009: 972).

Even though corpus results point to a different direction, Dutch children do not seem to rely on phonological characteristics of words when guessing or learning the correct labels of actions. How could this be so? One option is that Dutch verbs almost always occur with inflection – only the first singular form and the imperative consist of only the stem of the verb, all other forms come with a morphological marker. English verb morphology, on the other hand, is very poor: in the present tense paradigm, the only inflection occurs in third person singular. Dutch children may therefore rely much earlier and much more on morphology in categorizing novel words than English children. This conforms to differences between English and Dutch found by Erkelens (2009) in her replication of Mintz (2006). Although English 12-month-olds were found to use the frequent co-occurrence of words to categorize verbs, equivalent frequent word frames were not used by Dutch 12- or 16-month-old infants. However, the study on Dutch 16-month-olds revealed that they do use frequent *morpheme* frames to categorize nouns and verbs. Morphemes (like verbal inflection) might be a more reliable indicator of category than words, because in Dutch, adverbial particles like *dan* 'then' and *ook* 'also' often intervene between proposed frames like *ik X niet* 'I X not', but cannot intervene between morpheme frames like *hij X-t* 'he X-s' (Erkelens, 2009). The same morphological difference between English and Dutch might be the cause for the contrasting result in this study as well: for verbs, the morphological

cues might be so strong that the language learning child does not have to pay any attention to the phonological form of the stem.

This option also explains why the object effect is found in both the phonological guessing block and the identification block. The null effect on production might be caused by the difficulty of having to produce a word after hearing it only three times: raw scores on this block show that there is a floor effect which means that results are not very informative. There was a trend in the desired direction however in that consistent participants seemed to perform better than children in the independent condition, while inconsistent participants obtained the lowest scores. Therefore, it is likely that if more participants were tested, we would have found a significant effect of condition on this block also.

There is one last issue that needs to be discussed here. The disyllabicity advantage for object pictures could in principle also be due to the lexical neighbourhood differences between monosyllabic (verb-like) and disyllabic stimuli. Monosyllabic items all had a lexical neighbourhood (occurring words that differ only one phoneme with the pseudoword) that included both nouns and verbs and the number of items in the neighbourhood were relatively high. Disyllabic items generally had a smaller neighbourhood which consisted only of nouns. So, we might say that children could be relatively sure that disyllabic words were noun labels because they sounded like similar nouns, while for monosyllabic words similar words were ambiguous to category. If the established effect is related to children's lexicon in this way, however, a relation between vocabulary and score on this experiment should be expected. All participants in Experiment 3 were subjected to the Peabody Picture Vocabulary Test, and no correlation between scores could be established (for Block 1: $r = 0.188$, $p = 0.272$; Block 2: $r = 0.220$, $p = 0.197$; Block 3 (pct.): $r = 0.307$, $p = 0.069$).

6. General discussion

Various studies in the last decade have proposed that there are correlations between phonological characteristics of words and the grammatical categories verb and noun. These correlations could in theory be helpful or even essential in acquiring those categories. The main goal of this study was to find out whether one premise of this hypothesis is fulfilled for speakers of Dutch: are adults and more importantly children sensitive to the association between phonological characteristics and the grammatical categories nouns and verbs?

To this purpose, several experiments were carried out involving the presentation of novel words to two test groups of adult speakers of Dutch and two groups of children between 6- and 8-years. Experiment 1 established that a group of 31 speakers of Dutch were indeed

aware of the number of syllable difference between Dutch nouns and verbs, in accordance with the findings in Don and Erkelens (2008). A follow-up with 8-year-olds failed to find a result on initial categorization of novel items nor on learning those words, in contrast to an equivalent test with English 3- to 6-year-olds (Cassidy and Kelly, 2001). This was related to differences between the infinitival form of the verb in English and Dutch: where in English, the infinitive is equivalent to the stem, in Dutch the infinitive is formed by stem + *-en*.

In order to solve this problem of testing the phonological characteristic number of syllables of words, Experiment 2 investigated whether it was possible to establish a sensitivity to minor phonological cues which were found to be good predictors of category when taken together in a computational learning algorithm by Durieux and Gillis (2001). Using only disyllabic stimuli varying on these less reliable cues, no effect of phonological characteristics could be found in a test with 34 adults. It is possible that including more than 6 items per category would have revealed an effect of these minor cues. However, it seems likely that the high percentage of noun assignment for the verb-like stimuli was triggered by using only disyllabic items: the effect of having more than one syllable might have forced participants to interpret each word to be more noun-like than verb-like. It would be interesting to see whether an effect of minor cues could be established if only monosyllabic items were used.

Because of the null-result of Experiment 2, in Experiment 3 it was decided to go back to testing the number of syllables. By presenting the novel words in both noun- and verb-contexts, it was possible to test the phonological characteristics of stems, thus avoiding the infinitive-problem of Experiment 1. A test with 36 6- and 7-year-olds resulted in a significant effect of phonological characteristics on initial categorization and learning of novel words only in noun contexts. Remember that the positive finding in Experiment 1 was established on the basis of noun characteristics only as well: disyllabic stimuli did yield a noun response above chance, but participants judged monosyllabic stimuli to be noun-like or verb-like with equal probability.

It seems thus that Dutch participants are indeed sensitive to one phonological characteristic differing between nouns and verbs, the number of syllables, but only in one direction: that disyllabic words are most probably nouns, whereas monosyllabic words can be either nouns or verbs. This stands in contrast with experiments involving speakers of English, where novel monosyllabic items were more likely to be categorized as verbs (Cassidy and Kelly, 2001; Fitneva et al., 2009). Although corpus studies found that phonological cues to category are as strong in Dutch as they are in English (Monaghan et al., 2007) or even stronger (Durieux & Gillis, 2001), this does not imply that speakers of this language actually make use of these properties: results of this study show that they might not in the case of verbs.

But why would language learners disregard any lead to solve the problem of determining the category of a novel word when we know that human cognition makes use of any mechanism or strategy it has to its service (Chater & Oaksfield, 1999)? The most logical option is that there is another cue to category that is stronger and more reliable than the phonology of an item. In Dutch, verbs almost always occur with inflection, whereas in English they do not. Whereas English children seem to look at the surrounding words in order to find distributional cues for category, Dutch children have more success in category assignment by attending to the surrounding morphemes. Erkelens (2009) finds a similar result when addressing Mintz (2003, 2006)'s frequent word frame hypothesis on Dutch infants. Whereas Mintz finds that English 12- and 16-month-olds make use of frequent word frames in distinguishing between noun and verb use for novel items, Dutch 16-month-olds could not be found to do so. However, when presenting the Dutch infants with frequent *morpheme* frames, they could find a difference between correct and incorrect use of novel nouns and verbs.

According to Monaghan et al. (2007: 291), word frames for verbs are less reliable than for nouns: "[v]erbs tend to occur in a greater range of contexts [than nouns] and hence distributional information is more likely to benefit from supplementary word-internal information about syntactic category, such as phonological information." However, when there are sufficient cues in morphemes, there is no need for this supplementary word-internal information, which seems to be the case for Dutch verbs. Monaghan et al. (2007) have tested whether there is a difference on the strength of phonological cues when morphology is stripped away in a corpus of English speech and found that the effect remains robust. I expect that for a corpus of Dutch, stripping away phonological information of verbs results in lower scores of successful classification.

Participants in the current study were found to use phonological information of nouns to categorize novel items. It seems thus that there is a sensitivity to phonological cues when that information is reliable enough. It might still be the case that this information is use to bootstrap into the grammatical categories. Testing more children might have resulted in a significant effect for phonological characteristics of verbs as well; however, the null result on adults means that this is not very likely. The fact that there are correlations between grammatical category and phonological information means that in theory, these correlations will be considered by the human brain: in other areas of perception, human sensitivity to probabilistic information has been found even if it is redundant (e.g., Kelly, 1992: 350; Chater & Oaksfield, 1999). Although phonological knowledge appears to have a relatively weak impact, the ability to form quick hypotheses about new words is crucial for children's rapid and successful language development (Carey, 1978).

7. Conclusion

In this thesis I have attempted to find evidence for the phonological bootstrapping hypothesis in speakers of Dutch. In theory, children could make use of correlations between the phonological makeup of words and their distinctive use in language in order to acquire the different grammatical categories. In order to find out whether this bootstrapping into syntactic categories actually occurs, first, sensitivity to the relation between phonology and category should be established. Dutch adults did tend to associate the number of syllables of a novel word, but not less reliable phonological characteristics, with noun or verb membership. Following up on this result, a child-friendly task was developed in which object or action pictures were shown with two stimuli each, either containing phonological characteristics of nouns or of verbs. A number of 36 7-year-olds were found to choose according to phonological typicality only for object pictures. An effect of phonological characteristics for verbs could not be found.

This result is explained by the idea that children will use any cue that is available to them in solving the problem of category assignment, which can be distributional, prosodic, semantic, or phonological; for verbs, however, phonological information seems to be overshadowed by distributional information, because of the relatively rich morphological structure of Dutch. Sensitivity to this morphological structure has been established in children as young as 16 months (Erkelens, 2009). However, the positive result of nouns means that phonology might still be adopted in acquiring categories.

Acknowledgements

Writing this thesis would not have been possible without the support from two primary schools, the *Jozef en Mariaschool* in Spierdijk and the *Pancratiusschool* in Oosterblokker. My heartfelt gratitude to the pupils, their teachers, supporting staff and directors. I would also like to thank the participants in the two adult experiments (but most of them have already received my appreciation in the form of chocolate).

The experiments in this thesis have benefited enormously from the lab meetings at the Utrecht Institute of Linguistics under direction of Frank Wijnen and from the input of my supervisors there, Annemarie Kerkhoff and Elise de Bree. Most of all I would like to thank my primary supervisor Jan Don for his support and guidance during all stages of this research.

Finally, thank you to my friends and family for bearing with me during the writing of this thesis.

References

Adriaans, F. (2006). *PhonotacTools: A tool for computing phonotactic probabilities*. Unpublished manuscript. Utrecht, UiL-OTS.

Baayen, R.H. , Piepenbrock, R. & Van Rijn, H. (1993). The CELEX lexical database. Philadelphia, PE: Linguistic Data Consortium.

Brus, B. Th. & Voeten, M. J. M. (1972). *Een-minuut test. Vorm A en B.* Nijmegen: Berkhout Testmateriaal.

Busser, M. & R. Schröder (2006). *Jan en Noortje gaan naar zwemles en andere verhaaltjes*. Vianen [etc.]: The House of Books.

Busser, M. & R. Schröder (2007). *Het telboek van Jan en Noortje*. Vianen [etc.]: The House of Books.

Busser, M. & R. Schröder (2005). *De ettertjes*. Houten: Van Holkema & Warendorf.

Carey, S. (1978). The child as a word learner. In J. Bresnan, G. Miller & M. Halle (eds.), *Linguistic theory and psychological reality*, 264-293. Cambridge, MA: MIT Press.]

Cassidy, K. W. & M. H. Kelly (1991). Phonological information for grammatical category assignments. *Journal of Memory and Language, 30*, 348-369.

Cassidy, K. W. & M. H. Kelly (2001). Children's use of phonology to infer grammatical class in vocabulary learning. *Psychonomic Bulletin and Review, 8*, 519-523.

Clark, E. V. (2003). *First language acquisition*. Cambridge: Cambridge University Press.

Don, J. & M. Erkelens (2008). Possible phonological cues in categorial acquisition: Evidence from adult categorization. In U. Ansaldo, J. Don & R. Pfau (eds.), *Parts of Speech: Descriptive tools, theoretical constructs*, 670-682. John Benjamins: Amsterdam.

Don, J. & E. van Lier (forthcoming). Categorization and derivation in flexible and differentiated languages. In J. Rijkhoff & E. van Lier (eds.), *Flexible word classes: a typological study of underspecified parts-of-speech*. Oxford: Oxford University Press.

Dunn, L. M. & L. M. Dunn (2005). *Peabody Picture Vocabulary Test (PPVT)-III-NL*. Nederlandse versie: L. Schlichting. Lisse: Harcourt test Publishers.

Durieux, G. & S. Gillis (2001). Predicting grammatical classes from phonological cues: an empirical test. In J. Weissenborn & B. Höhle (eds.), *Approaches to bootstrapping: Phonological, lexical, syntactic and neurophysiological aspects of early language acquisition* (vol. 1), 189-229. Amsterdam: John Benjamins.

Echols, C. H. (2001). Contributions of prosody to infants' segmentation and representation of speech. In J. Weissenborn & B. Höhle (eds.), *Approaches to bootstrapping: Phonological, lexical, syntactic and neurophysiological aspects of early language acquisition* (vol. 1), 25-26. Amsterdam: John Benjamins.

Erkelens, M. A. (2009). *Learning to categorize verbs and nouns: Studies on Dutch.* Dissertation. Utrecht: LOT.

Erkelens, M. A., A. O. Kerkhoff & E. H. de Bree (in preparation). *Dutch 16-month-olds categorize nonsense words based on frequent morpheme frames.* Unpublished manuscript, Amsterdam, ACLC / Utrecht, UiL-OTS.

Ernestus, M. (2005). Analogical effects in reading Dutch verb forms. *Memory and Cognition, 33*, 1160-1173.

Fitneva, S. A., M. H. Christiansen & P. Monaghan (2009). From sound to syntax: phonological constraints on children's lexical categorization of new words. *Journal of Child Language, 36*, 967-997.

Fisher, C., & H. Tokura (1996). Prosody in speech to infants: Direct and indirect acoustic cues to syntactic structure. In J. L. Morgan & K. Demuth (eds.), *Signal to Syntax: Bootstrapping from Speech to Grammar in Early Acquisition*, 343-364. Mahwah, NJ: Erlbaum.

Francis, N., & H. Kucera (1982). *Frequency analysis of English usage: Lexicon and grammar.* Boston: Houghton Mifflin.

Gasser, M., N. Sethuraman & S. Hockema (2009). Iconicity in expressives: An empirical investigation. In S. Rice & J. Newman (eds.), *Experimental and empirical methods*, 1-18. Stanford, CA: CSLI Publications.

Ghyselinck, M., R. Custers & M. Brysbaert (2003). Age-of-acquisition ratings for 2332 Dutch words from 49 different categories. *Psychologica Belgica*, 43-3, 181-214.

Ghyselinck, M., W. De Moor & M. Brysbaert (2000). Age-of-acquisition ratings for 2816 Dutch four- and five-letter nouns. *Psychologica Belgica*, 40-2, 77-98.

Gleitman, L. (1990). The structural sources of verb meanings. *Language Acquisition, 1,* 3-55.

Gómez, R. & L. Lakusta (2004). A first step in form-based category abstraction by 12-month-old infants. *Developmental Science, 7*, 567-580.

Hengeveld, K., J. Rijkhoff & A. Siewierska (2004). Parts-of-speech systems and word order. *Journal of Linguistics, 40,* 527-570.

Hinton, L., J. Nichols & J. Ohala (eds.) (1994). *Sound symbolism.* Cambridge: Cambridge University Press.

Höhle, B., J. Weissenborn, D. Kiefer, A. Schulz & M. Schmitz (2004). Functional Elements in Infants' Speech Processing: The Role of Determiners in the Syntactic Categorization of Lexical Elements. *Infancy, 5*, 3, 341-353.

Kawamoto, A. H., W. T. Farrar IV. & M. Overbeek (1990). *Effect of syntactic context on naming bisyllabic words.* Poster presented at the 31st Annual Meeting of the Psychonomic Society, New Orleans, LA.

Kelly, M. H. (1988a). Phonological biases in grammatical category shifts. *Journal of Memory and Language, 27*, 343-358.

Kelly, M. H. (1992). Using sound to solve syntactic problems: The role of phonology in grammatical category assignments. *Psychological Review, 99*, 2, 349–364.

Kerkhoff, A. O., M. A. Erkelens, E. H. de Bree & F. Wijnen (in preparation). Dutch 16-month-olds categorize nonsense words based on frequent morpheme frames. (manuscript, ACLC/ UiL-OTS).

Kohnstamm, G. A.,A. M. Schaerlaekens & A. de Vries (1981). *Nieuwe streeflijst woordenschat voor 6-jarigen gebaseerd op onderzoek in Nederland en België*. Lisse: Swets & Zeitlinger BV.

Kohnstamm, G. A., M. Lejaegere & A. M. Schaerlaekens (1999). *Streeflijst woordenschat voor 6-jarigen*. Lisse: Swets & Zeitlinger BV.

Krikhaar, E.M., & F. Wijnen (1996). Children's categorization of novel verbs: syntactic cues and semantic bias. In E.V. Clark (ed.), *Proceedings of the Annual Child Language Research Forum*, vol. 27. Stanford, CA: CSLI Publications.

Kuhl, P. K., K. A. Williams, F. Lacerda, K. N. Stevens & B. Lindblom (1992). Linguistic experience alters phonetic perception in infants by 6 months of age. *Science, 255*, 606–608.

Landau, B., & L. R. Gleitman (1985). *Language and experience: Evidence from the blind child.* Cambridge, MA: Harvard University Press.

Langacker. R. (1987). *Foundations of cognitive grammar, Vol. 1.* Stanford, CA: Stanford University Press.

Maratsos, M., & M. Chalkley (1980). The internal language of children's syntax: the ontogenesis and representation of syntactic categories. In K. Nelson (ed.), *Children's Language* (volume 2), Gardner Press, New York. 127-214.

Maratsos, M. (1982). The child's construction of grammatical categories. In E. Wanner & L.R. Gleitman (eds.), *Language Acquisition: the State of the Art*, Cambridge University Press, Chicago. 240-266.

Maratsos, M. (1998). The acquisition of grammar. In D. Kuhn & R. Siegler (eds.), *Cognition, Perception, and Language* (volume 2 of 'Handbook of child psychology'). Hoboken, NJ: Wiley. 421-466

Mintz, T.H. (2003). Frequent frames as a cue for grammatical categories in child directed speech. *Cognition, 90*, 91-117.

Mintz, T.H. (2006). Finding the verbs: Distributional cues to categories available to young learners. In K. Hirsh-Pasek & R. M. Golinkoff (Eds.), *Action meets word: How children learn verbs*. Oxford: Oxford University Press.

Monaghan, P., N. Chater & M. H. Christiansen (2005). The differential contribution of phonological and distributional cues in grammatical categorization. *Cognition, 96*, 143-182.

Monaghan, P., M. H. Christiansen & N. Chater (2007). The phonological-distributional coherence hypothesis: Cross-linguistic evidence in language acquisition. *Cognitive Psychology*, 55, 259-305.

Morgan, J. L., & K. Demuth (1996). *Signal to Syntax: Bootstrapping from Speech to Grammar in Early Acquisition*, Mahwah, NJ: Erlbaum.

Mosel, U., & E. Hovdhaugen (1992). *Samoan reference grammar*. Oslo, Universitetsforlaget AS.

Nänny, M., & O. C. M. Fischer. *Form Miming Meaning: Iconicity in Language and Literature.* Amsterdam: Benjamins.

Nederlandse Taalunie (2004). *Corpus Gesproken Nederlands*.

Nygaard, L. C., A. E. Cook & L. L. Namy (2009). Sound to meaning correspondences facilitate word learning. *Cognition, 112*, 181-186.

Olguin, R., & M. Tomasello (1993). Twenty-five-month-old children do not have a grammatical category of verb. *Cognitive Development, 8*, 245-272.

Oller, J.W. (2005). Common ground between form and content: The pragmatic solution to the bootstrapping problem. *The Modern Language Journal, 89*, i, 92-114.

Piccin, T. B. & S. R. Waxman (in press). Why nouns trump verbs in word learning: New evidence from children and adults in the Human Simulation Paradigm. *Language Learning and Development.*

Pinker, S. (1984). *Language learnability and language development.* Cambridge, MA: Harvard University Press.

Pinker, S. (1987). *The bootstrapping problem in language acquisition.* Hillsdale, NJ: Lawrence Erlbaum.

Pinker, S. (1989). *Learnability and cognition: The acquisition of argument structure*. Cambridge, MA: MIT Press.

Polka, L., & J. Werker. Developmental changes in perception of non-native vowel contrasts. *Journal of Experimental Psychology: Human Perception and Performance*, *202*, 321-335.

Ramachandran, V. S., & E. M. Hubbard (2001). Synaesthesia: A window into perception, thought and language. *Journal of Consciousness Studies*, *8*, 3 - 34.

Redington, M., N. Chater & S. Finch (1998). Distributional information: a powerful cue for acquiring syntactic categories. *Cognitive Science, 22*, 425-469.

Rivera-Gaxiola, M., J. Silva-Pereyra & P. K. Kuhl (2005). Brain potentials to native and non-native speech contrasts in 7- and 11-month-old American infants. *Developmental Science, 8*, 162–172.

De Saussure, F. (1916/1959). *Cours de linguistique générale*. Paris: Payot. / 1966. *Course in general linguistics* (tr. W. Baskin). New York: McGraw-Hill.

Schaerlaekens, A. M., & S. Gillis (1987). *De taalverwerving van het kind*. Groningen: Wolters-Noordhoff.

Shi, R. S., J. L Morgan & P. Allopenna (1998). Phonological and acoustic bases for earliest grammatical category assignment: A cross-linguistic perspective. *Journal of Child Language, 25(1)*, 169-201.

Soreno, J. A., & A. Jongman (1990). Phonological and form class relations in the lexicon. *Journal of Psycholinguistic Research, 19*, 387-404.

Storkel, H. (2001). Learning new words: phonotactic probability in language development. *Journal of Speech, Language and Hearing Research, 44*, 1321-1337.

Storkel, H. (2003). Learning new words II: phonotactic probability in verb learning. *Journal of Speech, Language and Hearing Research, 46*, 1312-1323.

Tomasello, M. (1992). *First verbs: A case study of early grammatical development.* Cambridge: Cambridge University Press.

Tomasello, M., & R. Olguin (1993). Twenty-three-month-old children have a grammatical category of noun. *Cognitive Development, 8*, 451-464.

Tomasello, M. (2000). Do young children have adult syntactic competence? *Cognition, 74*, 209-253.

Trommelen, M. (1989). Lettergreepstruktuur en woordkategorie. *De Nieuwe Taalgids, 82*, 1, 64-77.

Van den Bos, K.P., Spelberg, H.C.L., Scheepstra, A.J.M. & De Vries, J.R. (1994). *De KLEPEL. Een test voor de leesvaardigheid van pseudo-woorden.* Nijmegen: Berkhout Testmateriaal.

Van Kampen, N.J. (1997). *First Steps in Wh-movement*. Delft: Eburon.

Wechsler, D. (2002). *Wechsler Intelligence Scale for Children, 3rd edition*. Nederlandse versie: WISC-III NL, door W. Kort, M. S. Schittekatte, M. Bosmans, E. L. Compaan, P. H. Dekker, G. Vermeir & P. Verhaeghe (2005). NIP dienstencentrum, Amsterdam.

Weissenborn, J. & B. Höhle (2001). Introduction. In Weissenborn, J., & B. Höhle (eds.), *Approaches to Bootstrapping*, pages? Amsterdam: John Benjamins.

Werker, J. F., & R. C. Tees, (1984). Cross-language speech perception: Evidence for

perceptual reorganization during the first year of life. *Infant Behavior and Development, 7*, 49–63.

Woodward, A. L. & E. M. Markman (1998). Early word learning. In D. Kuhn & R. Siegler (eds.), *Cognition, Perception, and Language* (volume 2 of 'Handbook of child psychology'), 371-420. Hoboken, NJ: Wiley.

Appendix

I: Experiment 1a.

I 1.1. Instruction

'Je ziet in de bijlage een lijst met nonsenswoorden. Het is de bedoeling dat jij zegt van elk woord of je denkt dat het de stam van een werkwoord is, of dat het de stam van een zelfstandig naamwoord is. Voor de duidelijkheid, de stam van een werkwoord is de vorm die je ziet na ik ... *; bijvoorbeeld ik zie, ik loop, ik ruik : zie, loop en ruik zijn hierbij de stammen. De stam van een naamwoord is de vorm die je ziet na* een ... *; bijvoorbeeld een boek, een kip, een lamp : boek, kip en lamp zijn hierbij de stammen.*

Je mag je keuze voor werkwoord of naamwoord aangeven door een kruisje in de betreffende kolom te zetten. Je moet voor elk woord één keuze maken. Denk er niet te lang over na, volg gewoon je eerste intuïtie.'

'In the attachment, you will find a list of pseudowords. I would like you to note for each word whether you think it is a verb stem or a noun stem. For the sake of clarity, the stem of a verb is the form that you see after *I* ...; for example, I see, I walk, I smell; see, walk and smell are the stems here. The stem of a noun is the form that you see after *a* ...; for example, a book, a chicken, a lamp; book, chicken and lamp are the stems here.

You can signify your verb or noun choice by putting a cross in the respective column. You have to make one choice for each word. Don't think too long, just follow your first intuition.'

I 1.2. Example of a fill-in form (version 1 of 4)

Naam: Moedertaal Nederlands? Ja/Nee
Datum: Dyslectisch? Ja/Nee
Geboortedatum:

Versie 1

		Werkwoord ("Ik ..")	Naamwoord ("Een .. "")
1	kloep		
2	boogst		
3	snoeg		
4	doelem		
5	pluig		
6	kellem		
7	donkam		
8	drauf		
9	krei		
10	giveno		
11	riele		
12	balter		
13	prellon		
14	ploren		
15	blaap		
16	meliens		
17	boele		
18	book		
19	smoza		
20	slamper		
21	frep		
22	fidong		
23	plif		
24	klui		
25	wadim		
26	sook		
27	krile		
28	kleest		
29	nado		
30	wirfem		
31	strempa		
32	taaf		
33	dieg		
34	pla		
35	pierst		
36	zierem		
37	bolee		
38	diel		
39	snef		
40	nort		

Opmerkingen:

I 2: Experiment 1b

I 2.1 Target word characteristics

Table 10 shows the target words used in Experiment 1b and a set of their characteristics: the CELEX-frequency (Baayen et al., 1993), the percentage of Dutch 6-year-olds who according to the Streeflijst know each word (Kohnstamm et al., 1981, 1999), and their estimated age of acquisition (AoA) where this information was available (Ghyselinck et al., 2000, 2003). In the list of verbs, it is also noted whether the verb is a change-of-state (CoS) verb as opposed to an action verb.

Scene	Nouns		CELEX-freq. (log)	Streeflijst '81	Streeflijst '99	AoA '00	AoA '03
9	*appel*	'apple'	1.2	98%	99%	3.7	4.0
7	*beker*	'mug'	1.2	96%	95%	5.2	4.6
5	*brommer*	'moped'	0.7	94%	89%		7.3
4	*egel*	'hedgehog'	0.5	82%	90%	6.2	5.9
6	*kikker*	'frog'	0.9	96%	97%		
3	*potlood*	'pencil'	1.1	100%	99%		4.4
8	*schildpad*	'turtle'	0.8	96%	92%		5.8
12	*theepot*	'teapot'	1.1	97%	99%		
11	*tijger*	'tigre'	0.8	92%	95%		6.2
10	*varken*	'pig'	1.4	97%	98%		4.8
1	*vogel*	'bird'	1.9	98%	96%	4.1	
2	*wortel*	'carrot'	1.6	93%	90%		4.9
Scene	**Verbs**		**CELEX-freq. (log)**	**Streeflijst '81**	**Streeflijst '99**	**AoA '00**	**AoA '03**
1	*buigen*	'to bow'	1.8	97%	96%		
2	*dansen*	'to dance'	1.8	100%	93%		
7	*drijven* (CoS)	'to float'	1.9	88%	92%		
8	*glijden* (CoS)	'to glide'	1.9	98%	89%		
9	*groeien* (CoS)	'to grow'	2.1	97%	100%		
3	*kleuren*	'to colour'	1.3	98%	98%		
4	*likken*	'to lick'	1.3	97%	93%		
5	*rijden*	'to ride'	2.3	100%	98%		
10	*rollen* (CoS)	'to roll'	1.7	100%	93%		
6	*springen*	'to jump'	2.1	94%	95%		
11	*vallen* (CoS)	'to fall'	2.8	100%	97%		
12	*zinken* (CoS)	'to sink'	1.3	78%	85%		

Table 10. Target words of Experiment 1b and their characteristics: verbs divided into Change of State (CoS) and action, translations, log frequencies, percentage of 6-year-olds reported to have acquired these words, and reported age of acquisition (AoA) if available.

Table 11 below presents the matching of nouns and verbs, the animacy of the subject in the scene (I: inanimate, A: animate), the verb type (action or change of state) and CELEX frequencies.

Stimnr	Scene	Animacy	Verbtype	CELEX freq.
1	vogel buigen	A	Action	1.9 < 1.8
2	wortel dansen	I	Action	1.6 < 1.8
3	potlood kleuren	I	Action	1.1 < 1.3
4	egel likken	A	Action	0.5 < 1.3
5	brommer rijden	I	Action	0.7 < 2.3
6	kikker springen	A	Action	0.9 < 2.1
7	beker drijven	I	CoS	1.2 < 1.9
8	schildpad glijden	A	CoS	0.8 < 1.9
9	appel groeien	I	CoS	1.2 < 2.1
10	varken rollen	A	CoS	1.4 < 1.7
11	tijger vallen	A	CoS	0.8 < 2.8
12	theepot zinken	I	CoS	1.1 < 1.3

Table 11. Scenes, animacy of objects, verb types and frequencies, Experiment 1b.

I 2.2 Pseudoword characteristics: phonotactic probability

Table 12 demonstrates the phonotactic probability (the likelihood of occurrence of sound sequences based on a corpus of spoken Dutch (Corpus Gesproken Nederlands: Goddijn & Binnenpoorte, 2003) using the database of Adriaans, 2006) for each of the pseudowords.

Pseudoword	Typicality	Pseudoword	Typicality
strempa	-15.807	fidong	-15.939
donkam	-16.412	smoza	-18.203
prellon	-15.830	blaap	-15.488
taaf	-16.852	wadim	-14.790
frep	-14.307	sook	-13.349
pluig	-17.954	krei	-14.664
Means	-16.194		-15.405

Table 12: Pseudowords and phonotactic probability items, Experiment 1b.

I 2.3 Pseudowords matched to target words

Table 13a and b show the matching of pseudowords and target words in Experiment 1b. Six items are matched to referents consistent (cons) with phonological characteristics of Dutch (disyllabic for noun (N) referents, monosyllabic for verb (V) referents) and six items are matched inconsistent with typical phonological characteristics.

Group 1							
Item	**Match N**	**Scene nr.**	**Meaning**	**Item**	**Match V**	**Scene**	**Meaning**
strempa	cons	9 (CoS)	appel	fidong	incons	2	dansen
donkam	cons	1	vogel	smoza	incons	8 (CoS)	glijden
prellon	cons	7 (CoS)	beker	blaap	incons	3	kleuren
taaf	incons	5	brommer	wadim	cons	10 (CoS)	rollen
frep	incons	11 (CoS)	tijger	sook	cons	6	springen
pluig	incons	4	egel	krei	cons	12 (CoS)	zinken

Table 13a. Pseudowords matched to target words for Group 1, Experiment 1b.

Group 2							
Item	**Match N**	**Scene nr.**	**Meaning**	**Item**	**Match V**	**Scene**	**Meaning**
fidong	cons	2	wortel	strempa	incons	9 (CoS)	groeien
smoza	cons	8 (CoS)	schildpad	donkam	incons	1	buigen
blaap	cons	3	potlood	prellon	incons	7 (CoS)	drijven
wadim	incons	10 (CoS)	varken	taaf	cons	5	rijden
sook	incons	6	kikker	frep	cons	11 (CoS)	vallen
krei	incons	12 (CoS)	theepot	pluig	cons	4	likken

Table 13b. Pseudowords matched to target words for Group 2, Experiment 1b.

II: Experiment 2

II 1 Instruction

*'Ik ga nu een paar niet-bestaande woorden voorlezen, en jij moet zeggen of je vindt dat het een werkwoord is (zoals ik **verstop** of ik **bedoel**) of een naamwoord (zoals een **kikker** of een **piloot**). Het is de bedoeling dat je het eerste kiest wat in je opkomt: werkwoord of naamwoord.'*

Right now, I am going to read to you some non-existing words, and you have to say whether you think it is a verb (like I **hide** or I **mean**) or a noun (like a **frog** or a **pilot**). I would like you to choose the first thing that comes into your head: verb or noun.

II 2 Example of a fill-in form (Version 1 of 4)

Naam: Moedertaal Nederlands? Ja/Nee
Datum: Dyslectisch? Ja/Nee
Geboortedatum:

Versie 1

		Werkwoord ("Ik ...")	Naamwoord ("Een ...")
1	prellon		
2	smozarp		
3	doelem		
4	tisoel		
5	bolee		
6	rideer		
7	strempa		
8	bronkasp		
9	giveno		
10	segie		
11	tikep		
12	prelbans		
13	meliens		
14	rieluf		
15	nado		
16	dranomp		
17	balter		
18	lebik		
19	slamper		
20	flibolt		

Opmerkingen:

III Experiment 3

III 1 Pseudoword characteristics

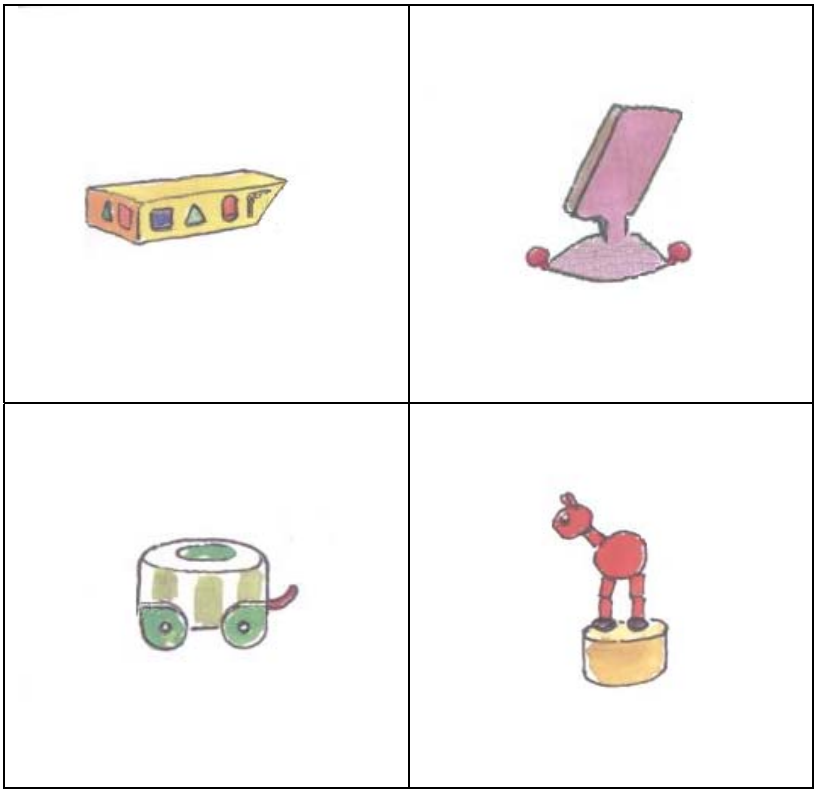
Table 14 presents the pseudowords in Experiment 3 per group, their phonotactic probability (typicality), the number of neighbours in CELEX differing one phoneme and the matching to pictures according to condition: consistent (C), inconsistent (I) or independent, neutral (N). In the independent condition, each word was matched once to an object, once to an action picture, so that each stimulus word would be tested consistently and inconsistently an equal number of times. Naturally, each child only learned one word per picture.

		Items	Typicality (log)	Neighbours	C	I	N
1 Items	**Verb-like**	ruip	-1,72	4	5	3	6/3
		ries	-1,63	5	7	4	7/3
		hook	-0,95	8	6	1	5/1
		zil	-1,41	7	8	2	8/1
	Noun-like	gater	-1,11	3	4	6	3/6
		banijn	-1,35	1	3	5	3/7
		wegel	-1,04	2	2	8	1/5
		bodee	-1,43	0	1	7	1/8
	Means		-1,329	3,75			
2 Items	**Verb-like**	voek	-1,30	7	6	4	4/5
		goop	-1,29	5	8	3	4/8
		hig	-1,39	2	5	2	2/6
		vijs	-1,64	3	7	1	2/7
	Noun-like	wamer	-1,10	3	3	5	5/4
		mapier	-1,26	2	4	7	8/4
		tinger	-1,10	1	1	6	6/2
		nado	-1,36	1	2	8	7/2
	Means		-1,300	3			
3 Fillers	**Verb-like**	wijg	-1,18	6	5/3	5/3	5/4/6/3
		weep	-1,20	1	7/8	3/5	3/7/7/2
		gol	-1,31	11	6/7	1/2	7/3/2/7
		guik	-1,52	3	8/2	8/7	2/6/1/5
	Noun-like	pineer	-1,27	1	4/6	4/6	4/5/3/6
		safel	-1,36	1	3/4	7/4	8/1/4/8
		fato	-1,42	1	2/1	2/8	6/2/5/1
		gappel	-1,01	1	1/5	6/1	1/8/8/4
	Means		-1,284	3,125			

Table 14. Pseudowords per group, phonotactic probability, number of neighbours and picture match.

III 2 Stimulus pictures

Objects (from left to right and top to bottom, picture number 1, 2, 3 and 4):



Actions (from left to right and top to bottom, picture number 5, 6, 7 and 8):