# Acquisition of Morpho-Phonology:
## the Dutch voicing alternation

Published by

LOT
Janskerkhof 13
3512 BL Utrecht
The Netherlands

phone: +31 30 253 6006
fax: +31 30 253 6406
e-mail: lot@let.uu.nl
http://www.lotschool.nl

Cover illustration: 'One pomegranate and two plums' by Oren Sadeh-Leicht.
Oil on canvas, 25 x 25 cm.

ISBN 978-90-78328-35-3
NUR 616

# ACQUISITION OF MORPHO-PHONOLOGY: THE DUTCH VOICING ALTERNATION

## Verwerving van Morfo-fonologie: de Nederlandse stemalternantie

(met een samenvatting in het Nederlands)

PROEFSCHRIFT

ter verkrijging van de graad van doctor
aan de Universiteit Utrecht
op gezag van de rector magnificus, prof. dr. W.H. Gispen,
ingevolge het besluit van het college voor promoties
in het openbaar te verdedigen
op maandag 27 augustus 2007
des middags te 4.15 uur

door

Annemarie Odilia Kerkhoff

geboren op 4 november 1974 te Nijmegen

Promotoren: prof. dr. W. Zonneveld
prof. dr. R.W.J. Kager

## Contents

**Acknowledgments**

This dissertation would not have been written without the help and support of my supervisors, René Kager and Wim Zonneveld. Thank you both for giving me the space and time I needed. I am particularly grateful to René for teaching me that I should not put aside problems as trivial too quickly. I also want to thank Wim for discussions and reading the manuscript so carefully.
The members of our project group also contributed to the end result. It was a pleasure to work with Suzanne van der Feest, thank you for your moral support and the fun times we had. Paula Fikkert was always an inspiration to me, thank you for your confidence and looking out for me. I have learned a lot from working with Tania Zamuner, who gave me experimental and statistical advice and proofread parts of this dissertation. I would also like to thank Anne Cutler, Elisabeth Johnson, and Dan Swingley for useful comments and interesting discussions during our 'project meetings' at the MPI.

I want to express my gratitude to the Dutch Organisation for Scientific Research NWO for funding this research. The Utrecht institute of Linguistics UiL-OTS has funded several courses and international conference visits. Among the staff at UiL-OTS there are several people who have helped me over the years. I would like to thank Jacqueline van Kampen, Mieke Trommelen and Janet Grijzenhout for their support. I also want to thank Hugo Quené, Bert Schouten, Maaike Schoorlemmer, Sergio Baauw and Mario van den Visser for their help. Finally, thanks to all the members of the Utrecht Dyslexia and SLI project (Petra van Alphen, Elise de Bree, Ellen Gerrits, Jan de Jong, Frank Wijnen, Carien Wilsenach and Marjolein van Woudenberg).

I want to thank all the children and adults who participated in the experiments, the children's parents and the staff of day-care centres and schools in Utrecht, Goor and Leiden (especially Cor van Rijswijk, director of the WG v/d Hulstschool and 'Janneke van de Flattekatten'). I am very grateful to my mother and several friends for helping me to find subjects. Thank you Hiske, Naomi and all the others.

Various people I have met during these years have also made a difference. I would like to thank Joe Pater, Paul Boersma, Harald Clahsen, Wolfgang Dressler and Steven Gillis for helpful comments. Professor Johan Taeldeman provided me with data and kindly agreed to meet in Ghent. Walter Daelemans has helped me understand more about computational modelling, while Emmanuel Keuleers deserves special thanks for showing an interest in voicing alternations and agreeing to work together.

My ‘phon & phon’ colleagues at UiL-OTS have played an important role. First, I want to thank my two roommates: Judith van Wijk for helping me with just about anything (including keeping track of important dates and events that I should have remembered myself), and Koen Sebregts for interesting discussions. Brigit van der Pas gave me advice and moral support, and even tested her own niece for me. Esther Janse deserves special thanks for being such a friendly, reliable and knowledgeable colleague. Willemijn Heeren and Maya van Rossum have kindly helped me with acoustic analyses and recording. I have also enjoyed spending time with Ivana Brasileiro, Paola Escudero, Anna Kijak and Shakuntala Mahanta. Four ‘non-phon’ colleagues also deserve mention; thank you Sharon Unsworth for critical comments, Elma Blom for discussions and collecting data, Nada Vasic and Nicole Grégoire for advice and help.

It was great to have so many colleagues and friends to share buildings, rooms, discussions and beers at Lokaal 9 with. Thank you Ines Antón-Mendéz, Anna Ashbury, Fabian Battaglini, Alexis Dimitriadis, Berit Gehrke, Nino Grillo, Silke Hamann, Judith Kamalski, Cem Keskin, Joke de Lange, Erica van Lente, Dimitra Papangeli, Anca Sevcenco, Giorgos Spathas, Roberta Tedeschi, Willemijn Vermaat, Evangelia Vlachou, Shalom Zuckerman and all the others. Among my fellow linguists I also want to thank Edith Schlag for being such an inspiring role model. Thanks also go to my LAIOO and Eurodoc colleagues, especially to my friend Ian Curry-Sumner for taking that memorable trip with me to Girona.

I am proud to have my two paranimphs at my side. Elise de Bree has been a wonderful colleague and friend, to whom I owe more than she probably realises. Thank you for asking me to work with you!

At the risk of sounding not very original, I want to thank Oren for everything; kafé België, retail therapy, explaining Chomskyan theory to me, cheering me up and making this period of my life a more valuable, joint experience. I am very lucky to have made a friend for life.

Finally, I want to thank my friends and family for their support: bedankt and moltes gràcies! My father told me to ‘hold on tight to my dreams’, which is why I dedicate this book to my parents. And Xavi, thank you so much for telling me to start writing, and making all the difference.

Utrecht, June 2007

# Chapter 1 Introduction

Morpho-phonological alternations are central to phonological theory, but little is known about how they are acquired. This thesis is concerned with the acquisition of voicing alternations in Dutch, as found in the singular - plural pair *bed* [bɛt] ~ *bedden* [bɛdən] 'beds'. The voicing alternation is due to the fact that the singular always ends in a voiceless obstruent (/t/) whereas the plural may contain a voiced obstruent (/d/). The current study is part of a larger research project, aimed at investigating the acquisition of phonological representations in the lexicon and the role of these representations in perception and production.[1] In this project, data from experiments and child language corpora were used to study the phonology of voicing in Dutch and to examine the role of phonological knowledge in acquisition (see, e.g., Zamuner (2006;2007) on infants' sensitivity to voicing and van der Feest (2007) on the voicing contrast in early lexical representations). The goal of the project was to ultimately link results on the acquisition of the voice contrast (in infants and children) to children's knowledge about voicing alternations, the topic of the current dissertation.

The term *alternation* refers to the phenomenon of a single morpheme having two or more alternative forms, depending on the phonological or morphological context in which it appears. The Dutch voicing alternation has been described as a classic example of a phonologically conditioned alternation, caused by final neutralisation of the voice contrast (Trommelen & Zonneveld 1979, Booij 1981, Berendsen 1983, Booij 1995). Although voicing is distinctive for alveolar (*t~d*) and bilabial (*p~b*) obstruents in Dutch (e.g. [tɑk] 'branch' ~ [dɑk] 'roof'), the contrast is neutralised in syllable-final position, as in the singular bɛ[t] 'bed' mentioned

[1] This project is entitled *The development of phonological representations for perception and production,* supported by a grant from the Dutch Organisation for Scientific Research (NWO 360-70-100) awarded to Paula Fikkert (Radboud University Nijmegen) and René Kager (Utrecht University). Researchers were Tania Zamuner, Suzanne van der Feest and the present author.

before.[2] Dutch has two productive plural suffixes, *-en* [-ə[(n)] and *-s* [-s], but only plurals formed with *-en* alternate.[3] Hence, in *-s* plurals a stem-final voiceless obstruent is maintained, as in *lafaard* ~ *lafaards* [lɑfart] ~ [lɑfarts] 'cowards'. In alternating *-en* plurals however, the voicing contrast surfaces in the plural, as in *bed* ~ *bedden* [bɛt] ~ [bɛdən] 'beds', see (1a-c) below for examples. Non-alternating pairs such as *pet* [pɛt] ~ *petten* [pɛtən] 'caps' show that there is no rule or constraint of intervocalic voicing in the language, see examples (2a-c).[4] The examples in (3) show that the voicing alternation is not restricted to the nominal paradigm, as it also occurs in *verbs* (3a-d) and *adjectives* (3e-g). Due to resyllabification, final devoicing does not occur before nominal and verbal plural *-en* or the schwa-initial suffixes *-er*, *-erd* and *-ig* (examples 1a-3h) and vowel-initial nominalising suffixes such as *-aard*, *-ing*, *-in* and *-ier* (examples 4a-d).

| | | | | | |
|---|---|---|---|---|---|
| 1) | a. | [bɛt] 'bed' | /bɛd/ + -en | → | [bɛ.dən] 'hands' |
| | b. | [wɛp] 'web' | /wɛb/ + -en | → | [wɛ.bən] 'webs' |
| | c. | [dif] 'thief' | /div/ + -en | → | [di.vən] 'thieves' |
| 2) | a. | [pɛt] 'cap' | /pɛt/ + -en | → | [pɛ.tən] 'caps' |
| | b. | [klɛp] 'cover' | /klɛp/ + -en | → | [klɛ.pən] 'covers' |
| | c. | [slɔf] 'slipper' | /slɔf/ + -en | → | [slɔ.fən] 'slippers' |
| 3) | a. | [blut] 'blood' | /blud/ + -en | → | [blu.dən] 'bleed-pl-inf' |
| | b. | [hɛp] 'have' | /hɛb/ + -en | → | [hɛ.bən] 'have-pl-inf' |
| | c. | [hɔut] 'hold' | /hɔud/ + -en | → | [hɔu.dən] 'hold-pl-inf' |
| | d. | [xef] 'give' | /xev/ + -er | → | [xe.vər] 'giver' |
| | e. | [xɔut] 'gold' | /xɔud/ + -en | → | [xɔu.dən] 'golden' |
| | f. | [bos] 'angry' | /boz/ + -er | → | [bo.zər] 'angrier' |
| | g. | [lif] 'sweet' | /liv/ + -erd | → | [li.vərt] 'sweetie' |
| | h. | [dɔns] 'down' | /dɔnz/ + -ig | → | [dɔn.zəx] 'fluffy' |

[2] This process applies to labial and coronal obstruents (/p/ ~ /b/, /t/ ~ /d/); velar stops do not have a voiced cognate in Dutch (except in loans such as *joggen* ~ *jog* [g] ~ [k] 'jog'). It also applies to fricatives (/f/ ~ /v/, /s/ ~ /z/ and /ɣ/ ~ /x/), although the fricative voicing contrast is weakening in large parts of the Netherlands and its status is debated (e.g. Gussenhoven & Bremmer 1983, Slis & van Heugten 1989, Ernestus 2000).
[3] Pronunciation of final *–n* is optional, although it is present in the orthography.
[4] In Dutch spelling, doubly spelled consonants are pronounced as singletons, but they indicate that the preceding vowel is lax (see Booij 1995 and Heemskerk & Zonneveld 2000 for rules of Dutch spelling-pronunciation correspondence). Note that in this dissertation, the terms 'short' vs. 'long' are used rather than 'lax' vs. 'tense', which capture phonological behaviour rather than phonetic implementation. However, nothing hinges on the use of this terminology.

| | | | | | |
|---|---|---|---|---|---|
| 4) | a. | [xreis] 'grey' | /xreiz/ + -aard | → | [xrei.zart] 'old man' |
| | b. | [lat] 'load-V' | /lad/ + -ing | → | [la.dɪŋ] 'load-N' |
| | c. | [fris] 'Frisian' | /friz/ + -in | → | [fri.zɪn] 'Frisian woman' |
| | d. | [beijart] 'carillon' | /beijard/ + -ier | → | [beijar.dir] 'carillon player' |

In contrast, the examples in (5) show that stem-final obstruents are voiceless before sonorant-initial stems in compounds (5a-b). Finally, the suffix *-achtig* 'like' is exceptional and behaves as if it induces a compound (5c).

| | | | | | |
|---|---|---|---|---|---|
| 5) | a. | [xɔut] 'gold' | /xɔud/ + ader | → | [xɔut.a.dər] 'gold vein' |
| | b. | [hɑnt] 'hand' | /hɑnd/ + rem | → | [hɑnt.rɛm] 'hand brake' |
| | c. | [dif] 'thief' | /div/ + -achtig | → | [dif.ɑx.təx] 'thievish' |

Due to regressive voicing assimilation, stem-final coda obstruents may surface as voiced before voiced stops (e.g. in compounds such as /klɑp/ + /dør/ [klɑbdør] 'swing door'). There are around 24 (mono-syllabic) minimal pairs that differ only in the underlying specification of the final obstruent. This difference is reflected in the spelling of both the singular and the plural (e.g. *graad* ~ *graden* 'degree' vs. *graat* ~ *graten* 'fish bones').

Voicing neutralisation in the singular presents a different task from neutralisation in the plural or differently derived form, as in American English flapping, which neutralises (at least for some speakers) the distinction between medial /t/ and /d/ (e.g. in *rider* and *writer*). A stem alternation also differs in important ways from an affix alternation such as that found in English plural formation, where the suffix alternates systematically between /z/, /s/ and /ɨz/, depending on the nature of the stem-final sound (e.g. *dogs*, *cats,* and *horses*). English-learning children do not need to change the representation of a singular; they need to learn that the plural suffix varies. Plausibly, Dutch children can be assumed to start from the neutralised singular (e.g. [bɛt] 'bed'), which is more frequent than the plural form in most cases. Crucially, the fact that there is a voicing alternation in the stem can only be learned on the basis of the plural form (e.g. [bɛdən] 'beds'), when the child notices the semantic overlap between the singular and the plural. More generally, the child could notice the overlap between a stem and a derived form in the paradigm (including cases such as *hond* [t] 'dog' ~ *hondehok* [d] 'doghouse' or *hand* [t] 'hand'~ *handig* [d] 'handy'). However, it is likely that plurals are among the first derived forms that children acquire. Under the hypothesis that there is a single morpheme for the word *bed*, children thus need to change the lexical representation of /bɛt/ to /bɛd/, whereas the suffix remains constant.

Dutch children have been observed to prefer non-alternating forms, producing errors in which plurals have voiceless intervocalic obstruents (e.g. *bɛ[t]ən 'beds'), see Kager (1999a) and Zonneveld (2004). However, the type of error in which voiced obstruents are produced has also been attested in spontaneous speech, e.g. *pɛ[d]ən 'caps', *bo[d]ən 'boats', *mʏn[d]ən 'coins' (anecdotal data from Emma at 3;0–3;5, reported in Kager 1999a:336). Generally, parents report such overgeneralisations, which are often remembered as having been a source of amusement (an oft-cited example being **muggenbulden* 'mosquito bites'). Furthermore, some of these errors seem to persist long after the first two years of age. These observations formed the starting point of the current research, where it was hoped that such 'overgeneralisations' of voicing could be elicited experimentally (for both existing words and novel words) and studied in a more systematic way. In addition, the nature and role of the input to the child was investigated (see also Zamuner 2007), to examine the evidence children receive for voicing neutralisation and alternations. In Dutch, assimilation in different morpho-phonological contexts produces substantial surface variation in stem and affix shapes. Moreover, final devoicing does not always occur across word boundaries, e.g. [hɛbɪk] for /hɛb + ɪk/ 'have I' (Ernestus 2000). Another source of variation concerns the neutralisation itself, as it has been argued that (utterance-)final neutralisation in Dutch is phonetically incomplete (Warner et al. 2004), which would mean that [t] in /bɛd/ is phonetically closer to [d] than [t] in /pɛt/ (e.g. in terms of duration). These issues will be discussed with respect to the representation of voicing alternations and the nature of the learning task.

At first sight, acquiring voicing alternations seems to be a challenging task for the Dutch child. From a linguistic point of view, moreover, a formulation of the nature of the task is not theory-independent. In deciding what constitutes knowledge of alternations, we need to address a fundamental debate in linguistics dealing with the representation of complex words in the lexicon. Furthermore, it is important to consider how adults may represent this knowledge, and whether they might differ from children. Thus, in order to answer the main question of this dissertation "when and how are voicing alternations acquired" it is necessary to define 'acquisition' or 'knowledge of alternations' more precisely, drawing upon theories of phonology, morphology, and their interaction or 'interface'. We shall see that the main opposition is that between *symbolic* rule- or constraint-based approaches on the one hand and *analogical* or usage-based approaches on the other. From this perspective it is therefore hoped that this dissertation will also contribute to the ongoing debate about how inflected words are represented in the mental lexicon.

Even though much (recent) work on (morpho-) phonology depends heavily on the theoretical construct of stem alternations, surprisingly few studies have investigated the acquisition of alternations. In addition, even though some authors present data from language acquisition (e.g. Bernhardt & Stemberger 1998), few studies have explored this topic using experimental methods (cf. MacWhinney 1978, Pater & Tessier 2003; 2005). The extent to which a morpho-phonological pattern applies to novel words can be investigated most fruitfully through experiments. Furthermore, the time course of acquisition and the link between production and comprehension are potentially relevant for linguistic theory. In recent work, the issue of *learnability* of alternations has received increasing attention (e.g. Boersma 1999, Tesar & Smolensky 2000, Boersma & Hayes 2001, Hayes 2004, Tesar & Prince 2004), but the proposed models are usually not informed by direct or extensive data from language acquisition. As acquisition of this type of morpho-phonological knowledge has not been investigated for Dutch, results will also provide more insight into the language-specific aspects of acquiring alternations.

## 1.1 Rules versus analogy

In rule- or constraint-based frameworks such as *SPE* (Chomsky & Halle 1968) and *Optimality Theory*, henceforth OT (Prince & Smolensky 1993), it is commonly assumed that alternating stems have abstract 'underlying' representations. In case of the Dutch voicing alternation, the underlying or lexical representation of a singular is said to contain a voiced final obstruent (e.g. /bɛd/), surfacing as [bɛt] due to final devoicing. Under this view, the plural is regularly formed by adding a suffix to the underlying form, e.g. /bɛd/+/ən/, followed by resyllabification (although we will see that the notion of *regularity* is not a straightforward one). This analysis is satisfactory from an economic point of view: [bɛt] and [bɛdən] are clearly semantically related, and there is no need to store (or memorise) the singular and plural form separately, since the surface form [bɛt] is fully predictable (i.e. it can be derived from a general phonological rule of Dutch, which holds without exception in both the derived and non-derived vocabulary). Moreover, an analysis which exploits the rule of final devoicing captures a phonological generalisation that not only holds for Dutch but is also widely attested cross-linguistically (e.g. in Russian, German, Polish, Catalan and Turkish).

Crucially, this view entails that the underlying voice specification of the final segment of a stem can only be deduced on the basis of the derived form. This has several implications for acquisition, since children must construct representations which are more abstract than the forms they hear. Although abstract underlying forms are arguably hard to learn, knowledge of final devoicing (e.g. 'obstruents are voiceless word-finally') is assumed to be acquired at an earlier stage, and is thought to be useful in the process of learning alternations. This touches on an important issue in current phonological theory, as many researchers now agree that much of morpho-phonology can be understood as accommodation to *phonotactic* requirements (Kisseberth 1970, Stampe 1973, Sommerstein 1974, Goldsmith 1993). The term phonotactics refers to the possible or probable sequences of sounds within a language (e.g. the initial sequence */tn/ is not allowed in Dutch in either onsets or codas), without regard to morphological structure. Researchers investigating the ability of infants to segment speech have shown that at least some phonotactic knowledge may be present at an early age. For instance, it was shown that 9-month-old English learning infants are sensitive to how (English) phonotactic sequences typically align with word boundaries (Mattys et al. 1999). This view entails that phonological rules (such as final devoicing) can be learned from distributional evidence, in the absence of detailed morphological analysis (Hayes 2004, Prince & Tesar 2004), to be discussed further in Chapter 6.

In formal frameworks, it is mostly assumed that children's errors may reflect phonological knowledge which is encoded in a grammar. For instance, assuming that children initially set up an incorrect (but surface-true) underlying form (/bɛt/), forms such as *[bɛtən] for adult [bɛdən] 'beds' are seen as 'regularisations', since there is no need to posit an underlying form that is different from the surface form. This has traditionally been described as 'Paradigm Uniformity', or Output-to-Output Faithfulness in more recent OT terminology (see Benua 1995; 1997, Bernhardt & Stemberger 1998, Burzio 1998, Kager 1999a, Benua 2000, Steriade 2000). Hence, children match the singular's voicing value in the plural, showing a preference for paradigms that share contextually invariant morphemes.

Forms such as *[pɛdən] for adult [pɛtən] 'caps' however, are 'overgeneralisations of voicing' and require a different explanation. These could be viewed as the result of the overgeneralisation of rules or constraints such as *intervocalic* (or *intersonorant*) *voicing* in the child's developing grammar (Stampe 1973). Even though such a rule is not active in adult Dutch (i.e. there are non-alternating forms such as [pɛt] ~ [pɛtən] 'caps'), the language learning child may set up hypotheses about possible rules (on the basis of [bɛt] ~ [bɛdən] 'beds') which may prove to be false at a later stage. Hence, it is possible that the child postulates a (lexically restricted) rule of intervocalic voicing. Devoicing of coda consonants and intervocalic voicing are both cross-linguistically frequent processes, rendering them likely candidates for universal constraints in an Optimality Theory framework (Kager 1999a:325). These processes are also phonetically natural (from an articulatory point of view), consistent with the view that constraints are phonetically grounded (see Archangeli & Pulleyblank 1994, Hayes 1999, Hayes 2004, Hayes, Kirchner & Steriade 2004). It is important to note that the issue of whether there is a unique underlying form for each morpheme is theoretically separate from the notion that the voicing alternation is conditioned by the phonology. For example, it has been proposed that constraint-based models should be able to accommodate more than a single lexical form (e.g. both /bɛt/ and /bɛd/), whose distribution is governed by constraints (see Burzio 1998, Kager 1999a). Generally, rule-based models predict a greater degree of consistency in children's behaviour, once the context of rule application is known. However, more recent constraint-based models have incorporated the notion of probabilistic behaviour (Boersma & Hayes 2001, Albright & Hayes 2003, Pater & Coetzee 2005), developed to specifically address the problem of learnability.

In analogical or usage-based models of language, explicit knowledge of derivational rules such as final devoicing is not assumed (Bybee 1985, Rumelhart & McClelland 1987, Skousen 1989; 1992, Derwing & Skousen 1994, Bybee 1995; 1998, Eddington 2000, Bybee 2001, Tomasello 2003). Rather, rules are seen as descriptions by linguists, which are not necessarily represented in a psychologically realistic way (i.e. in a mental grammar). Crucially, there is no separation between lexicon and grammar, and word formation proceeds by means of *analogy* to other words in the lexicon. In contrast to generative phonology, lexical representations are allowed to be redundant (i.e. contain predictable information and even phonetic detail). This view also entails that regular complex words may be stored as wholes in the lexicon. In Bybee's network model (1995, 2001), words are related to other words via sets of lexical connections between identical and similar phonological and semantic features. Morphological structure emerges from these connections in the form of *schemas*. Such schemas describe general phonological properties of a morphological class, and are used in organising and accessing the lexicon (see also Bybee & Slobin 1982). In such a model, a regular form may either be accessed from the lexicon or derived by applying a schema, depending on the (token) frequency of the form.

Even though from a linguist's point of view it is not economical or necessary to store both the singular *bed* and the plural *bedden* in the lexicon, it is conceivable that from a neuropsychological perspective, lexical storage is actually more efficient than computation (i.e. the rule-driven manipulation of symbols). Furthermore, one could argue that even if the voicing alternation reflects a valid generalisation about Dutch phonology (e.g. 'voiced obstruents do not occur in final position'), this does not automatically entail that voicing is present in underlying representations. In fact, these models deny the existence of underlying forms altogether. Given two surface phonological forms [bɛt] and [bɛdən], (implicit) knowledge of the alternation depends on the fact that these words form a *pair* (morphologically and semantically) and the observation that the alternation also occurs in other pairs. Crucially, knowledge of the alternation can be seen as entirely dependent on knowledge of surface forms of lexical items. In word formation, generalisations are possible but the productivity of the pattern depends on its (type) frequency and transparency (e.g. Bybee 1985). The pattern may be extended to novel forms, but it does not derive from a general rule or constraint ranking. Rather, it is a non-derivational generalisation across form classes in a particular lexicon. In order to use analogy, children need to be able to analyse stored words and abstract generalisations from them. Thus, even though plurals (e.g. both [bɛdən] and [pɛtən]) may be stored, this does not mean they are stored as unanalysed wholes. In analogical models, children's behaviour in an elicitation experiment such as the 'Wug-test' (see below) would be influenced by their knowledge of lexical frequency and statistics, rather than rules and segmental contexts. Generally, these models are better able to cope with variation, since these models are inherently probabilistic rather than deterministic.

Even though the above theories seem to take a very different approach, they could nevertheless lead to highly similar predictions, especially with regard to overgeneralisations of voicing in non-words. In fact, the distinction between rules and analogical generalisations may turn out to be very difficult to maintain. These issues will be discussed in more detail when we consider the predictions these models make in relation to the experiments reported in this dissertation.

## 1.2 Experimental evidence

In this dissertation, an attempt is made to test the claim that the phonology of a language enables learners to set up abstract lexical representations, using experimentally obtained evidence. Methods for studying the Dutch voicing alternation will mainly involve elicitation of plurals, providing children with both alternating and non-alternating words ([bɛt] 'bed' and [pɛt] 'cap') as well as non-words (e.g. [kɛt]), with the final obstruent placed in different phonological environments. A Wug-test was used to test *productivity* of the pattern for both children and adults (Berko 1958). Here, subjects need to derive the phonological properties of a plural on the basis of a novel singular. This experimental task matches the real-life situation in which a child has not heard or memorised an existing plural form yet. In forming a plural of a novel singular, there are potentially two well-formed realisations of a non-word such as [kɛt], i.e. [kɛtən] and [kɛdən]. A strategy of Paradigm Uniformity may lead the child to select a non-alternating form (e.g. [kɛt] ~ [kɛtən]). However, the choice between the two possible realisations may be also governed by a strategy of *phonological* generalisation. Hence, overgeneralisations of voicing (e.g. *[pɛdən] 'caps', [kɛdən]) could be driven by the application of phonological rules or constraints such as intervocalic voicing, reflecting children's developing grammars. A third possibility is that overgeneralisations are based on the characteristics of similar words in the lexicon, reflecting *analogical* generalisation.

One way of deciding between competing theories is by taking lexical frequency into account. Stimuli used for the elicitation experiments were monosyllabic words or non-words, with the final obstruent placed in three different environments (i.e. following short vowels, long vowels, or nasals). The frequency of occurrence of the alternation in these specific phonological contexts was studied using the CELEX lexical database (Baayen, Piepenbrock & Gulikers 1995). For instance, the distribution of the voicing alternation in Dutch is characterised by an accidental 'lexical gap'; *p~b* alternations do not occur after long vowels and nasal sonorants, e.g. lɑm[p] ~ *lɑm[b]ən (Zonneveld 1978:49). To investigate whether this might have an effect on overgeneralisations, non-words like [kɪmp] were constructed, for which non-alternations are expected. Here, phonological but not analogical generalisations might lead to voicing alternations (e.g. [kɪmbən]), as such generalisations show an effect of natural classes rather than lexical statistics. In sum, the distribution of alternating forms in Dutch will be shown to be valuable for predicting subjects' behaviour, and in deciding between different models of word-formation.

Another way of deciding between different theoretical approaches is by studying atypical populations. Theories that separate the lexicon (containing irregular forms) from the grammar (containing rules) make certain predictions regarding children with the disorder known as *Specific Language Impairment* (SLI). Proponents of these models claim that these children are unable to decompose (or have severe difficulty decomposing) complex words into their constituent parts. Thus, children with SLI have been found to use lexically-based strategies and 'explicit rule learning', e.g. "add -*s* to form a plural" (e.g. Gopnik & Crago 1991, Goad & Rebellati 1994). These claims lead to testable predictions in the domain of acquiring morpho-phonological alternations. For instance, SLI children are expected to treat alternating and non-alternating plurals the same, as both are presumably stored as wholes. Furthermore, overgeneralisation errors (e.g. *[bɛtən] 'beds', *[pɛdən] 'caps') are not predicted to occur for this group, and they are expected to be more sensitive to lexical frequency. To test such predictions, a study will be presented that compares the behaviour of typically developing children with that of SLI children (an earlier version of this study appeared as Kerkhoff & de Bree 2005).

The claim concerning the link between alternations and phonotactics will also be explored using an experimental approach. When children recognise that [bɛt] 'bed' and [bɛdən] 'beds' are related, knowledge of final devoicing is potentially useful for inferring an abstract lexical representation for *bed* (in combination with the knowledge that there is no intervocalic voicing in the language). As mentioned above, knowledge of final devoicing may depend on previously acquired knowledge of phonotactics. Either way, knowledge of final neutralisation would lead to the ability to derive singulars of newly-heard plurals. This was tested by administering a 'reverse' Wug-test to children, i.e. they were asked to form singulars from plural non-words (both [kɛtən] and [kɛdən]). If the rule of final devoicing is acquired in connection with knowledge of alternations, children should be equally good at producing singulars (i.e. [kɛt]) for both types of non-words (see also Zamuner, Kerkhoff & Fikkert 2006b).

The correct use of existing alternating pairs is not sufficient to conclude that the voicing alternation is acquired, as children could have memorised the singular and plural separately. Furthermore, it is impossible to tell whether a correct plural (such as [bɛdən] 'beds') results from whole form retrieval from the lexicon, or whether it may have been derived from a singular with an incorrect underlying form (i.e. /bɛt/) by means of intervocalic voicing. For this reason, elicitation of non-words is important, to determine the relationship between voicing in words vs. alternations in non-words. However, knowledge of alternations can also be tracked in *comprehension*, to see whether children accept (novel) alternating plurals. To test this, an 'acceptance task' was carried out, in which children were confronted with alternating forms of both words, including erroneous forms (e.g. *[bɛtən] 'beds' and *[pɛdən] 'hats') and non-words ([kɛdən]), see also the earlier report in Zamuner, Kerkhoff & Fikkert (2006a). In this way, it could be established whether children who do not produce voicing alternations might still accept them, showing a comprehension – production asymmetry. Theoretically, it is also possible that existing alternating forms are produced while novel alternating forms are not accepted, showing that productive knowledge of alternations is not conclusive. Furthermore, if children have acquired knowledge of voicing alternations, they may be expected to accept both [kɛtən] and [kɛdən] as plurals for [kɛt], whereas they should reject a place alternation such as [kɛt] ~ [kɛpən].

## 1.3 Outline of the dissertation

In order to answer the question of how and when voicing alternations are acquired it is necessary to first discuss theories of (adult) knowledge of morpho-phonological alternations and their implications for acquisition. In the second chapter, theoretical approaches to morpho-phonological alternations will be discussed, focusing on the Dutch voicing alternation in particular. Contributions and predictions of both generative and analogical models will be assessed, taking into account previous work on acquisition of alternations. The Dutch data are examined more closely in Chapter 3, focusing on the voicing contrast and the nature of the variation (i.e. voicing assimilation and incomplete neutralisation), as well as previous experimental findings. In Chapter 4 new corpus data will be presented, taken from an adult corpus (CELEX) and child corpora, which included the CLPF database (Fikkert 1994, Levelt 1994) and the van Kampen corpus (van Kampen 1997), available from CHILDES (MacWhinney 1991/2000). These data shed light on the frequency distribution of alternations and the nature of the input, as well as on early child productions (i.e. whether alternating forms occur, and to what extent overgeneralisation errors are attested in spontaneous speech). The *production* experiments (i.e. elicitation of novel and existing plurals) will be discussed in Chapter 5, taking into account results of (different groups of) children and adults. The *comprehension* experiments (i.e. elicitation of singulars and acceptability of novel plurals) will be discussed in Chapter 6. Finally, a summary and conclusion is provided in Chapter 7.

# Chapter 2 Theoretical Perspectives

In this chapter the theoretical aspects of morpho-phonological alternations will be discussed. The acquisition of morpho-phonology has been at the centre of the debate on whether complex words are derived by rules or created by analogy to other words in the lexicon. In most generative approaches, abstract lexical representations are subject to the phonological rule or constraint of final devoicing. In contrast, analogical models do not posit abstract underlying forms, leading to a different view of acquisition. After an introduction in section 2.1., important insights of generative phonology will be discussed in section 2.2, focusing on rule-based analyses and constraint-based Optimality Theory. In section 2.3, analogical models of language will be described, focusing on connectionist and usage-based models of language. Finally, the differences between the two main approaches and their general predictions will be summarised in section 2.4.

## 2.1 Introduction

Voicing alternations are prominent in phonological theory, and have been taken as a prime case in support of the notion that underlying representations are abstract. Early linguistic theory already expressed the idea that alternations constitute evidence for the abstract nature of lexical representations. In traditional phonological models, alternations involving final devoicing are seen as an important argument for the distinction between lexicon and grammar. This is because a *general* rule of final devoicing can successfully derive the singular form [bɛt] 'bed' from an abstract lexical representation /bɛd/, which means that there is no need to list both the surface form [bɛt] and the complex form [bɛdən] 'beds'. Rather, the plural [bɛdən] is derived straightforwardly from the same lexical representation /bɛd/ by a morphological rule which adds the suffix -/ə(n)/. The notion 'morpho-phonology' is linked to the distinction between *phonemes* and *morphemes*, which were traditionally appointed to different levels of the grammar (i.e. phonology and morphology). In distinguishing between types of alternations, linguists have traditionally taken into account the distinctive nature of alternating sounds, the phonetic motivation for the alternation and the relevance of lexical information (Kenstowicz 1994). Three types of alternations have traditionally been distinguished, i.e. *phonologically*, *morphologically* or *lexically* conditioned alternations with corresponding rule types.[1] Suppletive forms such as *go ~ went* could be said to constitute a fourth type of alternation, as there is only a meaning-based relation between the alternants, i.e. they cannot be *derived* from one another (or from a common underlying form) by any phonological process. Importantly, linguistic models differ in the way they treat types of allomorphy, as some account for all alternations by means of phonological rules, and others rely more heavily on lexical marking or storage for 'irregular' or 'subregular' patterns. As noted by Kiparsky (1996), the boundaries between rule types are not easily drawn. Traditionally, an important notion in deciding between different types of morpho-phonological alternations is the *productivity* of a process or pattern. All

---

[1] For instance, Kiparsky (1996:15) distinguished between a) phonological rules - allophonic rules and rules of phonetic implementation - b) morphophonological rules and c) allomorphy rules.

theories need to account for the fact that people may produce or comprehend structures that they have never heard before. Productivity thus relates to the extent to which a rule or process is applied to novel forms (e.g. new words, nonce words or loanwords).[2] A second, related notion is the *regularity* (or *generality*) of a process, related to the number of forms that it applies to, or the 'lack of exceptions'. It is difficult to define these terms in a theoretically neutral way; the term 'regular' is often equated with 'rule-based' (at least in the traditional view), and productivity is seen as a property of regular inflection (i.e. the 'scope' of a rule). A third relevant notion is *transparency*, i.e. the extent and location of the change in the inflected or derived form, reflecting semantic, phonological, or orthographic changes. Clearly, there is a relation between these three notions: productive, regular, and transparent processes are presumably 'easy' to learn and may, in psycholinguistic terms, apply 'on line'. In contrast, irregular, unproductive and non-transparent processes are 'hard' to learn and more likely to be represented as whole forms or 'exceptions' in the lexicon.

Using these three criteria, let us now turn to a description of the three main types of alternations, keeping in mind that the boundaries between the three types may turn out to be less clear-cut than presented here. The Dutch voicing alternation central to this study (leading to alternating pairs such as [bɛt] ~ [bɛdən] 'beds') is an example of a *phonologically conditioned* alternation occurring in the *stem* rather than the affix. Here, resyllabification ensures that /d/ of [bɛ.dən] is in onset position rather than the coda, which is why it escapes devoicing. Final devoicing is productive, as illustrated by the fact that non-advanced Dutch speakers of English tend to devoice final obstruents. Productivity is also attested for loanwords (e.g. [klʌp] 'club', [bɛnt] 'band', [wit] 'weed') and abbreviations (e.g. *bieb* [bip] for *bibliotheek* 'library'). The alternation reflects neutralisation of the voice contrast, which is considered a general or 'exceptionless' rule of Dutch, not restricted to a particular morpho-syntactic domain (i.e. it also occurs in the verbal and adjectival paradigm, see examples in Chapter 1). Even though final devoicing could be regarded as a

---

[2] Morphological productivity has also been taken to refer to the number of times a particular affix appears as part of any existing word, or the ratio of 'actual words' to 'possible words' (Aronoff 1976). Baayen (1989; 1991) has formalised this approach using corpus counts. However, this notion is close to the traditional notion of productivity, as the 'productivity index' is claimed to be a measure of the likelihood that a speaker will produce novel forms with a certain affix (see Aronoff & Anshen 1998, Hay and Baayen 2005).

regular process, it is not clear whether voicing alternations are themselves regular, to be discussed further below. The alternation is not phonologically transparent, as the voicing value is different in the derived form. However, it is both semantically and orthographically transparent, as the alternation is reflected in the spelling (i.e. *bed* ~ *bedden* 'bed(s)' vs. *pet* ~ *petten* 'hat(s)').[3] In the next section, the Dutch voicing alternation will be considered in more detail.

Another example of a phonologically conditioned alternation is the voicing alternation that occurs in the English plural suffix: /s/ occurs after voiceless obstruents (e.g. *cats* [kæts]) and /z/ after voiced obstruents (e.g. *dogs* [dɔgz]). Here, the alternants /z/ and /s/ are fully predictable, determined by the phonological character of the final segment of the noun stem. This allomorphy is productive, as it is extended to loanwords and non-words (such as nonce *wug* + *s* /wʌgz/) by both adults and children (Berko 1958, Derwing & Baker 1980). The allomorphy has been analysed as involving phonological rules of *voicing assimilation* (to prevent non-homorganic clusters such as [tz]) and *insertion* (of [ɨ] or [ə]) to avoid a geminate cluster (e.g. [zz] or [ss]). However, there is no *general* rule of voicing assimilation or insertion in English, even though the past tense suffix (/d/, /t/and /ɨd/) behaves in the same way. Hence, adjacent obstruents do not always agree in voicing (i.e. there are words like *jigsaw*, *website*, *lapdance*, etc.), see Lombardi (1991) for a discussion.

The Dutch diminutive suffix allomorphy is an example of a *morphologically conditioned* alternation. The diminutive suffix (whose basic form is generally proposed to be /tjə/) has five allomorphs, the distribution of which is dependent on phonological properties of the stem. For instance, the allomorph /kjə/ appears after noun stems ending in an unstressed syllable and a velar obstruent (e.g. *kóninkje* [konɪŋkjə] 'king-dim') and the /pjə/ form occurs after stems which end in unstressed syllables and bilabial nasals (e.g. *bézempje* [bezəmpjə] 'broom-dim'), which could be described as phonetically natural assimilation contexts. However, the distribution does not follow from *general* phonological rules of Dutch. One could derive *bezempje* from underlying /bezəm-tjə/, but there is no general rule or constraint that changes /t/ into [p] after /m/ (Trommelen 1983, Booij 2002:175). Here, rules have to refer to non-phonological properties such as

[3] Note that plosives (e.g. *bed* ~ *bedden* 'beds') are spelled differently from fricatives (e.g. *duif* ~ *duiven* 'dove(s)'), with the latter reflecting the effect of final devoicing in the singular.

the feature DIMINUTIVE, i.e. they would apply only in morpho-syntactically designated forms. Note that the English suffix alternation only applies in plural and past tense formation, indicating that it is also morphologically conditioned. However, the English suffix alternation does not have any exceptions, whereas Dutch diminutive formation does (e.g. forms like *bloempje* and *bloemetje* 'flower-dim' co-occur). Diminutive formation is less transparent than the Dutch voicing alternation and acquired relatively late (den Os & Harder 1987), although it is productive.

The choice between the two Dutch plural suffixes *-en* (e.g. *bedden* 'beds') and *-s* (e.g. *lepels* 'spoons') represents a similar case. Here, the lexical representation for the plural suffix must include the information that it has two alternants, even though their distribution may be governed by general constraints (related to sonority and the preference for a trochaic pattern, see van Haeringen (1947) and de Haas & Trommelen (1993)). Again, there are exceptions, e.g. both *lampionnen* and *lampions* 'Chinese lanterns' occur. Even though both plural suffixes are productive, there seems to be variability across subjects in the specific allomorph chosen for a specific non-word (e.g. Baayen et al. 2002, van Wijk 2007). We will return to the Dutch plural suffix and its acquisition below and in following chapters. As these examples show, morphologically conditioned alternations are generally associated with a greater amount of lexical variation.

The third type of alternation is *lexically conditioned,* typically viewed as *irregular* and *unproductive*. A well-known example is the voicing alternation in a small set of English plural pairs such as *knife ~ knives* and *wife ~ wives*. Again, the alternation does not reflect a general rule of English phonology, although the pattern could be called 'subregular' as it occurs in several words. Such patterns have also been called 'minor rules', examples of which include English past tense forms such as *break ~ broke* and *sing ~ sang*. The distinction between this type of alternation and the former is mainly based on the number of forms it applies to (or the number of exceptions that occur). In terms of transparency, these alternations are often similar to phonologically motivated alternations, compare English *wife* [f] ~ *wives* [v] and Dutch *vijf* [f] ~ *vijven* [v] 'fives'. A similar case is Dutch open syllable lengthening (OSL) in a number of plural forms, such as [slɔt] ~ [slo.tən] 'locks' and [wɛx] ~ [we.ɣən] 'roads' (see Zonneveld 1978, Booij 1995, Kager to appear).

English plurals that take an irregular suffix such as *ox ~ oxen* and *child ~ children* are even more lexically restricted (similar to Dutch plurals ending in -*eren*, such as [kɪnt] ~ [kɪndərən] 'children'). Another example of a lexically governed alternation is Spanish diphthongisation in verbs. Although many forms exhibit an alternation (e.g. *entendémos* [ɛntɛndɛmɔs] ~ *entiéndo* [ɛntjɛndo] 'we/I understand', many others do not (e.g. *comémos* [kɔmɛmɔs] ~ *cómo* [kɔmɔ] 'we/I eat'). This alternation is also morphologically conditioned, as it occurs only in the present tense and for certain conjugation classes. Nevertheless, general phonological rules have been proposed for this alternation (Harris 1985). Alternations that are both lexically and morpho-syntactically conditioned include French liaison, which occurs across word boundaries (e.g. *nous avons* [nuzavɔ̃] 'we have'). Lexically conditioned alternations are often associated with *lexicalisation*, a process of language change which obscures phonological processes that were regular and general in an earlier stage of the language. For instance, Dutch open syllable lengthening can be seen as a relic of OSL in Early Germanic. The extent to which diachronic processes shape alternations will be discussed in more detail when we consider analogical models of language.

The types of alternations introduced in this section are hard to distinguish on the basis of binary criteria such as regularity, productivity or transparency, and they seem to be a scale rather than belonging to different categories. Hence, most alternations have some degree of morphological and lexical involvement. However, even though the difference between types of alternations may be gradient rather than absolute, the three criteria outlined above are still important for current theories of morpho-phonology. In analogical models, the (lexical) frequency of alternations is another major factor, which is related to the notions of regularity and productivity (see section 2.3). In the next section, rule-based accounts of morpho-phonological alternations will be described in more detail, focusing on neutralisation of voicing in particular. The question of what needs to be acquired by the child will be discussed simultaneously.

## 2.2 Generative models

Generative linguistics has embraced Chomsky's (1965) idea that general, unconstrained learning mechanisms (such as *association*) are not sufficient for acquiring the abstract principles of language (i.e. the grammar). The famous argument from *poverty of the stimulus* (Chomsky's solution to "Plato's problem") entails that the input to the child is unreliable, ambiguous and contains little negative evidence, so that acquisition must be guided by innate principles in the form of *Universal Grammar* or UG (Chomsky 1980, Hornstein & Lightfoot 1981). In terms of language acquisition, this theory needs to explain how Universal Grammar is linked to some particular language (also called the *linking problem*), given that individual languages differ widely. The theory of generative phonology (e.g. Chomsky & Halle 1968, Halle & Mohanan 1985, Kenstowicz 1994) is based on the corresponding notion that the child needs to acquire the abstract adult phonological system, aided by universal principles. Systematic regularities or linguistically significant generalisations are distinguished from idiosyncratic features of a language. Thus, the grammar (containing rules) needs to be separate from the lexicon, and predictable information (which can be computed by rules) is not present in lexical representations. The lexicon was thought to contain only morphemes, which are concatenated by morphological rules.[4] Generative phonologists have further argued that linguistic description should have 'psychological reality'; a rule should reflect the (implicit) linguistic knowledge or *competence* of native speakers (see, e.g., Kiparsky 1975, Chomsky 1980:109, 1995:380). Finally, an important argument for positing a phonological rule is its phonetic plausibility or naturalness, reflected by its widespread occurrence in the world's languages. In Chomsky and Halle's influential *Sound Pattern of English* known as *SPE* (Chomsky & Halle 1968), morpho-phonology was not distinguished from pure phonology. In this approach, all surface forms of a morpheme are serially derived from one unique underlying form (e.g. /bɛd/), mapped to the (phonetic) surface representation (e.g. [bɛt]) by phonological rules of the grammar. Here, phonology and morphology are different grammatical components: after a word has been accessed from the lexicon, morphological rules (or *word formation rules)* combine stems and affixes (e.g. the operation *plural* has an abstract feature structure which is introduced by syntactic rules). The output of this process (e.g. the underlying form /bɛd-ən/) is then sent to the phonological component, which contains linearly ordered rules. In *SPE*, these phonological *rewrite* rules change segments (or feature values) in the phonological representations of words. Segments are lists of (binary) features or feature matrices and can be inserted, deleted or

[4] However, within the generative viewpoint it has been suggested that the lexicon includes complex words generated by *word formation rules* or *morpholexical* rules, even though their properties may be predictable. Such a lexicon could function in parallel to a more abstract one (see, e.g., Halle 1973, Anderson 1974, Kiparsky 1982a, and Inkelas 1990).

changed by rules. Rules have a specific *format* (A → B / X_Y), in which both the context (i.e. X_Y) and the segments undergoing the change (e.g. A and B) are minimally specified. In *SPE* (Chapter 9), each (binary) feature is associated with an unmarked and a marked value (e.g. the unmarked value for [voice] is [-voice]), correlating with typological facts and considerations of frequency (Chomsky & Halle 1986). The notion of 'markedness' as specification for a phonological distinction was first developed by Trubetzkoy (1929, 1939), who defined it according to notions of structural complexity.[5] In *SPE*, all types of alternations are derived by phonological rules. Even 'exceptional' patterns are handled by lexical marking for minor rules, which is why it could be called a *single mechanism* approach. For instance, a rule of regressive voicing assimilation is used to derive *wife* ~ *wives*, and a rule of 'Lowering Ablaut' derives *sang* from *sing*. In the Dutch example of open syllable lengthening or OSL (e.g. [slɔt] ~ [slo.tən] 'locks'), the word *slot* would be marked [+OSL] in the lexicon.[6]

It had been recognised by generative theorists that an alternative to abstract underlying forms would be to list each alternant in the lexical representation (of a stem or suffix). Thus, rules of *selection* would determine the choice between alternants (sometimes called the *morpheme alternant theory*). In case of the Dutch voicing alternation, one could list both allomorphs (e.g. /bɛt/ and /bɛd/) and assume a selection rule such as "Select the alternant with a final voiceless obstruent when the morpheme appears at the end of a word; otherwise, select the alternant with a voiced obstruent". However, as Kenstowicz & Kisseberth (1979:49) argue, such an analysis not only fails to capture the rule-governed nature of the voicing alternation, it also implies that *two* rules have to be assumed: a phonological rule of devoicing (as no word-final voiced obstruents are permitted in the language, e.g. in case of loanwords) and a morphological rule of allomorph selection.[7] Under this view, only a phonological statement captures both the fact that

[5] This was based on language-specific structures such as final neutralisation of the *t~d* contrast in German: "the structure and the functioning of the system tells us whether it is *d* that is mark-bearing (the mark being 'voice') or whether *t* is mark-bearing (the mark being 'tenseness')" (Trubetzkoy 1939:68).
[6] *Readjustment* rules were posited for truly suppletive allomorphy, similar to allomorphy rules in Aronoff's (1976) word-based model.
[7] Similarly, in Item-and-Arrangement models of morphology (Spencer 1991), rules determine the choice between two lexically listed allomorphs. Jackendoff (1975) has also argued that both representations can be stored in the lexicon, and related through *lexical redundancy rules* (see also Lieber 1981 and Cameron-Faulkner & Carstairs-mcCarthy 2000).

voicing alternations are predictable and the generalisation that no word-final voiced obstruents occur. In other words, the rule of final devoicing explains a *systematic gap* in the phonetic distribution of sounds. We will return to the topic of allomorph selection rules in the next section, when Optimality Theory is discussed.

Evidence from linguistic change has also been used to argue for rule-based analyses of voicing alternations. For instance, Kenstowicz & Kisseberth (1979:170) note that when an (optional) rule of word-final schwa deletion came to be operative in German, it resulted in devoicing of the final obstruent (e.g. in the dative singular *Hunde* 'dog' [hundə] → [hunt]). However, it is often not clear whether a synchronic phonological analysis reflects the grammatical knowledge of speakers. Hale (1973) offers an example from Proto-Polynesian, a language which is known to have lost final obstruents (i.e. all words end in a vowel). A modern Polynesian language such as Maori has alternations between the verb stem and the passive form of the following type: *awhi* ~ *awhitia* 'to embrace', *hopu* ~ *hopukia* 'to catch' and *aru* ~ *arumia* 'to follow'. In such a case, a generative phonologist would analyse the underlying form of the suffix as /-ia/, with the final consonant appearing in the underlying form of the stem (i.e. /awhit-/, /hopuk-/, /arum-/). The stem alternation can be elegantly described as deriving from the rule of final consonant deletion, without the need of listing (unpredictable) allomorphs (e.g. /-tia/, /-kia/, /-mia/). However, Hale found that Maori children must have *reanalysed* these consonants as belonging to the suffix, as one of the allomorphs (/-tia/) was regularised (e.g. in forming passives of nominal stems, adverbials or English loanwords), presumably because /-tia/ was more frequent than the others (cf. de Lacy 2004). A *lexical* account would explain these facts satisfactorily, because one could simply state that the child memorises the separate allomorphs of Maori passives.[8] Thus, even though historically these final consonants were part of the stem, there is no *synchronic* process reflecting the loss of final consonants. These observations led Hale (1973) to argue that there is a tendency in language acquisition to set up underlying forms that are surface-true. This argument has been put forward by others, although not always in connection to language acquisition. For instance, in Natural Generative Phonology,

---

[8] Similar findings have been reported for other Polynesian languages such as Hawaiian (in which one basic suffix remains), showing that reanalysis has taken place independently in each of the languages, see Kiparsky (1982b).

Venneman (1972) proposes a principle which states that the UR of a morpheme is identical with the phonetic alternant that appears in isolation (Kiparsky 1968, Smith 1973, Stampe 1973, Hooper [Bybee] 1976). Hence, Maori children would set up underlying forms that end in vowels, and Dutch children would set up underlying forms ending in voiceless obstruents (i.e. for both /pɛt/ 'cap' and /bɛt/ 'bed'). Hale (1973:33) posits the following constraint on underlying representations: "an underlying phonological representation of stems is disallowed if it violates a universal surface canonical pattern". In other words, Maori speakers do not postulate an underlying final consonant because no final consonant ever appears on the surface. Similarly, Kiparsky (1968) has formulated an Alternation Condition, which limits the abstractness of underlying representations to cases motivated by phonological alternations.[9]

It thus seems that an abstract analysis of the Dutch voicing alternation is not excluded, as there is sufficient surface variation (i.e. both /t/ and /d/ surface). Kenstowicz & Kisseberth (1979:174) note that it is generally difficult to choose between a phonological and a lexical analysis, and that 'external' evidence of the type described above (i.e. reanalysis) is important in deciding whether phonological rules have 'psychological reality'. In generative theory, a phonological analysis is always preferred over a lexical solution (listing of words or morphemes) unless there is evidence to the contrary. However, it is unclear what exactly constitutes evidence against abstract phonological representations in the adult or child lexicon. For instance, previous research investigating the psychological reality of English vowel alternations suggests that children's knowledge of these alternations relies heavily on the knowledge of orthography and spelling rules (Moskowitz 1973, Jaeger 1984; 1986, Ohala & Ohala 1986, Wang & Derwing 1986, Eddington 2001). In the next section, rule-based analyses of final devoicing are described in more detail, before turning to acquisition.

---

[9] This condition restricts the application of *absolute* neutralisation (which leads to underlying segments that never surface), which is different from *contextual* neutralisation (e.g. due to final devoicing), because its effects never lead to analogical reversal and are not productive. Kiparsky (1982b:36) states that "the only sense that can be made out of it is as a strategy of language acquisition which says that a learner analyzes a form "at face value" unless he has encountered variants of it which justify a more remote underlying representation".

### 2.2.1 A rule of final devoicing

Within the generative framework, the phenomenon of final devoicing has received much attention, as it appears to lend itself to a straightforward account. In introductory textbooks, a standard rule of final obstruent devoicing (for instance of the form [-sonorant] → [-voiced] / __ #) is often given as an example of a simple neutralisation rule (Venneman 1972, Kenstowicz & Kisseberth 1979, Trommelen & Zonneveld 1979, Katamba 1989, Booij 1995, Wiese 1996, Gussenhoven & Jacobs 1998, Roca & Johnson 1999, Jensen 2004). The rule of final devoicing thus converts final obstruents underlyingly specified as [+voice] into their [-voice] counterparts. However, upon closer inspection, the exact formulation of such a rule has not been generally agreed upon, as its interpretation depends on theories of underspecification and the nature of the feature involved.[10]

In rule-based approaches, it has been proposed that the lexicon contains only distinctive or contrastive features to which rules can refer. For instance, no rule refers to voicing in sonorants, as this feature is never distinctive for sonorants. Distinctive features are motivated by the existence of natural classes of segments, and they may be either *privative* or *binary*. Privative or single-valued features express the contrast between presence and absence (i.e. positive or 'nothing'), rather than a positive and negative value. When [voice] is taken to be a privative feature, the absence of [voice] is phonetically interpreted as 'voiceless'. Some analyses place constraints on the *licensing* of a privative [voice] feature or (its OT equivalent) Positional Faithfulness (e.g. Mester & Ito 1989, Lombardi 1991; 1999), which has also been proposed for Dutch (Iverson & Salmons 1995; 2003, Zonneveld to appear). Under this view, final devoicing can be analysed as a process involving *delinking* or *reduction*, which means that the feature [voice] is deleted from an underlying representation by delinking the laryngeal node of obstruents (resulting in loss of contrast).[11] More recently, a *binary* feature

[10] For instance, some authors have proposed features such as [+/-tense] or [+/-fortis] instead of (or in combination with) [+/-voice] (Brink 1975, Booij 1981).

[11] Such delinking originates in nonlinear frameworks employing feature geometry, in which a feature can be 'disconnected' from its superordinate node, which removes it from the phonetic realization of a segment (Clements 1985). Instead of a feature changing rule, a process of 'feature-value deletion' has also been proposed, see Charles-Luce (1985), Port & Crawford (1989), Steriade (1997), Ernestus (2000; 2003), Jansen (2004), and Piroth & Janker (2004). In frameworks that posit a binary feature [+/-voice], assimilation is often analysed as delinking

[±voice] has also been proposed, marking positive and negative values (Inkelas 1994, Inkelas, Orgun & Zoll 1997, Wetzels & Mascaró 2001). Importantly, predictable features tend to be cross-linguistically unmarked, and [-voice] is considered the default or unmarked value for obstruents (see Chapter 3 for further discussion).

The environment in which the final devoicing applies has been debated as well, as it can refer to (phonological) syllable boundaries as well as (morphological or prosodic) word boundaries (see Brockhaus 1995). Dutch phonologists have typically formulated rules of final devoicing (FD) that refer to the word end (#) as in (1) and (2) below, the latter accounting for the fact that obstruents may intervene between the obstruent that undergoes devoicing and the word end (e.g. *[lez-t] '(s)he reads'). Such a rule would apply vacuously most of the time, but would sometimes have to be applied twice (e.g. /hovd/ → [hoft] 'head') or even three times (e.g. in /ɔnt/-/hovd/-/d/ → [ɔnthoft] 'beheaded', in which final devoicing precedes degemination, see Zonneveld to appear). In (3), the rule refers to the syllable edge (σ), which has been proposed to account for the pronunciation of English loanwords such as *Sidney* ['sɪt.ni] or *Cambodia* [kɑm'bɔt.ja] (Booij 1977, Zonneveld 1983, Booij 1995).

(1) [-son] → [-voice] / ___ #

(2) [-son] → [-voice] / ___ $C_0$ #

(3) [-son] → [-voice] / ___ $C_0$ ]σ

It can be shown that any of these rules would correctly derive the surface representation (SR) [bɛt] if we take the form /bɛd/ to be the underlying representation (UR) as in (4). Deriving the plural from the incorrect underlying form /bɛt/ with a rule of intervocalic voicing (IVV) gives the incorrect result shown in (5) below.

| (4) | | | | | |
|---|---|---|---|---|---|
| | UR | /bɛd/ | /bɛ.dən/ | /pɛt/ | /pɛ.tən/ |
| | FD | t | n.a. | vac. | n.a. |
| | SR | [bɛt] | [bɛ.dən] | [pɛt] | [pɛ.tən] |

---

and *spreading* of [+voice]. This may lead to important generalisations, e.g. Wetzels & Mascaro (2001) propose a typology of voicing and devoicing in which intervocalic voicing across word boundaries only occurs for languages with word-final laryngeal neutralisation (such as Dutch).

| | | | | | |
|---|---|---|---|---|---|
| (5) | UR | /bɛt/ | /bɛ.tən/ | /pɛt/ | /pɛ.tən/ |
| | IVV | n.a. | d | n.a. | d |
| | SR | [bɛt] | [bɛ.dən] | [pɛt] | *[pɛ.dən] |

To ensure a correct outcome, Dutch regressive voicing assimilation also needs to be taken into account. In Dutch, a rule of regressive voicing assimilation (RVA) is assumed to apply in compounds such as *meetband* [db] 'measuring tape' and *breedband* [db] 'broadband', as shown in (6) below. A rule of final devoicing (FD) needs to be ordered *before* a rule of regressive assimilation, to avoid *[metbɑnt] and *[bretbɑnt], as in (7).[12]

| | | | |
|---|---|---|---|
| (6) | UR | /met/ + /bɑnd/ | /bred/+/bɑnd/ |
| | FD | /metbɑnt/ | /bretbɑnt/ |
| | RVA | /medbɑnt/ | /bredbɑnt/ |
| | SR | [medbɑnt] | [bredbɑnt] |

| | | | |
|---|---|---|---|
| (7) | UR | /met/ + /bɑnd/ | /bred/+/bɑnd/ |
| | RVA | /medbɑnt/ | n.a. |
| | FD | /metbɑnt/ | /bretbɑnt/ |
| | SR | *[metbɑnt] | *[bretbɑnt] |

To conclude, the examples in this section illustrate the view that in the adult system, the (language-specific) rule of final devoicing is part of an ordered set of rules which apply serially to underlying representations.

As stated above, the domain of application of the final devoicing rule (i.e. at syllable or word level) has given rise to debate. This is reflected in an influential generative approach to the relationship between phonology and morphology known as 'Lexical Phonology' (Kiparsky 1982b, Mohanan 1982, Halle & Mohanan 1985, Mohanan 1986), based on Siegel's (1977) theory of level ordering. The theory posits a *lexical* and a *postlexical* component of the grammar, where the output of the former provides the input to the latter. Phonological rules are divided into two types: lexical rules that interact with morphology, and postlexical rules that do not. In this model, phonology and

[12] See Zonneveld (1983), Trommelen & Zonneveld (1979), Berendsen (1983), Booij (1995), Grijzenhout & Krämer (1998), Grijzenhout (2000), Heemskerk & Zonneveld (2000) and Zonneveld (to appear) for analyses of Dutch final devoicing and voicing assimilation.

morphology are interwoven, in contrast to the earlier *SPE* notion that morphology applies *before* phonology. The notion of the *cycle* is important in lexical phonology (Chomsky & Halle 1968, Aronoff 1976). Within the lexicon, phonological rules apply to underived words (e.g. within the domain of the syllable or prosodic word), after which morphological rules can be applied. This then creates a new domain for phonological rules to apply, or a second *cycle*. Lexical rules are thus *cyclic* and have a number of other properties that distinguish them from postlexical rules. They are typically *phonemic, structure-preserving* (they do not introduce segments that do not occur in the underlying segment inventory) and *arbitrary,* in the sense that they may have exceptions. In contrast, postlexical rules are thought to be *automatic*, *allophonic* and *phonetically motivated* or *natural*. If necessary conditions are present, postlexical rules can apply without restrictions, after words have been inserted into a phrase or sentence.[13] They are not phonetic implementation rules however, as they can only refer to *phonological* elements (such as features and segments). Kiparsky (1982b) posits three levels at which phonology and morphology can apply, partly based on the notion of 'regularity'. The first level includes rules of derivational morphology that induce changes in stem phonology (including irregular patterns). The second level contains rules of regular derivational morphology and compounding. The third includes rules of regular inflectional morphology. Thus, the 'minor rules' for lexically conditioned allomorphy apply at level 1 (e.g. *sing ~ sang*), whereas phonological rules such as vowel insertion apply at level 2. Phonologically and morphologically conditioned alternations are thus assigned to the same level of the grammar (i.e. level 2).

The Dutch rule of final devoicing has traditionally been appointed to the lexical level (e.g. Booij 1985), whereas voicing assimilation has been assumed to apply at the postlexical level (e.g. Ruys & Trommelen 2003).[14] However, final devoicing clearly has a number of properties which would make it a candidate for a postlexical rule. Booij & Rubach (1987) have argued that Dutch final devoicing cannot be accounted for in the standard framework of Lexical Phonology, since it cannot be a cyclic rule while it cannot be postlexical either. It is argued that final devoicing applies to words before they are combined with a (clitic) pronoun (e.g. *vind 'er* [vɪn.tər] 'find

[13] But see, e.g., Ruys & Trommelen (2003) who invoke a domain-based approach by Nespor & Vogel (1986) to account for the variability of (post-lexical) Dutch voicing assimilation.
[14] But see Zonneveld (1983) on voicing homorganity in the lexicon.

her' or *vind ik* [vɪn.tɪk] 'find I') as opposed to a suffix, e.g. plural *vind-en* [vɪn.dən] 'find' or deverbal noun *vind-er* [vɪn.dər] 'finder' (see also Booij 1997, Grijzenhout & Krämer 2000).[15] Addition of the vowel-initial suffix thus triggers (cyclic) resyllabification, which in turn *bleeds* final devoicing (i.e. final devoicing does not apply, as in [vɪn.d-ər]). This means that final devoicing cannot itself be cyclic, as it must wait until all suffixes have been added or the 'wrong' form is derived (e.g. [vɪn.tər] 'finder'). This can be analysed as a case of opaque rule interaction, because final devoicing should not precede syllabification. However, the rule cannot be postlexical either because resyllabification across a word boundary does *not* bleed syllable-final devoicing.[16] These issues clearly have implications for language acquisition; even if the lexicon is innately structured such that level ordering constrains children's hypotheses about morphology, it is still unclear upon what evidence a child would assign a particular rule or affix to a particular level (see Anderson 1982). This problem is related to the fact that it is not clear to which 'level' of the grammar Dutch plurals are assigned to, i.e. whether they are considered regular or irregular. Let us therefore turn to models that have explicitly incorporated the notion of regularity.

Derivational theories have been criticised on several grounds from within generative theory. First, the excessive abstractness of *SPE* representations was considered problematic (Kiparsky 1968). The theory allows for features in underlying representations that never surface. For instance, to account for 'Trisyllabic Laxing' in forms such as *sane ~ sanity*, the form [sein] is itself derived from /sæ:n/, with a long vowel /æ:/ that never surfaces. As noted by Kiparsky, it is unclear how the child would arrive at such rules and representations, given that a record of diachronic change is not available. Thus, *SPE* does not impose constraints on the kind of relation that can exist between phonological and phonetic representations, e.g. by restricting the effects that phonological rules may have or the kinds of underlying forms that can be posited. Furthermore, the notion that all forms

---

[15] Experimental evidence supporting this observation (i.e. the pronunciation of forms like *vind* before clitics and suffixes) is given in Bauman (1995). Zonneveld (to appear) notes that forms like *moet ik* [mu.dɪk] 'must I' also occur (see Chapter 3).

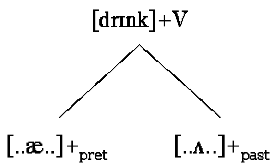
[16] To account for these facts, Booij and Rubach (1987) propose a third *postcyclic* component, where a rule can only apply once, to be posited *after* the cyclic lexical level and *before* the postlexical level. Grijzenhout & Krämer (1998) argue that the solution lies in incorporating the different prosodic structures of suffixes which trigger resyllabification, rather than in level ordering.

can be derived by ordered rules has been argued to complicate acquisition (see Dresher 1999, Pulleyblank & Turkel 2000), because the problem of distinguishing rule-based forms from exceptions is challenged if opaque analyses are allowed. Thus, the theory does not sufficiently constrain the choices available to the language learning child.

A *Dual Mechanism* model such as the *Words-and-Rules* model proposed by Pinker and others (Pinker & Prince 1988, Prasada & Pinker 1993, Marcus et al. 1995, Pinker 1998, Clahsen 1999, Pinker 1999) was partly intended to address these problems. In this model, the distinction between the lexicon and the grammar is maintained, but the model relies more heavily on storage than the *SPE* model. Here, the distinction between regular and irregular inflection is seen as a clear dichotomy; irregularly inflected words (e.g. *wife ~ wives, sing ~ sang, break ~ broke*) are stored in the lexicon and handled by associative memory, whereas *regular* or *rule-based* processes involve a symbolic operation which attaches a *default* suffix to a lexical category or variable (such as NOUN). A *pattern associator* in memory handles irregular patterns or 'minor rules', resembling *SPE* redundancy rules. A mechanism resembling the *Elsewhere Condition* (Kiparsky 1982b) is proposed, where the *general* rule is blocked by the *special* rule. In the Dual Mechanism model, a stored form (e.g. *oxen*) blocks the application of the default affix (e.g. **oxes*). Crucially, productivity or lexical frequency is claimed not to play a role in determining the regularity of a process. In later versions of the theory (e.g. Pinker 1999, Pinker & Prince 1994:331), it is assumed that *some* high frequency regular forms may be stored, e.g. when they rhyme with irregular neighbours (e.g. regular *blink* and irregular *drink*) or when the regular and irregular form both exist (e.g. *dived* and *dove*). However, as noted by Clahsen (1999:1052) "even though high frequency regulars may produce memory traces, the processing of regulars does not depend on stored representations" (see also Marcus et al. 1995:196). In this model, mixed verbs with regular affixation and irregular vowel alternations are stored as 'structured lexical entries'. Productivity of irregular patterns can be accounted for in terms of *default inheritance hierarchies*, according to which the different stem or root variants of a lexeme are stored as phonologically similar subnodes of structured lexical entries (Corbett &

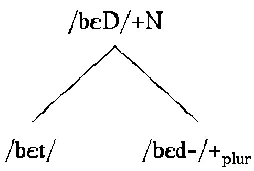
Fraser 1993, Wunderlich & Fabri 1995, Wunderlich 1996), see example (8) for the vowel alternation in *drink*.[17]

(8)




Since the [.æ.]$_{+pret}$ subnode occurs for several other verbs with *-ing-* and *-ri-* in the base entry (e.g. *ring, sing*), these subnodes constitute a *generalised lexical template*. Dutch vowel alternations (e.g. *stad ~ steden* [stɑt] ~ [stedən] 'cities') clearly belong to this class. In a similar vein, voicing alternations could be considered subentries of a lexical entry, as in (9).

(9)




Note that Pinker & Prince (1988) have stated that truly regular formations preserve stems unaltered. However, Berent, Pinker & Shimron (2002:489) argue on the basis of Hebrew that "nouns with both regular and irregular plurals may undergo stem changes or fail to undergo them" (e.g. *barak ~ brakim* 'lightning', has both a stem change and the 'regular' plural suffix -*im*). For now, we can conclude that even though the stem alternation in Dutch plurals may itself be irregular (i.e. require listing), the *-en* suffix may still be regular under the Dual Mechanism approach. Irregular forms are expected to be acquired in a different way by children, as they are influenced by similar forms in associative memory.

[17] Each node of a structured lexical entry is defined in terms of a phonological string and a morphological feature set; a subnode inherits all information from its mother, except for the features it replaces or adds.

Generally, the main drawback of the Words-and-Rules model (and any model that strictly separates stored irregulars from computed regulars) is the difficulty of determining which cases represent 'regular' rules, since 'exceptional' cases may also show productivity (e.g. *spling ~ splung*), whereas regulars have been found to be sensitive to phonological neighbourhoods and lexical frequency effects such as full-form priming, for English (e.g. Taft 1979, Sereno & Jongman 1997, Alegre & Gordon 1999, Gordon & Alegre 1999) and Dutch (Baayen, Dijkstra & Schreuder 1997, Bertram, Schreuder & Baayen 2000, Baayen et al. 2003). Proponents of Dual Mechanism models typically account for such effects by claiming that regulars 'resemble' irregulars in some way (Prasada & Pinker 1993, Marcus et al. 1995). However, every novel word resembles existing words to a certain extent, and as long as no formalisation of the notion 'resemblance' is offered productivity cannot be predicted.

More recently, Albright & Hayes (2003) have proposed an alternative to a Dual Mechanism approach, by advocating a model containing 'only rules'. This model resembles the earlier *SPE* model, since subregular forms are also derived by the grammar. However, irregular patterns can be lexically listed at the same time, preventing them from being regularised.[18] Importantly, this model is based on inductive learning of multiple *stochastic* rules instead of maximally general rules. Under this view, "learners posit rules that attempt to capture generalizations about all morphological processes, not just the largest or most productive ones" (Albright & Hayes 2003:154). For instance, Spanish diphthongisation (see section 2.1) was investigated experimentally by Albright, Andrade & Hayes (2001), who found that the alternation was moderately productive only for forms that occupy phonological 'islands of reliability' for diphthongisation. This term refers to subgeneralisations about phonological environments in which a morphological process is especially robust (see also Albright 2002). To capture the alternation, an inductive learner computed a grammar containing 3.346 rules. Rules with a higher adjusted reliability override competing rules, representing islands of reliability for diphthongisation (i.e. structural descriptions for verbs that diphthongise). The model also learned detailed rules for the "no-change" mapping, which was always judged better than the alternation, by the computer model as well as subjects. Albright et

[18] A 'Rules over Words' model is also proposed by Yang (2003). Here, only the irregular stems undergoing 'minor rules' are lexically listed but not the inflected forms.

al. (2001) argue that the alternation does not reflect a general phonological rule of Spanish, as knowledge of diphthongisation was highly sensitive to details of phonological shape. Contrary to claims of the Dual Mechanism model, they argue that there is no default pattern for Spanish diphthongisation, since even the 'no-change' pattern was found to be influenced by phonological islands of reliability. Finally, it is not easy to distinguish this model empirically from analogical models, to be discussed further in section 2.3. Before we turn to a more detailed comparison of rule-based models and analogical models, the acquisition of rules will first be discussed.

### 2.2.2 Acquisition of rules

It is a well-known observation (e.g. Marcus et al. 1992) that children often produce correct 'irregular' forms (such as *feet*, *broke* or Dutch *liep* 'walked') *before* they make 'regularisation' errors such as **foots*, **breaked* or Dutch **loopte* 'walked'. At a later stage they correct these mistakes again, which is why the phenomenon has been called the '*U-shaped learning curve*' (see Plunkett & Marchmann 1991, Marcus et al. 1992, Menn & Matthei 1992 and references therein). Such a learning curve is compatible with the idea that children first store unanalysed wholes (e.g. *feet*), and only produce 'errors' upon acquiring the relevant rule. This type of productivity has traditionally been taken as evidence for children's knowledge of rules. Under this view, the task for the Dutch language-learning child is to learn both the language-specific set of *morphological* rules (for instance, to add /-ən/ to form a plural, e.g. Ø → ən / [X___]$_{[+\text{PLUR}]}$), *phonological* rules (such as [-son] → [-voice] / ___ #), their *domain* and their respective *ordering*, i.e. the fact that final devoicing is ordered before the rule of regressive voicing assimilation. Moreover, children need to arrive at abstract underlying representations, e.g. /bɛd/ for cases of voicing alternations.

*SPE*-type frameworks do not offer a detailed solution of *how* acquisition might proceed, other than that *Universal Grammar* provides the set of universal *features* (such as [+/-voice]) and the *format* of rules (A → B / X_Y). Also, the principle of *rule ordering* is thought to be present (Chomsky 1976). An important learning strategy is known as the 'Subset Principle' (e.g. Pinker 1986), which expresses that the learner is essentially *conservative,* assuming the smallest possible inventory unless positive evidence is provided. Finally, a *Language Acquisition Devise* (LAD) is the proposed neurological mechanism that makes acquisition possible.[19] It is generally assumed that although words may be stored as wholes in the first stage of learning, the child will notice the semantic overlap between alternants and arrive at a more abstract lexical representation at a later stage. This type of learning thus implies that words that were once stored (e.g. [bɛt]) lose their memory trace as the lexical representation is *restructured*. Generally, generative theory holds that children's grammars are adjusted independently of their acquisition of lexical items.[20] Thus, a change to the grammar should have a uniform effect on all words (see Smith 1973 vs. Macken 1980).

The Words-and-Rules model bases its theory of learnability on Pinker (1984) and Pinker & Prince (1988), who subscribe to the *continuity hypothesis* (i.e. the continuity between child and adult grammars). UG may contain the *blocking principle*, the *subset principle,* the *Uniqueness Principle* (i.e. a child assumes one form per grammatical category) and operations such as *affixation* and the search for a unique default. All models of learnability need to address the induction problem that arises when several possible generalisations are consistent with the data. More specifically, all models face the question of how the child comes to know that exceptional patterns (such as *sing ~ sang*) are not general productive rules. For instance, how would a child know that a possible rule such as intervocalic voicing in Dutch is in fact an incorrect generalisation? On the basis of the singular - plural pair [bɛt] ~ [bɛdən] 'beds', a child acquiring Dutch could arrive at a rule such as in (10) below, after which rule collapsing could eventually lead to a more abstract class or domain, e.g. referring to an intervocalic position as in (11).

| (10) | (11) |
|---|---|
| Change t → d | Change t → d |
| Class: ɛ_ə | Class: V_V |

[19] Chomsky (1980:233) has stated "...then he [the child] will know the language compatible with his limited experience, though there will be no relation of generalization, abstractions, induction, habit formation, or the like that relates the system attained at the final state to the data of experience. The relation between experience and knowledge will be quite abstract."
[20] However, in Lexical Phonology the stored output of one level is the input to the next.

Crucially, a separate rule of suffixation (adding *-ən* to the stem) needs to apply *before* a change to the stem. Also note that there is no competing rule candidate which will lead the child to abandon (11), since the crucial pair [pɛt] ~ [pɛtən] 'hats' contains a no-change stem. However, note that the occurrence of [pɛtən] would weaken the support for (19). In this scenario, knowledge of the affix (-ən), combined with the rule of final devoicing should lead the child to change the underlying representation of the singular form /bɛt/ to /bɛd/, so that a rule such as (11) cannot apply anymore. However, it is unclear *how* (or upon what evidence) lexical representations are restructured by the child. Moreover, within the Words-and-Rules model, the regularity of the morphological operation (e.g. adding *-en* to the stem) depends on its status as the default inflection, which is not evident for Dutch (as there is a competing productive plural affix *-s*).[21] Finally, rule-based models hold that since there is no stored form available for a novel word (that can block rule application), any input form that meets a certain structural context is expected to be treated in the same way (i.e. rules apply across-the-board).

The models discussed in this section do not specifically address how the allomorphy of alternating stems or (default) affixes is acquired, other than stating that this is influenced by phonology. We will now consider some approaches to the acquisition of *phonology* based on generative theory (for an overview see also Fikkert (1995)). Researchers of child phonology face the problem of accounting for the difference between the child's utterances and adult forms, as it is not clear whether this difference is due to the child's grammar, the child's lexical representations, or 'child-specific' reasons (such as articulatory limitations). Moreover, children seem to perceive contrasts that they are not able to produce, illustrated by the well-known 'fish' phenomenon, where the child rejects her own pronunciation when imitated by adults (Smith 1973, Braine 1974). In early developmental theories such as Stampe's *Natural Phonology* (1969; 1973) and Smith (1973), it was thought that children (in a stage prior to learning morpho-phonological alternations) have 'adult output-like' representations, under the assumption that they have accurate perception. For instance, Smith (1973:139) reports data from early alternations that demonstrate that Amahl (2;2-2;4) may have had

---

[21] According to Pinker (1999:231), the two Dutch suffixes could both be considered defaults in phonologically defined domains, which undermines the notion of a unique default that is insensitive to phonology (see van Wijk 2007 for the acquisition of the Dutch plural).

adult-like underlying forms. In this case, the contrast between stops and fricatives was neutralised in *singular* forms (e.g. *cat* → [kæt] and *horse* → [ɔːt]), but a voicing contrast appears in the *plural* form (horses → [ɔːtid]). To account for the difference between child and adult surface forms, Smith (1973) derived the child's output form from the adult-like underlying form with a long series of child- and language-specific 'realisation rules'. Development was then thought to involve the 'unlearning' or simplification of rules. However, Smith noted that this led to the undesirable result that the child had more rules than adults and that many of the child-specific rules had the same purpose (mirroring Kisseberth's (1970) 'conspiracy problem'). Stampe (1973) was the first to derive outputs by a set of innate and universal 'natural' *processes*, which are (serially) ordered or *suppressed* as development progresses. According to Stampe (1973), *rules* are mostly phonetically unmotivated and language-specific (hence they must be learned), whereas *processes* are universal and innate, reflecting articulatory limitations. Hence, Stampe was among the first to distinguish alternations that are phonetically motivated from those that are morphologically or lexically conditioned (see also Venneman (1972) and Hooper (1976)). Final devoicing was seen as a candidate for a natural process, constraining early phonology. Crucially, processes can be *ordered* when there is a conflict between absolute and contextually restricted processes. For instance, it was recognised that phonetic pressure leads obstruents to be voiceless, *except* in certain environments (e.g. between vowels). Final devoicing errors are indeed frequent for children learning English, a language with final voiced obstruents (Velten 1943, Smith 1973, Stampe 1973, Donegan & Stampe 1979).[22] Stampe thus noted that children's realisations conform to final devoicing from the outset, i.e. they did not have to 'learn' the rule. It was argued that it would in fact be hard for the child *not* to devoice final obstruents; the opposite error (of voicing final obstruents) should not be attested.[23] This view thus entails that the child does not *learn* the rule of final devoicing, but rather fails to make a voicing distinction in final obstruents (due to articulatory limitations). According to Stampe, English speaking

[22] Effects of final devoicing have also been found when speakers of languages without codas learn a language such as English, e.g. Broselow et al. (1998) found that Mandarin learners of English go through a stage in which they devoice codas.
[23] Note however that there are reports of children who acquire contrasts first in non-prominent contexts: Amahl acquired the voice contrast first in final position but continued to merge it in initial position (Smith 1973, Dinnsen 1996).

children later learn to suppress this natural process. Grunwell (1987:226) notes that some English children continue to use voiceless obstruents in word-final position for some time after the age of 3;0. Donegan & Stampe (1979:134) describe naturally occurring alternations in English child speech (e.g. [hʌgɨ ~ hʌk] *hug* for Elisabeth). They conclude that "there is nothing to indicate that phonetically motivated alternations are governed by rules which are acquired" (Donegan & Stampe 1979:132).[24] As children learn to 'suppress' this process in English, there is evidence of variability (e.g. Elisabeth had overcome devoicing only in anterior obstruents as in *tub,* but not posterior ones as in *hug*). Importantly, MacWhinney (1978) has argued that the easiest alternations to learn should be those that "make use of the phonological tendencies of most young children". In other words, "morpho-phonological patterns that are in accord with natural phonological predispositions will enter early and will seldom lead to errors" (MacWhinney 1978:18). This means that errors like */kaetz/ are not likely to occur in spontaneous speech.[25] Similarly, Bernhardt & Stemberger (1998) describe early phonologically defined alternations such as the palatalisation that occurs between words ending in /t/ or /d/ and word-initial /j/ (e.g. need you /ni:d ju:/ [ni:dʒu]), which may be overgeneralised to all obstruents (e.g. love you [lʌvʒu:]). Morphologically or lexically conditioned alternations are different, in that the pronunciation of words with or without the alternation is mostly equally natural or easy (e.g. *wifes* vs. *wives*).

Generally, rule-based models such as *SPE* and the Words-and-Rules model have been criticised on several grounds. First, 'non-rules' cannot be formally excluded, as noted by Anderson (1981) and many others. Secondly, there are issues of abstractness, as underlying segments need never surface (Koutsoudas, Sanders & Noll 1974). Thirdly, rules do not express 'functional unity' (Kisseberth 1970, Smith 1973). Thus, many rules appear to satisfy cross-linguistic constraints (such as a ban on obstruent clusters), which could not be stated explicitly (also known as the 'conspiracy problem'). A

[24] Donegan and Stampe (1979:140) cite the Polish linguist Baudouin de Courtenay, who had noted as early as in 1895 that children *lose* variants of 'neophonetic' (i.e. phonetically motivated) alternations as their speech comes to resemble adult speech. He argued that a (voicing) alternation is not simply imitated by the child but develops 'independently' (Baudouin de Courtenay 1895: 209).

[25] Although MacWhinney (1978:71) reports that English children may voice stem-final consonants, such as with the non-word /trok/ + /z/ > /trogz/, indicating that faithfulness to the plural allomorph may sometimes be more important than stem faithfulness.

related problem was the *duplication* between *static* phonology (phonotactic constraints on lexical shape) and *dynamic* phonology or alternations (Sommerstein 1974, Kenstowicz & Kisseberth 1977, Goldsmith 1993). For instance, English plural suffix allomorphy (e.g. *cat-s* [khæts] not *[khætz]) is reflected in phonotactic restrictions: the word-final sequence [tz] is generally banned from the language (more generally, 'no word may end with a voiceless consonant followed by a voiced consonant').[26] Similarly, Dutch voicing alternations reflect the fact that voiced final obstruents are banned. It has been argued that this type of knowledge may also aid the acquisition of alternations, since the same constraints are involved (Hayes 2004, Tesar & Prince 2004). This topic will be explored further in Chapter 6. Third, the link between the nature of the *process* (e.g. final devoicing) and the nature of the *environment* that triggers it is not generally established in rule-based accounts. For instance, it is relevant to note that phonological processes like final devoicing apply in weak environments such as word-final position. Finally, linguists working on child language (e.g. Smith 1973, Stampe 1973) noted that rule-based theories do not capture parallels between child speech and rules found in the languages of the world. That there is such a parallel had already been noted by Jakobson (1941), who argued that the *order* in which children acquire sounds (e.g. fricatives before stops or voiced obstruents before voiceless obstruents) is reflected in typologies of the worlds' languages. These were called 'implicational universals': a language that allows voiced obstruents also allows voiceless ones. Constraint-based theories are able to express these kinds of conspiracies, which was an important argument for abandoning rule-based approaches in child language research (e.g. Pater 1999). Before turning to acquisition of constraints, we will first consider how alternations are handled in constraint-based Optimality Theory.

### 2.2.3 Constraints and Final Devoicing

Optimality Theory (OT) is a constraint-based version of generative phonology, first developed by Prince & Smolensky (Prince & Smolensky 1993; 2004). A central notion in OT is that generalisations may be *violated*,

[26] This is also known as "Harm's generalisation", which expresses that voiced obstruents must be closer to the syllabic nucleus than voiceless obstruents (Mester & Ito 1989, referring to Harms 1973, see also Lombardi 1991).

due to mutual competition. In OT, the mapping from underlying to surface representation is maintained, corresponding to the notions of *Input* and *Output*. The well-formedness of linguistic forms depends on an evaluation of candidate output forms by a set of minimally violable and inherently conflicting constraints, ranked in a language-specific way. Instead of deriving surface forms in a *serial* fashion by rules, *parallel* competing constraints are simultaneously active on output forms. Constraints (and their relative ranking) can be seen as a *filter* through which candidates (i.e. randomly generated modifications of lexical representations) are passed during production. The *optimal* output is the candidate which survives evaluation by the constraint hierarchy with the fewest number of violations. A central notion in Optimality Theory is *richness of the base*, which means that there are no language-specific restrictions on underlying representations. In addition to the set of constraints (*Con*), two functions have been proposed to be universal and innate, one for generating output candidates (*Gen*) and one for evaluating candidates (*Eval*). Two kinds of constraints have been proposed; on the one hand, *Faithfulness* constraints express the notion that the surface pronunciation of a word needs to be faithful to the underlying representation. These constraints ensure the preservation of contrast needed to encode meaning distinctions, and are inherently in conflict with *Markedness* constraints. The only violations of faithfulness allowed are those consistent with higher ranked Markedness constraints, expressing the grammatical pressure toward unmarked types of structure (Kager 1999a:4). Here, markedness is expressed directly in the grammar, reflecting typological and functional considerations. For instance, in case of the suffix allomorphy in English, phonological constraints filter multiple candidates (e.g. [nidd] and [nidəd]), and a general constraint against geminates would eliminate [nidd] (note that rule-based theories such as *SPE* had earlier posited a rule of vowel insertion). Finally, the principle of *Lexicon Optimisation* demands that as few constraints as possible should be violated in the mapping from input to output, leading to a lexical representation that is closest to the output (resembling Kiparsky's Alternation Condition).

OT resembles its predecessors in that it separates the grammar from the lexicon. It maintains one of the central premises of generative phonology in positing a unique underlying representation for alternating forms, except when they cannot be related by any general process. Alternations occur when two or more output forms share a single input form (or UR). Hence, at least one of the alternating forms is characterised by a *lack of faithfulness* to the underlying form of a morpheme. Faithfulness constraints must therefore be ranked low enough in order to be violated (i.e. below the triggering markedness constraint). Let us now turn to a discussion of how voicing alternations could be handled in OT.

There are two rival theories that have been proposed to account for voicing alternations due to final devoicing. First, a 'positional markedness' analysis using a constraint banning voicing in codas has been proposed, mainly based on German (Rubach 1989, Hall 1992, Brockhaus 1995, Itô & Mester 1997, Féry 2003).[27] Such a constraint is formulated in (12) below, which Kager (1999a) and Grijzenhout & Krämer (2000) adopt for Dutch. Under this view, neutralisation occurs when such a markedness constraint outranks a faithfulness constraint (i.e. IDENT-IO(voice) as in (13)), demanding that segments maintain their voicing specification between the underlying and surface representation.

(12) *VOICED-CODA: 'coda obstruents are voiceless'.
(13) IDENT-IO(VOICE): 'the value of the feature [voice] of an input segment must be preserved in its output correspondent'

Tableau (14) shows how the Dutch ranking leads to final devoicing in the singular form; the output [bɛd] violates the highest ranked constraint against voiced codas (indicated by *), even though it is faithful to the voicing specification of the input. A fatal violation (i.e. of the highest ranked constraint) is indicated as (*!).

| (14) Input: /bɛd/ | *Voiced-Coda | IDENT-IO (voice) |
|---|---|---|
| a. ☞ [bɛt] | | * |
| b. [bɛd] | *! | |

Final devoicing thus results when the markedness constraint on coda voicing outranks the faithfulness constraint. Note that a reversed ranking would lead to the English pattern with final voiced obstruents. A general or 'context-free' constraint reflecting the markedness of voiced obstruents (e.g. Voiced Obstruent Prohibition or VOP 'No obstruent must be voiced') is posited to

[27] Itô & Mester (1997; 2003) propose that this constraint is the result of a local conjuction, i.e. *VC results from [*VoiObs & *Coda]), banning voiced obstruents and codas.

account for languages which lack the voice contrast altogether. The (positional) markedness of voiced obstruents is reflected by the fact that neutralisation to the (marked) voiced member (e.g. /bɛd/ [bɛd] ~ /pɛt/ [pɛd]) or neutralisation in the (non-prominent) onset position (e.g. /bɛd/ [pɛd] ~ /pɛt/ [pɛt]) are not attested (Lombardi 1991, McMahon 2000).[28] Second, Lombardi (1991; 1995; 1999) has proposed a 'positional faithfulness' analysis, which *allows* features only in certain environments. Hence, a constraint such as 'IDOnsLar' expresses that onsets should be faithful to their underlying laryngeal or voicing specification (e.g. /d/ in [bɛ.dən]), even though codas are not (e.g. /t/ in [bɛt]), see Zonneveld (to appear) for analysis of Dutch. The constraint *Lar is equivalent to the VOP constraint discussed above, prohibiting laryngeal features. This constraint outranks general faithfulness to voicing (IDLar, equivalent to IDENT-IO[voice]), resulting in final devoicing (see tableaux (15) and (16)). Here, it is more important for onsets to be faithful to voicing, whereas codas are allowed to devoice due to high ranking *Lar.

| (15) Input: /bɛd/ | | | IDONSLAR | *LAR | IDLAR |
|---|---|---|---|---|---|
| a. | ☞ | bɛt | | * | * |
| b. | | bɛd | | **! | |
| (16) Input: /bɛd/ | | | IDONSLAR | *LAR | IDLAR |
| a. | ☞ | bɛ.dən | | ** | |
| b. | | bɛ.tən | *! | * | * |

Such an analysis will be shown to lead to different predictions for acquisition of alternations, a topic to which we turn now.

A potentially relevant constraint for the acquisition of the Dutch voicing alternation (at least from the point of view of the learner) is one that expresses the preference for intervocalic obstruents to be voiced (Smith 1973, Stampe 1973, Itô & Mester 1999, Kager 1999a, Hayes 2004, Hayes et al. 2004), formulated in (17) below.

[28] Apparent counterexamples discussed in Blevins (2004) are taken from aspiration languages, arguably involving a different feature (i.e. loss of [spread glottis] rather than [voice]). Yu (2004) claims that final voicing occurs in Lezgian, e.g. [pad] ~ [pata] 'side-ERG', although incomplete neutralisation is reported, and the alternation only occurs in certain monosyllabic nouns. Iverson (1982) reports neutralisation of voicing in initial position in the Lac Simon dialect of Northern Algonquin.

(17) INTER-V-VOICE (IVV): ‘Obstruents are voiced intervocalically’

A constraint such as (17), i.e. intervocalic, intersonorant or postnasal voicing is not active in the adult grammar of Dutch, i.e. it is ranked low. Tableaux (18) and (19) below show that given the correct underlying form, the Dutch ranking that was posited above will also lead to the desired result in case of alternating plural forms (taken from Kager 1999a):

| (18) Input: /bɛd-ən/ | *VOICED-Coda | IDENT-IO[voice] | *VOICE | IVV |
|---|---|---|---|---|
| a. ☞ bɛ.dən | | | ** | |
| b. bɛ.tən | | *! | * | * |
| (19) Input: /pɛt-ən/ | | | | |
| a. ☞ pɛ.tən | | | | * |
| b. pɛ.dən | | *! | * | |

Final devoicing and intervocalic voicing are both typologically frequent and articulatorily natural, rendering them likely candidates for universal constraints (Kager 1999a:325, Kenstowicz 1994:495). Partial voicing of plain obstruents in intervocalic position is phonetically natural and has been reported for many languages (e.g. Keating et al. 1983). For instance, a rule of intervocalic voicing has been proposed for Korean (Jun 1993; 1995, Hayes 1999), although its grammatical status is somewhat unclear (Cho & Keating 2001:183). Intervocalic voicing is also attested in child language, e.g. Hayes (1999) notes that Amahl (at 2;2) produces all stops in medial position as voiced, whereas they were devoiced word-finally (data from Smith 1973). According to Hayes, this reflects the phonetic grounding of the constraints (i.e. intervocalic voicing and final devoicing), which are initially high-ranked. Constraints banning voiceless obstruents after nasals have also been proposed (e.g. Itô, Mester & Padgett 1995, Pater 1999), as postnasal voicing seems to be frequent cross-linguistically and phonetically natural (Hayes & Stivers 1996, Hayes 1999, Myers 2002). For instance, postnasal voicing is a phonological process in Equadorian Quechua, whereas it is phonetically implemented in English (Hayes & Stivers 1996). Both Ferguson (1975:11) and Locke (1983:120) describe postnasal voicing in child language even though the input did not provide evidence for it (but see Ota 1999; 2003). In sum, children are likely to show a preference for both final and intervocalic or postnasal voicing, and such constraints may influence the acquisition of voicing alternations. In the next section, the acquisition of constraints will be discussed in more detail, focusing on voicing alternations.

### 2.2.4 Acquisition of constraints

In Optimality Theory, language acquisition and the issue of *learnability* has received considerable attention. In standard OT, children are thought to have the same constraints as adults, ranked in different ways. Under the nativist hypothesis, the child's task is to rank the (innately specified) constraints according to the grammar of the target language. The child starts from an initial state and demotes or promotes constraints under pressure from input forms, until a stable (adult) final state has been reached. There is continuity, as the initial and final grammar (as well as children's developing grammars) are governed by the same constraints. It has been argued that Markedness constraints are initially ranked above Faithfulness constraints (e.g. Smolensky 1996, Gnanadesikan 2004).[29] Such an *initial state* can be seen as a learning bias to choose the *maximally restrictive* grammar that is compatible with the input data, reminiscent of the *subset principle* discussed above (Pinker 1986). The advantage of this initial ranking can be illustrated by the acquisition of voicing phonotactics. If Markedness outranks Faithfulness in the initial state (e.g. *VOICEDCODA » IDENT-IO(voice)), it would lead to a restrictive grammar banning final voiced obstruents. The occurrence of a final voiced obstruent in the input would then trigger the child to change to a less restrictive grammar which allows final voicing (i.e. IDENT-IO(voice) » *VOICEDCODA). This would lead a child learning English to rerank the constraints (i.e. to be more permissive), whereas the hypothetical Dutch child does not need to change anything. However, if the less restrictive (*superset*) grammar had been the initial choice, allowing voiced codas, nothing in the input would lead the Dutch learning child to change that grammar, i.e. it would be overly permissive (Prince & Tesar 1999, Hayes 2004, Kager, Pater & Zonneveld 2004). To reach an adult grammar, markedness constraints need to be *demoted* during acquisition, upon positive evidence (e.g. that final voiced obstruents are allowed). It is assumed

[29] Hale and Reiss (1998) argue for an initial ranking of faithfulness over markedness, as they view initial lack of faithfulness to adult forms as resulting from articulatory difficulty only. Bernhardt and Stemberger (1998) view the initial ranking as partially random on the basis of variation between children.

that reranking of constraints is *error-driven*, i.e. when a child detects the difference between her output form and the adult form, she will rerank constraints to approximate the adult target. In standard OT, transitions in developmental stages (corresponding to constraint hierarchies) should be sudden and definitive without regressions, see Kager (1999a:298).

Another assumption in standard OT is that children's earliest underlying representations are taken to be the adult surface form (e.g. Gnanadesikan 2004), upon evidence that children's perception is more accurate than their production (Smith 1973, Ingram 1989, Dinnsen & Barlow 1998). When facing alternations, a child should be 'unfaithful' to the surface form, to posit an underlying form that differs from it. Thus, alternations should be taken as evidence that faithfulness is dominated. This means that high-ranking faithfulness constraints need to be demoted, instead of the general scenario in which markedness constraints are demoted. As noted by Tesar & Smolensky (2000), there is a language learning paradox involved in the acquisition of alternations: the choice of the grammar depends on how input forms are analysed, but the analysis of input forms depends on the chosen grammar. Hence, a child learning alternations first needs to decompose a complex form into a stem and affix, and then faces the problem of inferring both the correct underlying forms *and* the correct ranking between constraints. Smolensky (1996) has argued that the child will select the input representation that matches the adult output representation as the optimal input. In OT, this notion corresponds to *Lexicon Optimisation* (Prince & Smolensky 1993, Tesar & Smolensky 1998), which entails that the optimal candidate is the one that is most faithful to the perceived output. Of course, input representations need to be revised when morpho-phonological alternations are encountered: [bɛt] is the optimal input candidate for the perceived output [bɛt], but is assumed to be restructured when the child relates the perceived output [bɛdən] 'beds' to the perceived output [bɛt] 'bed'.

In an attempt to give a scenario for the acquisition of the Dutch alternation, Kager (1999a) observes that when the plural /bɛdən/ is heard, the Dutch-learning child could posit the underlying form /bɛt/ (i.e. a copy of the surface form of the singular) and arrive at a constraint ranking where intervocalic voicing (INTER-V-VOICE) dominates the faithfulness constraint IDENT-IO[voice]. Such a grammar would predict overgeneralisations of the type *[pɛ.dən] 'caps' rather than [pɛ.tən], as illustrated in tableau (20):

| (20) Input: /pɛt-ən/ | *Voiced-Coda | Inter-V-Voice | IDENT-IO[voice] |
|---|---|---|---|
| a. pɛ.tən | | *! | |
| b. ☞ pɛ.dən | | | * |

As noted in Chapter 1, such erroneous forms are attested for Dutch children. Kager (1999a) uses a ‘constraint demotion’ algorithm or CDA (Tesar & Smolensky 1998; 2000, Tesar 1996), and shows that it is theoretically possible to learn the correct grammar and underlying form, provided that the learner can go back and forth between estimating the underlying forms and estimating the constraint ranking. The learner receives a signal that the UR must be changed rather than the grammar, when she passes through the same grammar twice (see also Tesar & Prince 2004). Crucially, Kager argues that the principle of Lexicon Optimisation ensures that the incorrect underlying form /bɛt-ən/ can be restructured by the learner, since the alternative /bɛd-ən/ is available (copied from the surface form [bɛdən]). Importantly, a phonological approach to the acquisition of voicing alternations would predict that children will initially overgeneralise voicing in forms such as *[pedən] ‘caps’, due to high ranking constraints favouring intervocalic or postnasal voicing. This topic will be addressed further in the experimental chapters.

A second type of error that is predicted by phonological approaches is motivated by *Paradigm Uniformity* or maintaining a non-alternating paradigm. Crucially, such a strategy would result in errors such as *[betən] ‘beds’, in which the voicing value of the singular is matched in the plural. Presumably, uniformity towards *[bɛd] on the basis of the plural form [bɛdən] is blocked by knowledge of final neutralisation. The notion of Paradigm Uniformity in child language is well attested. An example involving ‘flapping’ in American English is provided by Bernhardt & Stemberger (1998). They describe a child who (from 2;0 to 3;8) realised taps in mono-morphemic words invariably as [d], as in *water* [waːdoʊ]. However, alveolar stops in bi-morphemic words were realised either as [t] or [d], depending on which appeared in other inflected forms. Hence, the child produced *[sɪtɪŋ] for [sɪɾɪŋ], on the model of [sɪt], leading Bernhardt & Stemberger to argue that “some faithfulness constraints are ranked higher than in the target adult language, for some children” (1998:636). Another example is reported in Kazazis (1969) who provides data from Marina, a four-year-old learning Greek. The sequence *[xe] (velar consonant before front vowel) was produced in [exete] ‘you-pl. have’ (adult [eçete]) on account of [exo] ‘I have’. In these cases, there seems to be a closer match between morphologically related forms in child than in adult phonology (see also Pater (2000). Constraints that drive uniformity (i.e. faithfulness between different surface forms of a morpheme) are presumably ranked *a priori* high (i.e. at the top of the hierarchy) by children acquiring language (McCarthy 1998, Hayes 2004).

In a recent OT approach, *Output-to-Output Correspondence* constraints are posited to account for PU effects. These constraints are extensions of ‘Correspondence Theory’ (McCarthy & Prince 1999), which is based on the phonological similarities between morphologically related words (Benua 1995; 1997; 2000).[30] Generally, phonological learners must have an inherent bias for high-ranking OO-Faith to prevent the acquisition of superset grammars (McCarthy 1998, Hayes 2004, Tessier 2006). However, in case of the Dutch voicing alternation, the form [bɛt] would be preferred to [bɛd] because the phonological constraint of Final Devoicing is ranked above Output-Output Correspondence constraints (Burzio 2000:63, see also Kager to appear). Kager (1996, 1999a:418-420, 1999b) further considers an extension of this model, in which the *distribution* of listed allomorphs (e.g. bɛt ~ bɛd) is governed by markedness constraints (for similar ideas see, e.g., McCarthy & Prince 1993, Mester 1994, Anttila 1997, Rubach & Booij 2001, Yip 2004). Positing faithfulness constraints between output forms (enforcing surface-to-surface similarity) clearly has implications for underlying representations. Hence, the ultimate consequence of such an approach is that underlying forms are dispensed with, as proposed by some (e.g. Burzio 1996; 1998, Steriade 1999, Burzio 2002).[31]

---

[30] In Benua’s Transderivational Correspondence Theory, constraints require a (morphologically derived) surface form not to deviate from the surface form of its (morphologically simplex) base, rather than from its underlying representation (Benua 1995; 1997; 2000). Other proposals involve constraints that require morphologically related lexical items (i.e. words, morphemes, stems or affixes) to resemble one another without assigning priority (i.e. symmetrically), also known as ‘Uniform Exponence’ (Kenstowicz 1996), ‘Paradigm Uniformity’ (Steriade 2000), ‘Anti-Allomorphy’ (Burzio 1996), or ‘Optimal Paradigms’ (McCarthy 2005), see also Downing et al. (2005).

[31] McCarthy (1999:385) proposes that the problematic Dutch interaction between final devoicing and resyllabification (Booij & Rubach 1987) can be analysed in this way, if the unexpected devoicing in [vɪn.tɪk] is interpreted as an effect of O-O Faithfulness to the base [vɪnt], which has undergone final devoicing (cf. Grijzenhout & Krämer 1999).

In recent proposals, universal and/or child specific constraints may *emerge* in acquisition, in response to *functional* articulatory and perceptual limitations (Menn 1980, Bernhardt & Stemberger 1998, Boersma 1998, Hale & Reiss 1998). Furthermore, constraints that are thought to be initially high-ranked, such as intervocalic or postnasal voicing (reflecting phonetic naturalness), have been argued to be 'created' by children through *inductive grounding* on the basis of the learner's experience in articulation and perception (Hayes 1999). In this view, an algorithm for inductive grounding permits the child to arrive at the appropriate set of formal phonological constraints, which also include (non-phonetic) Faithfulness constraints and constraints on Paradigm Uniformity.[32] Others have described markedness constraints as generalisations over the child's early lexicon (Fikkert & Levelt 2004). An argument in favour of emerging constraints are child-specific phenomena such as consonant harmony (Levelt 1994, Fikkert & Levelt 2002, Pater & Werle 2003, Fikkert & Levelt to appear).

Traditional rule-based approaches hold that rules are applied 'across-the-board' (i.e. to all items), as soon as the child proceeds from one grammatical state into the next. However, this does not account for the observation that a 'new' rule may be active in child phonology, even though existing pronunciations sometimes persist "as if output forms serve as independent lexical items" (Menn & Matthei 1992:213). Thus, theories of acquisition need to take into account variability and item-by-item learning. Standard OT does not incorporate factors such as lexical frequency or phonotactic probability, although partial orderings of constraints have been proposed to handle variation (Anttila & Cho 1998, Anttila 2002a). In *probabilistic* versions of OT, relative constraint rankings can be indeterminate to varying degrees. For instance, Boersma's (1998) Gradual Learning Algorithm (GLA) allows constraints to be reranked gradually rather than categorically. The GLA can deal with (noisy) input data with stochastic variation (Boersma 1997; 1998; 1999, Boersma & Hayes 2001) and it has been shown to generate realistic learning curves (Boersma & Levelt 2000, Curtin 2002), although see Keller & Asudeh (2002). Such functionally

---

[32] Note that emergent constraints could still be 'phonological', i.e. have *categorical* effects; purely phonetic or 'mechanical' effects are traditionally thought to be *gradient*. As Hayes (1999) proposes, even if (inductively learned) constraints reflect *functional* goodness, the grammar is biased towards *formal* simplicity and symmetry.

oriented approaches are closer to analogical models of language, which will be discussed in the next section (2.3).

### 2.2.5 Summary

The first part of this theoretical chapter described phonological theories that have been proposed to account for Dutch voicing alternations due to final devoicing. Generative models strictly separate a grammar (containing rules, parameter settings, or constraint rankings) from the lexicon. Recall that under rule-based approaches, acquisition of voicing alternations involves setting up the correct underlying abstract representation of a word (e.g. /bɛd/ 'bed'), such that it differs from the surface form that the child actually hears (e.g. [bɛt]). In Dual Mechanism models (Pinker 1999, Clahsen 1999) and recent rule-based models (Albright & Hayes 2003), alternating stems may be lexically listed even though regular suffixation applies. In Optimality Theory, there is an explicit link between previously acquired phonotactic knowledge and acquisition of final devoicing, which is subsequently used for the acquisition of morpho-phonological alternations.

Crucially, errors involving overgeneralisation of voicing (e.g. *[pɛdən] 'caps') could be taken as evidence that the child has overgeneralised certain (innate or emergent) phonological rules or constraints such as 'intervocalic voicing'. Most rule- and constraint-based approaches predict that these errors are common at the earliest stage of acquisition, when 'early' or 'natural' rules or constraints apply in certain phonologically determined contexts. In the acquisition process, it is expected that rules or high-ranked constraints may lead to overgeneralisations, but they should not do so unrestrictedly (i.e. the structural context of the rules or constraints should determine the outcome). Another view that has emerged from recent constraint-based theories of language acquisition is that constraints demanding 'paradigm-uniformity' (i.e. non-alternation) are initially ranked high in the child's grammar (e.g. Hayes 2004, Tessier 2006). This would lead children to produce errors of the opposite kind (*[bɛtən] 'beds'). As both types of errors are attested in spontaneous speech, these predictions will be tested further by eliciting both words and non-words (Chapter 5 and 6).

In the next section, analogical models of language will be discussed. Connectionist and usage-based approaches do not subscribe to the view of abstract representations, but nevertheless share some properties with more recent approaches within generative phonology. Different theoretical perspectives will be shown to lead to different assumptions about what constitutes knowledge of alternations, which in turn will lead to different predictions for acquisition.

## 2.3 Analogical models

The term analogy dates back to early structuralists such as Paul (1886) and de Saussure (1915/1969). It refers to the similarity between members of pairs or sets of linguistic forms that serves as a basis for the creation of other forms. Under this view, rather than constituting evidence for rules, the productivity that is found in child language (e.g. errors such as **breaked*) can be accounted for by analogy (or 'family resemblance') to other forms in the lexicon. Another view of analogy is that it is an initial basis for the acquisition of rules (i.e. as shown in section 2.2, some kind of analogy is needed to determine the conditions under which a rule applies). However, in generative theory, the individual instances or exemplars that led to the rule are discarded from memory as soon as the rule is established, whereas they remain active in analogical models. Most current analogical models recognise that words are analysed into constituents, even though they are stored in the lexicon. There is psycholinguistic evidence that decomposition plays a role in word recognition and speech production (e.g. Taft & Forster 1976, Taft 1979; 1994, Roelofs 1996, Levelt, Roelofs & Meyer 1999). Importantly, in order to account for people's ability to produce and understand new words, a mechanism of analogy is proposed that relates fully specified surface forms to each other, instead of to more abstract underlying forms (see also Cole 1995, Steriade 1995, Burzio 1996, Cole & Hualde 1998). As Skousen (1989:3) notes, the main problem with traditional notions of analogy is that there is "no limit to its use". However, current analogical models (e.g. Bybee 1985; 1988; 1995; 2001, Rumelhart & McClelland 1986, Stemberger 1985; 1994, Baayen et al. 1997, Baayen 2003, Skousen 1989; 1992, Eddington 2000, Langacker 2000, Bertram et al. 2000) use more refined and explicit definitions of analogy to model linguistic behaviour.

Importantly, analogy does not refer to the influence of a single 'similar' word, but is connected to the gang effect of similarly behaving words, which makes it a notion much closer to traditional rules. Another important tenet of analogical models is that phonological knowledge is not invariant or deterministic (such as in standard OT, where each input is mapped onto a single output) but rather probabilistic in nature. Instead of accounting for variation with mechanisms within the grammar (e.g. by differences in constraint rankings that can incorporate probabilities, as in Boersma (1998) or Anttila (1997, 2002b)), these models also take into account extra-grammatical and extra-linguistic factors such as meaning and speech style. Generalisations over forms are expressed as patterns based on phonetic and / or semantic similarities, which can be used for formation of new words. Patterns that apply to more items are stronger and more accessible, in contrast to rules or constraints in standard generative models, which are fixed or equally accessible no matter how many forms they apply to. Importantly, phonological and morphological knowledge resides in the network of associations between stored forms in the lexicon. The probabilistic rule model proposed by Albright & Hayes (2003) is similar to analogical models in that multiple generalisations can arise over output forms. Hence, it is argued that speakers do not capture certain regularities even if they do not violate the principles posited by Hale and Kiparsky, as discussed in section 2.2 (Skousen 1975).

With regard to language acquisition, proponents of analogical models hold that children construct mental representations for language on the basis of their linguistic experience (or input) and general-purpose learning mechanisms (e.g. Rumelhart & McClelland 1986, Elman et al. 1996). As we have seen, the 'poverty of the stimulus' argument entails that highly abstract structures of adult grammars cannot be learned with general-purpose mechanisms. Rather than reverting to innate knowledge, analogical models do not represent adult grammars with these abstract notions in the first place, leading to different assumptions about the nature of the acquisition task (e.g. Tomasello 2005). In the next sections, two types of analogical models will be discussed in more detail, i.e. connectionism and usage-based models. For reasons of space, the discussion excludes detailed descriptions of single-route analogical models such as the Analogical Modelling of Language or AML approach (Skousen 1989, Eddington 2000) or the Tilburg Memory-Based Learner TiMBL (Daelemans et al. 1993, Daelemans et al. 2004, Daelemans & van den Bosch 2005). Such models have successfully modelled human behaviour in recent psycholinguistic studies in Dutch (e.g. Baayen et al. 1997, Krott, Baayen & Schreuder 2001, Ernestus & Baayen 2004, Keuleers et al. 2007), to be discussed further in Chapter 3 (section 3.4) and Chapter 5. Before turning to usage-based theories, let us first consider connectionism, which was one of the first attempts to model language acquisition directly.

### 2.3.1 Connectionism

Connectionist or *neural network* models constitute a major theoretical alternative to rule-based theories (e.g. Rumelhart & McClelland 1986, Plunkett & Marchman 1991). The first model that was proposed is the influential 'Parallel Distributed Processing' (PDP) pattern associator network of Rumelhart & McClelland (1986), originally devised to model the acquisition of the English past tense. Like *SPE*, this model can be called a *single-mechanism* model, since it maps from input to output in a single step; e.g. all past tense forms (regular and irregular) were derived by direct modification of the stem. The mechanism consisted of an *encoding / decoding network* for converting phonological representations into featural representations (and back) and a *pattern associator* to link input units with output units, which have *connections* depending on activation values. Activation *weights* are stored and as the system 'learns', the weights change over time (resembling actual neuronal activation in the brain). Strong connections are associated with patterns that occur often, i.e. regular inflections such as *walk-walked*. Hence, the learning mechanism is sensitive to type frequency, as an increased number of types strengthen a pattern of morpho-phonological change. After a training phase, the model was shown to generalise to novel instances, handling both regular and irregular patterns. Crucially, the authors claimed that the network's behaviour seemed rule-based, even though "...[t]he rules themselves are not written in explicit form anywhere in the mechanism" (Rumelhart & McClelland 1986:217). Hence, a sub-symbolic view is favoured, in which probabilistic factors such as similarity and frequency interact in complex ways (Seidenberg & MacDonald 1999). This model is comparable to the stochastic rule model proposed by Albright & Hayes (2003) in that it involves inductive learning of detailed generalisations, rather than the most general (or default) rules. Connectionist models have been specifically designed to answer questions of learnability; i.e. whether language can be *learned* without access to innate ('hardwired') linguistic knowledge. Let us therefore first evaluate how the performance of such a model compares with child language acquisition, before we resume the arguments that have been made against this approach.

The Rumelhart & McClelland model succeeded in displaying behaviour typically found in child language, such as the 'U-shaped learning curve', in which correct forms are succeeded by 'regularisation' errors. According to the Dual Mechanism account (e.g. Pinker & Prince 1988), these errors occur when memory storage fails to block the application of the regular rule to an irregular stem. However, the connectionist model did not have access to the distinction between regular 'rule-based' forms (e.g. *walk ~ walked*) and irregular 'stored' forms (e.g. *break ~ broke*). The model's shift from correct to overregularised forms is driven directly by shifts in the environment, because irregular verbs tend to be high in frequency. In contrast, Pinker & Prince argued that the U-shaped curve emerges from an endogenous process, which is "a mechanism that makes categorical decisions about whether a hypothesized rule candidate is a genuine productive rule and about whether to apply it to a given verb". Thus, in a symbolic model, once the regular rule is discovered, it "doesn't care what's in the word or how often its contents were submitted previously for training; the concept of a stem itself is sufficient". These predictions are not borne out by data however, as it has been found that vocabulary size correlates well with overregularisation errors (Marchman & Bates 1994). According to the 'critical mass' hypothesis, the onset of abstractions (in the form of overregularisations) is triggered by some absolute size of exemplars (e.g. verb forms).

The Rumelhart & McClelland model has been criticised on other grounds (see Lachter & Bever 1988, Pinker & Prince 1988). For instance, it was argued that the network could handle the allomorphy of the past tense morpheme (i.e. /d/, /t/ and /ɨd/), but that it would not be able to represent the fact that the same process governs the plural suffix allomorphy. According to Pinker & Prince, the pattern of voicing assimilation responsible for the alternation must be factored out of morphology and 'stand on its own', as it is clear that "phonological and phonetic processes are entirely insensitive to morphology" (1988:107). Pinker & Prince further argue that "morphological localism destroys uniformity by preventing generalization across categories and by excluding inference based on larger-scale regularities. Thus it is inconsistent with the fact that the languages that people learn are shaped by these generalizations and inferences" (Pinker & Prince 1988:107). Clearly, final devoicing would be considered a prime example of such a generalisation. However, there is some evidence against the view of broad generalisations, as experimental results for diphthongisation in Spanish show. Here, the alternating pattern (*contámos* [kɔntamɔs] ~ *cuénto* [kwɛnto] 'we/I count' and *sentámos* [sɛntamɔs] ~ *siénto* [sjɛnto] 'we/I sit') can be described by a simple generalisation that collapses the changes /o/ → [wɛ] and /ɛ/ → [jɛ], expressing them as a single rule.

However, experimental results show that knowledge of diphthongisation is sensitive to environments specific to front and back vowels, i.e. a possible phonological generalisation (i.e. over both front and back vowels) is not made by speakers (Bybee & Pardo 1981, Albright et al. 2001).

Finally, Skousen (1989) argues that neither deterministic rule-based models nor connectionist models can capture variability in behaviour, i.e. the ability to switch from one rule of usage to another (by selecting a different outcome from the analogical set). A connectionist network will not display variability of this kind, as changes only occur after changes in the 'environment' (i.e. connection weights). Skousen (1989:85) reports several examples of variability in child language, e.g. the attempts of a boy (5;10) pronouncing the word *cliffs* in a single session: /klɪftəz/, /klɪfs/, /klɪvz/ and /klɪfs/, resembling the extensive variation found in sociolinguistic research (Labov 1972). This kind of variability is hard to handle even for models like stochastic OT. However, Bernhardt & Stemberger (1998) account for variability (e.g. in pronunciation of *sitting* with medial /t/ or /d/) by 'unstable' constraints, allowing the ranking to be more flexible than in standard OT. In this model, faithfulness constraints are indexed to 'activation levels', which can fluctuate because of 'cognitive resource limitations' (although constraints are still categorical). Stemberger (1994) has argued that Rumelhart and McClelland's model is an attempt to build (morphological) rules directly into the system (i.e. the base form is transformed into the past tense form), whereas a truly 'rule-less' model would derive morphology as an epiphenomenon or accidental by-product of lexical and phonological processing. Pinker & Prince (1988) have argued in a similar vein that the connectionist network is a strictly 'feed-forward' design, only generating past forms from stems (whereas evidence from backformations suggests that children can also retrieve stems given a past tense).

In sum, connectionist models view alternations in a different way from generative models of phonology, as they do not separate phonology from morphology or regular from irregular inflection. Abstract underlying forms are dispensed with: structure emerges from weighted connections between input and output. Connectionism is comparable to constraint-based approaches in that outputs are derived from inputs in a single step, although it does not distinguish grammar from the lexicon (i.e. there is no step from underlying to surface representation at all). Connectionist models do not take into account extra-linguistic or functional factors, perhaps because their focus has largely been on modelling certain aspects of linguistic behaviour such as the past tense. Connectionist models are like rule-based models in the sense that they are 'input-based', i.e. they do not allow for generalisations over output forms, which is an important tenet of the usage-based approach to language. The next section concludes the discussion of theoretical models by describing this approach in more detail. Bybee's network model will be shown to be particularly relevant to the acquisition of morpho-phonological alternations.

### 2.3.2 Usage-based theory

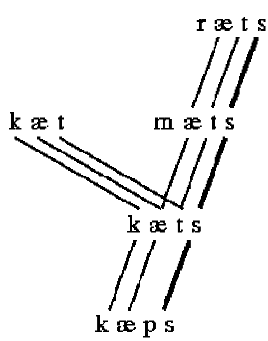
Usage-based models (e.g., Bybee & Slobin 1982, Bybee 1985, Langacker 1987; 1988; 1990, Croft 1991, Langacker 1991, Bybee 1995, Goldberg 1995, Tomasello 1998, Barlow & Kemmer 2000, Langacker 2000, Bybee 2001, Bybee & Hopper 2001, Croft 2001, Tomasello 2003, Croft & Cruse 2004) are largely compatible with connectionist models in the sense that grammatical generalisations are based on particular forms in the lexicon. Lexicon and grammar represent "two degrees of generalisation over the same memories and are thus strongly related to each other" (Pierrehumbert 2001:3). Generalisations are seen as 'emergent patterns', not explicit rules. However, usage-based models differ from connectionist models by explicitly incorporating functionalism and language use into the theory. The term *usage-based* refers to the proposition that "language structure emerges from language use, both historically and ontogenetically" (Tomasello 2003:327). Thus, the linguistic system is built from lexically specific instances; general representations may be abstracted, but they are activated together with specific instances. This approach has much in common with constructionist approaches; both study form-meaning pairings or 'constructions' directly, without distinguishing between 'competence' and 'performance' or 'core' and 'periphery' (e.g. Goldberg 2003). A central insight shared by usage-based models is that speakers do not just form generalisations about the relation between inputs and outputs, but also about the outputs themselves, so called *product-oriented* effects (Bybee 1995, 2001). As we have seen, in OT such effects could be handled by surface (or Output-to-Output) constraints in the phonology, as suggested by Burzio (1996, 2002) and others. In contrast, derivational rules express *source-oriented* generalisations, i.e. they change an input to a phonetic output, reflecting a single operation with a single set of conditions, unaffected by frequency. Similarly, the Rumelhart &

McClelland model forms generalisations over base and derived form. Albright & Hayes (2003) note that their rule-based model could not handle the fact that a substantial portion of their responses (22%) was product-oriented, e.g. subjects formed a past tense with [o] in response to a novel verb with [ɪ], even though this pattern is not attested for any real verb (it was apparently due to forms like *ride ~ rode* [ɑi] ~ [o]). Apparent 'single form analogies' were also subsumed under these product-oriented effects, and Albright & Hayes (2003) argue that an input-based model could accommodate these findings by allowing generalisation across multiple structural changes (i.e. comparing several changes), or by using surface constraints such as suggested by Burzio. In sum, rule-based models favouring formal simplicity run into problems when faced with actual behaviour of subjects (see also Zuraw 2000). In usage-based models, abstract *schemas* emerge from relations among words, which makes them product-oriented rather than source-oriented. In contrast to rule-based theories, there is no derivation of one schema from another. Bybee (1995:443) argues that a product-oriented schema (such as the one governing the *ring~rang* class) is more coherent phonologically (and at least as coherent semantically) as its counterpart 'source' schema. Lastly, in contrast to generative models, frequency of use determines which schemas emerge, i.e. frequency is not a constraint added to an otherwise structural account (as in Boersma 1998 or Bernhardt & Stemberger 1998).

In what follows, Bybee's network model will be discussed in more detail, as it is deemed most suitable to account for morpho-phonological alternations. In this model, inflected words are not stored as unanalysed wholes: they form connections with related words in the same paradigm and in other paradigms that show morphological similarities. Such a 'paradigmatic' view of morphology is shared by many traditional linguists (e.g. Aronoff 1976, Booij 2002). Crucially though, symbolic rules are maintained, leading to views more in line with Pinker's (1999) Dual Mechanism model. In contrast, Bybee proposes that schemas emerge from connections with varying lexical strengths. Connections or links between words express similarity in both form (phonological connections) and meaning (semantic connections); similarity in both is called a 'morphological' connection. Lexical strength depends on frequency of use; a word form that is used frequently enough to be stored independently is more *entrenched* in memory (or more *autonomous*) than an infrequent word (Langacker 1987, Bybee 1985; 1995; 2001). Thus, for both 'regular' and 'irregular' words, storage is ultimately a function of their token frequency. This is because high frequency items can be learned on their own terms, while lower frequency items are better learned in relation to existing items (Bybee 1995:429). For instance, a highly frequent noun plural such as 'boys' may be entrenched, whereas a noun plural schema is used to derive the plural of an infrequent form like 'cornice' (taken from Croft & Cruse 2004). This view also accounts for the well-known correlation between regularity and frequency, as high-frequency irregulars tend to resist regularisation. There is also experimental evidence for this relation, as Bybee & Slobin (1982) found that subjects were likely to regularise items with low token frequencies.

According to usage-based models, speakers make generalisations about regular *and* irregular forms in the form of a schema, which describes general phonological properties of a morphological class, used in organising and accessing the lexicon. Phonotactic generalisations are also argued to reflect frequency distributions in the existing lexicon (Bybee 2001). In contrast to symbolic rules or constraints that belong in a component separate from the lexicon, schemas have no existence independent of the lexical units from which they emerge. Figure 1 below (taken from Bybee 1995:429) illustrates how lexical connections yield word-internal morphological structure. In this figure, thicker lines represent stronger connections. Stems are connected to inflected forms, but inflected forms themselves also form connections to other inflected and non-inflected forms.

*Figure 1: A schema in Bybee's network model.*




This example shows how morphological structure *emerges* from the connections between identical and similar phonological and semantic features. Bybee (2001:98) argues that in the process of assigning morphological structure, meaning is more important than phonological shape. Thus, when there is only a phonological connection between words (e.g. in case of homophony), the connection is very weak, whereas a solely semantic connection (as in suppletive forms such as *go ~ went*) will yield a stronger effect (as evidenced by regularisations such as **goed* for *went*). To form a coherent schema, there must be enough resemblance between the types that contribute to the entrenchment of a schema (Bybee 1995:430). For instance, the three English past tense allomorphs each form a schema, showing identity of meaning and phonological family resemblance (reinforced by their complementary distribution), which may lead to the formation of a superordinate category such as [-ed/PAST] (Croft & Cruse 2004). Pairs such as *ring ~ rang* and *swing ~ swung* form a relatively coherent class, and are predicted to be mildly productive due to the their relatively high type frequency and low token frequency. This prediction is borne out, as has been noted before: Bybee & Slobin (1982:278) obtained novel past tenses such as *spling ~ splung*, *streak ~ struck* and *clink ~ clunck* in a sentence completion task (see also Bybee & Moder 1983, Prasada & Pinker 1993).

The notions *productivity*, *regularity* and *transparency* were argued to be important notions in determining types of morpho-phonological alternations. We have seen that *Dual Mechanism* models attempt to posit clear-cut boundaries, especially between regular and irregular forms. As noted in section 2.2, the existence of forms with a regular suffix and stem change (e.g. *feel ~ felt* or Dutch *bed ~ bedden*) suggest that regularity is on a continuum. In support of this, Marslen-Wilson et al. (1994) found in a cross-modal priming experiment that English mixed verbs produced a degree of priming intermediate to that of regulars and other classes of irregulars. In usage-based models, *productivity* of a pattern is directly related to its type frequency (e.g. Bybee 1985:133). Bybee proposes that type frequency is represented in the strength of the schema or the number of different word forms that are instances of a particular schema. Also, there is no clear-cut distinction between regular and irregular forms. A regular form may either be accessed from the lexicon or derived by applying a schema, depending on the token frequency of the form. In this sense, Bybee's network model could be said to be a *dual route* model, in which regularity (and the empirically related notion of productivity) is considered a gradient notion (see also Baayen et al. 1997, Bertram et al. 2000, Baayen 2003). A second determinant of productivity is whether or not the schema is relatively 'open', i.e. imposing the least phonological specificity on a stem. Bybee (1994) presents the following overview (21):

(21)

| | |
|---|---|
| Productivity: | i. lack of motivated (phonological and semantic) restrictions |
| | ii. high type frequency |
| Regularity: | iii. lack of arbitrary lexical idiosyncrasies |
| Transparency: | iv. lack of fusion |

High type frequency and arbitrary restrictions are not independent, because any restriction typically reduces type frequency. Bybee argues that these three properties are diachronically linked, whereas synchronically they are relatively independent of one another. For instance, the German plural -*s* has a low type frequency (around 2%) and an arbitrary lexical distribution, but the least fusion (i.e. it never has effects on the stem and is free of phonological restrictions). Thus, Bybee (1995) considers the -*s* plural a particularly open schema, which explains why it has a certain amount of productivity and applies to non-canonical nouns (such as derived nouns and proper names). Similarly, the Arabic sound plural (a suffix instead of an internal stem change) is considered an open schema with low type frequency (Bybee 1995:440-442, cf. McCarthy & Prince 1994). However, these are not default rules that apply to a symbolic category as claimed by the Dual Mechanism Words-and-Rules model. In fact, Bybee argues that this schema is not more productive than plural schemas with a high type frequency. This prediction seems to be born out for both German and Arabic. For instance, German children typically inflect novel words with -*e* or -*en* (e.g. MacWhinney 1978, Clahsen et al. 1995). Moreover, Köpcke (1988, 1998) found that the German -*s* plural applied mostly to nonce nouns that end in full vowels (i.e. resembling existing -*s* plurals such as *Autos*), suggesting that its application is influenced by the phonological shape of the stem.[33] In Arabic, both the 'default' sound plural and the iambic broken plural are overgeneralised by children (Bybee 1995:442), see also Plunkett & Nakisa (1997). Furthermore, it has been shown that 'default' schemas can occur

[33] See also Köpcke (1993; 1998) and Hahn & Nakisa (2000) for evidence against the German plural –*s* as default.

without high type frequency, if the non-default forms are a phonologically well-defined, relatively narrow class and the default forms are scattered across the remaining phonological space (Hare, Elman & Daugherty 1995:626). Generally, Bybee & Newman (1994) show there is no preference for morphologically transparent formations, provided that the fused formations have sufficient type frequency. However, a pattern cannot attain full productivity if there are phonological, semantic, or morphological restrictions on its applicability. Such an account would thus predict limited productivity for the Dutch pattern of voicing alternations if its type frequency is lower than the non-alternating pattern.

In contrast to connectionist models, where frequency refers to the mapping between base and derived form, Bybee assumes that the lexical strength of the derived form itself (as reflected by its token frequency) is important, noting that "*the higher the frequency of the derived form, the weaker the mapping between it and the basic form*" (Bybee 1995:432). High frequency irregulars are thus claimed to be more resistant to analogical change (i.e. regularisation) and are more likely to undergo paradigm splits (that may ultimately lead to suppletion, as in *go-went*). This means that it is important to consider the frequency of alternating forms in Dutch, which will be investigated further in Chapter 4. High frequency has two (seemingly contradicting) effects: it induces phonetic change such as reduction (due to 'automisation'), but it also protects items from analogical change or levelling (generalisations based on morpho-phonological processes in other forms). This 'conservatism' accounts for the fact that irregularity is mostly correlated with high frequency (e.g. *weep/wept* has a tendency to regularise to *weeped,* in contrast to high frequency *keep/kept*, see Bybee 1985).

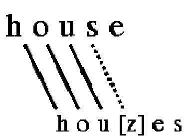
According to Bybee (2001:53), phonemes are sets of phonetically similar variants (or allophones) clustered in groups, constituting salient contextually determined prototypes (which are possibly strengthened by orthographic representation). Thus, instead of having one 'phonemic' representation, the abstracted prototype of a phonetic category is sensitive to context (e.g. speaking rate, speaker, and phonetic context). This is potentially relevant for voicing alternations, as it is unclear whether the [t] in *bed* [bɛt] 'bed', the [t] in pet [pɛt] 'cap' and the [t] in [tɑk] 'branch' are treated as the same phoneme in (child) representations. Usage-based models typically endorse *exemplar* models of representation, in which every token of experience is registered in memory (e.g. Johnson 1996, Goldinger 1997, Pierrehumbert 2001, Bybee 2002a). Exemplar models do involve abstraction, because phonetic exemplars are reorganised in favour of the more frequent types, leading to category formation. As Bybee (2002b:288) notes: "Various levels of abstraction emerge as exemplars are categorized by phonological and semantic similarity - morphemes, words, phrases, and constructions can all be seen as the result of the categorization of linguistic experiences". According to Bybee, categorisation thus involves a level of abstraction, resembling the construction of a prototype. A *schema* involves abstraction across exemplars to represent the central tendency and the variability within the category, i.e. categorisation proceeds on the basis of similarity to the stored prototype. The best exemplar of a word is the one with the highest frequency, accounting for the fact that sound change is phonetically gradual and affects different words at different rates (Bybee 2001, Pierrehumbert 2001).

In contrast, in mainstream generative theory, it is assumed that the effects of automatic neutralisation rules are *not* encoded in lexical representations, e.g. the /t/ in the singular form [bɛt] is not stored.[34] Nevertheless, it has been recognised that the rule of final devoicing can be lost in a language, whereas its effects can still be retained (Kenstowicz & Kisseberth 1979:212). The process of grammaticalisation generally involves loss of alternation and subsequent restructuring (see Plank 2000). For instance, Dutch adjectives like *bijdehante* [bɛidə'hɑntə] 'smart', literally 'by the hand' constitute evidence for levelling, since the plural of *hand* [hɑnt] is [hɑndən] (Booij 2002). An important question that naturally arises in this respect is why some forms are subject to levelling (a notion similar to Paradigm Uniformity). It is commonly assumed that there is a 'base form' for levelling, e.g. the nominative singular in case of nouns. However, a usage-based approach would predict that frequency effects can override this tendency. For instance, Albright (2004; 2005) notes that early Yiddish had alternations such as hɛl[t] ~ hɛl[d]ən 'heroes', which were levelled to hɛl[d] ~ hɛl[d]ən in modern Northeast Yiddish. Here, the direction of levelling (i.e. to a more marked paradigm) is determined by the plural base. Importantly, Albright notes coda voicing is only contrastive when there was paradigmatic pressure from the plural. According to Bybee, the vast majority of phonological alternations begin as phonetically motivated processes and gradually become more and more involved in morphology and the lexicon.

[34] Note that the problematic interaction of final devoicing and resyllabification is solved if the phonetic form of *vind* (i.e. with [t]) is stored (Baumann 1995).

They tend to lose their phonetic motivation and remain only as 'fossilized alternations' (Bybee 1994). Alternations within words can lexicalise more easily than alternations across word boundaries, which is why the latter are less common (Bybee 2001:143). This view also explains why phonological alternations may differ in their application depending on morphological context (Bybee 2001:102). For example, Bybee notes that intervocalic fricative voicing in Middle English gave rise to lexically restricted alternations in both nouns (*wife* ~ *wives*) and verbs (*leave* ~ *left*). These forms are 'morphologised', since the possessive /s/ does not trigger voicing (e.g. *my wi[v]e's car). In analogical levelling, the alternate that occurred in the more frequent form survives or 'regularises'. This has led to the preservation of voiced alternants for verbs (e.g. *bereaved* has replaced *bereft*), whereas in the (more recent) process of regularisation for nouns, the (frequent) singular form containing the voiceless fricative survives (e.g. *roofs* has replaced *rooves*, and *houses* is pronounced with intervocalic /s/ in some American dialects, Bybee 2001:101). Lexical strength of alternating forms could be depicted as in (22), where broken lines indicate shared features (rather than identity of segments), and font size indicates relative lexical strength (taken from Bybee 2001:116):

(22)




This example shows that even though the voicing alternation in English reflects the same 'phonological' process for both categories (something that could be captured by rule- or constraint-based theories), they are in fact independent of one another, and better accounted for in a morphologically based account. Bybee argues that the mechanism for regularisation is inherent to the model, i.e. there is not need to invoke 'ad hoc' OT constraints expressing Paradigm Uniformity.

It is clear that the usage-based approach could be adopted for Dutch alternating plurals, such that a lexical connection ties the singular and plural together as belonging to the same 'morpheme'. As Booij (2004:230) notes, the lexical connection between these words (e.g. [bɛt] and [bɛdən]) could be stronger than in cases of 'real' (i.e. lexically governed) allomorphy, which are not driven by a general phonological constraint such as final devoicing. Under the usage-based view, the difference between the two kinds of allomorphy is the *phonetic* nature of final devoicing, accounted for by functional considerations (see also Steriade 1997, Hayes 1999, Hayes, Kirchner & Steriade 2004, and Pierrehumbert 2001). Thus, the fact that devoicing applies regardless of morpho-syntactic category is to be expected in case of a phonetically natural process. In fact, one would expect exceptions to final devoicing to be restricted by morphology (which is indeed found for regional varieties of Dutch, e.g. van Oostendorp 2002).

As mentioned in the previous section, the productivity of morpho-phonological alternations such as Spanish diphthongisation has received considerable attention in experimental work. The diphthongisation pattern (e.g. in *contámos* [kɔntamɔs] ~ *cuénto* [kwɛnto] 'we/I count') has been described as a (lexically restricted) phonological rule in Spanish, predictable on the basis of stress (the occurrence of a diphthong in a stressed stem indicates a mid vowel in unstressed syllables). However, there are more non-alternating pairs, in which there is no diphthongisation (e.g. *flotar* 'to float'). Kernan & Blount (1966) found that both adults and children used the diphthong instead of the expected mid vowel in a nonce form (e.g. *suécha* ~ *suechó* instead of *sochó*). In a more elaborate study, Bybee & Pardo (1981) confirm that this common alternation was rarely extended to novel forms by Spanish speaking adults (see also Albright, Andrade & Hayes 2001). Moreover, productivity of diphthongisation was highly dependent on what morphological process is involved (see Eddington 1996). This outcome is expected if the alternation is a lexically specific pattern, not a rule that is independent of the lexical items to which it applies (Bybee & Pardo 1981:961, Bybee 2001:103). Diachronic evidence reveals that analogical shifts have taken place differently in different conjugations, and that the alternation is subject to levelling effects (occurring primarily in high-frequency forms). However, it is unclear whether diphthongisation is comparable to final devoicing, as it is commonly described as lexically conditioned.

Crucially, the view outlined above entails that lexical representations are not dependent on structural criteria but are determined by language use. In such a model, a rule like final devoicing is represented as a phonetically induced generalisation, while alternations are closely tied to the set of lexical items that exhibits them. Formally, there is no connection between the two, as there are no abstract underlying forms.

We will now turn to acquisition according to usage-based theory, to determine whether it leads to different predictions compared to rule-based models with respect to the acquisition of alternations.

### 2.3.3 Acquisition in usage-based theory

According to usage-based models, a theory of language and its acquisition cannot be separated from other (social-) cognitive processes such as perception, memory, categorisation, and analogy (e.g. Tomasello 2003). The *continuity hypothesis* is rejected by proponents of usage-based accounts, on the basis that "*child language is structured by much weaker and more local linguistic abstractions*" (Tomasello 2003:324). Tomasello argues that even though processes may be the same at different developmental stages, the actual structures and representations of child language are different. In a tradition dating back to Piaget (e.g. Piaget 1924), the usage-based or constructionist account does not rely on any form of innate linguistic knowledge.[35] As we have seen, usage-based and constructionist theories differ from connectionist models because of their *functional* approach. When the functional dimension is taken into account there is no 'linking problem', which derives from the exclusive focus on the formal dimension of language (Tomasello 2003; 2005). It is argued that there simply is no 'poverty of the stimulus' when the child's cognitive and semantic-pragmatic skills are taken into account, necessary for learning meaningful linguistic symbols (or form-function pairings). These skills include categorisation, analogy and statistical learning, rather than blind associations and inductive inferences with no conceptual understanding of linguistic function. For instance, infants have been shown to be able to use statistical learning to discover input patterns (Saffran, Aslin & Newport 1996, Maye, Werker & Gerken 2002). As in mainstream connectionist models, constructions are claimed to be learned on the basis of the input and general cognitive mechanisms (e.g. Langacker 1987, Langacker 1991, Goldberg 1995, Croft 2001, Goldberg 2003, Tomasello 2003; 2005).

Usage- or construction-based theory is comparable to connectionism in that it focuses on the child's learning and use of words, phrases, and

[35] Piaget (1924) rejected innatist views (such as Kant's *a priori* categories), arguing instead that categories or structures of thought are constructed.

expressions as concrete linguistic entities. Children (initially) make analogies across larger constructions, such as whole utterances (e.g. *I don't know*), rather than isolated words, morphemes or abstract categories: "when words occur together frequently, they begin to behave phonologically as if they constituted a single word" (Bybee 2001:161).[36] Under this view, the child gradually finds patterns in stored utterances (depending on their type and token frequency) and forms stronger and more abstract meaning-based representations (e.g. of words and constructions). Following Langacker (1987a) and Bybee (1995), Tomasello (2003) further claims that linguistic constructions can be accessed at several levels of abstraction simultaneously, i.e. children's comprehension and production of complex utterances can reflect simple retrieval of stored expressions or computation based on stored schemas. For instance, previous research concerning the acquisition of French 'liaison' (e.g. *nous [z] avons* 'we have') suggests that these alternations are learnt in a lexically specific way (Chevrot & Fayol 2001). This ties in which the usage-based view that French liaison is a collection of frequent *constructions* with specific liaison consonants that originated from phonetically conditioned final consonant deletion (Bybee 2001).

The acquisition of phonology is argued to be based on the gradual acquisition of more and more accurate phonetic detail in the production of words and phrases (e.g. Bybee 2001). Ferguson & Farwell (1975) have shown that children learn articulatory patterns by mastering particular words, suggesting that they are not universal and innate. Similarly, Werker & Fennell (2004) and Swingley (2004) have argued that the earliest words are stored with phonetic detail. Bybee (2001:54) argues that the first words are likely to be treated as holistic units, with children producing a wide and unsystematic range of phonetic elements. Vocabulary size then increases with 'phonemic analysis', by reusing the same elements in different words. Thus, "*children learn phonological sequences as parts of words, never independently of words*" (Bybee 2001:15). In generative models, regular forms that were once stored by the child must lose their memory trace, so that only irregular forms remain in the lexicon. The usage-based alternative is that the child stores all words (with and without affixes), and a network of associations among them begins to develop, as described above. Subparts

[36] For instance, in the earlier example of palatalisation in forms like *need you* (as noted by Bernhardt & Stemberger), it is important that palatalisation mostly occurred with the pronoun *you*, in what is arguably a small set of constructions.

with high type frequency (such as -*ed* in *played*, *spilled*, *wanted* or Dutch -*en* in *bloemen*, *sloten*, *bedden*) are reinforced because their occurrence in more combinations makes them more segmentable and accessible (i.e. high type frequency leads to greater 'analysability'), and the child builds up general schemas from local schemas. Words with high token frequencies are more 'autonomous' from the networks of associations, and can resist regularisation (parallel to diachronic development).

Importantly, in nativist approaches, children need to 'constrain' a tendency toward over-productivity (Pinker 1989). However, as recognised in Marcus et al 1992, overgeneralisation rates are generally low (less than 10% overall) and irregularisation errors (*flow* for *flew*, **brang*, **wope*, **talken*) are rare (Xu & Pinker 1995). In usage-based approaches, children are initially thought to be more conservative, repeating what they have heard. Children only start using language productively and creatively as they begin to construct categories and schemas based on their developing lexicon, mirroring frequencies in the input (e.g. Bates & MacWhinney 1987, van Valin 1991, Tomasello 1992, Bates & Goodman 1997, Tomasello 2003, Dabrowska 2006). For instance, Marchman & Bates (1994) propose that a certain number of exemplars (or a 'critical mass') is needed before the child can make abstract analogies (e.g., Bassano et al. 2004). Instead of the strict blocking mechanism proposed by modular theories (e.g., Marcus et al. 1992), functionalist models typically allow *competition* between productive forms (regularisations) and irregular forms (see Kuczaj 1977, Rumelhart & McClelland 1986, Maratsos 2000). As the child experiences more input, the irregular form gradually appears, 'winning' the implicit competition. Other researches have proposed a general cognitive process of *pre-emption* to restrict overgeneralisations. Here, the existence of a certain structure (e.g. an irregular form) enables the child to infer that an alternative structure (e.g. a regular form) is not appropriate (Tomasello 2005, Goldberg 2005).

In Bybee's network model, the gradience of schemas is reflected in the fact that individual types may be closer to or farther from the best exemplars of a category. This would lead speakers to exhibit probabilistic behaviour in assigning novel forms to one schema or another (Bybee 1995:27). The *degree* of overgeneralisation is argued to depend mainly on type frequency. Importantly, even generalisations that seem clear candidates for a source-oriented account, such as regular past tense affixation in English, may be conceptualised as product-oriented schemas. Thus, instead of "*add /t/, /d/ or /ɨd/ to a verb to form a Past*", a schema expresses that "*a*

*Past verb ends in /t/, /d/ or /ɨd/*". This would account for the frequent zero marking errors by children in case of verbs that already end in /t/ or /d/), as Bybee argues that these forms already fit a schema for past tense. Köpcke (1988; 1993; 1998) also finds evidence that English and German children use zero responses when stimuli already resemble a plural schema, arguing that "a schema-learning mechanism may underlie the acquisition of morphology, even when the end product of the learning process involves item-and-process rules" (Köpcke 1998:293). Similarly, Bybee & Slobin (1982) have argued that children in particular have a tendency to form product-oriented schemas: "we suggest that such schemas are precursors to source-oriented rules which change or add features" (Bybee & Slobin 1982:288). However, in some cases, generalisations remain product-oriented (e.g. in case of schemas such as those needed for *ring* ~ *rang*). These suggestions raise the interesting possibility that schemas are mainly used in acquisition, predicting different (i.e. rule-based) behaviour for adults. Note that the usage-based approach assumes that there are no rules for adults either: the reason that morphological operations mostly involve 'rule-like' affixation is argued to be diachronic in nature: morphology arises through *grammaticisation*, by which a previously independent word develops gradually into a grammatical morpheme or affix (Bybee 2001:129). However, it is possible that adults and children differ in the extent of overgeneralisations, partly because of differences in their lexicons and scope of generalisation (see also Pierrehumbert 2006).

To conclude, it is far from clear whether a lexical approach such as that described here is relevant to Dutch alternating plurals, which are traditionally argued to be conditioned by phonology only. If the Dutch voicing alternation behaves like lexically conditioned alternations, it is predicted that the pattern is productive only to a limited extent (since, similar to the Spanish alternation discussed above, non-alternations occur more frequently than alternations). Under such an approach, the role of the lexicon is predicted to be all important in learning alternation patterns (see, e.g., Beckman & Edwards 2000). Finally, Bybee's usage-based approach predicts that alternations are governed by morphological relations among surface forms of the paradigm, not by relations between individual surface forms and underlying forms (Bybee & Pardo 1981, Bybee 2001).

## 2.4 Summary and Discussion

The question of how alternations are acquired is not a straightforward one. The various theories discussed in the current chapter lead to very different claims about what it is that needs to be acquired. Traditional phonological theory (e.g. Chomsky & Halle 1968, Prince & Smolensky 1993) considers alternating forms to be evidence for the abstract nature of underlying lexical representations (e.g. /bɛd/ 'bed'). Alternating plurals such as [bɛdən] are considered *regular*, with the alternation fully determined by the grammar (containing phonological rules or constraints that change the singular rather than the plural). Children thus need to adjust their lexical representations and learn the appropriate rule orderings or constraint rankings. In *Dual Mechanism* models, alternating plurals could be described as *irregular* (involving associative learning), depending on the default suffix. However, it has been shown that it is not easy to draw a line between regular and irregular processes, and the Dutch voicing alternation seems to position itself somewhere along this continuum. Rather than strictly separating rule-based or regular processes from irregular processes, alternative models such as that advocated by Albright & Hayes (2003) propose inductive learning of multiple rules, which may cover subregularities. This type of model is closer to analogical models of language, in which the regularity of morpho-phonological processes is seen as a gradient notion, interwoven with factors such as productivity and transparency. Analogical models include connectionist and usage-based models, in which rules or constraints are replaced by generalisations based on surface forms in the lexicon. Crucially, in such models final devoicing has no independent status, apart from a phonetically driven generalisation. Hence, it does not *cause* children to restructure lexical representations in case of alternating plural forms. Instead, children and adults learn alternating forms solely on the basis of stored forms in their lexicon and the varying strength of the connections between them. Under this view, (prior) knowledge about voicing phonotactics or final devoicing is not necessary for learning morpho-phonological alternations. Instead of solely mapping singular forms to plural forms in order to deduce rules, the child is also sensitive to relations between inflected words and the relative frequencies within paradigms. Errors in which voicing is overgeneralised (e.g. *[pɛdən] 'caps') reflect analogy to existing words in the lexicon, on the basis of which a schema has been constructed. Although it is not clear what the precise domain of analogy is, usage-based models predict that patterns with a low type-frequency will have limited productivity. Therefore, the opposite error (e.g. *[bɛtən] 'beds') is predicted to be dominant if the non-alternating pattern is more frequent.

Deciding between symbolic vs. analogical approaches on the basis of experimental evidence is difficult, as it is clear that recent models are becoming more alike. Generally, traditional models of phonology predict that rules or constraints are consistently applied in certain contexts. Even though children may initially set up 'incorrect' rules or constraint orderings, productivity is expected to be stable (i.e. not variable) for a given set of items that have the same structural description (because a form either belongs or does not belong to a relevant category). However, we have seen that more recent versions of rule- or constraint-based theory (e.g. Albright & Hayes 2003, Boersma 2000) can handle some degree of indeterminacy or variability, which makes it more difficult to compare these models directly to analogical models. Specific predictions for the acquisition of voicing alternations will be discussed in the experimental chapters (5 and 6). First, the Dutch data will be described in more detail, in order to evaluate the nature of the task the Dutch child faces.

# Chapter 3 Voicing in Dutch

The question of how and when voicing alternations are acquired depends on theoretical assumptions about the nature of features, phonemes, rules, and underlying representations. In this chapter, the representation and acquisition of the Dutch voicing contrast will be discussed in more detail. Next, previous research into voicing assimilation and (incomplete) neutralisation of the voicing contrast will be discussed, which is important for determining the nature of the input that the Dutch child receives. Finally, an overview of previous experimental studies involving voicing alternations is given, focusing on studies on Dutch plural and past tense formation.

## 3.1 Introduction

Alternations occur when the phonetic shape of a word or morpheme varies, depending on its phonological or morphological environment. Even though the interface between phonology and morphology is central in the present study, the connection between phonetics and phonology also needs to be considered. In this respect, it is important to note that the voiced member of a pair (e.g. /d/) is considered to be 'marked' cross-linguistically (Greenberg 1966). For instance, the occurrence of voiced obstruents in languages implies the occurrence of voiceless stops, but not vice versa. Thus, many languages have only voiceless obstruents (e.g. Hawaiian), but no language has only voiced obstruents (Kenstowicz & Kisseberth 1979). Generally, the unmarked or preferred feature value is the most frequent one, and alternations are expected to make a word 'less marked'. The marked or 'unexpected' member is generally seen as the 'active' feature value, whereas the unmarked member (e.g. [-voice]) is 'underspecified'.[1] As discussed before, voicing tends to be marked in word- or syllable-final position. Hence, stops are frequently devoiced word-finally, but there are no languages in which voiceless stops are voiced in that position. On the other hand, voiced obstruents are generally allowed in initial position and even favoured in between vowels or after sonorants and nasals. Moreover, the degree of markedness of voicing may depend on the place of articulation of a segment (i.e. /b/ is less marked than /d/, see e.g. Ohala 1983). A voicing contrast thus differs from a place contrast in that it is more 'contextual'; cues to voicing differ according to the position the segment is in (Kenstowicz 1994, Steriade 1995). It has been argued that perceptually weak contrasts are especially likely to neutralise; lack of articulatory effort could result in poor voicing in final position, such that voiced obstruents tend to be misperceived as voiceless due to weak cues (e.g. Steriade 1997, Hume & Johnson 2001, Pierrehumbert 2001, Hayes et al. 2004). Let us now turn to the representation of the voicing contrast in Dutch, to determine the feature specifications which need to be acquired.

[1] Underspecification theory holds that unmarked, redundant or predictable features are absent from lexical representations (e.g. Archangeli 1988). The term underspecification can also refer to children's lexical representations, see Fikkert (1994) and Fikkert, Levelt & Zamuner (2005) for work on Dutch.

First, the phonetic aspects of the voicing contrast will be addressed in more detail.

## 3.2 The voicing contrast

Phonetically, voiced sounds are produced with vocal cord vibration; the main acoustic cue is voice onset time or VOT, indicating the difference in time (in milliseconds) between the release of a plosive and the onset of vocal cord vibration. The value of VOT is negative when voicing starts before the release of the plosive, i.e. there is vocal cord vibration during closure. This is generally referred to as 'prevoicing', assumed to be the major cue to voicing in Dutch (van Alphen 2004). Cross-linguistically, the realisation of the voice contrast varies considerably. This is illustrated for initial stops in three Germanic languages in Table 1 below (taken from Kager et al. to appear).[2] Here, voicing lead refers to prevoicing, short lag VOT indicates that voicing begins at the time of the release or shortly afterwards, and long lag VOT refers to a longer delay between the release and the onset of voicing, resulting in aspirated stops.

*Table 1: Average VOT values for Dutch, German, and English, with feature specifications according to a binary or privative interpretation.*

| | **Voicing lead** | **Short lag VOT** | **Long lag VOT** |
|---|---|---|---|
| **Dutch** | **-83** ms: b, d | **13** ms: p, t | |
| binary | [+voice] | [-voice] | |
| privative | [voice] | [ ] | |
| **German** | | **16** ms: b, d | **51** ms: p, t |
| binary | | [+voice] | [-voice] |
| privative | | [ ] | [sg] |
| **English** | | **32** ms: b, d | **59** ms: p, t |
| binary | | [+voice] | [-voice] |
| privative | | [ ] | [sg] |

This table shows that phonetically, German or English 'voiced' stops are within the range of Dutch voiceless stops (in initial position). The Dutch

2 These are average VOT values; coronal stops tend to have a longer VOT than labial stops (Lisker & Abramson 1964).

contrast is between voicing lead and short lag VOT (as in French and Spanish), while most other Germanic languages employ a contrast between short lag and long lag VOT (as in German and English).

As mentioned in Chapter 2, it has been proposed that the Dutch voicing contrast is captured by a *binary* feature [±voice], marking positive and negative values.[3] For instance, Wetzels & Mascaró (2001) claim that the negative value [-voice] is phonologically 'active' in Dutch (cf. Zonneveld to appear). However, it has also been proposed that prevoicing languages such as Dutch have a *privative* feature [voice], whereas aspiration languages such as German and English employ the feature [spread glottis] or [sg] (Iverson & Salmons 1995; 2003, Kager et al. to appear). This option is also referred to as the 'Multiple Feature Hypothesis, to be discussed further below. The acquisition of laryngeal phonology has also been investigated to address the question of featural specification, which will be discussed in the next section.

### 3.2.1 Acquisition of voicing

The voicing contrast is generally considered to emerge early in the speech of children even though its acoustic correlates may be complex, involving the timing of articulatory and laryngeal gestures (e.g. Bond & Wilson 1980). If features are universal and innate (e.g., Chomsky & Halle 1968, Stampe 1973, Prince & Smolensky 1997), it is expected that laryngeal features are acquired in an across-the-board manner. However, acquisition of voicing has been found to differ across language groups and children. Some authors have proposed that English learning children go through a stage in which they produce a voicing contrast that adults cannot perceive (Macken & Barton 1978; 1979, Scobbie et al. 2000). This covert or subphonemic contrast arises because children's voicing values for both voiced and voiceless stops fall within one of the adult categories, making it difficult for adult listeners to perceive the contrast. However, such a covert contrast is mostly attested in developmental-delayed or disordered children, and it is not clear whether it is a necessary stage for typically developing children (see Chapter 6 for further discussion). Generally, interpretation of VOT is difficult, as it tends to reflect the experimenter's bias in terms of target words and adult perceptual categories (Edwards & Shriberg 1983).

---

[3] Others express laryngeal features in more phonetic terms, e.g. Avery & Idsardi (2001).

Bernhardt & Stemberger (1998:436) argue that it is likely that young children initially produce voiceless unaspirated stops, proposing one default [-voice], which is not specified in lexical representations. However, as discussed before, both prevocalic and intervocalic voicing are common in child phonology (Smith 1973, Stampe 1973, Edwards & Shriberg 1983, Stoel-Gammon & Dunn 1985, Bernhardt & Stemberger 1998). Dinnsen (1996) also notes that children generally prefer voiceless obstruents in codas and voiced obstruents in onsets, leading him to argue for 'context-sensitive' underspecification (e.g. obstruents are specified for [+voice] in final position).

As perception usually precedes production in language acquisition, results from a perception study of the Dutch voicing contrast will first be discussed. Van der Feest (2007) tested young children's recognition of familiar words, using a split-screen preferential looking paradigm. Words contained 'mispronunciations', i.e. either the place or the voice feature of segments in the onset was changed. Results for voice mispronunciations show an asymmetry, as 24-month-olds did not accept voiced mispronunciations of voiceless words (e.g. *[bus] for [pus] *poes* 'cat'), but they did accept voiceless mispronunciations of voiced words (e.g. *[pom] for [bom] *boom* 'tree'). In contrast, 20-month-olds showed no mispronunciation effects, indicating that a voicing contrast had not been acquired. Van der Feest argues that at 24 months, voiceless stops (lacking prevoicing) are underspecified. A study with adults by van Alphen (2004) also showed that voiced segments only prime voiced words, whereas voiceless segments prime both voiced and voiceless words.[4] This was taken as evidence that Dutch speakers can deal with variation in the input, as segments can vary in the amount of prevoicing that is realised. Results by van der Feest thus replicate this effect for 24-month-old Dutch children, which is interpreted according to the FUL model of word recognition (Lahiri & Reetz 2002, Fikkert, Levelt & Zamuner 2005). In this model, lexical representations are assumed to be underspecified for features that display regular variation (e.g. coronal segments such as /t/ are underspecified for place of articulation as they typically assimilate to labial or velar, whereas labial segments /p/ are fully specified).

[4] This is in line with confusion matrices for Dutch (Smits et al. 2003), which predict that voiceless segments are more likely to be confused with voiced segments than vice versa.

Even if perception of the voicing contrast is in place, this does not entail that children can also produce the contrast. Kuijpers (1993ab) studied spontaneous productions of children between 1;5 and 3;9, and argues that the initial voicing contrast is acquired around the age of 3 (although young children often insert a schwa, which places the stop in intervocalic position). In a recent corpus study (Kager et al. to appear), productions from five Dutch children aged 1;1-2;11 (containing ~10,000 utterances) were taken from the CLPF database (Fikkert 1994, Levelt 1994), available through CHILDES (MacWhinney 1991/2000). This database contains naturalistic longitudinal data from 12 children acquiring Dutch, recorded between the ages of 1;0 and 2;11. Kager et al. show that Dutch children typically devoice initial obstruents. For instance, Tom produced [pɛt] for [bɛt] *bed* 'bed' at 1;5, and Robin produced [tir] for [dir] *dier* 'animal' at 1;10. In contrast, additional CHILDES data for German (Kerstin aged 1;3 - 3;4) and English (Seth aged 1;7 - 2;5) showed that these children produced more 'voiced' or plain unaspirated stops for voiceless targets. This confirmed earlier results from acquisition studies of prevoicing languages such as Spanish (Macken & Barton 1980) and Hindi (Davis 1995) and of aspiration languages such as English (Menn 1971, Smith 1973, Snow 1997) and German (Grijzenhout & Joppen-Hellwig 2002). Kager et al. (to appear) argue that error patterns of Dutch children do not necessarily reflect frequency. Thus, even though it is the case that voiceless stops are overall more frequent than voiced stops (Zamuner 2007), children attempt more voiced than voiceless targets. Kager et al. argue instead that error patterns can be accounted for by either articulatory factors (a phonetic account) or featural representations (a phonological account). In this respect, it is relevant to note that laryngeal contrasts seem to be acquired earlier in aspiration languages than in prevoicing languages. Hence, in initial position, the English contrast is acquired by the age of two (Macken & Barton 1978; 1979, Snow 1997), whereas the Dutch contrast is acquired around the age of three (e.g. Kuijpers 1993ab, Beers 1995, van der Feest 2007). For instance, Beers shows that children reached an accuracy level of 75% correct around age 2;11. Previous research (e.g. Davis 1995) has indicated a role of acoustic salience to account for this difference; prevoicing (voicing lead) is argued to be less salient than aspiration (long lag VOT). Also, both prevoicing and aspiration are articulatorily difficult, which may lead the Dutch child to 'devoice' and the German or English child to 'deaspirate'. This suggests that differences in acquisition are attributable to ease of perception and production. However, Kager et al. (to appear) point out that different feature representations across languages may also account for differences in acquisition. As noted before, cross-linguistic markedness is often reflected in language acquisition, as children tend to produce the least marked properties before more marked ones (Zamuner, Gerken & Hammond 2005a). If a single feature (e.g. [±voice] or privative [voice]) is active in all languages, one would expect children's errors to reflect the unmarked value (i.e. [-voice]). In contrast, a 'multiple feature hypothesis', according to which there are multiple language-dependent features (e.g. Iverson & Salmons 1995), correctly predicts that children's errors are different depending on the laryngeal features of the language they are exposed to. As has been argued by others (e.g. Lasky, Syrdal-Lasky & Klein 1975), acquisition rates seem to depend on the nature of the voicing contrast in the target language. Children acquiring languages with prevoicing such as Dutch display devoicing errors involving loss of [voice], whereas children acquiring aspiration languages such as German should produce 'voicing' or 'de-aspiration' errors, involving a loss of the marked feature [spread glottis]. Kager et al. (to appear) favour this explanation over an articulatory account, on the basis of English child data taken from CHILDES (Wilson & Peters 1988).[5] In sum, Kager et al. argue for a multiple feature account, under which a privative feature [voice] is active in Dutch (in all positions), in line with the theoretical account of Zonneveld (to appear).

In a larger study based on the same (CLPF) corpus containing data of 12 Dutch children (1;0 - 2;11) and one additional child Nora (1;3 - 3;0), Van der Feest (2004, 2007) also found that children initially produce voiceless stops, leading her to claim that children's early lexical representations are underspecified for the feature [voice]. Van der Feest shows that the coronal contrast (between /t/ and /d/) is acquired later than the labial contrast (between /p/ and /b/) in initial position, argued to be due to the input variation that children are exposed to in /d/-initial function words. Function words show variation in utterance-medial position, e.g. *wat is dat* 'what is that' is often pronounced as *wat is tat* (van Haeringen 1955, Zonneveld 1978, Ernestus 2000). Van der Feest argues that a carry-over

---

5 Devoicing errors in initial stops for Seth (1;7 - 4;1) were found to be triggered by a following voiceless obstruent (usually in the coda position of the word). However, this 'harmonic' effect was not attested for errors involving voicing of initial stops. This asymmetry was taken to reflect the laryngeal specification of voiceless segments, i.e. activity of [spread glottis].

effect leads Dutch children to devoice initial /d/. Alternatively, an articulatory account would predict that voicing is easier to produce in /b/ than in /d/, as closure of the mouth is complete (e.g. Van Alphen 2004). Van der Feest argues against this approach, given that /d/-initial content words are produced more accurately than /d/-initial function words (although the youngest children showed a difference between /d/ and /b/ even in content words). Note however that even an articulatory account does not necessarily exclude word frequency effects (e.g. Ferguson & Farwell 1975).

In sum, Dutch children have perceptive knowledge of the initial voicing contrast around the age of 2, whereas productive knowledge is in place around the age of 3.

The perception of the *medial* voicing contrast has only been investigated with older children. For instance, Kuijpers (1993a) has conducted an identification experiment with children aged 4;5, 6;4 and 12;2, using non-word stimuli (e.g. *patto* [pɑto] vs. *paddo* [pɑdo]). Results indicate that judgments of 4 and 6 yr. olds were very similar to those of 12 yr. olds or adults, although the slope of identification changes with age. The largest difference was found between the young children (4 and 6 yr. olds) and the older group (12 yr. olds and adults). Young children apparently have a larger area of ambiguity in the categorisation of voiced and voiceless stops, and need a longer silent interval to switch from perceived /d/ to /t/.

Turning to production, previous studies have indicated that the medial voicing contrast is also produced around the age of 3 (e.g. Kuijpers 1993ab, van der Feest 2007). The acoustic realisation of the medial voicing contrast is quite complex; Slis (1985) and Slis & Cohen (1969) found that as many as eight different acoustic parameters contribute to the intervocalic voicing distinction. Two prominent acoustic features are closure duration and burst duration, which are both relatively short for voiced stops. Clement (1991) found that Dutch children at a mean age of 2;3, 3;9 and 5;3 produce adult-like durational and spectral differences. Kuijpers replicated these results for children aged 4;5, 6;4 and 12;2, eliciting both bi-morphemic and mono-morphemic words with intervocalic /b/ (e.g. *krabben* 'crabs', *kabel* 'cable'), and /d/ (e.g. *bedden* 'beds', *vader* 'father'). Kuijpers found that the relative contrast between voiced and voiceless stops was stable across ages, i.e. the closure duration of a voiced stop was generally about half as long as that of a voiceless stop. However, 4;5-year-olds displayed more variability than older children, and mean values of closure durations for voiceless stops decrease with age. We will return to these results in Chapter 5, when acoustic measurements are discussed.

Once the voicing contrast is in place, the Dutch child also needs to acquire knowledge of final devoicing. Let us therefore turn to final devoicing and the variation that the child could be expected to encounter.

## 3.3 Final devoicing and the nature of variation

### 3.3.1 Voicing assimilation

Traditionally, it has been claimed that Dutch word-final obstruents are always voiceless, unless they occur before voiced stops (e.g. Trommelen & Zonneveld 1979, Booij 1995, Berendsen 1983, Zonneveld 1983).[6] Rule-based analyses have proposed that the rule of final devoicing applies after suffixation of vowel-initial suffixes, compounding, and cliticisation (see also Chapter 2). Examples from Chapter 1 are repeated in (23) and (24) below.

| | | | | |
|---|---|---|---|---|
| (23) | a. | [hɑnt] 'hand' | /hɑnd/ + -ən | [hɑn.dən] 'hands' |
| | b. | [dif] 'thief' | /div/ + -ən | [di.vən] 'thieves' |
| | c. | [blut] 'blood' | /blud/ + -ən | [blu.dən] 'bleed-pl-inf' |
| | c. | [hɛp] 'have-1st sg' | /hɛb/ + -ən | [hɛ.bən] 'have-pl-inf' |
| | d. | [xɔut] 'gold' | /xɔud/ + -ən | [xɔu.dən] 'golden' |
| (24) | a. | [dif] 'thief' | /div/ + -achtig | [dif.ɑx.təx] 'thievish' |
| | b. | [xɔut] 'gold' | /xɔud/ + ader | [xɔut.a.dər] 'gold vein' |
| | c. | [hɑnt] 'hand' | /hɑnd/ + rem | [hɑnt.rɛm] 'hand brake' |

Final devoicing thus neutralises the voicing distinction in singular nouns, e.g. /xrat/ 'fish bone' and /xrad/ 'degree' both surface as [xrɑt]. Languages with final devoicing often show a process of regressive voicing assimilation (RVA) in obstruent clusters, attested for Dutch, Polish, Russian, Catalan, and Sanskrit (e.g. Lombardi 1999). In Dutch, assimilation neutralises the distinction between words like *verwijdbaar* /vərwɛid-bar/ 'widen-able' (from *verwijden* 'widen') en *verwijtbaar* /vərwɛit-bar/ 'reproach-able' (from *verwijten* 'reproach'), see Ernestus (2000). Examples are provided in (25),

---

6 As mentioned in Chapter 1, final devoicing applies to loanwords (e.g. *club* [p]), although there are exceptions (e.g. *fez* [z] 'fez' and *gig* [g] 'gig', see Zonneveld to appear).

including progressive voicing assimilation (PVA) for fricatives (23d-e) and the Dutch past tense suffix (23f-g).[7]

| | | | | |
|---|---|---|---|---|
| (25) | a. | [hɑnt] 'hand' | /hɑnd/ + /pɑlm/ | [hɑntpɑlm] 'palm of the hand' |
| | b. | [blut] 'blood' | /blud/ + /bɑŋk/ | [bludbɑŋk] 'blood bank' |
| | c. | [klɑp] 'clap' | /klɑp/ + /dør/ | [klɑbdør] 'swing door' |
| | d. | [hɑnt] 'hand' | /hɑnd/ + /vɑt/ | [hɑntfɑt] 'hand grip' |
| | e. | [rat] 'advise' | /rad/ + /zam/ | [ratsam] 'advisable' |
| | f. | [krɑp] 'scratch' | /krɑb/+/də/ | [krɑbdə] 'scratched' |
| | g. | [klɑp] 'clap' | /klɑp/ + /də/ | [klɑptə] 'clapped', 'burst' |

Due to regressive voicing assimilation, obstruents in coda position may surface as voiced (as in [bludbɑŋk] and [klɑbdør]), which in effect obscures the effect of final devoicing. A specific ordering of rules has been proposed in the literature, in which final devoicing is followed by rules of voicing assimilation (e.g. Zonneveld 1983). Example (24a) shows that progressive assimilation needs to be ordered before regressive assimilation in forms like *handvat* 'handle, grip'. Forms like *handdoek* 'towel' (lit. 'hand-cloth') have been analysed as in (24b), with *degemination* applying after RVA.[8]

(26)

| | | | | |
|---|---|---|---|---|
| a. | UR | /hɑnd/+/vɑt/ | b. | /hɑnd/+/duk/ |
| | FD | /hɑntvɑt/ | | /hɑntduk/ |
| | PVA | /hɑntfɑt/ | | n.a. |
| | RVA | n.a. | | /hɑndduk/ |
| | SR | [hɑntfɑt] | | /hɑnduk/ |

Leaving aside the various phonological analyses of voicing assimilation, it is important to note that there seems to be considerable variation in the phonetic realisation of assimilation (e.g. Slis 1985; 1986). Not only do many authors report variability and uncertainty regarding observations, Ernestus

---

[7] Progressive voicing assimilation in case of the past tense suffix and clitics (/vɔnd di/ [vɔnti] 'found he') is unexpected under rule- or constraint-based analyses, see Zonneveld (1983), Berendsen (1983; 1986), Lombardi (1995; 1997), Grijzenhout & Krämer (2000) and Zonneveld (to appear) for analyses.
[8] Note that such examples of degemination in Dutch (e.g. hɑn(d)duk vs. *hɑn(t)tuk) should lead children to discover the direction of assimilation (Zonneveld, p.c.).

(2000:73-75) shows that speakers have different intuitions about the realisation of word-final stops before obstruent-initial function words. For instance, realisation of /d/-initial function words and clitics can vary when they attach to obstruent-final words, e.g. *loop dan* 'walk then' as [lobdɑn] or [loptɑn]. The clitic 'der' (a weak form of *haar* 'her') also behaves in this way, as in *ik krab haar niet* 'I do not scratch her' [krɑptər] / [krɑbdər] (see also Lahiri, Jongman & Sereno 1990, Ernestus & Baayen 2001). Zonneveld (1983:306) notes that this type of assimilation does not occur for /b/, indicating its relative strength in Dutch phonology. Hence, *ik ben* 'I am' is realised as [igbɛn], but never as *[ikpɛn]. In an analysis of dialectal variation, De Schutter & Taeldeman (1986) also note that regressive assimilation is generally stronger before /b/. This difference is important for the hypotheses that were formulated for the Wug-tests described in Chapter 5.

According to generative analyses, variation in regressive voice assimilation can be accounted for by syntactic boundaries or prosodic domains in the sense of Nespor & Vogel (1986), see Zonneveld (1983), Booij (1995) and Ruys & Trommelen (2003). For instance, Gussenhoven (1986) suggests that slow or formal speech is connected to the introduction of new Prosodic Word boundaries. However, experimental findings have not supported these claims, as Dutch voicing assimilation has been found to be partial, gradient and optional, displaying differences in both the realisation and the direction of assimilation (e.g. Slis 1986, Menert 1994, Ernestus 2000; 2003, Jansen 2004). These studies have typically found that the main factor influencing the realisation of segments as voiced or voiceless is speech rate. Ernestus (2000; 2003) has investigated a corpus of spoken Dutch (containing 'casual speech'), and argues that final devoicing and assimilation are phonetic in nature, following a neutralisation process which affects 'neutral' obstruents. This view correctly predicts that there is no influence of underlying voice specifications on the acoustic properties of neutral obstruents (supported by results of Jongman et al. 1992 and Bauman 1995). Thus, Ernestus (2000:276) argues that "there is no strong relation between classification of word-final stops as voiced or voiceless and the underlying [voice] specification of the stops". Ernestus, Lahey, Verhees & Baayen (2004) report frequency effects for Dutch regressive voicing assimilation, showing that there is less voicing in high frequency words. Hence, there was little assimilation in low frequency words, variable assimilation in mid frequency words, and mainly progressive assimilation in the highest frequency words. This tends to make these words more like mono-morphemic words, which typically contain voiceless clusters (Zonneveld 1983). Ernestus further reports extensive variation in the realisation of word-final obstruents before vowel-initial function words, such as *heb ik* /hɛb ɪk/ 'have I', realised as [hɛbɪk] or [hɛpɪk] and *heb daar* /hɛb dar/ 'have there' realised as [hɛptaːr] or [hɛbdaːr] (Ernestus 2000:57-58). When followed by a vowel, word-final obstruents may be realised as voiced in fast speech and as voiceless in slow speech. Apart from speech rate, Ernestus argues that the type of preceding segment, presence of stress, and sex and mood of the speaker play a role. Word-combinations involving clitics behave differently from other word-combinations, since they may be retrieved as single units from the lexicon. For instance, Ernestus found that the combination *heb ik* /hɛb ɪk/ 'have I' is realised as [hɛbɪk] in 97% of cases, whereas *dat ik* /dɑt ɪk/ 'that I' was realised as [dɑdɪk] in only 43% of cases (Ernestus 2000:236).[9] Following Bybee (1995; 2001), Ernestus points out that these word-combinations represent the acoustic form with the highest frequency of occurrence. This ties in well with the observation that highly frequent words tend to have more reduced realisations (Hay 2000, Bybee 2002b, Beckman & Pierrehumbert 2003). For instance, realisation of word-final obstruents as voiced or voiceless seems to depend on the frequency of a combination, as in /hœyt/ + /ɑrts/ → [hœytɑrts] 'dermatologist' vs. /tɑnd/ + /ɑrts/ → [tɑndɑrts] / [tɑndɑrts] 'dentist'.[10]

Kuijpers (1993ab) has investigated children's realisation of voicing assimilation and found differences in the direction of assimilation between children (6 and 12 yr olds) and adults, even though children showed the same rate of assimilation. All age groups showed progressive assimilation in case of fricatives (e.g. *broekzak* [ks] 'trouser pocket'), which may reflect the general weakness of the fricative contrast ([f] ~ [v] and [s] ~ [z]). For stops (e.g. *stropdas* [bd] 'tie' and *knap dier* [bd] 'clever animal'), assimilation could be either absent, progressive or regressive. Adults showed almost equal distribution of the three options, although regressive assimilation was predominant (43%).Both 6 and 12 yr old children showed more progressive

---

[9] Frequent combinations of this type are often contracted, e.g. as [hɛk] and [dɑk] respectively, which suggests they are stored in the lexicon (as noted by Booij 1985). The form *dat ik* 'that I' had an equal number of realisations as [datɪk], [dɑdɪk] or [dɑk], suggesting that the stored forms [dɑtɪk] and [dɑk] influence the general tendency to pronounce a voiced stop in this context.

[10] Note that some dialects maintain the underlying voicing distinction between words before clitics (e.g. West-Flemish *poot is* /pot ɪs/ [t] 'leg is' vs. *brood is* /brod ɪs/ [d] 'bread is').

assimilation (55%) than adults (31%). Kuijpers found no effect of linguistic context for children, whereas adults show more RVA in case of compounds (54%) than across word boundaries (33%). Taken together, these results indicate that assimilation is likely to be influenced by factors of phonetic implementation.

To conclude, it is unclear how a child arrives at the relevant rules, constraints, or generalisations for final devoicing and voicing assimilation, given the wide range of variation that is observed for Dutch. Word-final obstruents have been found to be variably realised as voiced or voiceless in a speaker- and context-sensitive manner, influenced by speech rate. This is problematic if the child needs to pick up on regularities from the input (e.g. "word-final obstruents are voiceless"), to be used for the subsequent acquisition of voicing alternations. In fact, the generalisation that final obstruents are voiceless only holds for words in isolation and utterance-final position (which is arguably the largest domain). However, as will be described in the next section, even in these positions the phonetic realisation of final devoicing has been claimed to be variable.

### 3.3.2 Incomplete neutralisation

Final devoicing has figured prominently in phonological theory, but as Plank (2000:175) states "no conceivable aspect of its phonology and phonetics is uncontroversial". Recently, even the assumption that final devoicing results in phonetically complete neutralisation has come under attack. Phonological processes are traditionally assumed to be categorical, i.e. their output should not be gradient. This thus touches on the distinction between phonology and phonetics, and neutralisation as a theoretical construct (see Hayes 1995).[11] As stated before, final devoicing is the prime example used to argue for abstract underlying representations, which cannot be deduced from isolated surface forms but can only be learned on the basis of alternating forms. Thus, Dutch minimal pairs such as the singulars *graad* /xrad/ 'degree' and *graat* /xrat/ 'fish bone' are considered phonetically identical (i.e. [xrat]) in phonological descriptions. Even though a (trained) listener usually cannot detect differences between neutralised and non-neutralised segments,

[11] Other cases of neutralisation have been disputed as well, e.g. intervocalic *flapping* involves vowel length differences (Dinnsen 1985) and is variably realised (e.g. Rimac & Smith 1984, de Jong 1998).

phonetic studies have shown that neutralisation of the voicing contrast may be incomplete.[12] For instance, Port & O'Dell (1985) had German speaker read a list of German words in isolation (*Rat* 'advice' vs. *Rad* 'wheel') and found that neutralised plosives show longer vowel durations and voicing into closure, but shorter releases than the non-neutralised segments. However, Fourakis & Iverson (1984) have pointed out that these results may be due to reading effects. Languages in which the underlying voicing distinction is not reflected by spelling have shown mixed results; incomplete neutralisation was found for Catalan (Dinnsen & Charles-Luce 1984), but not for Turkish (Kopkalli 1993). Charles-Luce (1985) has investigated the effect of word- vs. phrase-final position and found that neutralisation may be incomplete in some sentential contexts. Similarly, Charles-Luce (1993) found complete neutralisation for Catalan in utterance-final but not utterance-medial position, as well as an effect of semantically biasing (i.e. disambiguating) information in the sentence. Port & Crawford (1989) also found that incomplete neutralisation can be enhanced by the intent of the speaker, i.e. the effect is greater in a communicative context (when subjects were asked to read the sentence *Ich habe Rat gesagt; nicht Rad* 'I said *Rat* ('advice') not *Rad* ('wheel')), than when words were read in isolation. In sum, these studies have shown speaker- and context-dependent gradual incompleteness of final devoicing across languages.

Incomplete neutralisation has also been reported for Dutch. Warner, Jongman, Sereno & Kemps (2004) used pairs differing only in underlying /t/ and /d/, containing either short vowels (*rat* 'rat' vs. *rad* 'wheel') or long vowels (*noot* 'nut' vs. *nood* 'necessity'). Note however that for the majority of word pairs used, members belong to a different word class (e.g. *wet* 'law' vs. *wed* 'bet 1st sg') or different plural types ([rɑt] ~ [rɑtən] 'rats' vs. [rɑt] ~ [rɑdərən] 'wheels'), and some words do not have a corresponding plural or inflected form (e.g. *wat* 'what' or *Ad* 'proper name'). This lack of minimal pairs occurs for other languages as well, and illustrates the difficulty of

[12] Incomplete neutralisation has been reported for German (e.g. Dinnsen & Garcia-Zamora 1971, Charles-Luce 1985, Port, Mitleb & O'Dell 1981, Port & O'Dell 1985, Port & Crawford 1989, Piroth & Janker 2004), Russian (Chen 1970, Pye 1986), Polish (Slowiaczek & Dinnsen 1985, Slowiaczek & Szymanska 1989), Catalan (Dinnsen & Charles-Luce 1984, Charles-Luce & Dinnsen 1987, Charles-Luce 1993) and Dutch (Warner et al. 2004, Ernestus & Baayen in press). Complete neutralisation is reported for German (Fourakis & Iverson 1984), Catalan (Mascaro 1995), Polish (Jassem & Richter 1989) and Turkish (Kopkalli 1993), see Manaster Ramer (1996) and Port (1996) for discussion.

investigating effects of neutralisation (Port 1996). In a production task with 15 speakers, underlying voicing was shown to have an effect on vowel duration (3.5 ms), which is less than the difference found for German and Polish (10-15 ms). An effect of burst duration (9 ms) was observed for words with long vowels only. Baumann (1995) observed a difference in closure duration (6 ms) for one syntactic condition only, although her study was not specifically designed to test incomplete neutralisation effects. Finally, Ernestus & Baayen report an effect on burst duration (Ernestus & Baayen 2006), and an effect of closure duration (Ernestus & Baayen to appear). To test for effects of spelling, Warner et al. also tested homophones that differ in orthography rather than underlying voicing (e.g. *kleden* 'to dress-pl/inf' vs. *kleedden* 'dress-pl-past'). Similar durational differences were observed, although an effect of morpho-phonological structure cannot be ruled out here. Further support for the role of orthography comes from a study by Ernestus & Baayen (2006), who obtained sub-phonemic durational differences for non-words spelled with 'p, t' vs. 'b, d', presented as pseudo-verbs. However, it is possible that orthography leads speakers to construct underlying representations for non-words. Warner et al. suggest that sub-phonemic differences could be caused by a number of factors besides underlying voicing differences, including orthography, morphology, word frequency (citing Whalen 1991), or neighbourhood density (citing Wright 2002). In sum, there is some evidence for incomplete neutralisation in Dutch, although results are inconclusive and hard to interpret, partly due to the interfering effects of orthography and the low number of minimal pairs available (making it near impossible to control for factors such as word frequency). Finally, incomplete neutralisation may not be problematic if an 'archiphoneme' in the sense of Trubetzkoy (1939:71) is posited, since the realisation of /T/ could result in either one of the members ([t] or [d]) or an intermediate category (see Brockhaus 1995:238). In fact, many have argued that a 'feature changing' rule of final devoicing (that changes [+voice] into [-voice]), is incompatible with the attested effects of phonetic implementation (see Chapter 2). For instance, Ernestus (2000, 2003) claims that neutralisation affects both [+voice] and [-voice] obstruents: obstruents in coda position are considered *neutral*, which means that they are not specified as [-voice] or [+voice]. Instead, they are *unspecified* for [voice], which is why the final obstruent of /bɛT/ 'bed' and /pɛT/ 'cap' can be realised as either [t] or [d].[13]

Given that neutralisation may be incomplete for some speakers or contexts, the question naturally arises whether listeners could be sensitive to these small durational differences. Dinnsen (1985) argues that a difference in production which is not accompanied by a difference in perception could be due to a sound change in progress. A contrast is only truly non-neutralising if both production and perception are affected. Crucially, Port & Crawford (1989) and Port & O'Dell (1985) show that German listeners were able to discriminate pairs (e.g. *Rat* 'advice' vs. *Rad* 'wheel') in a forced choice task (60 - 80% of cases were correctly identified). Similarly, Warner et al. report that (some) Dutch listeners were able to distinguish minimal pairs differing in underlying voicing (e.g. *rat* 'rat' vs. *rad* 'wheel') with higher than chance accuracy, using either vowel duration or closure duration as a cue. Hence, listeners associate longer vowel duration or shorter closure durations with final /d/. However, as closure duration was not found to vary systematically with underlying voicing in speakers' productions, Warner et al. conclude that listeners' ability to use cues must be enhanced in a forced choice experiment. Alternatively, one would have to conclude that listeners use differences that are not reliably produced. Ernestus & Baayen (2006) show that listeners interpreted slightly voiced final stops (i.e. from pseudo-verbs spelled with final 'd, b' that had been read aloud) as underlyingly voiced more often than the completely voiceless stops (i.e. from pseudo-words spelled with 't, p'), which led them to argue that listeners use incomplete neutralisation. Ernestus & Baayen (in press) also asked subjects to inflect the pseudo-verbs, and argue that voicing was used as a subphonemic cue to past-tense formation. Thus, participants were more likely to choose the past tense realisation *-de* when final obstruents were perceived as weakly voiced. Similarly, Kemps, Ernestus, Schreuder & Baayen (2005) show that the acoustic realisation of a stem provides crucial prosodic cues to the listener, that can be used in morphological processing. On the other hand, no effects of incomplete neutralisation were found by Baumann (1995), who placed

[13] This is different from a 'delinking' account in which the underlyingly voiced obstruent loses its laryngeal node and emerges as a voiceless obstruent, as proposed by Lombardi (1991, 1995). It is also different from the underspecification account of the Turkish three-way contrast, in which only alternating final obstruents are underspecified (Inkelas 1994).

items (e.g. *raat* 'honeycomb' or *raad* 'council') in different sentential positions.

In a recent study, Ernestus & Baayen (to appear) asked subjects to rate the voicing of final obstruents on a five-point scale, presenting stimuli either as full words or as final rhymes of these words. Listeners rated unvoiced alternating obstruents as more voiced than the unvoiced non-alternating stimuli, and this effect was even stronger when listeners heard the full word. Thus, an underlyingly voiced obstruent (e.g. from *mand* 'basket') was rated as more voiced than when only the rhyme was presented (e.g. *and*). The authors conclude that the perception and interpretation of these obstruents was influenced by intraparadigmatic analogy, arising from stored inflectional exemplars. Importantly, these intraparadigmatic effects are much larger than incomplete neutralisation effects; they 'magnify the perceived differences between alternating and non-alternating obstruents, encouraging speakers to maintain these differences in their speech' (perhaps leading to paradigmatic levelling). Ernestus & Baayen (to appear) argue against the use of a rule that devoices an abstract underlying voiced obstruent, after which phonetic implementation rules partly voice it again. Instead, following Bybee (2001), they propose that both [bɛdən] and [bɛt] are stored in the lexicon. Due to lexical analogy with paradigm members, the production and perception of [bɛt] is directly influenced by the full form [bɛdən], which is activated as well. Further evidence for paradigmatic effects in word recognition was found in a lexical decision experiment (Ernestus & Baayen 2007). In this study, words were realised with incorrectly voiced final obstruents and played to listeners, who performed better on alternating words (e.g. /hɑnd/ as *[hɑnd] 'hand') than non-alternating words (e.g. /krɑnt/ as *[krɑnd] 'paper'), presumably because the former is supported by other forms in the morphological paradigm. Moreover, response latencies were influenced by the exact probability of paradigmatic voicing, based on the frequencies of different word forms in the paradigm. Finally, listeners performed better on [hɑnt] than on *[hɑnd] even though the latter matches the presumed underlying representation, leading Ernestus & Baayen (2007) to argue against an abstract lexical representation such as /hɑnd/.

Lahiri, Jongman & Sereno (1990) have investigated whether word recognition was influenced by surface phonetic or underlying phonological representations. Here, listeners were presented with a verbal stem followed by the clitic *der,* e.g. *kies der* /kiz dər/ 'choose her' was realised as [kistər] or [kizdər], which are both attested in Dutch (see also Zonneveld 1978). Non-alternating stems were also included, e.g. *kus der* /kœs dər/ 'kiss her', realised as [kœstər] or [kœzdər]. Subjects were then asked to perform auditory lexical decision to a target, which was the same verb stem in isolation (i.e. realised as [kis] 'choose' and [kœs] 'kiss'). Results indicate that subjects were faster when an alternating stem was preceded by a voiced prime (i.e. [kizdər]), or when a non-alternating stem was preceded by a voiceless prime (i.e. [kœstər]). Lahiri et al. take this evidence to argue against the idea that both variants of a form (e.g. [kis] and [kiz] or [bɛd] and [bɛt]) are stored in the lexicon. Instead, they claim that recognition is influenced by the abstract underlying phonological representation of the stem. However, as noted by Ernestus & Baayen (to appear), it is not clear whether effects are (partly) due to incomplete neutralisation. Moreover, intraparadigmatic effects (e.g. influence from the stored representation [kizən]) can also account for the results. Results from a study by Jongman, Sereno, Raaijmakers & Lahiri (1992) could be interpreted in this way as well. Here, the interaction between voicing and vowel length in Dutch was investigated, and vowel category boundaries were found to differ in pairs like /zɑt/ 'sat' - /zad/ 'seed' as compared to /stɑd/ 'city' - /stat/ 'state'. According to the authors, this indicates that the underlying voicing of the final obstruent affected listeners' perception of the vowels (although the differences were not significant).

As noted by Warner et al, it is not clear whether listeners actually make use of small sub-phonemic cues in perceiving natural speech. However, as Port & Crawford (1989) point out, a large number of small acoustic differences rather than a few large ones may be involved in making the distinction, as phonological feature involved is 'in no one-to-one correlation to any single phonetic feature or to any set of these' (Piroth & Janker 2004:103). More generally, Port & Leary (2005) and others have argued that the notion of a formal grammar with discrete, categorical and stable symbolic representations does not capture the gradient and probabilistic effect of rules, which seem to be phonetic in nature (see also Pierrehumbert 2001 and others). For instance, the English voicing contrast (e.g. between *bet* and *bed*) seems to be best characterised as a change in the vowel / consonant duration ratio, and such a relationship between voicing and vowel length has been attested for other Germanic languages as well (Port 1981). As noted in the previous chapter, stochastic OT (e.g. Boersma &

Hayes 2001) can deal with variation among discrete alternatives, although it does not take context or communicative intent into consideration.[14]

For our purposes, it is important to note that results on incomplete neutralisation suggest that surface forms (e.g. *graad* [xrat] 'degree' vs. *graat* [xrat] 'fish bone') may be discriminated on the basis of subtle acoustic cues. If the forms are actually phonetically different, it would not be necessary for the child to deduce abstract underlying forms (although results for Dutch are inconclusive). As noted by Hayes (1995), there is evidence that children (before the age of one year) may even be more sensitive to phonetic differences than adults (e.g. Werker & Tees 1984), which supports the possibility that children may be able to identify underlying forms directly from the surface form. Recall that underlying forms are not posited at all in analogical models of language, which replace deterministic rules or constraints by probabilistic generalisations based on related surface forms in the lexicon.

In the next section, experimental findings on the representation of Dutch plurals and past tense forms with voicing alternations in the adult mental lexicon will be presented, before we return to the question of how voicing alternations are acquired by children.

## 3.4 Psycholinguistic studies

In recent years there have been a number of experimental studies into the processing of morphologically complex words. Generally, there is evidence that factors such as productivity, frequency and (semantic and phonological) transparency play a role in the processing and representation of morphological information (e.g. McQueen & Cutler 1998). Taft (1979) has shown that there are two independent effects of frequency. First, a 'Base Frequency' effect reflects the frequency of the *lemma* or *lexeme* (i.e. the summed token frequency of *walk* and all its inflectional variants). The occurrence of such an effect means that the base form has been accessed during lexical access. According to *Dual Mechanism* models (e.g. Pinker 1999), this is the only effect of frequency that is predicted to occur for regularly inflected forms. However, effects of 'Surface Frequency' have also

[14] Gafos (2006) proposes a formal grammar based on nonlinear dynamics to capture the gradient and context-dependent aspects of final devoicing.

been observed for regular forms, reflecting the frequency of occurrence of the *wordform* or complex form itself (i.e. *walked*). This is generally interpreted as evidence that the inflected form is stored as a whole (e.g. Sereno & Jongman 1997, Alegre & Gordon 1999). As noted in Chapter 2, Baayen, Dijkstra & Schreuder (1997) propose a *parallel dual route* model, in which storage and computation operate in parallel. Thus, some words are accessed as wholes, whereas others are accessed via their morphemic constituents, mainly depending on their token frequency. Baayen, Dijkstra & Schreuder (1997) have argued that models of morphological processing should take language-specific properties such as distribution of affixes and word formation patterns in account, to determine the balance of storage and computation for any given language.

Baayen et al. (1997) conducted a visual lexical decision experiment on regular Dutch *-en* plurals, and found that response times for plural nouns depend on their surface frequency. This was taken as evidence for a whole-word route for nouns, which are claimed to have full form access representations. Evidence for *parsing* (i.e. decomposition) was only found for some singular-dominant plural nouns with low frequencies (e.g. *stoeten* 'processions'), 14% of which showed evidence for a parsing route. Note that alternating words were also included, which mostly had high stem frequencies (e.g. *eend* 'duck'). However, Baayen et al. (1997) do not report on these forms separately. Baayen et al. observe that the Dutch *-en* suffix is predominantly used as a verbal ending, creating a 'subcategorisation' conflict when this suffix is attached to nouns. They argue that many noun plurals are stored, to 'avoid the time-costly resolution of this subcategorisation conflict'. In contrast, the parsing route was found to be more efficient for verbal plurals. Furthermore, Baayen, Schreuder, de Jong & Krott (2002) conclude that the possible default status of an affix (*-en* or *-s*) as suggested by the Dual Mechanism model is irrelevant, i.e. it does not imply absence of storage in the mental lexicon (see also Baayen et al. 2003, McQueen, Dijkstra & Schreuder 2003, Hay & Baayen 2005).

In an experiment investigating regular past tense formation in Dutch, Ernestus & Baayen (2001, 2004) presented listeners with existing regular verbs. Subjects had to choose a past tense allomorph, i.e. *-te* or *-de*, which is traditionally assumed to depend on the underlying voice specification of the final segment (e.g. Trommelen & Zonneveld 1979:119, Booij 1995:61). Under this view, *-te* is added when the stem ends in /t/ or /p/, while *-de* is used when the stem ends in /d/ or /b/, and no analogical effects are expected to occur. However, Ernestus & Baayen found that regular past tense formation was influenced by phonological similarity structure in the lexicon. For instance, 11% of verbs ending in a bilabial stop are underlyingly voiced (e.g. *dubde* 'wavered'), while 44% of verbs ending in an alveolar stop take *-de* (e.g. *laadde* 'loaded'). Results showed that subjects created more unexpected past tense forms, produced more inconsistency errors, and responded more slowly when verbs were presented that had stronger analogical support for the unexpected form. Thus, subjects attached the 'wrong' suffix (e.g. **dubte* for *dubde* 'wavered'), when most words in the analogical gang of phonologically similar words favoured that suffix, especially in case of low frequency past tense forms. This resulted in many errors for /b/ words, which were inflected with *-te* in 36% of cases, while words with /d/ were inflected with *-te* in only 5% of cases. However, errors in which the suffix *-de* was attached to verbs ending in /p/ or /t/ were even less frequent, occurring in around 2% of cases. Ernestus & Baayen (2004) further show that these results are not due to subjects' unfamiliarity with the underlying voicing specification of the stem. Thus, when subjects heard the inflected verb plural (e.g. *wij krabben* 'we scratch') over headphones and were asked to push a button indicating the past tense suffix *-te* or *-de*, the same effects of analogy were observed, as well as an effect of type of obstruent and frequency of the past tense form. Also, subjects responded faster when a word required *-te*, indicating that 'alternating' verbs were more difficult. On the basis of these results, Ernestus & Baayen (2004) argue that systematic analogy plays a role in word formation, even for regular past tense formation for which a simple rule is available. Such a match between speaker's intuitions and patterns in the lexicon have been shown by others as well, see Eddington (1996, 2004), Albright (2002), Albright & Hayes (2003), Pierrehumbert (2006) and Krott, Baayen & Schreuder (2001) for the choice of Dutch linking elements.

Ernestus & Baayen (2003) have also investigated the question whether voiced underlying representations might be inferred for novel words with 'neutralised' segments. Listeners were presented with a pseudo-verb like *ik* ('I') *fat* [fɑt] or *ik* ('I') *dent* [dɛnt], and asked to write down a past tense form, for which either the suffix *-te* or *-de* could be used. Note that this experiment is very similar to the 'Wug-test' that will be described in Chapter 5, except that pseudo-verbs are used instead of pseudo-nouns, and it involves written rather than spoken responses. Ernestus & Baayen (2003) hypothesised that the choice between 'underlying' /t/ or /d/ could be based on several strategies. For instance, the choice could be random, reflect Paradigm Uniformity, or reflect the relative phonological strength of the phonemes (e.g. stops are more likely to be voiced than fricatives, and /b/ is less marked than /d/, both universally and for Dutch). Finally, the choice could reflect analogy with phonologically or phonetically similar words in the lexicon. Surprisingly, this leads to opposite expectations. On the basis of words in the CELEX lexical database (Baayen et al. 1995), fricatives are expected to be more often voiced than stops, and /d/ is expected to be more common than /b/. An analysis of the corpus data with analogical models like TiMBL (Daelemans et al. 2004) and CART (Breiman et al. 1984) indicated that the probability that a neutralised obstruent is underlyingly voiced depends on the type of final obstruent and the types of the preceding segments (to be discussed further in the next chapter). Results indicated that for 24% of the pseudo-words, subjects created a past tense form with -*de*, showing that they were not only guided by Paradigm Uniformity based on the final [t] in the signal. Responses were not random either, as probabilities were generally higher or lower than chance, and they did not reflect a markedness hierarchy. Instead, the probability that an obstruent was interpreted as underlyingly voiced was again predicted by similarity structure in the lexicon, as experimental results matched the lexical data set. Also, there was a general bias for voiceless segments, because non-alternating forms are overall more frequent than alternating forms. The authors show that the data could be modelled by various models, including stochastic OT and analogical modelling, with the latter providing the most economical explanation. Interestingly, token frequencies were not found to be relevant for the prediction of the underlying voice specification of final segments. These results will be compared with the current study in Chapter 5, in which Wug-test results involving the nominal plural with both children and adults will be described.

## 3.5 Summary and Discussion

Dutch children appear to have mastered the voicing contrast around the age of 3. During acquisition, initial obstruents are frequently devoiced (producing [tir] for [dir] 'animal'), and the contrast is acquired relatively late (e.g. in contrast to German children). Apart from possible methodological issues, this difference could be due to a number of factors. For instance, /t/ is overall more frequent in Dutch, and prevoicing is difficult from an articulatory point of view. Kager et al. (to appear) argue that Dutch children need to discover that [voice] rather than [sg] is the active feature in their language, with initial errors reflecting the unmarked value. Van der Feest (2007) argues that the feature [voice] is underspecified in early lexical representations, based on data from both production and perception.

Generally, children are exposed to a large amount of variation in Dutch voicing, as it has been shown that the realisation of final devoicing and voicing assimilation (both regressive and progressive) is speaker- and context dependent, reflecting non-grammatical factors like speech rate and frequency. Ernestus (1997, 2000, 2003) argues that the categorical status of voicing rules in general may need to be refined, given the large range of factors that affect the phonetic outcome. Even if final devoicing is consistent in some positions, there is evidence that when adults hear a neutralised segment (e.g. [t] in [bɛt]), there is activation of its alternating plural ([bɛdən]). This could be due to incomplete neutralisation in production (Warner et al. 2004, Ernestus & Baayen 2006), which may itself reflect effects of orthography. Alternatively, it could be due to the abstract underlying form (Lahiri et al. 1990, Jongman et al. 1992), or intraparadigmatic effects (Ernestus & Baayen to appear). This effect also ties in with psycholinguistic findings, as Baayen et al. (1997) show that Dutch nominal *-en* plurals are likely to be stored in the lexicon. Furthermore, Ernestus & Baayen (2004) show that when non-words are inflected, adults are sensitive to regularities in the lexicon and may produce alternating forms (for verbs) accordingly. Before we turn to the Wug-test experiments carried out with children and adults (Chapter 5), new corpus data on the frequency and distribution of voicing alternations and early child productions will be presented in the next chapter.

# Chapter 4 Corpus Data

To acquire knowledge about Dutch voicing alternations, children need to have both productive and perceptive knowledge about the Dutch voicing contrast, in combination with the ability to relate morphological pairs such as singulars and plurals. In this chapter, early acquisition of morpho-phonology will be discussed on the basis of corpus data. After a brief review of the literature on the acquisition of morpho-phonology (4.1), the relationship between the Dutch plural and alternations will be presented in section 4.2). In the next section, evidence for voicing alternations from the CELEX corpus will be presented (4.3). These corpus data will be used to assess the distribution and frequency of the voicing alternation, which is argued to be important for acquisition. Furthermore, the data will lead to testable hypotheses for children's behaviour in a Wug-test (discussed in Chapter 5). In section 4.4, data from child corpora (CHILDES) are presented, which were used to determine to what extent alternating forms occur in child-directed speech. Children's early productions were also studied, to assess whether and when alternations and overgeneralisation errors occur in acquisition. Finally, a summary and discussion of the results is provided in section 4.5, focusing on the relative importance of morphological and phonological factors.

## 4.1 Introduction

Children start to produce their first words by the age of one year, while morphological inflections emerge in the course of acquisition. At the earliest stage, children may produce semantically accurate forms (e.g. English *feet*), before overgeneralisation errors occur (e.g. **foots*), see Brown (1973). This type of productivity has been taken as evidence for children's knowledge of abstract categories, in the form of rules, constraints, analogies, or schemas. Generally, it has been suggested that the relative ease with which children learn morphology or morpho-phonological alternations across languages will vary according to the relative degree of frequency, phonological salience, transparency, and paradigmatic morphological richness (e.g. Slobin 1985, Bittner, Dressler & Kilani-Schoch 2003, Laaha & Gillis 2007, Laaha et al. 2007).[1] For instance, Derwing & Baker (1980) claim that frequency predicts the order of acquisition of the different English plural allomorphs. It has also been proposed that the easiest alternations to learn should be those that conform to with 'natural' phonological processes or 'phonotactic' predispositions (Stampe 1973; MacWhinney 1978). For instance, this applies to the voicing assimilation associated with the English plural suffix (*cats* vs. *dogs*). Also, there is evidence from corpus data that young children (before age 3;7) have knowledge of vowel alternations in European Portuguese, which are conditioned by the phonotactic pattern of vowel reduction (Fikkert & Freitas 2006). As was argued in Chapter 2, the boundary between types of alternations (i.e. phonologically, morphologically, or lexically conditioned) cannot always be clearly drawn. However, morphological and lexical involvement generally leads to lowered regularity, transparency, and productivity, which is related to age of acquisition. Thus, children's early language shows a preference for transparent inflection (e.g. MacWhinney 1978, Bernardt & Stemberger 1998). For instance, Laalo (2007) reports that Finnish children eliminate morpho-phonological alternations when forming diminutives by stem modification. Furthermore, lexically conditioned or irregular alternations such as *wife* ~ *wives* are not acquired until the age of

---

[1] Note that Laaha et al. (2007) consider 'slight' stem changes which involve a change in word-final devoicing and/or a vowel lengthening' as transparent, as opposed to stem changes which involve a stem vowel change and/or a consonant alternation.

10, possibly aided by knowledge of spelling (e.g. Bernhardt & Stemberger 1998; Pierrehumbert 2003). At the same time, it is often the case that irregular forms are highly frequent, which is why they may occur early in children's speech. As Marchman & Bates (1994) note, frequent irregular pairs (such as *run* ~ *ran*) are likely to dominate small lexicons. For instance, an analysis of early diminutive forms in two Hungarian children showed that two unproductive suffixes (e.g. *-u* and *-ó*) were produced first (Bodor & Barcza 2007). However, high frequency irregulars are also likely be unanalysed or rote-learned wholes (i.e. they are more entrenched in memory), which makes extraction of a morpho-phonological pattern harder (Bybee 2001). Previous studies show that when morpho-phonological alternations are overgeneralised or extended to novel forms, type frequency of the pattern is the most important predictor of productivity (e.g. Bybee & Pardo 1981, Bybee 2001).

Corpus data may reveal whether and when overgeneralisations occur in spontaneous speech. In case of Dutch voicing alternations, it can be determined to what extent errors of the type *bɛ[t]ən 'beds' or *pɛ[d]ən 'caps' occur. However, while errors such as **foots* are undoubtedly creative, it is unclear whether errors involving devoicing (i.e. *bɛ[t]ən 'beds') reflect the same kind of productivity. Whereas **foots* reflects morphological analysis (the addition of a morphological suffix to a stem), voicing errors involve the incorrect realisation of morphologically correct plurals, which may have been stored as wholes. For instance, voicing errors could also be produced in morphologically simplex or mono-morphemic forms, which would indicate that children have difficulties realising the voicing contrast. It is important to assess the frequency of alternations in the input to the child, both in terms of token frequency (i.e. how often a certain alternating form occurs) and type frequency (i.e. how often the alternating pattern occurs). Studying the frequency distribution of alternations also enables us to test more specific claims made by generative vs. analogical models of language. For instance, it is a priori unclear whether children are merely sensitive to the frequency of specific lexical items that display voicing alternations, or whether purely phonological factors also play a role. As discussed in Chapter 2, children may have a preference for voicing in certain phonological contexts (e.g. in intervocalic position), even when the lexical patterns of a language do not support this preference. Before turning to the frequency of voicing alternations, a more detailed overview of Dutch morphology will be provided in the next section, focusing on the nominal plural.

## 4.2 The Dutch plural and alternations

Dutch plurals are formed by attaching either /s/ or /ə(n)/ to the stem. The choice between the two productive suffixes depends on phonological characteristics (see van Haeringen 1947, de Haas & Trommelen 1993, Booij 1995, van Wijk 2007). There is a preference for *-en* when words end in a final obstruent or diphthong (e.g. *pet ~ petten* 'caps', *trui ~ truien* 'sweaters') and when words have final stress (e.g. *kameel ~ kamelen* 'camels'). A preference for *-s* occurs when words end in a (back) vowel or end in an unstressed syllable (e.g. *ski ~ ski's* 'skis', *lepel ~ lepels* 'spoons').[2] Plurals with voicing alternations only occur with obstruent-final stems suffixed with *-en* (e.g. *bed ~ bedden*), see a list of examples in (27) below.

(27) Regular *-en* plurals with voicing alternation

| | | |
|---|---|---|
| bed | [bɛt] ~ [bɛdən] | 'beds' |
| hoed | [hut] ~ [huden] | 'hats' |
| hand | [hɑnt] ~ [hɑndən] | 'hands' |
| web | [wɛp] ~ [wɛbən] | 'webs' |
| krab | [krɑp] ~ [krɑbən] | 'crabs' |
| neef | [nef] ~ [nevən] | 'nephews' |
| haas | [hɑs] ~ [hɑzən] | 'hares' |

Traditionally, irregular plurals include plurals with vowel alternations (e.g. sxɪ[p] ~ sx[e]pən 'ships') and those that end in the unproductive *-eren* (/ərən/) suffix (e.g. [ei]~ [eijərən] 'eggs'). In (28)-(30) below an overview of plurals with *-eren* is provided, which occur both with and without voicing alternations.[3]

(28) Irregular *-eren* plurals with voicing alternation

| | | |
|---|---|---|
| kind[4] | [kɪnt] ~ [kɪndərən] | 'children' |
| goed | [xut] ~ [xuderen] | 'goods' |
| rund | [rʌnt] ~ [rʌndərən] | 'cattle' |
| lied | [lit] ~ [lidərən] | 'songs' |

[2] Some regular plurals can take either *-n* or *-s*, such as those ending in schwa (e.g. the plural of *ronde* 'round' can be either *ronden* or *rondes*).
[3] Historically, *-eren* plurals are doubly marked (i.e. there was an older plural *kinder*).
[4] This list is not exhaustive due to compounds such as *kleinkind* 'grandchild' or *landgoed* 'estate', indicated by (-).

| | | | |
|---|---|---|---|
| | kalf | [kɑlf] ~ [kɑlvərən] | 'calves' |
| | gemoed | [xəmut] ~ [xəmudərən] | 'emotions' |
| (29) | Irregular -*eren* plurals with voicing and vowel alternation | | |
| | blad | [blɑt] ~ [bladərən] | 'leaves' |
| | rad | [rɑt] ~ [radərən] | 'wheels' |
| | gelid | [gəlɪt] ~ [gəledərən] | 'ranks' |
| (30) | Irregular -*eren* plurals without voicing alternation | | |
| | volk | [vɔlk] ~ [vɔlkərən][5] | 'peoples' |
| | lam | [lɑm] ~ [lɑmərən] | 'lambs' |
| | ei | [ei] ~ [eijərən] | 'eggs' |
| | been[6] | [ben] ~ [bendərən] | 'bones' |
| | hoen | [hun] ~ [hundərən] | 'hens' |

These data show that when irregular stems (taking -*eren*) end in /d/ they always co-occur with voicing alternations (i.e. there are no stems with /t/) in this class, indicating a relationship between voicing alternations and irregularity. Furthermore, there is a tendency for voicing alternations to co-occur with vowel alternations, to be discussed further below (see footnote 11).[7] Examples of patterns of voicing and vowel alternations are shown in (31) below.

| (31) | | | Irregular suffix | Vowel alternation | Voicing alternation |
|---|---|---|---|---|---|
| rat | [rɑt] ~ [rɑtən] | 'rats' | - | - | - |
| raat | [rat] ~ [ratən] | 'honeycombs' | - | - | - |
| raad | [rat] ~ [radən] | 'councils' | - | - | + |
| gat | [ɣɑt] ~ [ɣatən] | 'holes' | - | + | - |
| rad | [rɑt] ~ [radən] | 'wheels' | - | + | + |
| stad | [stɑt] ~ [stedən] | 'cities' | - | + | + |
| kind | [kɪnt] ~ [kɪndərən] | 'children' | + | - | + |
| rad | [rɑt] ~ [radərən] | 'wheels' | + | + | + |

[5] The plural [vɔlkən] 'peoples' is also possible.
[6] Note that the final two plurals in (2) have an inserted /d/, e.g. *been* ~ *beenderen* 'bones'.
[7] There are four types of vowel alternations that occur both for forms with and without voicing alternations, i.e. [ɑ] ~ [a], [ɔ] ~ [o], [ɪ] ~ [e], and [ɛ] ~ [e]. Additionally, [ɛɪ] ~ [e] (e.g. *mogelijkheid* ~ *mogelijkheden* 'possibilities') and [ɑ] ~ [e] (*stad* ~ *steden* 'cities') only occur for stems in /d/.

This overview shows that there are different degrees of transparency in Dutch plurals, which are likely to be related to degrees of regularity and productivity. Thus, fully transparent, non-alternating pairs that take the regular *-en* suffix, such as *rat ~ ratten* 'rats', could be posited at one end of a scale, whereas pairs such as *rad ~ raderen* 'wheels' are at the other end. On such a scale, alternating pairs such as *raad ~ raden* 'councils' could be said to take an intermediary position. Even though they are not fully irregular, they could behave differently from fully regular pairs. Finally, alternating forms are not always produced with a voiced obstruent. Intervocalic /d/ frequently undergoes weakening in Dutch, such as in /xujə/ for /xudə/ *goede* 'good-infl.' or /auwə/ for / audə/ *oude* 'old-infl' (see Zonneveld 1978 for an extensive discussion). Frequent adjectives and verbs may undergo glide formation, which means that they are less reliable as evidence for voicing alternations.[8]

In sum, voicing alternations in Dutch stems often co-occur with irregularities such as the suffix *-eren*, vowel alternations and weakening (glide formation), suggesting that the voicing alternation may itself reflect irregularity. If this is the case, words with voicing alternations are expected to be relatively frequent, as irregularity and frequency tend to correlate (see also Chapter 3). Before we turn to the acquisition of plurals based on corpus studies, the frequency of alternating words will be described in more detail in the next section.

## 4.3 Frequency of voicing alternations

According to some models, lexical frequency plays a central role in the acquisition of morpho-phonological alternations (see Chapter 2). To assess the role of the lexicon, the frequency of voicing alternations will first be explored. Generally, stem or *lemma* frequency (summing the frequency of all word forms, e.g. *hand*, *handje*, *handen*, *handjes*) is distinguished from surface or *wordform* frequency, which is the frequency of one of the word forms (e.g. *handen).* For the present corpus study the Dutch CELEX lexical database was used, based on written language in a variety of text types and

[8] CELEX data show that words that may be realised with glides tend to be highly frequent (e.g. there are 22 adjectival stems for which weakening is common, representing 5% of alternating types but 50% of tokens).

containing around 42 million words (Burnage 1990, Baayen et al. 1995).[9] To provide a general estimate of the probability of occurrence of voicing alternations, all abstract stems ending in /p, b, t, d/ were extracted from the corpus. Importantly, fricatives are not considered in the present study because the voicing contrast is less relevant for fricatives, given word-initial realisations (the contrast is not made in the Western part of the Netherlands) and the small set of minimal pairs (Ernestus 2000). For detailed information about the data selection see Appendix A. Only words (i.e. nouns, verbs, and adjectives) with a frequency above zero were selected, resulting in 7,955 complex or bi-morphemic words (84,531 tokens). To explore the distribution of alternating words, frequency counts were restricted to the phonological contexts that were used for the experimental stimuli (Chapter 5). In this way, the corpus data could provide testable hypotheses for the experiments. The count was restricted to stems ending in a. two types of final obstruent (alveolar T /t,d/ and labial P /p,b/) and b. three types of rhyme (short vowel V_ (/ɑ,ɛ,ɪ,ɔ,œ/), long vowel VV_ (/a,e,o,ø/) and the high tense vowels /i,y,u/) and a vowel followed by a nasal N_ (/m,n/). Hence, stems in which the final obstruent is preceded by an obstruent or liquid were not considered (e.g. *beest* 'animal', *paard* 'horse'), and no such stimuli were used in the elicitation experiments.[10] Plurals with vowel alternations (e.g. *stad ~ steden* 'cities') were excluded, as they do not form transparent pairs.[11] This count resulted in 4,224 words (42,853 tokens), containing 2,131 stems (29,313 tokens), 85% of which have an alveolar obstruent (i.e. /t/ or /d/). Table 2 shows the number of alternating and non-alternating stems in this set together with the proportion of stems and the mean token frequency (note that highly frequent stems with vowel alternations are excluded, e.g. /pɑd/ ~ /padən/ 'roads').

[9] Web-based CELEX was used for these counts, see http://www.mpi.nl/world/celex/.
[10] This restriction was partly based on the notion that stems ending in /Rd/ (liquid and /d/) and /Nd/ (nasal and /d/) are expected to behave in a similar way (from a phonological perspective).
[11] It was suggested in the previous section that voicing alternations and vowel alternations tend to co-occur. This is borne out by the data: CELEX contains 40 stems (frequency above zero) which undergo vowel alternations in the plural but no voicing alternation (788 tokens), vs. 265 stems that undergo both vowel and voicing alternations (2746 tokens). When words ending in -*heid* '-ity, '-ness' are excluded, 73 stems (1113 tokens) remain.

*Table 2: Number of stems and mean frequency (sd) (CELEX).*

| | **Non-alternating** | | | **Alternating** | | |
|---|---|---|---|---|---|---|
| | /t/ | /p/ | Total | /d/ | /b/ | Total |
| Type | 918 | 319 | 1237 | 883 | 7 | 890 |
| Token | 8,297 | 3,365 | 11,662 | 16,731 | 676 | 17,407 |
| Mean | 9 (31) | 11 (31) | 9 (31) | 19 (149) | 97 (252) | 20 (150) |

Table 2 shows that alternating stems are more frequent than non-alternating stems. An ANOVA with factors Alternation (Non-alternation, Alternation) and Obstruent (T, P, where 'T' stands for stems ending in /t/ and /d/ and 'P' stands for stems ending in /p/ and /b/) was conducted with mean frequency as dependent variable. This analysis shows that overall, alternating stems are more frequent than non-alternating stems ($F(1)=6{,}297$, $p=.012$) and stems ending in T are more frequent than stems ending in P ($F(1)=4{,}241$, $p=.040$). There was also an interaction ($F(1)=3{,}920$, $p=.048$), which means that the effect of alternation is larger for labials. As shown in Table 2, there are very few /b/ types, and their high token frequency derives mainly from a single type (i.e. *heb ~ hebben* 'have', which accounts for 98 % of /b/ tokens).

The distribution of alternations was determined by computing the frequencies of stems with an underlying voiced final segment as a percentage of the summed type or token frequencies of all stems (i.e. both alternating and non-alternating). The result shows that there are generally fewer alternating than non-alternating *types* (42%), although alternating stems account for more than half of the total number of *tokens* (60%).[12] Hence, alternations are associated with lower type frequency and high token frequency. To illustrate, the twenty most frequent words in this corpus (representing 55% of all tokens) are listed in (32) below (alternating words are bold).

[12] When all rhymes are considered (e.g. including *beest* 'animal' and *paard* 'horse', proportions are 38% of types and 48% of tokens.

(32)

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| 1. | **had** 'had' | V | 3946 | 11. | **kind** 'child' | N | 454 |
| 2. | **goed** 'good' | A | 1071 | 12. | **deed** 'did' | V | 444 |
| 3. | **hebben** 'have' | V | 1054 | 13. | zat 'sat' | V | 412 |
| 4. | **tijd** 'time' | N | 958 | 14. | liet 'let' | V | 387 |
| 5. | **hadden** 'had' | V | 851 | 15. | groot 'big' | A | 386 |
| 6. | **heb** 'have' | V | 667 | 16. | **handen** 'hands' | N | 377 |
| 7. | **hand** 'hand' | N | 645 | 17. | moet 'must' | V | 375 |
| 8. | **stond** 'stood' | V | 645 | 18. | laten 'let' | V | 371 |
| 9. | **kinderen** 'children' | N | 484 | 19. | liep 'walked' | V | 358 |
| 10. | grote 'big' | A | 480 | 20. | **goede** 'good' | A | 340 |

This list shows that alternating nouns, verbs and adjectives are among the most frequent words, although there is only one /b/ alternation (i.e. *heb* ~ *hebben* 'have'). Finally, Table 3 shows the relative number of alternating stems in each word class.

*Table 3: Number of voicing alternations per word class (%) (CELEX).*

| | **Types** | | **Tokens** | |
|---|---|---|---|---|
| | t ~ d | p ~ b | t ~ d | p ~ b |
| Noun | 34.6 | 0.6 | 41.1 | 0.05 |
| Verb | 13.8 | 0.2 | 34.6 | 3.8 |
| Adjective | 50.8 | 0.0 | 20.2 | 0.0 |
| Total | 99.2 | 0.8 | 96.1 | 3.9 |

Table 3 shows that nouns are associated with a relatively high number of *t~d* alternations compared to verbs (but not adjectives), and that nouns account for the majority of tokens.[13]

Now that the overall frequency of alternating words has been established, let us turn to the distribution of voicing alternations according to rhyme. This is relevant for determining whether children are sensitive to phonological context. In Table 4, the three rhyme types used for the corpus search are indicated by 'V_' for short vowels (e.g. *bedden* 'beds'), 'V:_' for

[13] When mean token frequencies are considered, *t~d* alternations are more frequent than non-alternations for verbs (49 vs. 16) and nouns (24 vs. 8) but not adjectives (8 vs. 9).

long vowels (e.g. *laden* 'load') and N_ for nasals (e.g. *handen* 'hands'). There are six phonological environments in total, as the proportion of alternating stems is shown separately for T (*t~d*) and P (*p~b*).

*Table 4: Overall distribution of voicing alternations (%) (CELEX).*

| | **Type Frequency** | | **Token Frequency** | |
|---|---|---|---|---|
| | t ~ d | p ~ b | t ~ d | p ~ b |
| V_ | 2.8 | 0.8 | 24.6 | 3.9 |
| V:_ | 29.3 | 0.0 | 29.6 | 0.0 |
| N_ | 67.1 | 0.0 | 42.0 | 0.0 |
| Total | 99.2 | 0.8 | 96.1 | 3.9 |

The first thing to notice is that there is a labial-alveolar split for contexts following long vowels and nasals (indicated by shading). This indicates a 'lexical gap', i.e. *p~b* alternations occur after short vowels (e.g. [hɛp] ~ [hɛbən] 'have'), but there are virtually no pairs such as [ap] ~ *[abən], compare [ap] ~ [apən] 'monkeys'.[14] As noted before, /b/ alternations are mostly due to a single type and have a high mean frequency.[15] Given that there are exceptions (and no loanword adaptations), this lexical gap seems to be 'accidental' and not motivated by phonotactic restrictions. This is shown by the distribution of mono-morphemic words with voiced intervocalic obstruents, which do occur in these lexical gap environments (e.g. *globe* 'globe'), see Table 5 below.

[14] Exceptions are three adjectives and one noun of Greek origin ending in the noun- and adjective forming suffix *–foob* 'phobe' (e.g. *claustrofoob ~ claustrofobe* 'claustrophobic'), as noted by Zonneveld (1978:49), who also provides *Zwaab ~ Zwaben* 'Swabian', and argues that these exceptions belong to the [-NATIVE] portion of the Dutch lexicon (acquired in adulthood).
[15] The frequency of /b/ after short vowels is 97 vs. 10 for /p/ after short vowels. Note that /d/ is also frequent after short vowels (with a mean frequency of 165 vs. 6), whereas the difference is less extreme for long vowels (21 vs. 12) or nasals (12 vs. 8).

*Table 5: Distribution of medial voicing in mono-morphemic words (%) (CELEX).*

| | **Type frequency** | | | | **Token frequency** | | | |
|---|---|---|---|---|---|---|---|---|
| | /t/ | /d/ | /p/ | /b/ | /t/ | /d/ | /p/ | /b/ |
| V_ | 3.8 | 3.5 | 2.9 | 2.5 | 2.6 | 1.9 | 0.8 | 0.5 |
| V:_ | 18.8 | 29.7 | 6.8 | 3.3 | 28.2 | 31.6 | 4.8 | 1.1 |
| N_ | 6.6 | 19.6 | 1.7 | 0.8 | 4.1 | 22.6 | 0.6 | 1.1 |
| Total | 29.2 | 52.8 | 11.4 | 6.6 | 34.9 | 56.1 | 6.2 | 2.7 |

Table 5 shows that mono-morphemic words with medial /b/ occur in all environments, although they are generally less common than mono-morphemic forms with /p/.[16] Clearly, the two 'lexical gap' environments behave similarly when complex words are considered, which is presumably related to the length of the rhyme.[17] Also note that medial /d/ is generally more frequent than medial /t/ in mono-morphemic words. If the count is restricted to nouns (e.g. *moeder* 'mother'), words with /d/ account for half of the types (50.1%) and tokens (52.6%).

The second thing to notice about the distribution in Table 4 is that most *t~d* alternations occur in the nasal context (e.g. *handen* 'hands'). On the basis of type frequency, most overgeneralisations are thus expected to occur for this environment, for which the overall probability of voicing based on alternations is 28% (597/2131).[18] It is not clear how word class will affect children's knowledge of alternating patterns. However, it is likely that subjects' behaviour in a Wug-test is mainly influenced by words within the same class (and only nouns were used for the experiments described in Chapter 5 and 6). Table 6 therefore shows the number of alternating and non-alternating nouns and their mean frequencies (comparable to Table 2).

[16] However, words with /mb/ (e.g. *december* 'december') are more common than words with /mp/ (e.g. *drempel* 'threshold').

[17] Both long vowels and vowel-nasal sequences are bimoraic (constituting two moras), see Hayes (1989) and Kager (1989) for Dutch.

[18] This pattern is mirrored in simplex word; words like *tante* 'aunt' are relatively uncommon, whereas there are many high frequent words with postnasal /d/, e.g. *ander* 'other'.

*Table 6: Number of nouns and mean frequency (sd) (CELEX).*

| | **Non-alternating** | | | **Alternating** | | |
|---|---|---|---|---|---|---|
| | /t/ | /p/ | Total | /d/ | /b/ | Total |
| Types | 670 | 251 | 921 | 308 | 5 | 313 |
| Tokens | 5,305 | 2,202 | 7,507 | 7,201 | 8 | 7,209 |
| Mean | 8 (23) | 9 (21) | 8 (23) | 23 (82) | 2 (1) | 23 (81) |

This table shows that the alternating nouns are more frequent than non-alternating nouns, although this only applies to alveolar stems (F(1)=26,825, p<.001). Finally, the probability of voicing is somewhat lower for nouns (26%), although alternations still account for half (50%) of tokens.[19] Table 7 shows the distribution of alternations for nouns per phonological environment.

*Table 7: Distribution of voicing alternations for nouns (%) (CELEX).*

| | **Type Frequency** | | **Token Frequency** | |
|---|---|---|---|---|
| | t ~ d | p ~ b | t ~ d | p ~ b |
| V_ | 4.5 | 1.6 | 4.2 | 0.1 |
| V:_ | 42.2 | 0.0 | 35.4 | 0.0 |
| N_ | 51.8 | 0.0 | 60.3 | 0.0 |
| Total | 98.4 | 1.6 | 99.9 | 0.1 |

This distribution confirms that there are virtually no nouns with /b/, whereas nouns also show a tendency for postnasal voicing of /t/.

It is also relevant to consider singular and plural nouns separately. For instance, the singular *bed* occurs 284 times in the corpus, whereas the plural *bedden* 'beds' has a frequency of 12 (for comparison, the singular *pet* has a frequency of 16 whereas the plural *petten* 'caps' occurs only twice). Hence, the relatively low frequency of some plurals could affect the acquisition of alternations, which are presumably learned through pairs. Separate frequencies for singulars and plurals are provided in Table 8.

[19] When all rhymes are considered (e.g. including *paard* 'horse') the proportions are 27% of types and 42% of tokens.

*Table 8: Number of types and mean frequency of singulars and plurals (CELEX).*

| | **Singular** | | **Plural** | | **Total** | |
|---|---|---|---|---|---|---|
| | types | mean (sd) | types | mean (sd) | types | mean (sd) |
| Non-alternating | 921 | 8 (23) | 575 | 5 (13) | 1,496 | 7 (20) |
| Alternating | 313 | 23 (81) | 179 | 13 (49) | 492 | 20 (71) |
| Total | 1,234 | 12 (46) | 754 | 7 (27) | 1,988 | 10 (40) |

Table 8 again shows that there are fewer alternating nouns, which are associated with a higher mean frequency. An ANOVA with independent factors Alternation (alternating or non-alternating) and Inflection (singular or plural) was performed, to assess the difference in mean frequency. The result showed that alternations are more frequent than non-alternations ($F(1) = 30{,}393$, $p<.001$), and singulars are more frequent than plurals ($F(1) = 9{,}272$, $p=.002$). There was also a trend towards an interaction, which means that alternating singulars may be relatively more frequent than non-alternating singulars ($F(1)=3{,}075$, $p=.080$).[20] These results underscore the importance of considering the frequency of both singulars and plurals, as children presumably learn alternations on the basis of *pairs*. To illustrate, the ten most frequent singular and plural forms in the database are shown in (33) below (alternating words are bold, stars indicate the occurrence of a pair).

[20] The interaction between Alternation and Inflection is significant when all words are taken into account, i.e. including words with a frequency of zero ($F(1) = 8{,}988$, $p<.003$).

(33)

| Singulars | | | | Plurals | | | |
|---|---|---|---|---|---|---|---|
| 1. | **tijd*** | 'time' | 958 | 1. | **kinderen*** | 'children' | 484 |
| 2. | **hand*** | 'hand' | 645 | 2. | **handen*** | 'hands' | 377 |
| 3. | **kind*** | 'child' | 454 | 3. | **maanden** | 'months' | 155 |
| 4. | **grond** | 'ground' | 321 | 4. | voeten | 'feet' | 129 |
| 5 | **land*** | 'land' | 298 | 5 | **vrienden** | 'friends' | 125 |
| 6. | **bed** | 'bed' | 284 | 6. | **landen*** | 'countries' | 114 |
| 7. | staat | 'state' | 273 | 7. | lippen | 'lips' | 102 |
| 8. | kant | 'side' | 235 | 8. | groepen | 'groups' | 91 |
| 9. | **mond** | 'mouth' | 220 | 9. | **tanden** | 'teeth' | 74 |
| 10 | feit | 'fact' | 203 | 10 | **tijden*** | 'times' | 62 |

This list shows that there are four pairs among the most frequent alternating words (including the irregular *kind ~ kinderen*), while there are no non-alternating pairs. Furthermore, even though alternating singulars may be highly frequent, there are plurals which are more frequent than their singular (e.g. *tanden* 'teeth'). This is relevant for acquisition, as these plurals are likely to be among the first plurals to be acquired (as opposed to plurals of words that are dominant in their singular form, such as *mond* 'mouth'). However, highly frequent plurals are also more likely to be stored as a whole, i.e. they are less likely to be acquired in relation to the singular.

In sum, adult corpus data (which could be taken to reflect the input) show that alternating plurals tend to be associated with a high token frequency, and that most of the evidence for alternations comes from nominal *t~d* alternations after nasals (e.g. *handen* 'hands'). On the other hand, there is little evidence for *p~b* alternations in general, and there is a 'lexical gap' for such alternations after long vowels and nasals. The frequency data provided in this section will be used to investigate whether children are sensitive to the distribution and frequency of voicing alternations in the input. Before we turn to children's productions however, it is useful to investigate whether the input to the child is similar to the adult corpus data. The next section therefore discusses child corpus data, both in terms of the input (4.3.3) and the child's own productions (4.3.4).

## 4.4 Child corpus data

### 4.4.1 Introduction

Traditionally, child language acquisition has been investigated using parental diary notes and recorded spontaneous speech. Morphological productivity is cross-linguistically attested, although it is more obvious for languages with rich morphology (Brown 1973). Typically, morphological overregularisation errors appear when children are around 3 years of age, and continue into school-age years. Based on corpus studies of languages that require both affixal inflection and stem alternations in the inflected form (i.e. Russian, Polish, Spanish, French, German and Hebrew), Slobin (1985) claims that stem modifications are difficult, as children generally add an affix only (see also Orsolini, Fanari & Bowles 1998 for Italian). MacWhinney (1978) provides data from Hungarian, German and French diary data, showing some evidence for the incorrect use of the root allomorph, as in French **buver* for *boire* 'to drink' (from forms such as *buvons* 'we drink'). On the basis of spontaneous speech data of 15 Spanish children (1;7–4;7), Clahsen, Aveledo & Roca (2002) found that irregular stems may be over-extended (e.g. **cayí* for *caí* 'I fell' at 2;1), although the large majority of stem errors consisted of regularisations (e.g. **sabo* for *sé* 'I know', **poniste* for *pusiste* 'you put-past'). Errors in which non-alternating stems were used instead of diphthongised stems were also common (e.g. **juga* for *juega* 's/he/it plays' at 2;2), at a rate of 18% (similar to the 14% errors for irregular past tense forms). In contrast, diphthongised stems were never over-extended. Finally, Clahsen et al. found that the overall error rate for irregular verbs was only 4.6%.[21] In a large corpus study of the English past tense based on data from 25 children, Marcus et al. (1992) report a mean overregularisation rate of 4.2%. Generally, overgeneralisations of irregular patterns are rare (e.g. **truck* for *tricked*), while overgeneralisations of regular patterns (e.g. **foots*) do not occur often either (Cazden 1968, Kuczaj 1977, Marcus et al. 1992, Xu & Pinker 1995). Overregularisations have been found to be sensitive to frequency and similarity, as more errors occur for low-frequency irregular verbs (Bybee & Slobin 1982, Pinker 1984).

[21] Error rates are defined as the number of overregularisations divided by the total number of irregular productions (including overregularisations and correct productions).

Plural formation is argued to be one of the first inflectional systems acquired, although there may be cross-linguistic differences (Brown 1973, de Villiers & de Villiers 1973). Overgeneralisations of English plurals (**foots*) are reported to occur from around 1;8 until 2;8 (MacWhinney 1978). Examples from other languages show a similar picture, e.g. Slobin (1977) claims that Turkish children can inflect nouns productively by age 2. MacWhinney (1987:57) reports data from German diary studies that suggest that the first plurals occur around 2;0. However, the complex system of German plural morphology is not fully acquired until age 5 (Clahsen et al. 1992). In general, few errors with German plurals are found in spontaneous data, although overgeneralisations of both -*e*(*n*) and -*s* occur, e.g. **Rädern* for *Räder* 'wheels' and **Manns* for *Männer* 'men' by Simone at 2;8 (Clahsen et al. 1992). As found for English, such overgeneralisations are rare in spontaneous speech, ranging from 4% to around 10%.

With respect to morpho-phonological voicing alternations, MacWhinney (1987:2) notes that errors reflecting uniformity (e.g. *[waifs] for [waivz] *wives*) are common, comparable to Dutch *[bɛtən] 'beds'. However, the voicing alternation in English is much more restricted than in Dutch, applying to a very small number of nouns. The acquisition of voicing alternations has not been investigated systematically for either German or Dutch. Before turning to Dutch child corpus data, previous corpus studies on the acquisition of the Dutch plural (rather than the alternation) will first be described.

### 4.4.2 Acquisition of the Dutch plural

Before turning to plurals with voicing alternations, it is important to consider the acquisition of the plural suffix. For Dutch, Schaerlaekens & Gillis (1987) report that functionally appropriate plural forms appear between the ages of 2;7 and 3;1 (see also de Houwer 1990). However, earlier occurrences of plurals are noted, which may be either imitations or rote forms. For instance, Zonneveld (2004) shows that singular - plural pairs occurred as early as 1;6 based on diary notes of his daughter Nina, and at the age of 1;8 based on recordings of spontaneous speech. Schaerlaekens & Gillis (1987) and Schaerlaekens (1980) report that around the age of 2;0–2;6, (Flemish) Dutch speaking children produce a limited number of correct plurals, which are usually plural dominant nouns like *schoenen* 'shoes' and *ogen* 'eyes'. Children's early plural forms include both -*en* and -*s* plurals, although De Houwer (1990) notes that early *-s* plurals are limited to diminutives (also noted by Zonneveld 2004). At a later stage, Schaerlaekens (1980:153) reports that one affix is used for most plurals, although overgeneralisations of both *-s* and *-en* occur: **visses* 'fish', **schoenes* 'shoes' (Katelijn at 2;8), **twee voets* 'feet' (Gerrit at 3;1), **auto-e* 'cars' (Erik at 3;2), see also Schaerlaekens & Gillis (1987:138). Generally, 4-year-olds are reported not to produce errors for regular forms in spontaneous speech (de Vleeshauwer 1986). However, irregular forms (e.g. undergoing stem vowel changes) remain difficult throughout primary school, until around the age of 6 (Schaerlaekens 1980). For instance, regularisations such as *schippen* for *schepen* 'ships' (from [sxɪp] ~ [sxepən]), *blatten* for *bladeren* 'leaves' (from blɑ[t] ~ bla[d]ərən) and *vatten* for *vaten* 'barrels' (from [vɑt] ~ [vatən]) are reported for this age group. However, there has not been any detailed investigation of voicing alternations, and previous studies do not report whether these plurals are considered to be regular, or taken to belong to the class of irregulars. Before we turn to early productions of plurals that involve voicing alternations, the input to the child will be investigated.

### 4.4.3 Evidence for voicing alternations

Within the present research project (see Chapter 1), Zamuner (2006ab) has studied the distribution of voicing alternations in child directed speech on the basis of the van de Weijer corpus, and found that there are more non-alternating pairs than alternating pairs in the input to the child (Zamuner 2006).[22] In the present study, analyses of a different corpus of child-directed speech are presented, with the aim of studying a wider age range. To this end, the 'van Kampen' corpus (van Kampen 1997) was used, which includes the speech of a mother and daughter Sarah, recorded when the child was aged between 1;6 and 5;2.[23] To determine to what extent alternating forms

---

[22] This corpus (van de Weijer 1998) includes 18 days of adult-, child- and infant-directed speech recorded when the child was aged 2;6-2;9 and the infant was aged 0;6-0;9.

[23] The van Kampen corpus (van Kampen 1994) is available through CHILDES (the Child Language Data Exchange System, see MacWhinney 2000) and contains naturalistic data of two Dutch speaking girls, Sarah and Laura. Sarah was studied from 1;6.16 to 5;2.13, resulting in 50 45-minute recordings. Sarah's MLU ranged from 1.1 at the start of the recordings (at age 1;6) to 6.0 (at age 4;9). Laura was studied from the age of 1;9.4 to 5;6.12, resulting in 72 45-minute recordings. Recordings were made once or twice every month, in an unstructured home setting by the mother (van Kampen).

occur in the input to the child, the mother's speech was first analysed. Counts of words in the input are comparable to the CELEX counts presented in section 4.3, representing target words, rather than actual realisations. In this corpus, errors are indicated orthographically, which makes the corpus less reliable for phonological analyses. Thus, any variation in the mother's speech such as contracted forms (e.g. /hɛbɪk/ 'have I'), voicing assimilation (e.g. /hɛbdɑt/ 'have that'), deletion of final /t,d/ (e.g. *ik vind* [vɪnt] 'I think' as [ɪkfɪn]), glide formation or deletion of /n/, is not considered (see Chapter 3). To illustrate, the input to Sarah until age 3;4 (33 recordings) contained 122 instances of the form *vind* 'find', exactly half of which were transcribed as *vin(d)*, i.e. pronounced without final /t/. Still, it is useful to compare this corpus to the adult CELEX corpus, to determine whether a child (up until age five) receives the same kind of evidence for alternations.

A set of 570 bi-morphemic words (8708 tokens) with medial and / or final obstruents (/t/, /d/, /b/, /d/) was extracted from the corpus. Arguably, the clearest evidence for a child who needs to acquire knowledge of alternations is the occurrence of pairs. Overall, the input to Sarah contained 91 *pairs* (61 non-alternating vs. 30 alternating).[24] This confirms previous corpus findings by Zamuner, in that there are roughly twice as many non-alternating pairs in the input.

To directly compare the input data to the CELEX frequencies, pairs with rhymes in which the final obstruent was preceded by a liquid (e.g. *paarden* 'horses') or an obstruent (e.g. *beesten* 'animals') were excluded. After restricting the counts to the relevant environments (i.e. short vowels, long vowels and nasals), the corpus contained 237 stems (5182 tokens), for which the proportion of alternating stems was 38% of types (90/237) and 48% of tokens (2463/5182), which is similar to CELEX (see Table 2).[25] The twenty most frequent words in Sarah's input (representing 61% of all tokens) are listed in (34) below (alternating words are bold), comparable to the list in (32).

[24] There were 14 additional pairs like *hoorde ~ hoort* or *kent ~ kende* 'knew', in which the past tense suffix [də] alternates with the present tense and participle ending [t] (possibly constituting evidence for voicing alternations).
[25] When all rhymes are considered the proportion is 36% of types and 45% of tokens.

(34)

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| 1. | **heb** 'have'* | V | 758 | 11. | moeten 'must'* | V | 108 |
| 2. | moet 'must'* | V | 686 | 12. | grote 'big' | A | 89 |
| 3. | **goed** 'good' | A | 555 | 13. | **had** 'had' | V | 88 |
| 4. | zitten 'sit'* | V | 373 | 14. | slapen 'sleep' | V | 88 |
| 5. | **hebben** 'have'* | V | 294 | 15. | **hond** 'dog' | N | 86 |
| 6. | weet 'know' | V | 287 | 16. | **grond** 'floor' | N | 72 |
| 7. | zit 'sit'* | V | 274 | 17. | laat 'late' | A | 72 |
| 8. | **vind(t)** 'find' | V | 229 | 18. | kapot 'broken' | A | 71 |
| 9. | eten 'eat, food' | V/N | 187 | 19. | **bed** 'bed' | N | 63 |
| 10. | heet 'hot/is called' | A/V | 123 | 20. | kant 'side' | N | 62 |

The only alternating pair in this list is the frequent *heb ~ hebben*, although alternating *singulars* are among the twenty most frequent words (as was found for CELEX, see Table 2).[26] The distribution of voicing alternations in each phonological context is shown in Table 9, comparable to Table 4 for CELEX.

*Table 9: Overall distribution of voicing alternations (%) in child-directed speech (CHILDES).*

| | **Type Frequency** | | **Token Frequency** | |
|---|---|---|---|---|
| | t ~ d | p ~ b | t ~ d | p ~ b |
| V_ | 13.3 | 3.3 | 9.4 | 31.3 |
| V:_ | 34.4 | 0.0 | 32.1 | 0.0 |
| N_ | 48.9 | 0.0 | 27.2 | 0.0 |
| Total | 96.7 | 3.3 | 68.7 | 31.3 |

Results are again similar to the CELEX data (Table 4), as there is a 'lexical gap' for *p~b* alternations and alternations after nasals have the highest type frequency. Note that the relative number of *t~d* alternations after short vowels is somewhat higher than in CELEX. In Table 10, the distribution of alternations for nouns on the basis of child-directed speech is shown, comparable to Table 7 for CELEX.

[26] Adjectives commonly produced with a glide (e.g. [xujə] 'good' ) were highly frequent, i.e. they accounted for 80% (35/44) of tokens of inflected alternating adjectives.

*Table 10: Distribution of voicing alternations for nouns (%) in child-directed speech (CHILDES).*

| | **Type Frequency** | | **Token Frequency** | |
|---|---|---|---|---|
| | t ~ d | p ~ b | t ~ d | p ~ b |
| V_ | 17.3 | 3.8 | 22.0 | 1.9 |
| V:_ | 25.0 | 0.0 | 15.9 | 0.0 |
| N_ | 53.8 | 0.0 | 60.2 | 0.0 |
| Total | 96.2 | 3.8 | 98.1 | 1.9 |

This table shows that *t~d* alternations mostly occur after nasals, as was found for CELEX.[27] There is hardly any evidence for *p~b* alternations, i.e. there were only two singulars with /b/ (*krab* 'crab' and *slab* 'bib') and no plurals. The probability of voicing alternation for nouns is relatively high in child-directed speech, i.e. 34% (52/154) vs. 26% in CELEX. The relative frequency of singulars and plurals is shown in Table 11 (comparable to Table 8).

*Table 11: Number of types and mean frequency of singulars and plurals in child-directed speech (CHILDES).*

| | **Singular** | | **Plural** | | **Total** | |
|---|---|---|---|---|---|---|
| | types | mean (sd) | types | mean (sd) | types | mean (sd) |
| Non-alternating | 102 | 6.7 (11) | 33 | 4.1 (7) | 135 | 6.1 (10) |
| Alternating | 52 | 12.4 (19) | 13 | 10.5 (17) | 65 | 12.0 (19) |
| Total | 154 | 8.6 (14) | 46 | 5.9 (11) | 200 | 8.0 (14) |

An ANOVA showed that alternating nouns are more frequent than non-alternating nouns ($F(1) = 26{,}479$, $p<.001$), and singulars are generally more frequent than plurals ($F(1) = 42{,}435$, $p<.001$). An interaction ($F(1) = 8{,}919$, $p=.003$) was also present, which means that alternating singulars are relatively more frequent than non-alternating singulars (as found for CELEX, see Table 8). However as argued before, the number of pairs in the

[27] Mono-morphemic words with medial /d/ (e.g. *modder* 'mud', *minder* 'less') are more frequent than medial /t/ (e.g. *water* 'water', *lente* 'spring'), as was found for CELEX. Simplex words with /d/ make up 64% of types and 60% of tokens. Words with /nd/ represent 79% of types and 91% of tokens.

input is important, as children presumably benefit most from (frequently) hearing both forms. An example of a pair is shown in (35) below (where 'M' stands for the mother's speech and 'S' stands for the child Sarah).

(35)

| | | |
|---|---|---|
| M: | o, misschien ligt ie [: hij] wel in bed. | *Oh maybe he lies in bed.* |
| M: | of niet? | *Or not?* |
| S: | ik zie (h)et niet. | *I don't see it.* |
| M: | weet je nog die mooie bedden, (.)? | *Remember those nice beds(.)?* |

Even though here, the plural occurred together with the singular, Sarah's mother produced the singular *bed* 63 times, while the plural *bedden* was used only twice. A list of the most frequent singulars and plurals is provided in (36) below (alternating words are bold and stars mark the occurrence of pairs), comparable to the list in
(33).

(36)

| **Singulars** | | | | **Plurals** | | | |
|---|---|---|---|---|---|---|---|
| 1. | **hond** | 'dog' | 86 | 1. | **handen*** | 'hands' | 50 |
| 2. | **grond** | 'ground' | 72 | 2. | voeten | 'feet' | 42 |
| 3. | **bed** | 'child' | 63 | 3. | **kinderen** | 'children' | 41 |
| 4. | kant | 'side' | 62 | 4. | **tanden** | 'teeth' | 27 |
| 5 | **hand*** | 'hand' | 50 | 5 | kippen* | 'chickens' | 14 |
| 6. | **bad** | 'bath' | 44 | 6. | olifanten | 'elephants' | 9 |
| 7. | pop* | 'doll' | 43 | 7. | poppen* | 'dolls' | 8 |
| 8. | **mond** | 'mouth' | 42 | 8. | tomaten | 'tomatoes' | 6 |
| 9. | kip* | 'chicken' | 40 | 9. | ruiten | 'diamonds' | 5 |
| 10 | sap | 'juice' | 36 | 10 | **potloden** | 'pencils' | 4 |

This list shows that alternating singulars and plurals are among the most frequent words in the input. The three most frequent plurals (*handen* 'hands', *kinderen* 'children', *tanden* 'teeth') are all *t~d* alternations after nasals, accounting for 44% (118/271) of all plural tokens.[28] As found in

[28] The two most frequent alternating plurals in the contexts that were not considered were *paarden* 'horses' (7) and *woorden* 'words' (5).

CELEX, *tanden* 'teeth' is more frequent than its singular (27 vs. 6), and the same applies to *potloden* 'pencils' (4 vs. 2). Similarly, some frequent alternating words only occur as singulars (e.g. *mond* 'mouth'). This shows that not all alternating forms are equally good evidence for the alternating pattern. The list shows only one alternating pair (with equal frequencies of singular and plural, i.e. *hand ~ handen* 'hands'). To explore this further, the number of *pairs* in the input was counted (based on the restricted corpus). This yielded a total of 74 pairs (50 non-alternating vs. 24 alternating), including 16 non-alternating noun pairs vs. 12 alternating noun pairs.[29] An overview of the number and mean frequency of noun pairs is provided in Table 12.

*Table 12: Number and mean frequency (sd) of noun pairs in child-directed speech (CHILDES).*

| | **Types** | **Singular** | **Plural** | **Total** |
|---|---|---|---|---|
| Non-alternating | 16 | 16.3 (19) | 5.8 (10) | 11.0 (16) |
| Alternating | 12 | 22.1 (29) | 11.3 (17) | 16.7 (24) |
| Total | 28 | 18.8 (23) | 8.1 (14) | 13.4 (20) |

This table shows that the input to the child contained fewer alternating pairs than non-alternating pairs, even though alternating singulars and plurals were generally more frequent than non-alternating singulars and plurals. This shows that child-directed speech provides more evidence for the non-alternating pattern in terms of number of pairs, whereas alternating plurals are more likely to be stored than non-alternating plurals.

When these data were compared to the input to Sarah's sister Laura (see fn. 23), no important differences were found. In the Wijnen corpus, (Wijnen & Elbers 1993)[30] the input to the child Niek only contained four non-alternating noun pairs (*beesten* 'animals', *voeten* 'feet', *kanten* 'sides' and *gaten* 'holes', 10 plural tokens) and three alternating noun pairs (*handen*

[29] Here, *potlood* 'pencil' and *kleurpotlood* 'coloured pencil' are counted separately. The frequencies of two forms were summed for adjectives (e.g. *grote* 'big' and *groter* 'bigger') and past tense forms (e.g. *zette* 'put-Past' and *zetten* 'put-Past-Plur'). The pair *krab ~ krabben* is left out of the analysis: the singular *krab* 'crab' is always used as a noun but the only occurrence of *krabben* is as a verb 'scratch'.

[30] The Wijnen corpus is available through CHILDES and contains unstructured child–father interactions, with a total of 31 hours of speech directed to the child Niek (2;7-3;10).

'hands', *tanden* 'teeth' and *landen* 'countries', 28 plural tokens). When all noun pairs in the input to the three children were compared (Sarah, Laura and Niek), results show that alternating plurals are more frequent than non-alternating plurals, ranging from 60 - 70% of the total number of plural *tokens*.

Finally, the role of diminutives was considered, as these are well known to be used frequently in child-directed speech. Recall that diminutives are formed by a suffix -/(t)jə/ [cə] and always take an -*s* plural. This means that no voicing alternation occurs, e.g. *eend* [ent] ~ *eenden* [endən] 'ducks' vs. *eendje(s)* [encə] 'duckling'. Note that some plurals are more frequent in their diminutive form (e.g. in the input to Sarah, *eendjes* 'ducks' occurred 8 times whereas the plural *eenden* occurred only 3 times). To determine the relative number of diminutives, all noun pairs in the child corpus were compared to those same noun pairs in CELEX, together with their diminutive singulars and plurals. Table 12 shows the mean number of diminutives as a portion of the total number of singulars or plurals in Sarah's speech and CELEX.

*Table 13: Proportion of diminutives in child-directed speech and CELEX.*

| | **Singular** | | **Plural** | |
|---|---|---|---|---|
| | CELEX | Child-directed | CELEX | Child-directed |
| Non-alternating | 0.05 | 0.09 | 0.05 | 0.25 |
| Alternating | 0.04 | 0.26 | 0.05 | 0.20 |
| Total | 0.04 | 0.19 | 0.05 | 0.22 |

These results confirm that there are more diminutives in child-directed speech than in adult-directed speech (i.e. CELEX), for similar findings see Souman & Gillis (2007). For singulars, the number of diminutives in the child's input tends to be higher for alternating forms than for non-alternating forms.[31] Nevertheless, the increased use of diminutives in child-directed speech may have affected alternating words more than non-alternating words, which weakens the evidence for alternations. A study by Wijnen, Krikhaar & Den Os (1994) showed that individual mothers differ in the number of diminutives produced, which means that the effect may be

[31] There was an effect of Corpus (CELEX, CHILDES) ($F(1)=7{,}162$, $p<.010$), but no significant effect of Alternation (non-alternating, alternating) ($p=.097$), or Corpus x Alternation ($p=.076$).

enhanced for some children. This finding is relevant because it is possible that children benefit from exposure to all words in the inflectional paradigm (e.g. *eend*, *eenden*, *eendje* and *eendjes*) in equal measure.

As mentioned in the first chapter, there are potentially informative alternating derivational pairs, such as compounds (e.g. *hond* [t] 'dog' ~ *hondehok* [d] 'doghouse') or other inflected forms (e.g. *hand* [t] 'hand'~ *handig* [d] 'handy' or *tijd* ~ *tijdig* 'timely'), see also the examples in Chapter 1. In Table 14, an overview of such forms from the input to Sarah is presented, together with diminutive forms, in which voicing is not present (e.g. *handje* [c] 'hand-dim' and *handjes* [c] 'hands-dim').

*Table 14: Paradigmatically related forms with and without voicing in child-directed speech (CHILDES).*

| **Base** | **Diminutive [t]/[c]** | | **Related forms [d]** | |
|---|---|---|---|---|
| *bed* 'bed' | *bedje* | 1 | | |
| *hond* 'hond' | *hondje* | 53 | | |
| | *hondjes* | 4 | | |
| *hand* 'hand' | *handje* | 12 | *handig(e)* 'handy' | 2 |
| | *handjes* | 7 | *handdoek* 'towel' | 1 |
| *tand* 'tooth' | | | *tandenborstel* 'toothbrush' | 2 |
| *kind* 'child' | *kindje* | 38 | *kindertjes* 'children-dim' | 7 |
| | *kindjes* | 2 | *kinderkamer* 'child room' | 2 |
| *eend* 'duck' | *eendje* | 28 | *eendekuiken* 'duck chick' | 1 |
| | *eendjes* | 8 | | |
| *vriend* 'friend' | *vriendje* | 3 | *vriendinnetje(s)* 'fem-dim' | 12 |
| | *vriendjes* | 1 | | |
| *paard* 'horse' | *paardje* | 13 | *paardebloem* 'dandelion' | 1 |
| | *paardjes* | 3 | *paardestaart* 'ponytail' | 1 |
| Total | | 96 | | 29 |

This table shows that paradigmatically related forms with medial voicing are relatively infrequent. For the three most frequent alternating plurals (*handen* 'hands', *kinderen* 'children', *tanden* 'teeth'), there is little evidence for voicing except from plurals. Furthermore, it is unclear whether children (or even adults) detect a semantic relationship between members such as *hand* ~ *handig* 'handy' or *vriend* ~ *vriendinnetjes* 'friends-fem-dim'. It is also clear that diminutives are overall more frequent than paradigmatically related forms with voicing. However, it is unclear whether such forms influence children's knowledge of the voicing alternation in singular-plural pairs.

The next section focuses on children's *productions* of voicing alternations. After a brief introduction, Sarah's target forms will first be compared to the input (i.e. based on CELEX and CHILDES). Next, an error analysis of children's realisations (from the 'van Kampen' and 'CLPF' corpus) will be presented.

### 4.4.4 Voicing alternations in early productions

As mentioned in Chapter 1, overgeneralisations of voicing have been attested in spontaneous speech. Kager (1999a:336) reports plurals such as *pɛ[d]ən for *petten* 'caps' and *olifɑn[d]ən 'elephants' for 3-year-old Emma (based on observations by a parent). Zonneveld (1978:119 fn. 29) has observed the plural *slooien* for *sloten* 'ditches' (from singular *sloot* [slot]) in child speech, presumably from a plural with underlying /d/. Furthermore, parents often notice overgeneralisations of voicing, e.g. *[xrodə] for *grote* 'big' for Jacob between age 4;0 - 4;6 and **schatkaarden* for *schatkaarten* 'treasure maps' for Sam at 3;8.[32] Devoicing errors also occur, as in **schouderblaten* [blatən] 'shoulder blades' for *schouderbladen* [bladən] (from schouderblad [blɑt]).[33] Furthermore, spelling errors are reported to occur, see van de Hulsbeek (2005).[34]

Aside from these observations, more systematic data are available in the form of diary studies and recorded speech corpora. These data are necessary, as it is possible that parents (and linguists) tend to notice certain (unexpected) errors more than others. For instance, the 'Nina corpus' (Zonneveld 2004) contains diary notes of the spontaneous speech of a girl aged 1;6 to 4;5, as well as monthly recordings from 1;5 - 3;9. Zonneveld (2004) notes that Nina's first plurals often occur earlier than singulars, e.g. *handen* [hɑndə] occurs at 2;0 whereas *hand* [hɑnt] occurs at 2;2. Errors in which obstruents are devoiced include **paate* for *paarden* at 1;11, **hante* for

[32] I am grateful to Elma Blom who provided these data.
[33] Observed by the author, from the television programme 'Praatjesmakers'.
[34] Dutch children learn to read in school after the age of 6. Spontaneous spelling errors occur in both directions. For instance, 6-year-old Joyce writes *klatsrift* for *kladschrift* 'exercise book' and *plandjes* for *plantjes* 'plants-dim', based on the author's own observations from a child's diary.

*handen* at 2;10 (which was realised correctly at 2;0), **nijwpaate* for *nijlpaarden* at 2;4 and **schiltpatte* for *schildpadden* at 2;8. There is also some evidence for overgeneralisation of voicing in the imitation *toeda* [tudə] for *voeten* at 1;9. Zonneveld (2004:19) interprets this as intervocalic voicing because of its early occurrence.

Let us now turn to new child corpus data to study both the early realisation of alternating forms and any occurrences of overgeneralisations of voicing in more detail. First, the van Kampen corpus will be investigated further to determine whether Sarah's productions are in line with the corpus data discussed so far (i.e. from CELEX, CHILDES and van de Weijer).

#### *4.4.4.1 Productions in the van Kampen corpus (1;6-5;2)*

As described in the previous section, the input to Sarah contained a small number of highly frequent alternating word pairs. It is therefore interesting to see whether the child attempts to produce alternating words to the same extent. Furthermore, it is also relevant to determine to what extent errors or overgeneralisations occur (although results should be interpreted with caution as errors were transcribed orthographically).[35] A total of 406 bi-morphemic words (5304 tokens) produced by Sarah was extracted from the van Kampen corpus. Excluding words with liquids or obstruents preceding the final obstruent, the corpus contained 159 stems (2897 tokens). The twenty most frequent words in the 'output' (65% of tokens) are listed in (37).

[35] Data in the van Kampen corpus were transcribed by two independent transcribers, although errors reported could be an underestimation of the total number of errors.

(37)

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| 1. | **heb** 'have' | V | 697 | 11. | laat 'late' | A | 72 |
| 2. | moet 'must' | V | 594 | 12. | kapot 'broken' | A | 69 |
| 3. | weet 'know' | V | 320 | 13. | heet 'hot, is called' | A/V | 65 |
| 4. | zitten 'sit' | V | 187 | 14. | grote 'big' | A | 54 |
| 5. | **hebben** 'have' | V | 139 | 15. | opeten 'eat up' | V | 53 |
| 6. | zit 'sit' | V | 136 | 16. | **had** 'had' | V | 51 |
| 7. | eten 'eat, food' | V/N | 118 | 17. | **hond** 'dog' | N | 50 |
| 8. | **goed** 'good' | A | 117 | 18. | **bad** 'bath' | N | 47 |
| 9. | slapen 'sleep' | V | 86 | 19. | **bed** 'bed' | N | 47 |
| 10. | **vind** 'find' | V | 82 | 20. | moeten 'must' | V | 45 |

This list is almost identical to the input from the mother (see list 32), containing only one pair (*heb* ~ *hebben* 'have') and some frequent alternating singular nouns.

The probability of voiced targets in Sarah's speech will be discussed only briefly, as they include both correct and incorrect realisations. The overall proportion of voicing alternations was very similar (i.e. 40% of types vs. 38% in the input and 41% in CELEX), whereas the probability of *t~d* alternations for nouns is relatively high compared to the adult corpus (i.e. 39% vs. 34% in the input and 25% in CELEX). The distribution of targets shows the same lexical gap for *p~b* alternations (the only noun with /b/ was *krab* 'crab', no plural occurred). Finally, the nasal environment is strongest for *t~d* alternations generally (21% of types vs. 19% in the input and 29% in CELEX), and nouns in particular (22% of types vs. 18% in the input and 13% in CELEX).[36] The most frequent singular and plural targets in Sarah's speech are provided in (38) below (alternating words are bold).

[36] The distribution of mono-morphemic words shows that medial /b/ is infrequent whereas medial /d/ is more frequent than /t/. Words with /d/ constitute 63% of types and 55% of tokens, words with /nd/ represent 85% of types and 90% of tokens (see fn. 27).

(38)

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| 1. | **hond** | 'dog' | 50 | 1. | **handen** | 'hands' | 24 |
| 2. | **bad** | 'bath' | 47 | 2. | **kinderen** | 'children' | 22 |
| 3. | **bed** | 'bed' | 47 | 3. | **tanden** | ' teeth' | 14 |
| 4. | kip | 'chicken' | 44 | 4. | voeten | 'feet' | 12 |
| 5. | pop | 'doll' | 41 | 5. | apen | 'monkeys' | 9 |
| 6. | **grond** | 'ground' | 39 | 6. | ruiten | 'diamonds' | 6 |
| 7. | **kind** | 'child' | 28 | 7. | olifanten | 'elephants' | 5 |
| 8. | sap | 'juice' | 23 | 8. | schapen | 'sheep' | 4 |
| 9. | kwartet | 'quartet' | 21 | 9. | **potloden** | 'pencils' | 3 |
| 10. | aap | 'monkey' | 19 | 10. | poppen | 'dolls' | 2 |

This list shows that alternating singulars and plurals are also among the most frequent words in the child's speech. The first three plurals account for 50% of all plural tokens (60/119), which shows that alternating plurals are relatively frequent in a smaller size lexicon. However, Sarah's speech contained more non-alternating pairs, as was found for the input (i.e. 35 non-alternating and 14 alternating).[37] An overview of noun pairs is provided in Table 15 (see Table 12 for the input).

*Table 15: Number and mean frequency of noun pairs in child speech (CHILDES).*

| | **Types** | **Singular** | **Plural** | **Total** |
|---|---|---|---|---|
| Non-alternating | 13 | 11.2 (11) | 4.0 (3) | 7.6 (9) |
| Alternating | 6 | 9.1 (10) | 10.5 (11) | 9.8 (10) |
| Total | 19 | 10.6 (11) | 6.1 (7) | 8.3 (9) |

Again, there were fewer alternating pairs, while alternating plurals (but not singulars) tend to be more frequent.[38] Table 16 shows the age at which Sarah produces correct singulars or plurals for the first time, while the age in brackets is the age after which no errors occurred. The age at which a (de)voicing error occurred (i.e. *[hɑntən] 'hands') is also provided.

[37] When all rhymes are considered there were 48 non-alternating pairs and 21 alternating pairs.
[38] The six alternating pairs were produced for *handen* 'hands', *kinderen* 'children', *tanden* 'teeth', *vrienden* 'friends', *eenden* 'ducks', *schildpadden* 'turtles'), excluding *paarden* 'horses', *nijlpaarden* 'hippos' and *woorden* 'words'.

*Table 16: Age at which alternating words are produced by Sarah (CHILDES).*

| | **Tokens** | **Singular** | **Plural** | **Devoicing** |
|---|---|---|---|---|
| handen /hɑndən/ 'hands' | 24 | 2;3 | 2;4 (2;8) | 2;5 |
| kinderen /kɪndərən/ 'children' | 22 | 2;9 | 2;6 (2;9) | - |
| tanden /tɑndən/ 'teeth' | 14 | 4;9 | 2;11 | |
| woorden /wordən/ 'words' | 5 | 4;3 | 4;3 | |
| potloden /potlodən/ 'pencils' | 3 | - | 2;1 | |
| nijlpaarden /nɛilpardən/ 'hippos' | 2 | 2;9 | 2;9 | 2;9 |
| paarden /pardən/ 'horses' | 2 | 1;8 (2;4) | 3;4 | 2;10 |
| eenden /endən/ 'ducks' | 1 | 1;6 (1;7) | 2;7 | |
| vrienden /vrindən/ 'friends' | 1 | 4;11 | 3;7 | |
| landen /lɑndən/ 'countries' | 1 | - | 3;10 | |
| schildpadden /sxilpɑdən/ 'turtles' | 1 | 2;6 | 3;5 | |

These data show that the three most frequent plurals (in the nasal environment) tend to be produced early (before the age of 3), at around the same time or even before the singular form. As found by Zonneveld (2004), these first plurals may initially be realised faithfully, even though later realisations may be incorrect. Also note that some plurals appear before singulars (e.g. the plural-dominant *tanden* 'teeth'), which suggests they may be initially unanalysed. The most common error for alternating plurals was deletion of /d/. For instance, the plural *tanden* 'teeth' was first realised as *tanne* at 1;9 (note that it was mostly produced with the verb *poetsen* 'brush'). Similarly, *kinderen* 'children' was sometimes realised as *kinne* (3/22) and *handen* 'hands' was realised as *hanne* or *hamme* (5/23).[39] Deletion errors also occurred for mono-morphemic words with /nd/, e.g. *ander* 'other' is realised as *anner* or *ander* between 1;10 and 3;2 (16/28 deletions), which means that error rates may be higher than for bi-morphemic words. The same pattern of results is found for *onder* 'under' (3/9) and *onderbroek* 'underpants' (4/8), produced correctly at 2;9 and 2;10 respectively.[40] The

39 The corresponding singulars were realised correctly, except for *eet* for *eend* at 1;6. Also, the singular *paard* was realised variably as *pa, paat, paajt* or *paard* until the age of 2;4.

40 Other errors include *pannestoel* for *paddestoel* at 2;9, which is produced correctly at 2;3. Errors for /t/ occur only at the earliest stage, i.e. [fufə] for *voeten* [vutən] 'feet' at 1;6 and 1;7, but

word *vlinder* ‘butterfly’ was always realised correctly (6/6), possibly because of its late appearance (at 2;11). It thus seems that /d/ is often deleted after nasals in both bi-morphemic and mono-morphemic words, whereas plurals with /nt/ never underwent deletion of /t/ (e.g. *olifanten* ‘elephants’ appears at 2;3 (5/5 correct) and *kranten* ‘newspapers’ appears at 2;0 (2/2 correct)). Mono-morphemic words with /nt/ are mostly realised correctly, although there are few tokens.[41] The only *p~b* alternation involved the frequent verb *hebben* ‘have’, which appeared early but was variably realised (Sarah produces both **hemme* and *hebbe* until 2;5, with one occurrence of **heppe* at 1;10). Table 16 further shows that voicing errors were rare and mainly occurred for plurals with a liquid preceding /t/. An example of a devoicing error (*koute* for *koude* ‘cold’ at 2;0) is provided in 39. [42]

(39)

| | | |
|---|---|---|
| M: | want zo heb je kouwe [: koude] voeten | ‘like this you have cold feet’ |
| S: | (i)k heb niet koute [: koude] voete(n). | ‘I don’t have cold feet’ |
| M: | <(i)k heb niet koute [: koude] voeten> ["]. | ‘I don’t have cold feet’ |
| M: | je hebt wel koute [: koude] voeten. | ‘You do have cold feet’ |
| S: | nee. | No. |
| M: | koude [!] voeten. | ‘Cold (!) feet’ |
| S: | nee. | No. |
| M: | ja. | Yes. |

Here, the mother realises the form with a glide (*kouwe)* and imitates the incorrect form *koute* before she stresses the form with /d/. This is particularly strong evidence that the child does not merely imitate plurals.

here the initial obstruent is equally affected. Lohuis-Weber & Zonneveld (1996) have studied the spontaneous speech of Joost, recorded from age 1;8 to 2;11, resulting in a total of 65 15-minute sessions. They report similar errors, e.g. [pɑna] for *panda* ‘panda’ at 2;2 and [panɛne] for *kalender* ‘calendar’ at 2;5. However, *onderdoor* ‘underneath’ was realised as [ɔdəjdoj] at 2;9, indicating that /n/ may also be deleted. At that age, intervocalic /d/ was correctly realised, as in [nimnade] for *limonade* ‘lemonade’. Note that Smith (1973) also reports deletion of /d/ rather than /t/ for Amahl.

[41] Sarah produces *klinniklaas* or *klinklaas* for *sinterklaas* ‘St Nicholas’ until 2;2, but no errors occur from the age of 2;5. The word *panters* ‘panters’ appears at 2;4 (2/2 correct), and *tante* ‘aunt’ is always realised correctly (10/10) but only appears at 4;8.

[42] In the example, round brackets ( ) indicate missing parts, inserted for clarification of the target. Intended words or meanings are supplied between square brackets [ ]. Words are underlined by the author.

In sum, there is evidence for Paradigm Uniformity in these early realisations of voicing alternations (see also Zonneveld 2004, who reports consistent errors of this type for Nina). However, in the same way that /d/ is often deleted after nasals, the five voicing errors found for Sarah could reflect devoicing of /d/ on the basis of a correct target form (note that the mono-morphemic word *puzzel* 'puzzle' was produced as *pussel* at 2;4). Most importantly, voicing errors (e.g. **hanten*) occur at a relatively young age (i.e. between 1;10 and 2;10). Such voicing errors can be compared to overgeneralisations such as weakly inflected verb forms **geefte* or **geefde* for the strong verb form *geven* ~ *gaf* 'gave', which occurred at a later stage (i.e. at 4;3 and 4;9 respectively).[43] The fact that such 'morphological' errors tend to occur later suggests that Paradigm Uniformity may not always reflect the same type of productivity. Hence, errors such as *[hɑntən] could reflect an overgeneralisation on the basis of a singular with final /t/, but they could also reflect children's incorrect realisation of /d/. This is especially likely if they occur early, affecting highly frequent plurals such as /hɑndən/ 'hands'. As Sarah's data do not provide much evidence for this error type, we will return to this issue when the CLPF corpus is discussed in the next section.

Finally, even though young children might be expected to have a preference for intervocalic voicing (see Chapter 2), errors in which voicing was overgeneralised (e.g. *[olifɑndən] 'elephants') were not observed (for either Sarah, Laura or Niek). However, there is anecdotal evidence for errors of this kind, and it is possible that this tendency may manifest itself at an earlier age (or in a larger corpus). In the next section, data from younger children from the CLPF corpus will therefore be discussed, which has the added advantage of a narrow phonetic transcription.

#### *4.4.4.2 Plural realisations in the CLPF database (1;0 - 2;11)*

The CLPF database contains naturalistic longitudinal data from 12 children, who were between 1;0 and 1;11 at the first recording sessions, and between

[43] The earliest overgeneralisations of this kind involve past participles like *gekreegt* for *gekregen* 'got' at 2;5, **gekoopt* for *gekocht* 'bought' at 2;7, **opgedrinkt* for *opgedronken* 'drank up' at 2;8 and **geschrijft* for *geschreven* 'wrote' at 3;0. Thus, there is evidence for morphological analysis before the age of 3. However, such errors mostly occur later, i.e. **gespringt* for *gesprongen* 'jumped' at 3;11, **opgespringt* for *opgesprongen* 'jumped on' at 4;3, **koopten* for *kochten* 'bought' at 4;3, **sluipte* for *sloop* 'stole', 'crept' at 4;8, **zoekten* for *zochten* 'sought' at 4;8, **gebrengt* for *gebracht* 'brought' at 4;9 and **driede* for *derde* 'third' at 4;11.

1;10 and 2;11 at the last session. The corpus contains a total number of 37,180 word tokens in approximately 20,00 utterances (see Fikkert 1994; Levelt 1994, Taelman 2004). As mentioned in Chapter 4 (section 4.2.1.), van der Feest (2007) has studied children's realisations of voicing on the basis of the CLPF corpus. In medial position, children's errors mainly involve devoicing, conforming to the pattern found for initial stops (van der Feest 2007; Kager et al. to appear). It was found that 35% of medial /b/ targets were produced as /p/ (202/578), whereas the reverse was found in 3.4% of cases (44/1307). Devoicing of medial /d/ was found in 28% (81/285) of cases, whereas /d/ was only produced for only 1.5% of /t/ targets (23/1552). For instance, Robin produces [hopɔs] for *robot* /robot/ 'robot' at 2;4, and Robin produces [bodəsan] for *boterham* [botər(h)ɑm] 'slice of bread' at 1;10. In medial position, children produced more voiceless targets overall, which is in line with the input frequencies found in the Van de Weijer corpus (see Zamuner 2006). These counts are based on all voiced word-medial stops in the corpus, including those in mono-morphemic words such as *kabouter* 'gnome' and *baby* 'baby'. Van der Feest found that only 4,8% (41/863) of targets were produced in bi-morphemic words, which revealed two overgeneralisations of voicing; [habə] for /sxapən/ 'sheep' by Tirza at 2;0 and [xodə] for /xrotə/ 'big' by Leon at 2;7. Importantly, she reports that bi-morphemic words were not produced more or less accurately than the mono-morphemic words and compounds.

For the present study, all noun plurals were extracted from the CLPF corpus (which did not include any targets with medial /b/, as was found for Sarah). When the results for the two plosives are considered, children attempt more non-alternating plurals than alternating plurals (24 vs. 9 types).[44] Also, non-alternating nouns were more frequent overall (122 vs. 111 tokens). However, if only /t/ and /d/ targets are considered (12 /t/ vs. 9 /d/), /d/ was more frequent than /t/ (66 vs. 43 tokens), as was found for Sarah.[45] Alternating plural targets only occurred in the following three contexts; a. nasals (*kinderen* 'children', *tanden* 'teeth', *eenden* 'ducks, *honden* 'dogs'), b.

[44] All types were counted, i.e. both *kleurpotloden* 'coloured pencils' and *potloden* 'pencils', *nijlpaarden* 'hippos' and *paarden* 'horses', *kokosnoten* 'coconuts' and *noten* 'nuts'.

[45] Some forms that may have been intended as plurals were realised without a suffix. For instance, Catootje produces [tijɛi pa:t] as well as [te pa:ʁm] for *twee paarden* /twe pa:rdən/ 'two horses' at 2;1. Noortje produces [tɛina tant] for *kleine tanden* /klɛinə tɑndən/ 'little teeth' at 2;7. David produces [pɔtilo:] for *potloden* /pɔtlodən/ twice at 1;11. These targets were not counted as there is no schwa-like ending.

long vowels (*potloden* 'pencils', *kleurpotloden* 'coloured pencils', *broden* 'breads'), and c. the liquid /r/ (*paarden* 'horses' and *nijlpaarden* 'hippos'). Hence, most plurals occurred in the nasal context, mirroring previous results for the input (from CELEX and CHILDES).

Only the three oldest children realised voicing in alternating plurals. For instance, Enzo produced [kɪndəri] for [kɪndərən] 'children' at 1;11, and David produced [pɔtlo:lidə] for *potloden* [pɔtlodən] 'pencils' at 1;11. Note that these three children also had the highest initial vocabulary sizes (Taelman 2004:68).[46] However, these early productions were often variably realised, as noted for the van Kampen corpus in the previous section. For instance, David also produced [pɔtloʊə], [pɔtlolə] and [pɔtilo:] at that age, as well as [kupɔtlojə] and [kupɔtlolə] for *kleurpotloden* [klœrpɔtlodən] 'coloured pencils'. Enzo produces [tɑmputsi] for *tandenpoetsen* [tɑndəputsə] 'to brush (one's) teeth' at 2;1, [kɑndəputsɛ] at 2;2, and [tɑndə] for *tanden* 'teeth' at 2;3. Finally, Leon produces [kɪndəʀə] *kinderen* 'children' and *[tɑnə] for *tanden* [tɑndən] at 2;0.22, while [ɔlifɑntə] 'elephants' was produced at 2;0.10. If the child had the correct target form /tɑndə/, it is unclear why he deleted the stop in [tɑnə] but not in [kɪndəʀə] (note that the same logic applies in case the target /tɑndə/ is derived from the singular */tɑnt/). As the child should be able to form a plural [tɑntə] (as in [ɔlifɑntə] 'elephants'), the form [tɑnə] shows that he is likely to have been aware of the voiced stop in /tɑndən/. These examples show that voicing is realised correctly in some items whereas the voiced stop is deleted in others (i.e. showing item-by-item learning). Furthermore, there is considerable intra-subject variation, e.g. Leon correctly realises both [hɔndə] 'dogs' and [hɔntə] at 2;8, only to revert to [hɔndə] again in the same session. Also, he produces [paʀə], [paʀəʀən] and [paʀdər] for *paarden* /pardən/ in one session (at 2;4).

Results further show that even though more targets were voiced, only 10 out of 66 tokens with /d/ were produced correctly, while 41 out of 43 targets with /t/ were correct. To investigate this in more detail, all target obstruents were scored as either correctly realised (including realisations with aspiration) or classified according to three error types (deletions,

[46] While the mean vocabulary size was 49 (range 1-159), these three children are well above that average (i.e. Leon 92, Enzo 125 and David 159).

substitutions, or voicing errors). The resulting error analysis is shown in Table 17.[47]

*Table 17: Errors for noun plurals with /t/ and /d/ in child speech (CLPF).*

| | | Correct | Deletion | Substitution | Voicing | Total |
|---|---|---|---|---|---|---|
| /t/ | V:_ | 29 | 0 | 1 | 0 | 30 |
| | N_ | 12 | 1 | 0 | 0 | 13 |
| | R_ | 0 | 0 | 0 | 0 | 0 |
| | Total | 41 | 1 | 1 | 0 | 43 |
| /d/ | V:_ | 1 | 0 | 7 | 1 | 9 |
| | N_ | 8 | 29 | 2 | 6 | 45 |
| | R_ | 1 | 2 | 0 | 9 | 12 |
| | Total | 10 | 31 | 9 | 16 | 66 |

The first observation is that plurals in the nasal context were most often correct, indicating that these are the first plurals that children produce correctly. However, the most frequent error also occurred most for these plurals, i.e. deletion of /d/ constituted 78% (29/37) of errors in the nasal context (e.g. [tɑnə] for *tanden* by Catootje at 2;2, [ɛinə] for *eenden* by Tom at 1;6 and [parə] for *paarden* by Leon at 2;5), while 94% (29/31) of deletion errors were post-nasal. Note that this error is rare for targets with /t/, i.e. the only occurrence found in the database is [ɔfa:na] for *olifanten* [olifɑntən] 'elephants' by Tom at 1;6, at which time he also produces [ɛinə] for *eenden*. This suggests that Tom may have had a target */olifɑndən/, which would constitute the only (indirect) evidence for overgeneralisation of voicing found in the corpus. The large difference in deletions for /t/ and /d/ suggests that the voicing contrast (i.e. between /nd/ and /nt/) was perceived, although /d/ is also difficult to perceive in this environment. Hence, errors may have reflected misperception rather than deletion.

Secondly, substitutions occurred in 16% (9/56) of cases, mostly affecting targets in the long vowel context: 88% (7/8) of errors in that context are substitutions, while 78% (7/9) of substitution are in the long

[47] Non-alternating targets were also restricted to these contexts, excluding the items *ratten* 'rats' and *beesten* 'beasts'. Both items occurred only once, the latter was incorrectly realised as [epθə] (by Robin at 1;10).

vowel context.[48] For example, Tom produces [pɔloʒə] for *potloden* /pɔtlodə/ 'pencils' at 1;11, and David produces [bo:lə bɑkə] for *broden bakken* 'bake breads' at 1;11. David also produces [bo:tən bɑkən] for *broden bakken* at that age, which indicates that there is substitution rather than insertion of /l/, and that the target was /d/ rather than /t/ (leading to a devoicing error in case of [bo:tən]). These substitutions are likely to be due to "legal" glide formation in the language, as in [ro:jə] for *rode* 'red'.[49]

Third, devoicing of /d/ occurred in 29% of cases (16/56) for 7 out of 12 children, with the earliest occurrence at 1;10. As was noted for Sarah in the previous section, devoicing mostly affected targets in the liquid context, i.e. 82% (9/11) of errors in that context were devoicing errors (see Table 18), while 38% (6/16) of devoicing errors occurred after nasals.[50] An overview of devoicing errors is provided in Table 18 below, in which only the first occurrence for each child is provided.[51] The number of overgeneralisations is indicated in brackets after each name, and the age at which the first and last error is produced is also indicated, with the age at final recording indicated in brackets.

*Table 18: Devoicing errors for noun plurals (CLPF).*

| **Child** | **Age period** | **Target** | **Realisation** |
|---|---|---|---|
| Robin (1) | 1;10 (2;4) | *eenden* /endən/ 'ducks' | [e:tə] |
| David (2) | 1;11 – 2;2 (2;3) | *broden* /brodən/ 'breads' | [bo:tən] |
| Enzo (1) | 2;1 (2;6) | *nijlpaarden* /nɛilpardən/ 'hippos' | [nɛipa:tə] |
| Tirza (2) | 2;1 – 2;2 (2;6) | *paarden* /pardən/ 'horses' | [pæ:tə] |
| Catootje (1) | 2;2 (2;7) | *paarden* /pardən/ 'horses' | [pautɔ] |
| Noortje (7) | 2;7 – 2;11 (2;11) | *paarden* /pardən/ 'horses' | [pata] |
| Leon (1) | 2;8 (2;8) | *honden* /hɔndən/ 'dogs' | [hɔntə] |

[48] The only substitutions that occurred with nasals are two productions of an affricate [tantsa] for *tanden* by Noortje at 2;7, a period in which she also produces devoicing errors.
[49] The only substitution of /t/ that occurred is an affricate [potjə] for *poten* 'paws' by Jarmo at 2;1. This is likely to have been a diminutive form as it was preceded by *kleine* 'little'.
[50] Typically, *paarden* [pardən] 'horses' is realised as *[pa:tə], i.e. /r/ is not realised by any of the children.
[51] The total number of types for which overgeneralisations were produced is 6 (the only form not listed in the table is *tanden* 'teeth', realised as [tɑntɑ] by Noortje at 2;7).

As stated before, these errors could be taken to indicate that plurals have been formed productively, on the basis of the singular form with /t/. However, it is also possible that the child's voiced target form was devoiced in the course of production. After all, van der Feest (2007) shows that medial /d/ was devoiced in 28% of cases (on the basis of the same corpus, taking into account both mono-morphemic and bi-morphemic forms). The fact that stops are more likely to be deleted or substituted when they are underlyingly voiced in the adult form suggests that a voiced stop may be present in the child's form. This would mean that the low number of correctly voiced plurals (15%) partly reflects the fact that voicing is articulatorily more difficult, rather than the fact that children do not know that these plural forms have a voiced medial obstruent. This also means that devoicing errors (e.g. [hɔntə] 'dogs') need not necessarily reflect Paradigm Uniformity, as devoicing could be a strategy to deal with voiced obstruents. The fact that devoicing mostly occurred after liquids while deletion mostly occurred after nasals suggests that phonological (or phonetic) factors play a role, which may be rooted in perception and/or production.[52] Hence, the CLPF data show that 'devoicing' errors occur until age 2;11 (the final age of recording), while results from the van Kampen corpus (recorded until 5;2) suggest that such errors are not likely to occur after the age of 3 (Sarah's 'devoicing' errors occurred until 2;10 and deletion of /d/ occurred until 3;2, for both mono-morphemic and bi-morphemic words).

In sum, children's earliest productions of plurals show that underlying /d/ is treated differently from underlying /t/. In combination with the fact that alternating targets have a higher frequency than non-alternating targets, this suggests that children may initially store these plurals and delete or devoice /d/ in the course of production. Alternatively, it is possible that plurals with /d/ are derived from a singular with an abstract underlying final /d/, which would undergo suffixation and subsequent deletion. As the empirical evidence is inconclusive, considerations of abstractness might lead one to favour the first scenario, in which the plural is retrieved from the lexicon. Devoicing errors may also reflect productive

[52] A cluster such as /nd/ is characterised by continuous voicing, and the timing of the onset of orality and the release of the obstruent is subtle. In case of /nt/, there is a discontinuity in voicing that coincides with the consonant transition. Neu (1980) presents evidence that there is more /t/-deletion in English after alveolar fricatives and nasals ([s], [z], [n]) than after non-alveolars, see also van Hout (1989) for Dutch.

overgeneralisations on the basis of a singular. As this is most likely to happen to words whose plurals are of low frequency, differences in devoicing rates could also reflect effects of word frequency, accounting for the effects of item-by-item learning. Recall that plurals in the nasal context are very frequent, which might account for the fact that devoicing errors occur more often in the liquid context. Frequency also seems to interact with 'age of acquisition', as deletions of /d/ mostly occur for frequent plurals, which tend to appear early. Hence, several factors are likely to play a role in the early production of voicing alternations, related to perception and production of the voicing contrast as well as lexical frequency.

Finally, only one voicing error (i.e. /p/ transcribed as [b]) for noun plurals was found in the CLPF database (i.e. *[sxabə] 'sheep' at 2;0), and the only evidence for voicing of /t/ is the early occurrence of *[ɔfa:na] for *olifanten* [olifɑntən] 'elephants' at 1;6. Also, Leon produced the adjective *[xodə] for *grote* 'big' at 2;7, together with correct non-alternating plurals (e.g. [ɔlifɑntə] at 2;0). Similarly, van der Feest (2007) shows that voicing errors (medial /d/ for /t/) were only produced in 1.5% of cases. Taken together, there is little evidence for a rule of intervocalic or postnasal voicing of 'underlying' voiceless obstruents applying across-the-board. However, it is possible that elicitation experiments with non-words may reveal a different pattern. This possibility will be examined further in the next chapter.

## 4.5 Summary and discussion

In this chapter, the frequency of voicing alternations has been explored, argued to be important for morpho-phonological acquisition. For instance, analogical models predict that the productivity of a pattern is related to its type frequency, whereas high token frequency is connected to irregularity and storage. Even though it is a matter of debate whether voicing alternations are regular or irregular, many alternating plurals co-occur with vowel alternations or the irregular suffix *-eren,* suggesting that voicing alternations are somewhere on a continuum. Furthermore, corpus data show that alternating words tend to have a higher token frequency and lower type frequency than non-alternating words. It was argued that frequent plurals such as *tanden* 'teeth' and *handen* 'hands' are likely to be stored as wholes. Moreover, alternating singulars were shown to be relatively frequent in relation to the plural form. This indicates that the connection between singulars and plurals may be less strong in case of alternating words, which is relevant in analogical models that stress the paradigmatic connection between words (e.g. Bybee 2001, see also Corbett et al. 2001). As noted in Chapter 2, Bybee assumes that when a form is more frequent, the relation to its base (or uninflected form) is weaker. Thus, it is possible that even though some alternating plurals are frequent, singulars and plurals are more 'independent' or autonomous. This ties in with the finding that the *relative* frequency of inflected words and their bases is important in morphological processing (Hay 2001, Baayen 2003, Hay & Baayen 2005). For instance, complex forms which are more frequent than their bases (as in *government,* with base *govern*) are less likely to be parsed than complex forms which are less frequent than their bases (e.g. *discernment*, with base *discern*). This is potentially relevant to the current research, as the relative frequency of the singular stem in relation to the plural form could affect the ability to form generalisations on the basis of a pair.

The CELEX data show that alternating stems constitute around 40% of types and 60% of tokens in the relevant contexts (those used for the Wug-test stimuli), although the probability of voicing for nouns is weaker (25% of types). When all rhyme types are taken into account, the probability of voicing is somewhat lower (29%). This replicates earlier findings by Ernestus & Baayen (2003, 2004), who report an overall probability of voicing of 25% for /t/. On the other hand, there is very little evidence for *p~b* alternations overall, except for the highly frequent pair *heb ~ hebben* 'have'. In case of a novel word, children need to infer the lexical representation from a neutralised singular. Although Ernestus & Baayen (2003) found that adults are sensitive to factors such as the quality of the final obstruent, it is not clear whether the same applies to children. For instance, voicing alternations for fricatives are common, e.g. there is a probability of 99% for *f ~ v* alternations after long vowels, as in *neef ~ neven* 'nephews' (Ernestus & Baayen 2004). Hence, the possibility that children may generalise from fricatives to plosives such as /p/ cannot be excluded.

A corpus of child-directed speech (the 'van Kampen' corpus) was shown to contain a similar number of alternating types (e.g. 33% of nouns). Alternating plurals were among the most frequent words in the input to Sarah, but the input contained twice as many non-alternating pairs than alternating pairs. Child-directed speech was found to contain more diminutives than CELEX, which may affect alternating pairs more than non-alternating pairs. Results concerning the occurrence of alternating forms in the input may be an overestimation, as some words are variably realised. As noted in Chapter 4, word-final devoicing does not always apply (e.g. in the combination *heb ik* [hɛbɪk] 'have I'). Furthermore, final obstruents are not always realised (as in *ik vin* for *ik vind* 'I think'), and there is an optional process of glide formation for medial voiced obstruents in Dutch (CELEX data indicate that a third of alternating types have possible glides).

In conclusion, it is not clear whether there is sufficient evidence for voicing alternations in the input to the child, taking into account the low number of alternating pairs and the small number of highly frequent alternating plurals (such as *handen* 'hands' and *kinderen* 'children'). Rather than constituting evidence for an alternating pattern, the fact that these plurals are frequent makes them more likely to be stored as unanalysed wholes. Moreover, these frequent plurals are most likely to undergo reduction (i.e. deletion of /d/). Such frequency effects on production have been found, e.g. high-frequency words tend to be produced with shorter durations and more segmental reduction (e.g. Whalen 1991). Also, Bybee (2000, 2002b) found that the rate of deletion of final /t/ and /d/ in regular English past tense verbs was higher for high-frequency verbs than for low-frequency verbs. It is likely that this process may also affect Dutch plurals in the input, as frequent words in casual speech show both assimilation and reduction (Ernestus 2000, Ernestus et al. 2004). Similarly, highly frequent adjectives are most likely to undergo weakening (i.e. /d/ produced as /j/ or /w/). Generally, the input contained up to twice as many non-alternating pairs as alternating pairs. Assuming that these corpora are representative of the input a child is exposed to, children predominantly hear non-alternating forms in the ambient language. This means that there may not be enough evidence for the child to extract a pattern, or to relate the process of final neutralisation to morpho-phonological alternations (see Chapter 6). Even if a pattern can be extracted, it is not clear whether the child would extend it to novel words, as the overall type frequency of alternations is lower than 50% (see Chapter 5).

Previous data on the acquisition of the Dutch plural suffix indicates that plurals are formed productively by Dutch children around the age of 2. The data discussed in this chapter do not reveal whether children have productive knowledge of voicing alternations. This is because singulars and plurals that are correctly produced could have been processed without any morphological analysis, reflecting storage of either form. Similarly, devoicing errors need not reflect Paradigm Uniformity, but may reflect children's difficulty to produce voicing. For instance, Sarah's data show that /nd/

clusters are difficult until she is around three years of age, whereas /nt/ cluster are acquired earlier. Realisations from younger children based on the CLPF corpus (1;0 - 2;11) show that /d/ after nasals is often deleted in plurals (e.g. [hɔnə] for /hɔndən/ 'hands').[53] This result could be taken to support the presence of a voicing contrast, but it could also reflect the fact that [d] is difficult to perceive in this environment. The results further indicate that error rates are similar or higher for mono-morphemic words, as was found by van der Feest (2007). Thus, errors such as [hɑnə] for /handən/ 'hands' co-occur with errors such as [ɑnər] for /ɑndər/ 'other'. Note however that child corpora contained few mono-morphemic targets (the realisation of bi-morphemic and mono-morphemic words will be discussed further in Chapter 5). Furthermore, the CLPF corpus showed a devoicing rate of 29% for /d/ in noun plurals, which is similar to the overall percentage of devoicing for medial /d/ found by van der Feest (2007) on the basis of the same corpus (28%). Note that this is a much higher percentage than the mean overgeneralisation rate reported for spontaneous speech in the literature (which is around 4 to 10%). The latter rate is more in line with the sporadic occurrence of truly morphological overgeneralisations such as **geefte* for *gaf* 'gave', which are typically found at a later stage. Taken together, this suggests that errors such as *[bɛtən] 'beds' may reflect both morphological and phonological factors (i.e. the difficulty of producing voiced obstruents).

The corpus data discussed in this chapter are relevant for testing the different expectations that may be formulated on the basis of phonological and analogical models. The distribution of voicing was found to be influenced by phonological context or rhyme, e.g. /t/ is most likely to be voiced after nasals (e.g. *hand* ~ *handen* 'hands'). Both phonological and analogical models might predict that overgeneralisations and alternations for non-words in a Wug-test will predominantly be found in this context (e.g. for the novel form [flɑnt]). Hence, postnasal voicing is not only phonetically natural and typologically common, it is also supported in the target language. However, there is a 'lexical gap' for /b/ after long vowels and nasals, which does not reflect functional factors related to production or perception. Importantly, if children are sensitive to effects of lexical analogy, they are not expected to show overgeneralisations of voicing in these

[53] Note that deletion of /d/ after nasals is a regular phonological process in Afrikaans, resulting in pairs such as *hond* ~ *hone* 'dogs' (Wissing 1971).

environments. For example, when a non-word like [bɛmp] is presented to the child in a Wug-test, the expected plural is [bɛmpən] rather than [bɛmbən], as no alternating words occur in this environment. On the other hand, if children are sensitive to purely phonological factors, voicing may occur in 'lexical gap' environments such as postnasal /p/, as natural classes and markedness are expected to play a role (e.g. leading to alternations for both [flɑnt]) and [bɛmp]). Models in which lexical frequency plays a role do not predict generalisations to be formed on the basis of words with an extremely high token frequency (e.g. Bybee 2001). As most of the evidence for /b/ stems from the single verbal pair [hɛp] ~ [hɛbən] 'to have', analogical overgeneralisations for /b/ are not likely to occur. More generally, the low type frequency and high token frequency of alternating plurals predicts that the pattern will not be very productive.

These predictions will be discussed in more detail in the next chapter, when the elicitation experiments are presented. Even though there is little evidence for overgeneralisations of voicing in spontaneous speech, elicitation of plurals for non-words may reveal a different picture, as children do not know the voicing specification of novel stems.

# Chapter 5 Production Experiments

This chapter discusses experimental results concerning Dutch children's productive knowledge of voicing alternations. First, previous experimental studies on the acquisition of morphology will be discussed in section 5.1, focusing on the nominal plural. Results of two classic 'Wug-test' studies involving the elicitation of plurals of existing words and non-words with children will be discussed in section 5.2. These studies are aimed at determining the role of *Paradigm Uniformity* in children's word formation, and investigate whether children are influenced by phonological or analogical generalisations when they form novel plurals. Experiment I (5.2.3) involves children in three different age groups (3-, 5- and 7-year-olds).[1] Experiment II (5.2.4) involves a group of 5-year-olds diagnosed with Specific Language Impairment (SLI) and an age-matched control group.[2] Finally, Experiment III (5.2.5) involves a group of adult controls. A general discussion of the results will be provided in section 5.3.

---

[1] An earlier version of this work has appeared as Kerkhoff (2004). Acquisition of Voicing alternations. In: S. Baauw and N. J. van Kampen (eds.). *Proceedings of GALA 2003*. Utrecht: LOT, 269-280.

[2] An earlier version of this work has appeared as Kerkhoff & de Bree (2005). Acquisition of Morpho-phonology in Children with Specific Language Impairment and Typically Developing Children. In: A. Kerkhoff et al. (eds.). *UiL-OTS Yearbook 2004*. Utrecht: LOT, 37-51.

## 5.1 Introduction

In the previous chapter, we have seen that Dutch children acquiring voicing alternations may produce two types of errors in spontaneous speech. First, errors such as *[bɛtən] for *bedden* /bɛdən/ 'beds' may reflect children's difficulty to produce a voicing contrast. However, previous findings indicate that Dutch children have acquired the medial voicing contrast by the age of three. Such errors could also reflect a strategy of *Paradigm Uniformity*, since a singular form (e.g. /bɛt/) always ends in a voiceless obstruent. As described in the previous chapter, 'devoicing' errors in a Dutch corpus of plurals (e.g. *[hontən] 'dogs') reached 29%. Elicitation studies using existing words generally indicate a higher rate of overgeneralisations than observed in corpus studies (e.g. Clahsen et al. 1992).

Secondly, children may occasionally produce forms such as *[pɛdən] for /pɛtən/ 'caps', reflecting a process of intervocalic voicing. Such errors may be due to an early rule or constraint, but may also reflect analogy to other words (e.g. /bɛdən/ 'beds'). Corpus research (see Chapter 4) showed that such errors are rare in spontaneous speech. The elicitation experiments in this chapter will provide more information on children's knowledge of existing plurals, as well as on the productivity of the alternating pattern in case of non-words.[3] Generally, children make fewer errors on existing words compared to non-words, due to effects of lexical storage. For this reason, the elicitation experiment takes the form of a 'Wug-test' using non-words (Berko 1958), which is more informative about children's strategies for word formation.

The Wug-test has traditionally been used to study children's productive knowledge of morpho-phonology (e.g. Berko 1958). The most frequently studied area is the English past tense, but other languages include Hungarian (MacWhinney 1978), Spanish (Kernan & Blount 1966, Bybee & Pardo 1981), Tagalog (Zuraw 2000), Dutch (van Wijk 2007) and artificial languages involving novel suffixes (Tessier 2006). Previous studies have

[3] Neijt et al. (2006) found some evidence for postnasal and intervocalic voicing in Dutch, asking children who were learning to write to complete words (e.g. write 't' or 'd' in *olifan_en* 'elephants'). In this experiment, children used incorrect 'd' (as in **olifanden*) more often than incorrect 't' (as in **liefte* 'love'). These findings could be partly due to hypercorrection, because children have to learn that /t/ is sometimes spelled as 'd' (as in *hond* 'dog').

found that children are capable of inflecting novel roots around the age of three (e.g. Berko 1958, Kernan & Blount 1966, Bryant & Anisfeld 1969, Bybee & Slobin 1982, Derwing & Baker 1980, Marcus et al. 1992). As described in the previous chapters, frequency (and the related notions of regularity and transparency) and phonological naturalness are predicted to play a role in the acquisition of morpho-phonology. For instance, Hungarian final vowel lengthening was applied productively to non-words as early as 2;6 (MacWhinney 1978). Although MacWhinney cites this as an example of an early alternation that is not necessarily in accord with 'natural predispositions', final lengthening appears to be phonetically natural (e.g. Edwards, Beckman & Fletcher 1991, Myers & Hansen to appear). In the first 'Wug-test' study, Berko (1958) asked children (ages 4 to 7) to form plurals from novel singulars such as "wug". Results show that even though the English suffix alternation (*cats* vs. *dogs*) is acquired early, the third alternant (i.e. /ɨz/ in *horses*) is not acquired until around age six. Similarly, adding the past tense suffix [ɨd] to a verb base that already ends in /t,d/ begins around age six, and is established for some clusters by age eight for nonce words (Derwing & Baker 1980, Bybee & Slobin 1982). Previous research has found moderate to strong correlations between vowel insertion in plural and vowel insertion in the past tense (MacWhinney 1978:73). In Berko's experiment, the only stem alternation occurred for the non-word *heaf*, which was pluralised as *heafs* and *heaves* (analogous to *leaf* ~ *leaves*). The Spanish plural suffix -*s* shows a similar process of insertion as that of English, with -*es* appearing after consonants and -*s* after vowels. Kernan & Blount (1966) show that children (from 5 to 10) had difficulty with the -*es* suffix, which was produced correctly less than half the time for non-words (i.e. the suffix was deleted or produced as -*s*). This shows that children are not always very proficient when they are confronted with novel words, and may not have internalised certain morpho-phonological patterns, even though they may be regular. Furthermore, children may be able to inflect novel words even though they have difficulties with allomorph selection.

In MacWhinney's early model of the acquisition of morpho-phonology, rote-learned forms or *amalgams* are distinguished from *combinations* (i.e. forms derived by rule) and the use of *analogy*. In MacWhinney's (1978) study of German plural formation, children in three age groups (3-4, 4;2-6 and 11-12) were asked to inflect words and non-words. Note that MacWhinney (1987) restricts the term analogy to immediate 'priming' effects, such as when the plural of *scarf* is elicited before the non-word *narf*, which could trigger *narves*. This priming effect as well as an interaction between age and priming was found to be nearly significant in the German study. Thus, older children were found to be slightly more sensitive to priming when forming analogous plurals (i.e. plurals that are formed with the same suffix and umlauting as the previous word plural). However, MacWhinney (1978) found no overall priming effect in his study with Hungarian children. In this study, evidence for a strategy of analogy was found for both 2-year-olds and 6-year-olds, but not 3- and 4-year-olds (who were claimed to have a more rule-based approach). MacWhinney concludes that analogy is available as a strategy at all ages but that its use depends on the structure of the language (1978:69). However, the criteria for deciding whether the child is using combination (rules) or analogy are not clearly defined.[4] As described in Chapter 2, both source-oriented and product-oriented processes may account for children's productivity, and there is no direct evidence that adults or children actually manipulate rules and rule symbols (e.g. MacWhinney et al. 1989).

There are no previous experimental studies that have specifically addressed the acquisition of voicing alternations. In MacWhinney's German study results for alternating plurals are not reported separately.[5] However, it seems that there were two cases of zero marking (for *Pferd* and *Norb*), whereas the other items were inflected with *-e*. Since plurals were only checked for affixes, it is unclear whether voicing errors occurred (e.g. **Pferte*). As MacWhinney acknowledges, these errors are likely to have gone unnoticed, since the experimenter was not tuned to listen to correct realisation of voicing. However, there is one reported case of a voicing overgeneralisation (*Nerden*), which occurred even though the non-word *Nerd* was not in a primed position (i.e. it was not elicited after *Pferd*).[6] It is possible that the other three non-words did not lead to overgeneralisations

---

[4] MacWhinney (1978) states that "when the child produces correct real forms but cannot generate nonce forms there is evidence for *rote*, when the child produces erroneous real forms and regularised nonce forms there is evidence for *combination* (usually with somewhat longer response latencies), and when the child produces correct real forms and nonce forms of many different shapes there is evidence for *analogy* (increased by priming)."

[5] Items included four voicing alternations: *Hand* ~ *Hände* 'hands', *Korb* ~ *Körbe* 'baskets', *Kind* ~ *Kinder* 'children' and *Pferd* ~ *Pferde* 'horses'. Note that two of these plurals have an accompanying vowel alternation, rendering them less transparent. It is not clear whether MacWhinney (1978) considers German plurals with voicing alternations (*Hund* – *Hunde* 'dog') as regular. Corresponding non-words were *Gand, Norb, Dind* and *Nerd*.

[6] These data were provided by Brian MacWhinney, p.c.

of voicing because their real word analogs have either vowel alternations or an *-er* suffix. The fact that voicing alternations have not received much attention in previous experimental studies is an important empirical gap, since voicing alternations are considered to be among the clearest cases of phonological alternations, and are well studied in the phonological literature. Before we turn to experiments eliciting Dutch voicing alternations, the next section will briefly discuss previous experimental findings on the acquisition of Dutch plural formation.

### 5.1.1 Acquisition of the Dutch plural

As discussed in previous chapters, there are two regular plural suffixes in Dutch, *-en* and *-s*, with a distribution that is largely determined by phonological factors. However, there are exceptions. Corpus studies indicate that children use the appropriate plural suffix in spontaneous speech by age 4 (see Chapter 4). In experimental studies of acquisition of the Dutch plural, it is often assumed that *-en* is the 'elsewhere' suffix, whereas the child needs to acquire a rule to apply *-s* after a schwa and sonorant consonant (e.g. den Os & Harder 1987, Snow et al. 1980). Snow et al. (1980) investigated a group of first and second language learners of Dutch using both words and non-words, and conclude that 7-year-olds do not show fully adult-like performance on diminutive and plural formation. In a study with children aged between 4 and 12, den Os & Harder (1987) also found that errors occurred for both suffixes, although *-en* was used correctly before *-s* was. From the age of 8, existing words with *-en* (*tent* 'tent' and *vis* 'fish') were produced without errors, whereas errors for *-s* plurals (*lepel* 'spoon' and *toren* 'tower') continued until the age of 10. The higher error rate for *-s* seems to be due to the item *toren* 'tower', which was often left uninflected as the singular stem already resembles a plural in *-en*.[7] Den Os & Harder also elicited novel words, which were mostly inflected with *-en*. From age 8, children produced *-en* plurals for more than 90% of novel words (e.g. *praan*,

[7] This phenomenon is similar to 'backformations' that are reported in the literature, reflecting a tendency to avoid double-marking (called 'affix checking' in MacWhinney 1978, see also Bernhardt & Stemberger 1998:475). In Dutch, singulars such as *varken* 'pig' or *jongen* 'boy' (which take an *-s* plural) are sometimes mis-analysed by children as plurals, leading to non-existent singulars such as **vark* and **jong* (see also Snow et al. 1980:546). Snow et al. and van Wijk 2007 find a similar effect for non-words ending in *-s*, e.g. *keps* was uninflected or *-s* was deleted before adding *-en*.

*zos*). This percentage is also high for 4- and 6-year-olds (average 74%), although it drops for 7-year-olds (55%). For non-words that require *-s* (*gapel*, *poden*), the 90% correct level was reached only at 11, while 4- to 6-year-olds applied *-s* in around 25% of cases.

These results suggest that relatively young children (aged 4-7) may not have fully acquired plural formation in Dutch, due to the fact that there are two competing affixes. A study by van Wijk (2007) was specifically aimed at finding out whether children use a default suffix, as predicted by Pinker's Words-and-Rules model (and possibly a 'double' default as suggested by Pinker 1999:231). In a Wug-test, non-word plural responses were considered 'pure' overgeneralisations when subjects used a suffix which was not predicted for a certain non-word on the basis of either phonological factors, or type frequency of highly similar neighbours (e.g. /woey/ ~ /woeys/). Surprisingly, the number of such overgeneralisations increased for 3-, 4- and 5-year olds. As van Wijk takes these overgeneralisations to reflect 'default' use, she concludes that children's use of a default increased rather than decreased with age, contrary to expectations of the Words-and-Rules model. In contrast, the number of 'phonological' overgeneralisations (i.e. those that seem motivated by a phonological factor such as sonority) decreased with age. Moreover, whereas 3-year-olds use mainly *-en* in pure overgeneralisations (in 80-100% of cases), older children and adults seem to shift to *-s* (in around 60% of cases). Van Wijk further shows that frequency of use has an effect on overgeneralisations, since both affixes show a similar 'relative overgeneralisation rate'.[8] Hence, the increase with age of *-s* in overgeneralisations coincides with a general increase in the use of *-s* plurals. In sum, the claim that children search for a unique default affix is not supported, and van Wijk proposes a model for Dutch in which the terms 'regular' and 'default' are not synonymous. Recall that according to the Words-and-Rules model, Dutch *-en* plurals (i.e. both non-alternating and alternating plurals) could be considered *irregular* if *-s* is the default suffix. We can conclude that such a scenario is not supported by experimental evidence (van Wijk 2007). In the present study, non-words are expected to be inflected with *-en*, the suffix that appears after obstruents and is presumably acquired by the youngest children.

---

[8] This rate is calculated by dividing the number of overgeneralizations of an affix by the number of productions with that affix (a method taken from Dabrowska 2001).

The first elicitation experiment will be discussed in the next section, after a description of the experimental set-up and hypotheses.

## 5.2 Elicitation experiments: Wug-test

The elicitation experiments of this dissertation were set up to access Dutch children's knowledge of existing alternating plurals as well as their productive knowledge of the voicing alternation. As stated before, the underlying voice specification of the final segment of a singular stem can only be deduced on the basis of alternations, which poses a problem for acquisition. If only the singular's surface form is known, the plural may contain either a voiced or an unvoiced consonant. To derive the phonological properties of a plural on the basis of a singular, children may assume several strategies. Firstly, a child might always opt for a non-alternating form when presented with a novel form (e.g. [kɛt] ~ [kɛtən]). Alternatively, the choice for one of the possible realisations ([kɛtən] or [kɛdən]) might be random, based on some phonological generalisation of the stem-context or on the characteristics of phonologically similar words in the lexicon (to be discussed further below).

As discussed earlier, the distribution of alternating forms is not entirely unpredictable. Previous results by Ernestus & Baayen (2001) suggested that the probability that a neutralised obstruent is underlyingly voiced is governed by a number of factors, such as the quality of the stem-final obstruent and the preceding vowel. The corpus results presented in Chapter 4 also show that alternations are more frequent in some segmental environments. For instance, an analysis of CELEX and child-directed speech revealed that *t~d* alternations are most frequent in the nasal environment (e.g. *handen* /hɑndən/ 'hands'). The corpus analysis also showed that alternating words are associated with a high token frequency. Furthermore, Dutch corpora show that there is a 'lexical gap' for /b/ after long vowels and sonorant consonants (such as nasals). However, there are also very few alternating words with /b/ after short vowels. The word and non-word items used in the present study were derived from the same six phonological environments that were used in the corpus study (see section 4.3), to test predictions of phonological models vs. analogical models. In the next section, the method and non-word stimuli will be presented first, after which the hypotheses are discussed further.

### 5.2.1 Method

A set of phonotactically legal non-words was created for the Wug-test, consisting of 24 monosyllabic items in the relevant phonological environments. In this way, it was possible to determine the effect of rhyme (V_, V:_, N_) and final obstruent (P, T). The non-words were chosen in such a way that neither of the two possible 'inflected' forms was an existing Dutch word. Most items resembled existing words if one phoneme was altered, e.g. *ket* [kɛt] has neighbours such as *pet* [pɛt] 'cap', *bed* [bɛt] 'bed' and *kat* [kɑt] 'cat'). This could not have been avoided, as monosyllabic words generally occupy a densely populated area in the lexicon (i.e. with many neighbours). The items did not include 'minimal' pairs (e.g. *taap* and *taat*) because this was thought to be too confusing for young children. The full list of non-word items is shown in Appendix I. In each phonological environment, there were two words with exactly the same rhyme (e.g. /jɪt/ and /mɪt/). This was done to test whether children would be consistent in a certain phonological context. Phonological models generally predict that children are able to group together items that end in a certain rhyme (such as a vowel, nasal and alveolar obstruent), using natural classes. However, in case children would not be consistent in a certain rhyme category (containing four items), two items with the same rhyme (e.g. /ɪt/) would be expected to be treated the same. In models that employ analogy to existing words, differences between such same-rhyme items could arise due the probabilistic nature of word formation processes. This prediction will be discussed in more detail when the results are reported (in section 6.2.4.4). In every pair of same-rhyme non-words, one item was chosen to resemble an existing mono-morphemic form with a voiced medial obstruent (i.e. the final item for each category). For example, the non-word *jit* [jɪt] is paired with the non-word *mit* [mɪt], for which an alternation would result in the existing word *midden* [mɪdən] 'middle'. This manipulation was mainly devised for the adult control experiment, to see whether alternating forms could be 'provoked' by such items. Note that items in the lexical gap environments were also chosen to resemble mono-morphemic forms (e.g. *gloop* [xlop] was chosen because of the existing word *globe* [xlobə] 'globe'). However, children are not expected to be sensitive to this manipulation as the mono-morphemic words are low-frequent and typically 'learned'. This topic will be discussed further when the results for the adult controls are presented in section 6.4.

### 5.2.2 Hypotheses

Let us now turn to what can be predicted to happen if a child is confronted with a novel form (e.g. [kɛt]) and is asked to form a plural. A first hypothesis is that Paradigm Uniformity or stem-to-stem faithfulness may lead children to match the singular's voicing value in the plural, producing non-alternations only (e.g. [kɛt] ~ [kɛtən]). Strict Paradigm Uniformity such as that advanced by Hayes (2004) predicts that voicing alternations for non-words (e.g. [kɛt] ~ [kɛdən]) or errors such as *[pɛdən] 'caps' will not occur at all in children's productions. A weaker version of this prediction is that voicing alternations will occur less often than would be expected on the basis of the general likelihood that words alternate in the language. Dutch corpus data (see Chapter 4), do no provide much evidence for errors such as *[pɛdən] 'caps'. However, the scenario may be different for non-words, as their plural form cannot have been memorised. Generally, younger children are expected to rely on Paradigm Uniformity more often than older children (see Chapter 2). This should also be apparent from their behaviour on alternating words, as they should produce more plural errors such as *[bɛtən] 'beds' than older children.

When alternations are produced for novel singulars, a logical possibility is that they will be produced randomly across phonological environment (i.e. by *random selection*). This would mean that the type of final obstruent or rhyme (i.e. the phonological environments chosen for the non-word stimuli) will not affect children's behaviour.

A third hypothesis is that children will be guided by *phonological generalisations* reflecting phonological rules or constraints, which may be phonetically grounded. As discussed earlier, rule- or constraint-based models assume that children's phonology reflects natural classes and cross-linguistic markedness rather than lexical statistics. As discussed in previous chapters, postnasal voicing and intersonorant or intervocalic voicing occur both in languages of the world and child language. Under this view, voicing is predicted to occur in intervocalic position or after nasals. Such a phonological preference is expected to affect /b/ more than /d/, as it is stronger both universally (Ohala 1983:195) and for Dutch (based on assimilation data, as noted by Zonneveld 1983, see Chapter 3). Languages such as Arabic even have /b/ but not /p/ (Ladefoged & Maddieson 1996), although there are also languages that lack voiced obstruents altogether (see Lombardi 1991). Note that /b/ is also more likely to occur on the basis of phonetic factors, as voicing is easier to produce in /b/ than /d/ (see van Alphen 2004 for Dutch). As shown by van der Feest (2007), Dutch children acquire /b/ before /d/, and /p/ was more often erroneously voiced than /t/. This suggests that phonological factors might be grounded in phonetics, as discussed in Chapter 2. Furthermore, long vowels are expected to favour voicing more than short vowels, which is often explained as a compensation in duration between the vowel and consonant (e.g. for English, see Hayes 1989). Importantly, such a strategy of phonological generalisation should result in a different distribution of voicing alternations than the distribution actually found in the ambient language. For instance, non-word alternations could occur more often for /b/, (e.g. [dɑp] ~ [dɑbən] > [slɑt] ~ [slɑdən]). Moreover, postnasal or intervocalic voicing could lead to alternations in 'lexical gap' environments (e.g. [dep] ~ [debən]), even though there is no evidence for such alternations in the child's input. Importantly, natural classes such as long vowels and nasals are expected to pattern together. Traditional rule-based models posit deterministic phonological grammars, which produce the same output for a given input. Such a model would predict that children's behaviour is consistent within a particular phonological environment (e.g. alternations are produced after nasals and long vowels). Moreover, items with the same rhyme should be treated the same (e.g. [jɪt] = [mɪt]). A related prediction is that children's behaviour should be stable within a short period of time. However, as discussed in Chapter 2, more recent models have incorporated probabilistic rules or constraints, which can handle variation. Thirdly, knowledge of existing plurals is not predicted to affect (early) phonological overgeneralisations of voicing in non-words. Finally, voicing alternations for non-words and errors for words (e.g. *[pɛdən] 'caps') are expected to decrease with age, as children learn that Dutch has no process of intervocalic or postnasal voicing.

A fourth hypothesis is that children rely on an *analogical generalisation*, reflecting the frequency of existing patterns in their lexicon. This view entails that children are sensitive to the distribution of voicing alternations in the input, and word-formation proceeds according to similarity- or exemplar-based analogy. First, the fact that alternations are associated with a high token frequency and low type frequency would predict that the pattern may not be very productive. However, an analogy-based strategy could result in overgeneralisations of voicing for words or non-words. The distribution of alternating forms would be expected to mirror the distribution of these forms in Dutch. Here, type rather than token frequency of similar words is expected to influence children's behaviour. As discussed in Chapter 2, analogical or usage-based models that predict an influence of frequency assume that type frequency (or the number of times a particular pattern is encountered) determines productivity, whereas items with high token frequency should be more entrenched or autonomous.

Note that phonological and analogical models predict the same outcome in a number of cases. For instance, both phonological and analogical generalisations predict more alternations after long vowels than after short vowels (e.g. [fedən] > [kɛdən]). Hence, alternations after long vowels are phonetically and phonologically preferred, as well as more frequent in the input. Interestingly however, in other cases, analogical generalisations would result in the opposite pattern of results predicted by phonological models. More specifically, under analogy, voicing alternations are expected to occur for /t/ rather than /p/, and are expected to be most frequent after nasals, mirroring the input (e.g. [jɔndən] > [kɛdən]). Moreover, it is expected that alternations will not occur in lexical gap environments (e.g. *[debən]). A second prediction in analogical models is that overgeneralisations reflect children's lexical knowledge. This means that overgeneralisations of voicing may *increase* with age rather than decrease, under the assumption that children's knowledge of alternating plurals increases with age. Generally, analogical models would predict that children's performance on alternating words correlates with the frequency of existing plural forms, as plurals are stored in the mental lexicon.

To test these predictions, a Wug-test (Berko 1958) was carried out for this dissertation, in which children were asked to pluralise the set of non-words described above (see Appendix I). Furthermore, plurals of a set of existing words were also elicited. The first experiment will be described in the next section.

### 5.2.3 Elicitation Experiment I: age groups

#### *5.2.3.1 Subjects*

A total of 60 children (36 girls and 24 boys) participated in the experiment. Two boys (aged 3;2 and 3;9) failed to complete the test and their results were consequently left out of the analysis. The remaining 58 subjects can be divided into three age groups. The first age group will be referred to as '3-year-olds' and consisted of 24 children (15 girls, 9 boys) who attended day care. They had an average age of 3;4, with ages ranging from 2;9 to 3;11. The group of '5-year-olds' consisted of 19 children (8 girls and 11 boys), with an average age of 5;2, and ages ranging from 4;0 to 6;2. The group of '7-year-olds' consisted of 15 children (10 girls and 5 boys), with an average age of 7;2, and ages ranging from 6;9 - 7;8. The latter two groups were in separate elementary school classes. Subjects were tested in an isolated room in their school.

### *5.2.3.2 Materials*

The set of non-words consisted of the 24 items in the relevant phonological environments (see Table 19 for examples and Appendix I for the complete set of non-words).

*Table 19: Non-words in six phonological environments.*

| | **T** | **P** |
|---|---|---|
| V_ | /kɛt/ | /tɛp/ |
| V:_ | /klat/ | /tap/ |
| N_ | /flɑnt/ | /bɛmp/ |

The set of words consisted of 8 non-alternating singulars and 8 alternating singulars in the same contexts (see Table 20 for examples).

*Table 20: Words in six phonological environments.*

| | **T** | | **P** | |
|---|---|---|---|---|
| | Non-alternations | Alternations | Non-alternations | Alternations |
| V_ | /pɛt/ 'cap' | /bɛd/ 'bed' | /kɪp/ 'chicken' | /wɛb/ 'web' |
| V:_ | /vut/ 'foot' | /hud/ 'hat' | /ap/ 'monkey' | * |
| N_ | /olifɑnt/ 'elephant' | /hɑnd/ 'hand' | /lɑmp/ 'lamp' | * |

There were only two words with /b/ (*web* 'web' and *krab* 'crab'), as there are no alternating words in the other two rhyme contexts (i.e. after long vowels and nasals). Moreover, apart from the frequent verb *heb* ~ *hebben* 'have', these two forms are the only two alternating /b/ plurals that children are likely to know. Generally, the test-words were chosen because they were familiar to children (at least in their singular form) and easy to depict.[9] In Appendix I, the set of words is shown together with the CELEX frequency of the singular and plural form and Age of Acquisition ratings.

The average (CELEX) frequency of the set of alternating words was higher than that of the set of non-alternating words, both for singulars (136 vs. 23) and plurals (56 vs. 24). However, this was mainly due to the high frequency of the items *handen* 'hands' and *honden* 'dogs'. In contrast, the *median* frequency of plurals of alternating test-words was lower (3 vs. 10), even though the median frequency of singulars was higher (21 vs. 14). This result was also found in Chapter 4, and indicates that there is a relationship between frequency and alternations. Hence, the frequency of word items could not be matched, and a t-test revealed that alternating and non-alternating words differ significantly in the token frequency of the singular ($t(14)=-1.37$, $p=0.13$). However, words did not differ in the frequency of the plural form ($t(14)=-0.66$, n.s.). The frequency of the test-words and its relation to the results are discussed further below.

The set of stimuli included 6 mono-morphemic words with voiceless medial stops (e.g. *panter* 'panter', *appel* 'apple') and 6 mono-morphemic words with voiced medial stops (e.g. *ladder* 'ladder', *Dribbel* 'Spot'), see Appendix I. These were included to check whether children would produce a voice contrast for mono-morphemic forms, in case they did not produce one for bi-morphemic forms.

A set of 8 additional filler items was added to the list of items (e.g. *vis* 'fish', *ballon* 'balloon', *tafel* 'table'). These items were both mono- and bi-syllabic and had different final consonants, with some requiring an *-s* plural. Two filler items were existing irregularly formed plurals (*koe ~ koeien* 'cows' and *ei ~ eieren* 'eggs').[10] The fillers were added to prevent children from noticing the alternating plurals and using them as direct analogs for the non-words. It was also done to make the task more varied, as test-items all end in /t/ or /p/. A set of 10 additional filler items was added for older children (e.g. *paraplu* 'umbrella', *leeuw* 'lion', *vliegtuig* 'airplane'). For the youngest children, the number of test-items was more restricted due to time

---

[9] Available AoA 'Age of Acquisition' ratings for singulars (Ghyselinck, de Moor & Brysbaert 2000) and ratings from the *streeflijst voor 4- tot 6-jarigen* 'target list vocabulary for 4- to 6-year-olds' (Damhuis et al. 1992) are provided (see Appendix I for more information). The only word that did not occur in the *streeflijst* was *krab* 'crab'.

[10] There is a small number of 'irregular' plurals with glide insertion in Dutch, among which *koe ~ koeien* /ku/ ~ /kujən/ 'cows' (see Zonneveld 1978).

limitations (caused by a lower attention span, apparent from previous literature and a pilot test).

A pilot session with six children included more word items, to test whether the pictures sufficed to elicit plurals. After the pilot study some pictures were changed and some words were removed.[11] There were also some changes in non-word test items, in order to create two items with the same rhyme (including one 'high analogy' item), as described above. The results of the pilot study were included in the overall results, as the changed items had similar rhymes.[12]

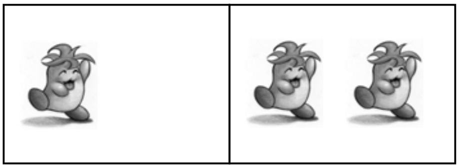
The final test thus contained a total of at least 60 stimuli, depending on the age group. The list of items was constructed in such a way that non-words were separated by fillers or words ending in a different obstruent. Also, items ending in a certain consonant (/t/ or /p/) were not placed directly after one another, and alternating forms did not directly precede non-words. The lists were presented to subjects in two opposite orders, to control for effects of shyness and familiarity with the test routine (for the initial items) and fatigue or boredom (for the final items).

#### *5.2.3.3 Procedure*

To elicit plurals of non-words, pictures of fantasy animals were presented to the child in a PowerPoint slide show. A recorded version of each non-word was inserted, to ensure that all subjects heard the same stimulus. To this end, a recording was made of a female speaker who read the list of non-words aloud, making sure there was an audible release (the non-words were always spelled with *t* or *p*).[13] The stimuli were recorded in a soundproof room by means of a DAT-recorder Aiwa HD S100 and a Sony microphone ECM MS957. The recordings were stored as .wav files (sample rate: 48 KHz) on a

[11] Four words were removed from the test (*brood* 'bread', *pad* 'toad', *kat* 'cat' and *krant* 'newspaper'), as most children responded with a related word (i.e. *boterham* 'slice of bread', '*kikker*' 'frog', *poes* 'cat' and *boek* 'book'). For 21 subjects in the older age groups the word *olifant* 'elephant' was replaced by *tent* 'tent', because the former often lead to diminutive formation. Finally, *hemd* [hɛmt] 'shirt' was often transformed it into [hɛmpen], which changed the coda.
[12] For non-words, *plamp* was replaced by *fomp*, *trep* by *tep* (because of its likeness to *trap* 'stairs'), *tuip* by *boop*, *fot* by *mit*, *ment* by *gint*, and *mip* by *zwap* (in order to obtain non-words with an identical rhyme to one of the other non-words). The only overgeneralization that occurred for these forms was [trɛp] ~ [trɛbən] by Femke (3;4), which was added to the total number of overgeneralisations for the non-word *tep* [tɛp].
[13] I am grateful to Maya van Rossum for data recording.

laptop by means of the speech analysis package *Praat* (Boersma & Weenink 1996). Upon presentation of the first picture, the pre-recorded non-word was played at least twice, in a sentence context provided by the experimenter ("*Dit is een* …", "*This is a* …"). The child was always encouraged to repeat the stimulus. A second (identical) picture then appeared on the screen, upon which the experimenter would prompt the child to form a plural ("*Nu zijn er twee. Er zijn nu twee* ..", "*Now there are two. There are two…?"*). An example of a non-word picture used is given below.[14]



To elicit plurals of existing words, children were shown pictures in the same way and encouraged to first name the stimulus. For both word and non-word stimuli, the child was encouraged to repeat each singular item until it was correctly produced. The test sessions were tape-recorded, and the plural forms supplied by the child were transcribed later. When two or more responses were given, the final one was included in the analysis.

All stimuli were transcribed by the experimenter and a portion of the data was rated by five independent transcribers. These included responses that were deemed difficult to judge, which is why the youngest children are overrepresented in the sample, as well as 'lexical gap alternations'. In total, 80 stimuli were judged by the five listeners, who were asked to indicate "Voiceless", "Voiced" or "Don't know" for each sound file (recorded in *Praat*, .wav format, Sample rate 22.05 kHz). The transcribers were blind to the original transcription, although the target item was presented in its singular form. If stimuli that were originally transcribed as voiced were rated as voiced by fewer than four out of six transcribers, the score was changed.

### *5.2.3.4 Results*

In this section, the results of the auditory analysis will be discussed. First, the mono-morphemic words were transcribed, to ascertain whether children

[14] Taken from the *neopets* website, see also van Wijk (2007).

were capable of making a voicing contrast. As children were not very familiar with most of the mono-morphemic words with voiced medial obstruents, there were not enough tokens to compare mono-morphemic and bi-morphemic words systematically. However, most children produced /d/ and /b/ correctly in the words *vlinder* /vlɪndər/ 'butterfly' and *Dribbel* /drɪbəl/ 'Spot'. As was found for the corpus data, deletions of /d/ occurred for some of the youngest children (e.g. [vlɪnər] by Emma 2;9 and Rubie 3;11). Devoicing errors occurred for /d/ (e.g. [vlɪntər] by Rutger 3;8 and Lisanne 4;1, [mɔtər] for *modder* /mɔdər/ 'mud' by Femke 3;4) as well as /b/ ([drɪpəl] by Noortje 3;5). Errors in which medial obstruents were voiced did not occur in the data (e.g. *[ɑbəl] for *appel* /ɑpəl/ 'apple'). A separate experiment comparing bi-morphemic and mono-morphemic words more systematically was carried out (reported in Zamuner et al. 2006b), results of which will be discussed at the end of this section.

All plural word responses were transcribed separately, and scored as either "Correct" or "Incorrect" (a voiceless or voiced realisation of a medial obstruent, depending on whether the word was alternating or not). Plural non-word responses were scored as "Non-alternating" or "Alternating", corresponding to a voiceless or voiced realisation of the medial obstruent. Items with voicing errors in onsets were not excluded, e.g. [bɛmp] ~ [pɛmpə] by Bente (3;5) or [dɪnt] ~ [tɪndə] by Helge (5;4).[15] Similarly, items with a changed vowel were included as long as the rhyme category was the same (e.g. [jɪt] ~ [jɛtə] by Sammy 4;9), even when the resulting form may have been intended as an existing word (e.g. [jɪt] ~ [wɪtə] 'white' by Jeffrey 4;4).

Other possible responses included: (i) Unmarked plural forms or bare stems, i.e. a repetition of the singular (e.g. [jɪt]), (ii) a stem change or change in the nucleus or coda (e.g. [jɪt] ~ [jɪpə]), (iii) formation of *-s* plurals (e.g. [jɪt] ~ [jɪts]), (iv) missing responses because the child did not respond at all, created a diminutive form (e.g. [jɪt] ~ [jɪtjəs]) or substituted the item with another word (e.g. [jɪtvoxəls]). The results will be discussed separately for words and non-words.

---

[15] Children's plural realisations are mostly transcribed without final *-n*, which was present for only a minority of (mostly older) children (who would sometimes clearly pronounce final -ən, stressing the suffix). Recall that pronouncing final *-n* is optional in Dutch.

**Words**

First, the results for non-alternating words are summarised for all age groups in Table 21 (with the two most relevant responses shaded).

*Table 21: Results for non-alternating words in numbers (%), Exp. I.*

| | 3-yr-olds | 5-yr-olds | 7-yr-olds | All |
|---|---|---|---|---|
| Correct [pɛtən] | 152 (71.1%) | 122 (80.3%) | 118 (98.3%) | 392 (81.0%) |
| Voicing error *[pɛdən] | 6 (2.8%) | 7 (4.6%) | 2 (1.7%) | 15 (3.1%) |
| Bare stem | 22 (10.4%) | 6 (3.9%) | 0 | 28 (5.8%) |
| Stem change | 0 | 0 | 0 | 0 |
| S-plural | 0 | 6 (4%) | 0 | 6 (1.2%) |
| Missing | 32 (15.1%) | 11 (7.2%) | 0 | 43 (8.9%) |

The data show that overgeneralisations of the type *[pɛdən] 'caps' occur in 3.1% of cases (15/484). Recall that van der Feest (2007) observed a similar percentage of medial voicing errors for words in the CLPF database, as discussed in Chapter 4. However, the spontaneous voicing errors observed by van der Feest occurred for /b/ (3.4%) rather than /d/ (1.5%), whereas the present voicing errors mostly occurred for /d/. In fact, the only /b/ error was observed in a 'lexical gap' environment, by a child in the youngest age group (i.e. *[abə] *monkeys* by Femke at 3;4). Voicing errors were mostly produced for the item *[pɛdən] 'caps' (8/58), followed by *[vudən] 'feet' (3/58), *[tɛndən] 'tents' (2/42) and *[olifɑndən] 'elephants' (1/36). This pattern of results shows that errors are not likely to be influenced by phonological factors, as most errors were produced for /t/ after short vowels.[16] Moreover, the CELEX frequency of plurals (see Appendix I) indicates that the plurals of these words are associated with the lowest frequency (e.g. *petten* 'caps' has a plural frequency of 2), indicating that children are most likely to produce such errors if they are not familiar with the plural. However, this is not always the case, as *voeten* 'feet' is associated with high plural frequency (i.e. 129), suggesting that phonological factors also play a role.

To determine effects of age on children's performance on non-alternating words, the variable Group (3-year-olds, 5-year-olds, 7-year-olds)

[16] It is important to note that this error often co-occurred with the opposite error, e.g. Chiara (4;2) produces both [hɔndən] 'dogs', *[tɛndən] 'tents' and *[hɑnten] 'hands'.

was entered in a Univariate ANOVA, which showed a main effect on score correct ($F_1(2)=7{,}512$, $p<.001$, $\eta_p^2=.215$). Tukey Post-hoc tests revealed that the 3-year-olds and the 5-year-olds differed from the 7-year-olds ($p<.001$ and $p=.043$ respectively), although the two youngest age groups did not differ from each other. The data in Table 21 show that the younger children's poorer performance was due to a higher number of missing responses (including diminutives) and bare stems.[17] Importantly, overgeneralisations of voicing (e.g. *[pɛdən] 'caps') were produced in all groups, by a total of ten children (three 3-year-olds, five 5-year-olds and two 7-year-olds). The 5-year-olds produced most overgeneralisations, which indicates that these errors do not necessarily reflect an early voicing rule. The results for alternating words are summarised in Table 22 below.

*Table 22: Results for alternating words in numbers (%), Exp. I.*

| | **3-yr-olds** | **5-yr-olds** | **7-yr-olds** | **All** |
|---|---|---|---|---|
| Correct [bɛdən] | 52 (27.1%) | 64 (42.1%) | 69 (57.5%) | 185 (39.9%) |
| Voicing error *[bɛtən] | 81 (42.2%) | 63 (41.4%) | 49 (40.8%) | 193 (41.6%) |
| Bare stem | 19 (9.9%) | 5 (3.3%) | 0 | 24 (5.2%) |
| Stem change | 5 (2.6%) | 0 | 0 | 5 (1.1%) |
| S-plural | 0 | 4 (2.6%) | 1 (0.8%) | 5 (1.1%) |
| Missing | 35 (18.2%) | 16 (10.5%) | 1 (0.8%) | 52 (11.2%) |

This table shows that the overall performance on alternating plurals is much lower than on non-alternating plurals (40% vs. 81%). Even 7-year-olds only reach a 58% correct score on alternating words, indicating that children were not very familiar with these plurals (even though they are the most frequent alternating words, and children were familiar with the singulars). Regularisations of the type *[bɛtən] *beds* occurred in 42% of all cases, persisting into the oldest age group.

Again, the results were entered in an ANOVA, which indicated a main effect of Group on score correct ($F_1(2)=8.182$, $p<.001$, $\eta_p^2=.126$). This difference was due to the fact that 3-year-olds produced fewer correct plurals compared to the 7-year-olds (Tukey, $p<.001$). This difference was not due to

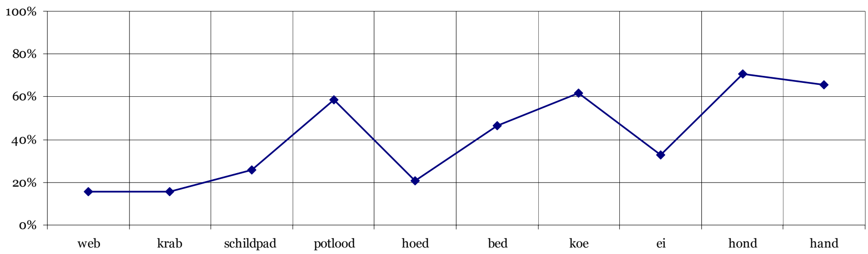
[17] Note that only 5-year-olds produced unexpected *-s* plurals for existing words, e.g. *[aps] 'monkeys' by Maxime (6;1) and *[pɛts] 'caps' by Helge (5;4). However, it is sometimes difficult to distinguish a /t/ with a loud release from /ts/, which occurred for the youngest age group.

children's overgeneralisation rate (e.g. the production of errors such as *[bɛtən] 'beds'), which is similar across age groups, i.e. around 40%. The youngest children did produce more bare stems and missing responses (replicating an effect frequently found in the literature). Also, only the youngest children produce stem changes for alternating words, involving deletions of /d/ in five cases (i.e. [hɑnə] 'hands' by Emma (2;9), who also produced *[hɔntə] 'dogs', and both [hɑnə] 'hands' and [hɔnə] 'dogs' by Robin (3;4) and Esmée (3;9)). Note that this error occurred twice for the mono-morphemic word *vlinder* 'butterfly' (i.e. [vlɪnər], while devoicing errors *[vlɪntər] also occurred twice. As was found in the corpus studies described in Chapter 4, errors in which /t/ was deleted never occurred. Indeed, there were no stem change errors at all for non-alternating words. As was found for non-alternating words, -*s* plurals were only produced by two 5-year-olds.

The overall results show that alternating plurals were correct in 40% of cases, whereas non-alternating plurals were correct in 82% of cases. To compare performance on alternating and non-alternating words, the independent variable Alternation (non-alternating, alternating) and Group (3-year-olds, 5-year-olds, 7-year-olds) were entered in a multivariate ANOVA. This analysis shows a main effect of Group ($F_1(2)=15.415$, $p<.001$, $\eta_p^2=.22$, $F_2(2)=8.975$, $p<.001$, $\eta_p^2=.30$), reflecting the fact that older children performed better than younger children. There was also a main effect of Alternation ($F_1(1)=97.243$, $p<.001$, $\eta_p^2=.47$, $F_2(1)=56.777$, $p<.001$, $\eta_p^2=.72$), which means that alternating words were more difficult than non-alternating words. There was no interaction between Group and Alternation, which means that alternating words were more difficult than non-alternating words for all groups. It thus seems that children are poorer at inflecting alternating words, even though their singulars are associated with a higher token frequency. This indicates that children may not have known the alternating plural form.

In Figure 2 below, performance on existing plurals (including the irregulars *koe* /ku/ ~ /kujən/ 'cows' and *ei* /ei/ ~ /eijərən/ 'eggs') is shown, with words ordered according to the CELEX frequency of the plural form (see Appendix I).

*Figure 2: Performance correct on alternating and irregular plurals (Exp. I).*




| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| -Vb | -Vb | -Vd | -V:d | -V:d | -Vd | irr | irr | -Nd | -Nd |

Generally, this figure shows children's performance improves as the frequency of alternating plurals increases. Alternating plurals with /d/ after nasals were most often correct. This was to be expected on the basis of the corpus data (Chapter 4), which revealed that these plurals are among the most frequent in the input. When results of the three groups are combined, performance ranges from 16% correct for low frequent *krabben* 'crabs' to 71% correct for *honden* 'dogs' (the three age groups scored 50%, 84% and 87% correct respectively). In comparison, children performed well on non-alternating plurals (ranging from 64% for *olifanten* 'elephants' to 93% for *kippen* 'chickens'). The three alternating plurals that received the highest scores were *potloden* 'pencils', *handen* 'hands' and *honden* 'dogs'. Both *krabben* 'crabs' and *webben* 'webs' have a plural frequency of zero, while *handen* 'hands' and *honden* 'dogs' have the highest frequency (see Appendix I). While the plural *potloden* 'pencils' has a lower CELEX frequency, it is clearly a word that is well-known to children (see also appendix I), and it was frequent in the input to Sarah (Chapter 5).

To investigate the relation between performance and frequency, non-parametric correlations were performed. Results show that there was no significant correlation between children's performance correct and the CELEX frequency of singulars (p=.156), whereas there was a moderate correlation with the frequency of plurals (ρ=.635, p=.008). This was due to a strong correlation between performance on alternating words and the frequency of the plural (ρ=.836, p=.010), whereas there was no correlation between performance on non-alternating words and the frequency of the plural (p=.69).[18] The frequency effect found for alternating plurals suggests that they are likely to be retrieved as wholes, which is expected if they are treated as irregulars (as suggested in Chapter 5).

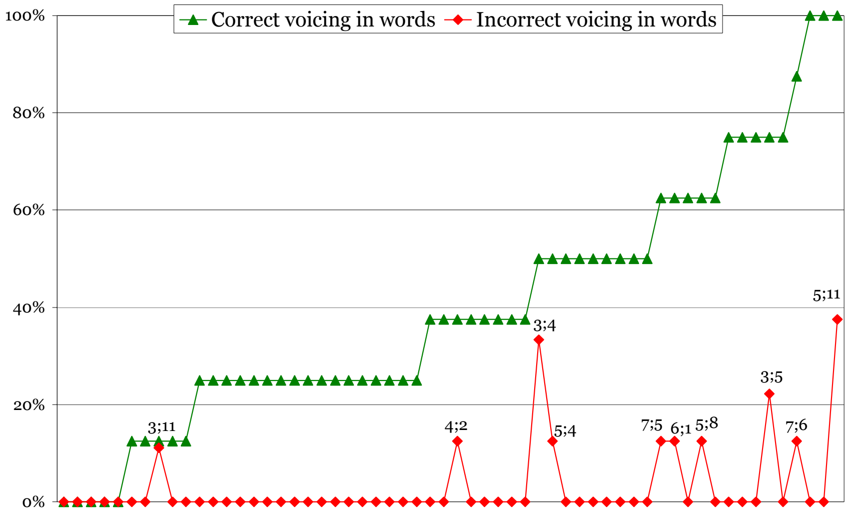
Finally, Figure 2 shows that performance on irregular *koeien* 'cows' is better than on most of the alternating plurals (62% correct), which indicates that frequency is more important than irregularity for this form. On the other hand, children performed worse on the irregular plural *eieren* /eijərən/ 'eggs' than expected on the basis of its frequency (33% correct), producing regularised [eije]. Hence, the irregular suffix *-eren* presumably affected children's performance. This result could be taken as an indication that alternating plurals occupy a position in between regular and irregular forms, as argued in Chapter 4.

In sum, elicitation of plurals of existing words shows that children have not fully acquired existing alternating plurals yet, and that the frequency of the plural form plays a role.

As was found in the corpus study, overgeneralisations of voicing (e.g. *[pɛdən] 'caps') are rare, and produced more often by 5-year-olds than 3-year-olds (see Table 22). Also, errors were produced for words with /t/, with the exception of one 'lexical gap' word [ap] ~ [abə] 'monkeys'. The remaining fourteen errors were made for words with /t/, i.e. *[pɛdən] 'caps' (14%), *[olifɑndən] 'elephants' (5%) and *[vudən] 'feet' (5%). The relationship between correct voicing in words (e.g. [bɛdən] 'beds') and incorrect voicing in words (e.g. *[pɛdən] 'caps') is shown in Figure 3 (results are ranked according to performance correct on words, and the age of the child is indicated only when alternations were produced).

[18] There was no significant correlation between performance and the frequency of plurals in the input to Sarah (taken from the van Kampen corpus), which reveals that the input to a particular child is less reliable than (adult) CELEX data (ρ = .560, p=.074).

*Figure 3: Correct and incorrect voicing in words (Exp. I).*




This figure shows that errors such as *[pɛdən] ‘caps’ do not only occur for the youngest children. Instead, error rates increase as children’s performance on alternating words improves. This suggests that errors are linked to children’s familiarity with alternating plurals. Results for non-words are discussed in the next section.

### Non-words

Non-words were scored as ‘Non-alternating’ or ‘Alternating’ (i.e. plurals with a voiceless or voiced medial obstruent) or belonging to one of four other response categories (bare stems, stem changes, -*s* plurals and missing responses), see Table 23 below.

*Table 23: Results for non-words in numbers (%), Exp. I.*

| | 3-yr-olds | 5-yr-olds | 7-yr-olds | All |
|---|---|---|---|---|
| Non-alternation | 374 (64.9%) | 336 (73.7%) | 345 (95.8%) | 1055 (75.8%) |
| Alternation | 20 (3.5%) | 16 (3.5%) | 7 (1.9%) | 43 (3.1%) |
| Bare stems | 86 (14.9%) | 29 (6.4%) | 1 (0.3%) | 116 (8.3%) |
| Stem change | 4 (0.7%) | 4 (0.9%) | 1 (0.3%) | 9 (0.6%) |
| S-plurals | 0 | 23 (5.0%) | 0 | 23 (1.7%) |
| Missing | 92 (16.0%) | 48 (10.5%) | 6 (1.7%) | 146 (10.5%) |

This table shows that most non-words were inflected without voicing alternations. The percentage of non-alternating forms increases with age, up to 95.8% for the 7-year-olds. However, around a third (31%) of all children (18/58, 11 girls and 7 boys) produced at least one voicing alternation for non-words (e.g. [kɛt] ~ [kɛdən]), in a total of 3% of cases or 4% of inflected *-en* plurals.[19] In total, 43 alternations were produced for 16 (out of 24) non-words, ranging from 1 to 6 per non-word. Also note that the overall percentage of alternations for non-words (3.1%) is the same overall percentage that was found for voicing errors in words (e.g. *[pɛdən] 'caps').

Children produced only few stem changes, involving mainly substitutions, i.e. [slɑt] ~ [sɑpə] or [dɪnt] ~ [dɪŋkə] by Lisanne (4;1), including instances in which the non-word may have been substituted for an existing word (e.g. [tap] ~ [takə] 'tasks' by Rabie (5;9).[20] As found for words, only 5-year-olds produced *-s* plurals for non-words, which is in line with earlier finding (e.g. van Wijk 2007), which show that this affix appears late.[21] Still, the *-s* plural is unexpected given the lexical neighbours of these non-words (i.e. monosyllables ending in plosives). Interestingly, some 5-year-olds had a tendency to produce mainly *-s* plurals for non-words, and many *-s* plurals for existing words.[22] Furthermore, some children devoiced the initial consonant

[19] The number of children producing alternations is similar for the three age groups (29% of 3-year-olds or 7/24, 32% of 5-year-olds or 6/19 and 33% of 7-year-olds or 5/15).

[20] Children sometimes remarked on the fact that non-words resembled words (e.g. *klaten* and *praten* 'talk', *jont* and *hond* 'dog', *zoot* and *zout* 'salt', *jit* and *wit* 'white'), including words with different final obstruents (e.g. *slat* and *slak* 'snail', *gop* and *god* 'god'). Lieke (3;9) remarks that *flant* resembles *klanten* 'customers' but produces [flɑndə]. Chiara (4;2) remarks that *jont* resembles *hond* but produces [jɔnte] (and [hɔndə] 'dogs').

[21] One doubly marked plural was produced for a non-word, i.e. [dɪnt] ~ [dɪntsə] by Rachel (4;0).

[22] Maxime (6;1) produced only *-s* for non-words, some *-s* plurals for words (e.g. **schaaps* 'sheep', **hoeds* 'hats'), as well as some correct *–en* plurals (e.g. *kippen* 'chickens', *honden* 'dogs',

and produced a voiced medial, e.g. [dɪnt] ~ [tɪndə] by Femke (3;4) and Helge (5;4).[23]

To further determine age effects, the independent variable Group (3-year-olds, 5-year-olds, 7-year-olds) was entered in a Univariate ANOVA, with the number of non-alternating forms (i.e. the default form [kɛtən]) as dependent variable. Results show a main effect on the number of non-alternating forms ($F_1(2)=7{,}648$, $p<.001$, $\eta_p^2 = .22$), indicating that older children produced more non-alternating forms than younger children (which is mostly due to the fact that young children produce more bare stems). Tukey post-hoc tests show that both 3-year-olds and 5-year-olds differed from the 7-year-olds (p=.001 and p=.028 respectively). Hence, it is not the case that young children rely more on a strategy of Paradigm Uniformity. The difference is likely to be due to the number of missing responses and bare stems produced by younger children.[24] Finally, the number of voicing alternations for non-words (i.e. [kɛdən]) is generally low, although the two youngest age groups produced more alternations than the 7-year-olds. Importantly however, results show that the 3-year-olds did not produce more alternations than the 5-year-olds. The number of alternations produced per child ranges from 1 to 7 per child. One could argue that children who only produced an alternation once (n = 9) do not show rule-like behaviour. Conversely, if we take the production of two or more alternations as evidence for rule-like behaviour, such behaviour is found at all ages, or at least up until the age of six.

Children were not found to be consistent within a certain rhyme. For instance, [dɑp] ~ [dɑbə] was produced by Emma (2;9), Robin (6;1) and Amy (7;5), whereas [zwɑp] ~ [zwɑbə] was produced by Rubie (3;11), Njana (5;11) and Eline (7;6). This result is not compatible with models that would predict

---

*potloden* 'pencils'). For the youngest children, it was sometimes difficult to distinguish an –*s* plural from an unmarked plural (due to the fact that final /t/ may be produced with extra release) or diminutives (e.g. [kɛtjə]) from palatalisation of /t/.

[23] To determine whether the onset of a non-word may have influenced the number of alternations, three pairs of non-words were compared which differed in the voicing value of the onset. In all cases, the item with a voiced onset /d/ was more often subject to a voicing alternation than the item with a voiceless onset /t/ (i.e. *deep* > *taap*, *dint* > *gint*, *dap* > *tep*), which reveals a possible effect of the onset (see also Kager et al. to appear). However, no alternations occurred for items with initial /b/ (e.g. *boop*, *bemp*).

[24] The production of unmarked forms or singulars seems to be partly an effect of shyness and unfamiliarity with the task, because they tend to occur at the beginning of the test sessions rather than the end.

across-the-board behaviour and are more in line with probabilistic models. However, it is possible that evidence for 'rule-like' behaviour is found within a certain rhyme category or phonological environment. For instance, 7 out of 18 children produced more than one alternation in a certain phonological environment, while three children produced alternations in 3 out of 4 non-words. These three children are shown in Table 24, together with the results for two additional children who produced non-word alternations as well as overgeneralisations of voicing in words (in a specific environment).

*Table 24: Alternations for non-words and overgeneralisations for words in a specific phonological environment (Exp. I).*

| Child | Age | Environment | Productions |
|---|---|---|---|
| Femke (F) | 3;4 | N_/t/ | [dɪndə], [flɑndə], [jɔndə]<br>*[krɑndə] 'newspapers'[25] |
| Noortje (F) | 3;5 | V_/t/ | [jidə], [slɑdə]<br>*[pɛdə] 'caps' |
| Helge (M) | 5;4 | N_/t/ | [dɪndə], [flɑndə], [jɔndə]<br>*[tɛndə ]'tents' |
| Maarten (M) | 5;8 | V_/t/ | [kɛdə], [slɑdə]<br>*[pɛdə] 'caps' |
| Njana (F) | 5;11 | N_/t/ | [flɑndə], [jɔndə]<br>*[olifɑndə] 'elephants' |
| | | V_/t/ | [kɛdə], [mɪdə], [slɑdə]<br>*[pɛdə] 'caps' |

It thus seems that these children show rule-like behaviour for these two environments (/t/ after short vowels and nasals), which may be due to an effect of *phonological generalisation*. However, three of these children are past age five, which means that the effect is not 'early'. Also, the environment in which most of these alternations are produced is also the environment for which most alternations are predicted to occur on the basis of Dutch noun plurals (i.e. /d/ after nasals). Alternations for /d/ after short vowels were also produced by three children, but two of them were nearly six years old. Moreover, these children also produced [t] for non-words in these

[25] This item was replaced by *olifant* 'elephant' in the second test version.

environments, which is unexpected under categorical, rule-based accounts. For instance, Maarten produced non-alternating forms for the majority of non-word items, including [jɪtən] and [mɪtən]. Likewise, Njana produced [dɪntən], [xɪntən] and [jɪtən]. The only 3-year-old who produced relatively many alternations for /t/ after short vowels is the youngest child, Noortje (3;5), who also produced [kɑtə] 'cats' and [kɛte]. Hence, if her behaviour reflects an early rule of intervocalic voicing, it was not applied across-the-board. Moreover, phonological rules are expected to favour voicing after long vowels rather than short vowels.

Let us now turn to the distribution of the alternations that were produced for non-words (e.g. [kɛdən]). The distribution across phonological environments is shown for each age group in Table 25 below (lexical gap environments are indicated by shading).

*Table 25: Distribution of non-word alternations in numbers (%), Exp. I.*

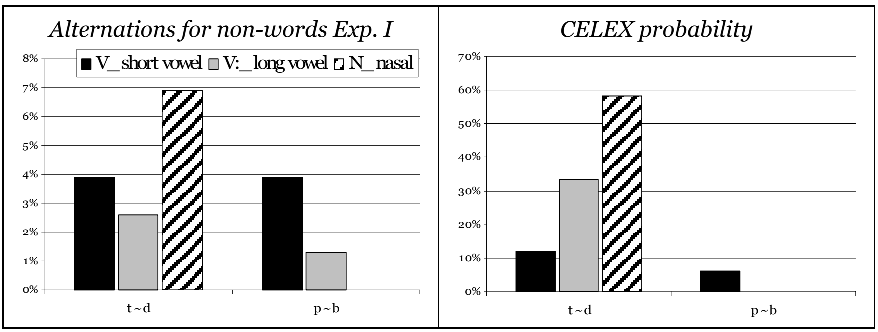
| | **3-year-olds** | | **5-year-olds** | | **7-year-olds** | |
|---|---|---|---|---|---|---|
| | t ~ d | p ~ b | t ~ d | p ~ b | t ~ d | p ~ b |
| V_ | 4 (4.2%) | 4 (4.2%) | 5 (6.6%) | 3 (3.9%) | 0 | 2 (3.3%) |
| V:_ | 4 (4.2%) | 2 (2.1%) | 1 (1.3%) | 1 (1.3%) | 1 (1.7%) | 0 |
| N_ | 6 (6.3%) | 0 | 6 (7.9%) | 0 | 4 (6.7%) | 0 |
| Total | 14 (4.9%) | 6 (2.1%) | 12 (5.3%) | 4 (1.8%) | 5 (2.8%) | 2 (1.1%) |

These results show that children produce more alternations for /t/ than for /p/, and more alternations after short vowels than long vowels. Also, children are sensitive to the fact that *p~b* alternations do not occur after nasals. However, 'lexical gap' alternations were produced after long vowels, i.e. Noortje (3;5) produces [debə] and [tabə], and Sven (4;0) produces [debə], which are not expected on the basis of the input.

An ANOVA was carried out to determine whether alternations occurred randomly (i.e. by *random selection*). Independent variables were Final Obstruent (P, T) and Rhyme (V_, V:_, N_), constituting the six phonological environments under consideration. A Univariate ANOVA with overall percentage of voiced responses as dependent variable yielded an effect of Final Obstruent ($F(1)=8{,}004$, $p=.011$, $\eta_p^2=.31$), reflecting the fact that *t~d* alternations were more frequent than *p~b* alternations. Moreover, an interaction between Final Obstruent and Rhyme was found ($F(2)=4{,}732$, $p=.022$, $\eta_p^2=.35$), reflecting the fact that no *p~b* alternations were found in one of the 'lexical gap' environments (i.e. after nasals). Hence, the distribution of alternations is not random, but depends on the phonological environment of the non-word item.[26] In sum, children's behaviour was influenced by the final consonant and rhyme of the non-word item.

To enable a comparison between CELEX frequency and the experimental data, probability of voicing was computed based on the type frequency of precisely those rhymes that were used for the experimental stimuli. When this count is restricted to monosyllabic words, the overall likelihood of voicing is 26%, as was found for all nouns in CELEX (see Chapter 5). The resulting graph is shown in Figure 4, together with alternations produced in the Wug-test (note that scales are different).

*Figure 4: Alternations for non-words and probability of voicing based on type frequency of monosyllabic nouns with identical rhymes (Exp. I).*




This figure shows that children produced more *t~d* alternations after short vowels than expected on the basis of the input (at least when compared to long vowels), although this effect is not present for 7-year-olds (Table 25). Also, children produce fewer *t~d* alternations after long vowels than expected, which might be due to the fact that these items are often produced with glides. However, children did not produce glides for non-words in the Wug-test. Finally, not only are *p~b* alternations after long vowels unexpected, children produced more alternations for /b/ than expected on the basis of overall type frequency. To explore this further, Table 26 shows

[26] When the factor Group (3-year-olds, 5-year-olds, 7-year-olds) was also taken into account, the effects of Final Obstruent and Final Obstruent x Rhyme remained significant, although there was no main effect of Group and there were no interactions with Group.

the proportion of alternations per rhyme for both singular and plural nouns, together with the average token frequency of plurals. It was deemed important to calculate the probability of voicing for plural nouns separately, as the data in Chapter 5 showed that plurals may be less common than singulars (e.g. plurals such as *webben* 'webs' or *krabben 'crabs*' had a frequency of zero). Rhymes in the 'lexical gap' environment are shaded.

*Table 26: Alternations for non-words (Exp. I), probability of voicing for singular and plural nouns (%), and average token frequency of plurals.*

| Rhyme | Alternation | Prob. sg. | Prob. pl. | Avg. token freq. pl. |
|---|---|---|---|---|
| /ɛmp/ | 0 | 0 (0/1) | 0 (0/1) | 0 |
| /ɪmp/ | 0 | 0 (0/1) | 0 (0/0) | 0 |
| /ɔmp/* | 0 | 0 (0/5) | 0 (0/5) | 0 |
| /op/* | 0 | 0 (0/4) | 0 (0/4) | 0 |
| /ap/ | 1.7 | 0 (0/7) | 0 (0/6) | 0 |
| /ep/ | 3.4 | 0 (0/9) | 0 (0/5) | 0 |
| /ɪt/* | 1.7 | 0 (0/8) | 0 (0/2) | 0 |
| /ɔp/ | 1.7 | 0 (0/13) | 0 (0/11) | 0 |
| /ɑp/ | 5.2 | 7.7 (1/13) | 0 (0/9) | 0 |
| /ɛt/ | 6.9 | 12.5 (1/8) | 14.3 (1/7) | 12 |
| /ɛp/ | 3.4 | 16.7 (1/6) | 0 (0/2) | 0 |
| /et/ | 0 | 18.2 (2/11) | 25.0 (1/4) | 1 |
| /ɑt/ | 5.2 | 27.3 (3/11) | 27.3 (3/11) | 2 |
| /ot/ | 2.6 | 33.3 (5/15) | 25.0 (3/12) | 3 |
| /ɪnt/* | 4.3 | 33.3 (3/9) | 40.0 (2/5) | 245 |
| /at/ | 5.2 | 43.8 (7/16) | 46.7 (7/15) | 10 |
| /ɔnt/ | 8.6 | 57.1 (8/14) | 80.0 (8/10) | 14 |
| /ɑnt/ | 10.3 | 70.0 (14/20) | 65.0 (13/20) | 53 |

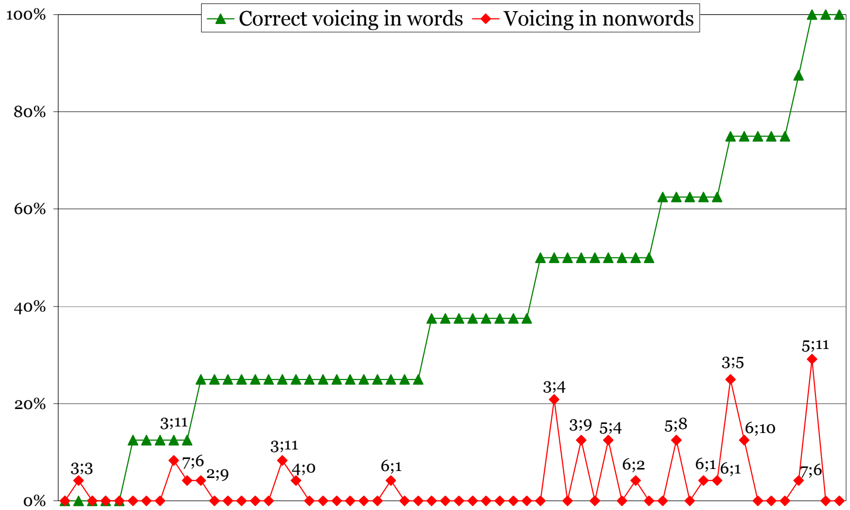
* The results for two items with identical rhymes are combined (e.g. *tomp* and *fomp*).

This table shows that most alternations were produced for the two rhymes which are most likely to alternate on the basis of similar monosyllabic words (i.e. the test-items with /ɑnt/ and /ɔnt/ rather than all items in this phonological environment). Hence, frequency effects can explain why items within a certain phonological environment (e.g. /t/ after nasals) are treated differently. Also note that the probability of voicing is above 50% for these two items. However, there is no perfect match between the two measures, e.g. there were fewer alternations for the item ending in /et/ (i.e. *feet*) than expected on the basis of its rhyme. There could be a variety of reasons for such a result, including chance effects, the influence of a certain monomorphemic form (e.g. *veter* 'shoestring'), the influence of a word that is very common in child speech (e.g. *eten* 'to eat'), etc. Generally however, token frequency does not seem to affect results directly.

Nonparametric correlations were performed to further assess the relationship between the distribution of children's responses (i.e. the percentage of voiced responses, computed over the total number of responses) and the probability of voicing based on monosyllabic words from CELEX. Results show a high correlation between alternation rate and the probability of voicing based on singular nouns with identical rhymes, both for the probability based on type frequency ($\rho=.745$, $p<.001$) and token frequency ($\rho=.786$, $p<.001$). This indicates that children are sensitive to the lexical distribution of voicing when they produce alternations for non-words. Surprisingly, probabilities based on summed token frequencies show a somewhat higher correlation, suggesting that both type and token frequency play an important role. Finally, children are most sensitive to the frequency of singular nouns, as the correlation with probability based on plural nouns is lower (i.e. $\rho=.679$, $p=.002$ for types and $\rho=.696$, $p<.001$ for tokens). The frequency effect is also strongest within a paradigm, as correlations with probabilities based on all word types are lower (e.g. $\rho=.661$, $p=.003$ for all singular types).

To further determine the relation between alternations for non-words and lexical knowledge, Figure 5 shows non-word alternations (e.g. [kɛdən]) together with the percentage of correctly produced alternating plurals (e.g. [bɛdən] 'beds'). Results are shown per child; only when non-word alternations were produced is the child's age indicated. As in Figure 3, results are ranked according to performance correct on words, and the age of the child is indicated only when alternations were produced.

*Figure 5: Correct voicing in words and voicing in non-words (Exp. I).*




First, this figure confirms that older children tend to produce more correct alternating forms (see Table 22), although there is no strong correlation of performance and age. Interestingly, alternations for non-words (e.g. [kɛdən]) seem to increase with knowledge of alternating plurals rather than age, such that the occurrence of non-word alternations implies knowledge of alternations. The reverse is not true, as there are children who produce correct voicing alternations for words but no alternations for non-words.[27] Nevertheless, this pattern of results is expected if children are influenced by the words in their lexicon rather than by phonological preferences.

However, recall that lexical gap alternations (e.g. [tap] ~ [tabə]) were produced by two children (Noortje and Sven). [28] Such responses may reflect

[27] There is one exception, i.e. Pim (3;3) produces [tɛp] ~ [tɛbə] and no alternating plurals. However, this child produced only five plural forms. Generally, it seems that /b/ alternations were not directly linked to knowledge of words with /b/, e.g. Robin (6;1) produces [tɛp] ~ [tɛbə] and correct plurals with /d/ but no plurals with /b/. In contrast, Njana (5;11) produces only /d/ for non-words, although she knows /b/ plurals.

[28] Note that some productions proved to be difficult to judge in terms of voicing. For the subset of 80 non-word items presented to the six listeners, there were 7 items for which the original transcription as 'voiced' was changed. These were judged to be voiced by only three out of six listeners (3 items) or two out of six listeners (4 items). Note that these items were often judged as 'Don't know', rather than unvoiced. There were three cases that were judged to be voiced by only four out of six listeners, including one lexical gap alternation (i.e. [debən] by Sven 4;0). For

merely noise, as no child consistently produced alternations in these environments (whereas there are children who consistently produced *-s* plurals). The first child, Noortje (3;5), had a high score correct on alternating words, including the two plurals with /b/ (i.e. /krabə/ 'crabs' and /wɛbə/ 'webs'). This child also produced four /d/ alternations for non-words and two overgeneralisations for words (i.e. *[pɛdə] 'caps' and *[vudə] 'feet'). Hence, Noortje was the only 3-year-old who had acquired many alternating plurals, which suggests that her non-word alternations do not merely reflect a phonetic voicing effect. However, it is possible that 'correct' alternating plurals are also due to a voicing overgeneralisation (although Noortje devoices medial /b/ in mono-morphemic *Dribbel*). In general, children who produced many alternations for non-words were found to be advanced in forming plurals. For instance, the six children who produced most alternations for non-words did not produce any bare stems. Sven (4;0) also produced a lexical gap alternation, as well as the correct plurals [hɔndə] 'dogs' and [pɔtlodə] 'pencils'. However, this child did not produce any other voicing alternations for non-words or overgeneralisations in words. In sum, the occurrence of 'lexical gap' alternations is rare, although at least for one child it may reflect a genuine overgeneralisation on the basis of her lexicon. On the other hand, it may also reflect an overextension of a phonological rule of intervocalic voicing. We will return to this issue in the discussion (section 5.2.3.5).

A remaining question concerns the stability of children's behaviour over time. Generally, children showed variable behaviour in a single test session. For instance, an erroneous response was given (e.g. *[hɑntə]) after which the correct form was used spontaneously. This seems to point towards a task effect, which leads children to produce more non-alternating plurals than they would do in spontaneous speech.[29] However, the opposite pattern

---

Noortje, 100% agreement was reached on [debən], whereas [tabən] was judged to be voiced by five out of six raters (one "Don't know"). The average agreement on items originally transcribed as voiced was 92%. One item originally transcribed as unvoiced was judged to be voiced by two listeners (i.e. [tapən] by Gijs 3;11), and nine other items were judged to be voiced by only one listener. The fact that disagreement arose relatively often in 'lexical gap' cases may further point towards a gradient or phonetic effect of voicing for intervocalic /b/. Alternatively, transcribers may have been influenced by the lexical gap.

[29] E.g. Tiemen (3;8) produced *[bɛtə], but when he was asked 'are they chairs'? he responded with "*nee het zijn bedden*" 'no they are beds'. Another sign of a task effect (or hypercorrection) is *[paraplydən] for *paraplus* /paraplys/ 'umbrellas'. Children may also be consistent, e.g. Tiemen (3;8) produced **krappen*, both as a response and spontaneously (*waar wonen *krappen*

also occurred.[30] To determine whether children were consistent across two sessions, three children were randomly selected for a re-test two weeks after the first session (in the reversed order). In Table 27, alternations from session I are provided together with results for those same items in the session II.

*Table 27: Alternations for non-words and overgeneralisations for words in two sessions.*

| Child | Stimulus | Session I | Session II |
|---|---|---|---|
| Njana (5;11.13 / 5;11.27) | words | *pɛ[d]ən 'caps' | same |
| | | *olifɑn[d]ən 'elephants' | same |
| | | *vu[d]ən 'feet' | missing |
| | non-words | mɪdən | same |
| | | slɑdən | same |
| | | kɛdən | kɛtən |
| | | flandən | flantən |
| | | jɔndən | jɔntən |
| | | zodən | zotən |
| | | zwɑbən | same |
| Maarten (5;8.8 / 5;8.23) | words | *pɛ[d]ən 'caps' | same |
| | non-words | kɛdən | same |
| | | slɑdən | same |
| | | xɔbən | same |
| | | zwɑpən | zwɑbən |
| | | tɛpən | tɛbən |
| Robin (6;1.19 / 6;2.14) | words | *hu[t]ən 'hats' | hu[d]ən |
| | | *krɑ[p]ən 'crabs' | same |
| | non-words | dɑbən | dapən |

The results show that two out of three children produced fewer alternations in the second session, whereas one child produced more instances (adding two /b/ alternations). The latter shows that such /b/ alternations are not necessarily associated with an early phonological effect. Furthermore, these

*eigenlijk*? "Where do crabs actually live"?). Bare stems were also used spontaneously (e.g. Chiara 4;2: *sommige klaat zijn blauw* 'some *klaat* are blue').

[30] Noortje (3;5) produced *[vudə] for *voeten* 'feet' before using the correct form spontaneously.

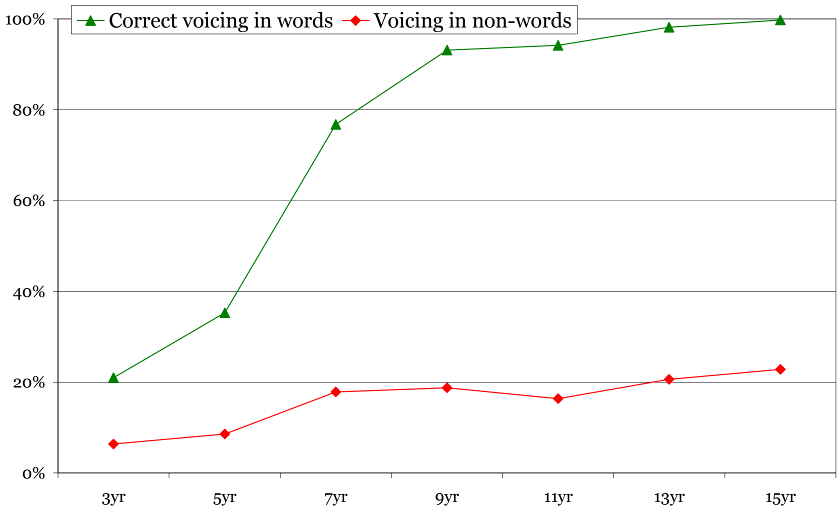
results show that targets of children's rules or schemas are apparently not stable over a two week period. Thus, even though phonological environments are identical, the items undergoing alternations may change. Another possibility is that long-term memory representations had been formed, which may have increased children's reliance on Paradigm Uniformity. As discussed in Chapter 2, rule-based models incorporating probabilistic rules or constraint rankings (e.g. Albright & Hayes 2003) would be able to account for this type of variability. As analogical models are inherently probabilistic, they would also be able to handle these data.

A second outcome of this experiment is that overgeneralisations of voicing for words and alternations for non-words were not principally produced by the youngest children. To further determine the effect of age, data from older children were taken from previously unpublished Dutch Wug-test data from four studies with Belgian (i.e. Flemish-Dutch speaking) children. These studies which will be referred to as the "Ghent studies" (data from Gekiere 1983, Minnebo 1983, Vervenne 1983).[31] In these studies, each experimenter tested 7 groups of 14 children with ages ranging from 3 to 15, using the same set of words and non-words to elicit plurals. The set of words contained 15 items that were similar to the present stimuli in terms of rhyme category (e.g. *pet* 'cap' and *bed* 'bed'), as well as 11 words with a liquid preceding the final obstruent (e.g. *tulp* 'tulip', *paard* 'horse'). Similarly, the set of 12 non-words contained 6 'similar' items (e.g. *tep*, *peit*) and 6 items with liquids (e.g. *kilp*, *leert*). Results for the similar words show that the mean score correct on alternating words for 3-year-olds was 16% (35/218), while 5-year-olds scored 27% (66/248) correct.[32] The only overgeneralisations produced for words were one occurrence each of **liben* 'lips', **poben* 'dolls' and **lamben* 'lamps' in the results of the 5-year-olds. Generally, the Ghent results confirm the present results, in that 3-year-olds produced fewer alternations than 5-year-olds, for both words (0 vs. 3.2%) and non-words (4.3% vs. 8.0%). The only exception is the non-word *vlap*, whose plural is voiced more often by 3-year-olds than 5-year-olds (although numbers are generally low). Results matched the present Wug-test data, as

[31] I am grateful to professor Johan Taeldeman of Ghent University, who gave me the opportunity to analyse unpublished data from MA theses. The selection and combination of data as well as the interpretation of results are my own.
[32] Note that data reflect the auditory judgment of the experimenters, as no judgment data from other sources or acoustic measurements are provided.

alternations were only produced for non-words with /p/ after short vowels (*tep* /tɛp/ 20%, *vlap* /vlɑp/ 22%) and /t/ after long vowels or diphthong (*peit* /peit/ 26%), in contrast to non-words with /t/ after short vowels (*lut* /lʌt/ 0%, *let* /lɛt/ 1.0%) and a lexical gap item (*kleup* /kløp/ 0%). Voicing alternations for words and non-words for all age groups are shown in Figure 6 below, comparable to Figure 5.

*Figure 6: Correct voicing in words and voicing in non-words Ghent studies.*



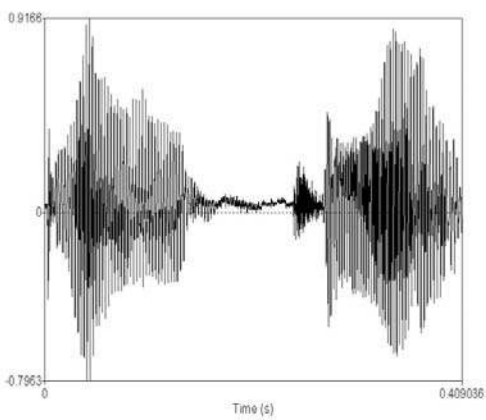

This figure shows that the number of alternations for non-words increases until the age of 7, after which there is a plateau (with an average of 14% overall). As has been observed before, it seems that overgeneralisations of voicing increase as children score better on existing alternating plurals. As Figure 6 shows, an increase of alternations for non-words is observed at age 7 (17%), which could be due to 7-year-olds' familiarity with spelling (i.e. *bed* is spelled with *d*). A fourth study comprised in the Ghent studies (i.e. using the same stimuli) included a group of fourteen 4-your-olds, who show a high overgeneralisation rate for the item **pedden* 'caps' (25%), data taken from Van Doorne (1992). However, these data generally show the same trend as observed before, in that 5-year-olds produce more overgeneralisations of voicing than younger children. Generally, results mirror the first Wug-test results, except that more voicing alternations were produced. This could be due to a difference between Belgian (Flemish Dutch) and Dutch (Northern Dutch) children or a difference in the non-word stimuli used.

Before we turn to a general discussion of Experiment I in 5.2.3.5, acoustic measurements will be provided in the next section, to provide more information on the realisation of the voicing contrast.

**Acoustic measurements**

Two prominent acoustic features of the Dutch voicing contrast are closure duration and burst duration, which are both relatively long for voiceless stops (e.g. Slis & Cohen 1969). Figure 7 below shows the word *petten* [pɛtən] 'caps', with clearly visible closure and burst durations. In the present analysis, closure duration (CD) and burst duration (BD) were taken as acoustic variables; the beginning of closure was measured by the change in formant structure and amplitude, the end was marked at the release burst or the characteristic periodicity and amplitude of the vocalic segment. The analyses were performed using *Praat*.

*Figure 7: The waveform of petten 'caps'.*




As described in Chapter 3, Kuijpers (1993ab) found that children (ages 4;5, 6;4 and 12;2) produced durational differences for the medial voicing contrast that were similar to those of adults. As there are only five correct /b/ tokens for the 3- and 4-year-olds, an analysis per age group is not possible for the *p~b* contrast. Table 28 shows the group results for /t/ and /d/, taking only correct words into account. For each group, 60 tokens were selected on the basis of the quality of the sound file (30 voiceless, 30 voiced).

*Table 28: Mean closure and burst duration for medial /t/ and /d/in ms (sd), Exp. I.*

| | **3-yr-olds** | | **5-yr-olds** | | **7-yr-olds** | |
|---|---|---|---|---|---|---|
| | /t/ | /d/ | /t/ | /d/ | /t/ | /d/ |
| CD | 111 (63) | 55 (33) | 97 (18) | 57 (16) | 60 (14) | 43 (10) |
| BD | 34 | 11 | 23 | 18 | 30 | 12 |

The results show that the acoustic measurements generally match the auditory analysis, i.e. children produced a reliable voicing contrast. These data are consistent with those of Kuijpers (1993ab), who found that durations for voiceless stops generally decrease with age. Previous studies indicate that individual differences are large, especially for the youngest children. Combined with the low number of voiced tokens per child, this means that a comparison of the voicing contrast for individual children is difficult. For instance, the closure duration of /p/ in [apə] 'monkeys' was 20 ms longer than the duration of /b/ in [tabə] (both produced by Noortje at 3;5). However, such a result is not conclusive considering the large variation displayed in child speech. In sum, group results show that children were able to produce a voicing contrast, even though it is not clear whether this applies to all individual children. Children did not produce enough voicing alternations for non-words to compare voiced and voiceless realisations (e.g. [tabən] vs. [tapən]). Furthermore, there were too few mono-morphemic forms (e.g. /obər/ 'waiter') to compare them systematically to bi-morphemic words (e.g. /sxapən/ 'sheep'). Before we turn to a discussion of the general results, the results of a separate experiment investigating younger children's production of intervocalic voicing in mono-morphemic and bi-morphemic words will be discussed in the next section.

### Bi-morphemic vs. mono-morphemic words

In order to investigate the acquisition of the voicing contrast further, results from a study by Zamuner, Kerkhoff & Fikkert (2006b) will be discussed.[33] In

[33] This work was carried out in collaboration with Tania Zamuner and Paula Fikkert, supported by NWO grants 275-75-001 awarded to Tania Zamuner, 016-024-009 awarded to Paula Fikkert, and 360-70-100 awarded to Paula Fikkert and René Kager. Children were tested at the KindertaalLab at the Radboud University Nijmegen, and children were recruited through the MPI Baby Research Center, funded by the Spinoza Project awarded to Dr. Anne Cutler. I am

this study, a production experiment was carried out, to determine how children under the age of four produce non-alternating versus alternating words, and how children produce complex (or bi-morphemic) words versus simplex (or mono-morphemic) words. Generally, if children do not produce voicing in bi-morphemic words, realisation of voicing in mono-morphemic words would show that this is not due their inability to produce a voicing contrast. To this end, a Picture Naming Task and Word Imitation Task were used, restricted to words with intervocalic /t/ and /d/ (i.e. there are too few Dutch nouns that are familiar to young children with /b/). A further comparison was made between children's elicited versus imitated productions, to investigate whether perceptual or articulatory limitations play a role in how children produce /t/ and /d/. Children in two different age groups (18 2;6-year-olds aged between 2;5-2;8 with a mean age of 2;7 and 18 3;6-year-olds aged between 3;6-3;8 with a mean age of 3;7) were tested on four non-alternating (e.g. *petten* 'caps') and four alternating nouns with /d/ (e.g. *bedden* 'beds'). In addition, four mono-morphemic nouns with medial /t/ (e.g. *water* 'water') and four mono-morphemic nouns with medial /d/ were added (e.g. *ridder* 'knight').

Results from the Picture Naming Task showed that children can produce both intervocalic /d/ and /t/. However, intervocalic voiced stops were more likely to be produced in mono-morphemic forms than in bi-morphemic forms. This effect was stronger for the older children than for the younger, who generally produced fewer intervocalic voiced stops. Hence, for the older children /d/ was more accurate in the mono-morphemic than in the bi-morphemic context. This interaction between voicing and morphology had not been previously found, as children's spontaneous speech does not provide enough examples of both word types. At age 3;6, only four participants produced alternations with /d/ in bi-morphemic words in elicited productions, showing that children were not very successful at producing alternating words while voicing was produced more accurately in mono-morphemic words. Finally, an effect of lexical development was found, as children with a smaller receptive vocabulary (as measured by the Dutch version of the N-CDI) were more likely to produce only mono-morphemic

grateful for assistance and support by Ellen Westrek, Marleen van der Avoird, Esther Janse, Erik Jan van der Torre and Elise de Bree.

words correctly.[34] Vocabulary scores for the older age group were at ceiling, and revealed no relationship with performance.

Results for the Word Imitation task showed that the youngest children performed worse on /d/ words than on /t/ words, whereas there was no such effect for the older group. Hence, younger children's difficulty with medial /d/ seems at least partly based on perceptual or articulatory difficulties (as was also suggested by the corpus study described in Chapter 4). It is not likely that children's difficulty with /d/ only reflects the higher frequency of medial /t/ compared to /d/, as this would not explain the difference between /d/ in the bi-morphemic versus /d/ in the mono-morphemic context. Zamuner et al. conclude that medial /d/ must have been acquired at 3;6, extending results from Kuijpers (1993ab), who tested children from the age of 4;5. Hence, the authors find that a reliable medial voicing contrast is produced at a younger age than previously found.[35] Finally, an acoustic analysis of children's elicited productions did not provide evidence for the hypothesis that children maintain a covert voicing contrast (i.e. between /t/ in *[betən] versus /t/ in [pɛtən]), which has been noted in other studies of voicing acquisition for English (e.g. Macken & Barton 1979, Scobbie et al. 2000).

In conclusion, these results are similar to the results found in the first Wug-test, showing that children are poor at producing alternating plurals even though they are capable of producing a voicing contrast. Results by Zamuner et al. (2006b) confirm that the youngest children in the Wug-test experiment were able to produce the contrast, as they were all older than 3;6. Even though this study only addressed the alveolar contrast, the labial contrast will also be assumed to have been acquired as it is generally acquired earlier (see, e.g. van der Feest 2007). The overall results will now be discussed, addressing the four hypotheses of section 5.2.2 above.

---

[34] For this measure parents completed the Dutch version of the N-CDI or MacArthur Communicative Development Inventory, scores were standardised as developmental norms are for children up to 30 months (Zink & Lejaegere 2002).

[35] Zamuner et al. adopted criteria from Sanders (1972), according to whom a segment is 'customarily produced' if over 50% of children produce a segment correctly and 'acquired' when over 90% of children produce a segment correctly (see Brown 1973 for similar criteria).

#### *5.2.3.5 Discussion*

The experiments described in the previous section involved the elicitation of plurals with three groups of children (3-year-olds, 5-year-olds, and 7-year-olds). Let us now turn to the hypotheses that were formulated in section 5.2.2.

The first hypothesis (formulated in 5.2.2) entailed that children would rely on a strategy of *Paradigm Uniformity*, which should lead to non-alternating plurals. Results for words show that such errors (e.g. *[bɛtən] 'beds') were produced in equal numbers by all age groups, i.e. around 40%. Recall that the corpus study based on the CLPF database showed that around 29% of /d/ plural targets were devoiced (see Chapter 4). Similarly, van der Feest (2007) found that young children's spontaneous devoicing rate for medial /d/ was around 28%. As the elicitation task shows a higher percentage of errors, it is likely that the task induces children to rely more on Paradigm Uniformity. However, the fact that the three age groups show a similar overgeneralisation rate for words (e.g. *[bɛtən] 'beds') indicates that younger children do not rely more on this strategy than older children. Note that the lack of alternations could also reflect the relative unmarkedness of voiceless obstruents. The results show that the majority of children strictly kept to non-alternating forms. On the other hand, there is evidence against a strategy of Paradigm Uniformity, as children produced overgeneralisations of the type *[pɛdən] 'caps' in 3% of cases. Moreover, around a third of all children produced alternations for non-words (e.g. [kɛt] ~ [kɛdən]) in 3% of cases. These results suggest that Paradigm Uniformity is not the only factor that guides children's plural formation.

Alternations for non-words were not produced randomly (i.e. by *random selection*, hypothesis 2), as children's behaviour was found to be influenced by the final consonant and rhyme of the non-word item. This then means that children's behaviour was guided by either phonological or analogical generalisations, which brings us to the remaining two hypotheses.

The third hypothesis was that children are guided by a strategy of *phonological generalisation*. Under this view, contextual markedness constraints like intervocalic or postnasal voicing are prominent in children's early grammars, leading to voicing alternations that do not necessarily reflect the ambient language. This theory correctly predicts that alternations occur in 'lexical gap' environments (e.g. [lɑmp] ~ [lɑmbə] ' lamps' and [tap] ~ [tabə]), which are unexpected on the basis of the input. Moreover, two children produced alternations for non-words in a 'lexical gap' environment (e.g. [tap] ~ [tabən]), which are not expected on the basis of Dutch plurals. The occurrence of these unexpected voicing alternations could be accounted for by an analysis that posits an early (and false) rule of intervocalic voicing. Such a rule would be expected to disappear with time, as the child develops both the correct underlying representations and the correct grammar. This account ties in with the observation that these alternations were produced by children in the youngest age group (but keep in mind that there were only three cases for non-words). However, as long vowels are expected to pattern with nasals, it is unclear why children did not produce alternations for /b/ after nasals (reflecting postnasal voicing, e.g. [kɪmp] ~ [kɪmbən]). Children also produced more *t~d* alternations after short vowels than expected on the basis of the input. However, a phonological generalisation taking into account markedness considerations would lead to alternations after long rather than short vowels. Alternatively, an input-driven account would explain this finding by examining differences between adult and child corpora. For instance, the input to Sarah (see Chapter 4) contained relatively more *t~d* alternations after short vowels than the adult corpus. Finally, even if overgeneralisations reflect rule-based behaviour, children did not behave consistently within a certain phonological environment (e.g. alternations were produced for /flɑnt/ but not /dɪnt/, or for /jɪt/ but not /mɪt/). This variability shows that rules were not applied across-the-board, which is more in line with models that incorporate probabilistic rules or constraints. Furthermore, there may be an effect of phonetic implementation for /b/ rather than /d/, which would account for alternations of /b/ after short and long vowels. Recall that the phonetic pressure to realise voicing in between vowels is greater for /b/ due to articulatory, acoustic, and auditory factors. Also, the CLPF corpus showed that young children (up to age 2;11) produce more errors in which /p/ is realised as /b/, and few errors in which /t/ is produced as /d/ (van der Feest 2007). In contrast, the present results point towards an overwhelming tendency for /d/. Hence, /b/ alternations may reflect 'early' phonetically grounded phonological effects, which ties in with the observation that lexical gap alternations occur mostly for the youngest group.

The fourth hypothesis held that children are sensitive to analogy with lexical items when they inflect novel nouns (i.e. *analogical generalisation*). Such an account predicts correctly that the alternating pattern is not very productive, as the strength of a morphological pattern is related to its type frequency (Bybee 1995). The type frequency of the Dutch alternating pattern is indeed lower than the non-alternating pattern, and alternations are associated with high token frequency. Secondly, when children do extend the alternating pattern to non-words, they were influenced by the probability of voicing based on similar words in Dutch. Hence, alternations were produced for 4.5% of non-words with /t/ (31/696) and 1.7% of non-words with /p/ (12/696), reflecting the relative type frequency of alternations (see Chapter 4). Furthermore, no alternations occurred for /p/ after nasals (a 'lexical gap' environment), while children produced most alternations for non-words with /t/ after nasals (e.g. [jɔndən] > [kɛdən]). This pattern of results reflects the fact that alternations are most common in this environment, while children did not generalise over a 'natural' class (such as nasals). Importantly, results show a correlation between the probability of voicing in the input and the number of alternations produced. Also, there was a relationship between children's knowledge of alternating words and the occurrence of alternations for non-words (see Figure 5). Hence, children who produced many correct alternating words (e.g. [bɛt] ~ [bɛdən] 'beds') also produced most alternations for non-words (e.g. [kɛt] ~ [kɛdən]). Even though the former could also reflect an overgeneralisation (e.g. [bɛdən] on the basis of */bɛt/), this is not likely as children tended to produce overgeneralisation errors for low-frequent plurals (e.g. */pɛdən/ 'caps'). Furthermore, performance on alternating plurals correlated with the frequency of the plural form. Results thus suggest that some children had formed a semi-productive schema on the basis of alternating plurals in their lexicon.

Note that it was not the case that alternations for non-words were produced due to a direct 'priming' effect, as alternating words were never placed immediately before an analogous non-word (e.g. [kɛt] after /bɛd/ 'bed'). This suggests that exemplars in the lexicon rather than a child's age predicts the number of voicing alternations for non-words. Hence, 3- and 5-year-olds produced the same number of alternations for non-words, and 5-year-olds produced more overgeneralisations of voicing for words (e.g. *[pɛdən] 'caps'). Furthermore, three out of five children who produced more than one alternation in a certain environment were between age 5 and 6, suggesting that this behaviour does not reflect an early rule-based strategy.[36] As discussed in Chapter 2, previous research indicates a relationship between children's overgeneralisations and vocabulary size (Marchman & Bates 1994, Derwing & Skousen 1994, cf. Marcus et al 1992:99). The current results show that 7-year-olds produce fewer alternations, whereas there is a 'peak' at around age five (although results from the Ghent studies show a peak at age 7). This suggests that overgeneralisations do not continue to increase. This could be due to the fact that in the early stages of acquiring voicing alternations, these plurals dominate the lexicon more than at later stages. As shown in Chapter 4, alternating plurals are relatively frequent in child-directed speech. Such an effect may also account for the unexpected 'lexical gap' alternations after long vowels produced in the first Wug-test. As these alternations were mainly produced by a child who was familiar with *p~b* alternations in existing words (e.g. [krɑbən] 'crabs'), these alternations could have occurred on the basis of her lexicon.

In sum, the first Wug-test experiment provides little evidence supporting purely phonological models, although the youngest children produced some voicing alternations that are unexpected on the basis of Dutch. Children seem to overgeneralise voicing in certain environments (e.g. /b/ after short vowels), although it is as yet unclear whether this is a child-specific phenomenon. This issue will be addressed in section 5.2.5, which contains adult Wug-test data. The experiment provided more evidence for analogical models, as children's behaviour was found to correlate with lexical factors. Let us now turn to the second Wug-test experiment, which investigates these issues comparing a group of language impaired children to a group of typically developing children.

## 5.2.4 Elicitation Experiment II: SLI and TD children

### *5.2.4.1 Introduction*

The Wug-test methodology has also been used to investigate productive knowledge in children with Specific Language Impairments (Oetting & Rice 1993, Goad & Rebellati 1994, Leonard et al. 1997). The current study was

[36] Note that two of the youngest children who produced many correct plurals as well as overgeneralisations (Noortje and Femke) participated in the pilot study. This might not have been a coincidence, as teachers tend to select children with large vocabularies first.

aimed at assessing knowledge of the voicing alternation in children with Specific Language Impairment (SLI) compared to a control group of Typically Developing children (TD).[37] Children with SLI have been found to be delayed in both their phonological and morphological development (see de Bree 2007 for Dutch).[38] Many researchers hold that SLI reflects a delay in language development rather than a deviance, as children's behaviour is often found to be qualitatively similar to that of younger children (see Bishop 1997; Leonard 1997). Importantly, grammatical morphology seems to be more delayed than any other area of language development, and overt plural marking has been found to be poor (Oetting & Rice 1993, Leonard et al. 1997, Niemi 1999, Conti-Ramsden 2003), see de Jong (1999) and van Alphen et al. (2004) for Dutch. For instance, previous studies have reported an increase in zero marking or the use of bare stems for language-impaired children (e.g. Clahsen et al. 1992). However, plural inflection seems to be relatively intact compared to verb inflection (Clahsen et al. 1992, Rice & Oetting 1993, Bishop 1994, Leonard, Salameh & Hansson 2001, Marshall 2004). Furthermore, SLI children have been found to show poorer performance on regular versus irregular past tense marking in English. These findings have led proponents of rule-based or Dual Mechanism models to argue that SLI is a grammar-specific deficit (e.g. Gopnik 1990, Gopnik & Crago 1991, Gopnik & Goad 1997, Clahsen 1999, van der Lely & Christian 2000, van der Lely & Ullman 2001). Under this view, a grammatical deficit causes an inability to decompose inflectionally complex words into their constituent parts. Hence, SLI children only rely on analogy with existing words to extend patterns to new forms. Furthermore, they use lexically based strategies instead of rules to compensate for impaired inflectional morphology, such as learning plurals by rote. For instance, some studies have shown frequency effects for regularly inflected plural nouns and past tense verbs for children with SLI

[37] This study was carried out in collaboration with Elise de Bree, with assistance from other researchers from the Dyslexia project at the Research Institute of Language and Speech (UiL-OTS) at Utrecht University ('Early language development in Specific Language Impairment and Dyslexia: a prospective and comparative study'), supervised by Frank Wijnen. A group of children at risk of dyslexia also participated in this study, results for this group are discussed in de Bree (2007). In this dissertation, children at risk of dyslexia are compared to children with SLI and a control group. A subset of the results discussed in this dissertation has also appeared in Kerkhoff & de Bree (2005) and de Bree (2007) as they are based on a joint study.

[38] A Specific Language Impairment is diagnosed in the absence of other cognitive or neurological defects, although there is an ongoing debate to what extent language can be dissociated from other cognitive domains (e.g. Gopnik & Crago 1991, Leonard 1998).

but not typically developing children (Oetting & Rice 1993, Oetting & Horohov 1997, van der Lely & Ullman 2001). Previous studies have also considered morpho-phonological alternations. For instance, Goad & Rebelatti (1994) claim that plural realisation of English SLI children was based on explicit rule-learning (add -*s* to create a plural). This led to slow and effortful realisation of the plural and prevented children from using alternating suffix forms (/z/ and /ɨz/) in the appropriate phonological contexts. According to these authors, SLI children are unable to use the (phonological) rules component to master the suffix alternation. This claim was further supported by the fact that they found a difference in performance between words and non-words. Conti-Ramsden & Windfuhr (2002) also found that 3-year-old children with SLI had more difficulty marking stems with the /ɨz/ allomorph than their age-matched peers. A similar result was found for SLI children as old as thirteen by Marshall (2004). However, results for morphologically rich languages such as Italian (Bortolini, Leonard & Caselli 1998) and Spanish (Clahsen et al. 2002) suggest that SLI children may produce overgeneralisations of verbal stems or suffixes, e.g. Spanish *[ponistə] for /pusistə/ 'you put-past' on the basis of *poner* (e.g. /ponɛs/ 'you put').

In contrast, some have argued that SLI is characterised by a processing deficit rather than a grammatical deficit (e.g. Bishop 1997, Leonard 1997), which may be related to the ability to detect rapid temporal changes in auditory stimuli (Tallal et al. 1996). For instance, according to the "surface account" (e.g. Leonard et al. 1997), processing limitations interfere with the perception of relatively brief grammatical morphemes. Moreover, as discussed in Chapter 2, proponents of analogical models have proposed links between the rate of lexical learning and grammatical development or generalisation behaviour (Plunkett & Marchman 1993, Marchman & Bates 1994, Marchman et al. 1997). SLI children show protracted lexical learning as well as a tendency to memorise individual forms, which is associated with later and less efficient generalisation. Importantly, the difference between regular and irregular inflection may not reflect a grammatical disorder, but has been argued to be due to effects of phonological complexity, which affect regular inflection more than irregular inflection in English (e.g. Joanisse & Seidenberg 1998; Marshall 2004). For instance, regular past tense forms are more likely to have word-final clusters than irregulars (*walked* /kt/ vs. *ran*). Bernhardt & Stemberger (1998:482) also point towards this interaction of phonology and morphology, i.e. the absence of a morphological suffix may be motivated by phonological constraints or articulatory difficulties, rather than an impairment of inflectional morphology. As SLI children show more instances of phonological simplification compared to both age-matched and MLU-matched controls (e.g. Beers 1995, Aguilar-Mediavilla, Sanz-Torrent & Serra-Raventós 2002), this may have adverse effects on morphological abilities. For example, final consonant (cluster) omission, reduction or substitution have been found to interfere with production of the grammatical morphemes for plurals, possessives, regular past tense and regular third-person singular (e.g. Bortolini & Leonard 2000, Owen, Dromi & Leonard 2001). Evidence from connectionist simulations also support the importance of phonological representations for the acquisition of morphology (Joanisse & Seidenberg 1998).

Finally, it is possible that children with SLI have more difficulties with realising the voicing contrast, as was found for younger typically developing children (see Chapter 4). Previous studies have shown that SLI children exhibit motoric difficulties, already manifest at the babbling stage (e.g. Whitehurst et al. 1991). These motoric difficulties persist; for example, Edwards et al. (1999) found that phonologically disordered children (4;4) used less controlled (or 'ballistic') gestures from stops to vowels in comparison to their age peers. In terms of the ability to produce the voicing contrast, several studies have shown that, compared to other phonological and articulatory abilities (such as place and manner contrasts), voicing in word-initial and prevocalic position is relatively advanced in American children with SLI (Ingram 1981, Leonard et al. 1985, Forrest & Morrisette 1999). Others claim that language-delayed children's control of the acoustic-phonetic details of the initial and final voicing contrast is less mature than that of normally developing children (Bond & Wilson 1980). As the intervocalic voicing contrast has not been investigated in Dutch children with SLI, a more detailed phonetic analysis of children's productions will be provided in the results section. Let us now turn to predictions for the present study.

#### *5.2.4.2 Predictions*

Based on previous findings, the expectation is that children with SLI will have more difficulty with plural marking than typically developing (TD) children. Under a rule-based approach, the absence of a morphological rule should leave SLI children unable (or less able) to inflect non-words or produce overregularisations (such as *[bɛtən] 'beds' or *[pɛdən] 'caps'). Moreover, if all inflected words are treated as unanalysed wholes by SLI children, alternating plurals should not be more difficult than non-alternating plurals (given that SLI children have no difficulty producing voiced obstruents). SLI children might even be expected to perform better on alternating plurals than on non-alternating plurals, since no knowledge of the alternation is needed to memorise inflected forms, and alternating plurals are frequent. Such a scenario would then predict a high score for existing plurals but a low score on inflected non-words. Finally, a rule-based account would predict that the performance of SLI children on words should show a greater effect of lexical frequency than that of the control group, as they presumably store words as unanalysed wholes.

On the other hand, if children store all plurals but are still able to derive generalisations from them (as would be evidenced by their ability to inflect non-words), alternating plurals are expected to be more difficult for both groups of children. If both groups of children store alternating plurals alongside their singular and rely on analogy to form new words, unimpaired children are expected to be better at extracting generalisations from these words, resulting in a better performance on both types of words. As alternating words were shown to be more difficult than non-alternating words for typically developing children, this effect may even be enhanced for SLI children. Finally, analogical or usage-based models would predict that lexical frequency affects the performance of both groups of children. It is possible that TD children are even more sensitive to frequency, contrary to expectations of rule-based or Dual Mechanism models.

For non-words, predictions differ. Under a rule-based account, SLI children are expected to perform worse on inflecting non-words (compared to words), since they do not have access to a morphological rule. Also, if SLI children rely on explicit rule learning (i.e. add *-en* to create a plural), no alternations (e.g. [kɛt] ~ [kɛdən]) are expected for non-words. If occurrences of voicing alternations in lexical gap environments (e.g. [dep] ~ [debən]) are taken to reflect a phonological rule mechanism, they should not occur for this group of children. Furthermore, if such rules reflect a *phonological generalisation*, rules should apply across-the-board for both groups of children (i.e. consistently within a phonological context). On the other hand, if children are guided by analogy to existing words, typically developing children might be more sensitive to the distribution of alternations in Dutch. As proposed by some analogical models, the ability to form abstract generalisations on the basis of a lexicon may itself be more impaired in SLI, which means that they might show more unexpected alternations.

The present study mirrors the first study, in which plurals of both words and non-words were elicited, except that only a subset of items was used due to time constraints (as this task was part of a larger test battery presented to the children).

#### *5.2.4.3 Subjects*

Two groups of 5-year-olds participated in the experiment; a control group of 27 typically developing children (14 girls, 13 boys) with a mean age of 5;1 (61 months, sd 3 months) and a group of 24 children (4 girls, 20 boys) diagnosed with SLI with a mean age of 5;2 (62 months, sd 5 months). There were more boys in the group of SLI children, which is in line with findings that more boys than girls are diagnosed with language impairment (Leonard 1998). SLI children were recruited through speech therapists and *ESM* schools (*Ernstige Spraak en taalMoeilijkheden* 'serious speech and language difficulties) across the Netherlands. These are schools that provide full-time specialised education programmes for children with severe speech and language difficulties from 3 to 12 years. These children had been classified as language-impaired after extensive multidisciplinary assessment of their verbal and non-verbal abilities by certified speech pathologists. Furthermore, the usual exclusion criteria applied, in that children did not have a primary perceptual disorder, a neurological deficit or hearing problems, and reached IQ scores within the normal range (cf. Leonard 1998). Non-verbal IQ scores were measured through the SON-R (*Snijders-Oomen niet-verbale intelligentietest*, Snijders, Tellegen & Laros (1988)). SLI children had a lower mean IQ score on the SON-R than the control children (101 vs. 117), although both scores are above average. Furthermore, de Bree (2007) reports a difference in maternal education levels for the two groups, which could have affected children's vocabulary input and linguistic development (e.g. Hoff-Ginsberg 1991, Dollaghan et al. 1999). The control children were matched in terms of chronological age and were contacted via day-care centres in Utrecht. All children were monolingual native speakers of Dutch. For more information on subject selection and results concerning children at risk for dyslexia see de Bree (2007).

### *5.2.4.4 Materials*

The stimuli set consisted of 12 non-words, which was a subset of the 24 non-words described in 4.2.5.1. The stimuli set was restricted because of time limitations, as this task was part of a larger test battery in which both speech-language skills and IQ were tested. However, items still matched all six environments chosen for the study, yielding two items in each environment (e.g. [kɛt] and [slɑt], [dɑp] and [xɔp]). In the same contexts, 16 high-frequency words were chosen, 8 of which were non-alternating (e.g. [pɛt] 'hat', [kɪp] 'chicken') and 8 of which were alternating (e.g. [bɛt] 'bed', [krɑp] 'crab'). The most frequent words were selected in each category, to maximise the likelihood that the children would know them. The singular of the non-word was pre-recorded by a female native speaker of Dutch to ensure adequate and consistent presentation. Words and non-words were mixed and five fillers were added to the list, which included words which take an *-s* plural (*vlinder* 'butterfly', *paraplu* 'umbrella') and irregular plurals (e.g. *ei~eieren* 'eggs', *koe ~ koeien* 'cows'). The task was administered in two different orders to control for effects of fatigue and task novelty. A full list of stimuli is presented in Appendix II.

### *5.2.4.5 Procedure*

The task was administered in two different orders to control for effects of fatigue and task novelty. Children were tested in the language lab at the Research Institute for Language and Speech (UiL-OTS) at Utrecht University, or at their school. Testing took place in a quiet room. The task was fifth in a session that included other language tests as well as IQ measures. Plurals were elicited through the presentation of pictures in a PowerPoint slide show. For existing words, the child saw a picture of the object and had to name it. A second picture of the same object then appeared and the child had to complete the sentence "Now there are two ... ". For non-words, the child saw a fantasy animal and its name was presented through a loudspeaker. The non-words had been pre-recorded to ensure adequate and consistent production (including an audible release). The child had to repeat the name (until it was repeated correctly) and was then presented with the second image and asked to form the plural in the same way. Data were recorded on DAT (Tascam DA-P1) through a sensitive microphone (Crown PZM-185). Children's realisations of both the singulars and the plurals were converted into .wav files. Auditory transcriptions of all utterances and acoustic analysis of the children's plosives were then made by the two investigators (independently), and a portion of the data was judged by an additional 3 raters, with agreement reaching 90%. The items that were classified as voiced or voiceless by at least 4 out of 5 raters were used for further acoustic analysis, using the same method as in the previous experiment.

*5.2.4.6 Results*

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

**Words**

Results for non-alternating words are summarised in Table 29 below (comparable to Table 21 for Experiment I). As in the first experiment, other responses included missing responses (including diminutives such as [bɛtjəs] 'beds'*)*, bare stems, stem changes and *-s* plurals.

*Table 29: Results for non-alternating words in numbers (%), Exp. II.*

| | TD | SLI |
|---|---|---|
| Correct [pɛtən] | 167 (87.9%) | 116 (69.0%) |
| Voicing error *[pɛdən] | 9 (4.7%) | 6 (3.6%) |
| Bare stem | 10 (5%) | 30 (17.9%) |
| Stem change | 1 (1%) | 9 (5.4%) |
| S-plural | 1 (1%) | 2 (1.2%) |
| Missing | 2 (1.1%) | 5 (3.0%) |

These results show that TD children are very similar to the 5-year-olds in the previous study, who produced voicing errors of the type *[pɛdən] 'caps' in 4.6% of cases. However, this type of error was produced somewhat less often by SLI children. To investigate group differences in performance on words, the independent variable Group (TD, SLI) was entered in a Univariate ANOVA, which showed a main effect on score correct ($F_1(1)=8{,}322$, $p=.006$), $\eta_p^2=.145$). This means that the control group performed better on words than the SLI group, which was due to their increased number of bare stems. Also, SLI children made more stem change errors, involving substitutions and cluster simplifications (e.g. [kɛŋkə] for /tɛntən/ 'tents', [lape] for /lɑmpən/, [ofilatə] and [ohana] for /olifɑntən/ 'elephants'. In contrast, there was only one stem change in the TD group, i.e. [olifɑxtə] for /olifɑntən/ 'elephants' by Max. Results for alternating words are shown in Table 30 below (comparable to Table 22 for Experiment I).

*Table 30: Results for alternating words in numbers (%), Exp. II.*

| | **TD** | **SLI** |
|---|---|---|
| Correct [bɛdən] | 106 (49.1%) | 31 (16.6%) |
| Voicing error *[bɛtən] | 96 (44.4%) | 100 (53.5%) |
| Bare stem | 9 (4.2%) | 27 (14.4%) |
| Stem change | 2 (0.9%) | 15 (8.0%) |
| S-plural | 1 (0.9%) | 6 (4.3%) |
| Missing | 1 (0.5%) | 6 (3.2%) |

These data show that results for typically developing controls are again very similar to results of the first experiment, in which 5-year-olds produced voicing errors for alternating words in 41.4% of cases. However, SLI children show a higher number of voicing errors for alternating words (54%).

To investigate the performance on alternating words, an ANOVA on score correct was performed using Group (TD, SLI) as independent variable. There was a main effect of Group ($F_1(1)=29{,}629$, $p<.001$, $\eta_p^2=.377$), $F_2(1)=5{,}909$, $p=.022$, $\eta_p^2=.174$). Again, the SLI group performed worse on alternating plurals, which is due to an increased use of bare stems and stem change errors.[39] A total of fifteen stem change errors were made by SLI children, involving mostly deletion of /d/ after nasals (i.e. 9 cases of [hɑnə] or [hɔnə] for /hɑndən/ 'hands' or /hɔndən/ 'dogs'). As was argued in the previous chapter, such errors suggest that children were aware of intervocalic /d/, but may also point towards misperception. The remaining errors involved substitutions (e.g. [hɑŋkə] for /hɑndən/ 'hands', [pɔtnono] and [pɔtjojə] for /pɔtlodən/. For the control group only two stem change errors occurred, i.e.

[39] For instance, one child with SLI produces [pɔtlot] 'pencil' twice. When the experimenter points out that these are *potloden*, he replies "that's what I said", after which he produces [pɔtlot] again (showing that the unmarked form was intended as a plural).

[wɛtə] and [wɛə] for /wɛbən/ 'webs'. Interestingly, the two groups show a similar percentage of 'regularisation' errors (i.e. producing errors such as *[bɛtən]), with a slightly higher number for SLI children. Hence, even though SLI children are poorer at producing correct alternating plurals, they are capable of inflecting these words.

To assess the effect of word frequency, correlations were measured between performance on a word (% correct) and the frequency of its plural form in the CELEX database. Results show that correlations with frequency of alternating plurals are strong for the control group ($\rho$=.874, p=.005) and moderate for SLI children ($\rho$=.667, p=.071). Such a result is not expected if SLI children are more sensitive to lexical frequency effects. The frequency effect was due to the alternating plurals, as there was no correlation if only non-alternating plurals were taken into consideration. This is not surprising if alternating plurals are stored in the lexicon, and replicates results from the first elicitation experiment.

In order to compare children's performance on non-alternating and alternating words directly, the independent variables Group (TD, SLI) and Alternation (Non-alternating, Alternating) were entered in a Univariate ANOVA. The dependent variable (score correct on words) reflected the number of correct answers as a percentage of the total number of inflected words (i.e. counting only correct plurals such as [pɛtən] and [bɛdən] and incorrect plurals such as *[pɛdən] and *[bɛtən]). The results show a main effect of Group ($F_1(1)=19{,}484$, $p<.001$, $\eta_p^2=.17$), which means that TD children performed better on words than SLI children. There was also a large main effect of Alternation ($F_1(1)=301{,}275$, $p<.001$, $\eta_p^2=.76$), which means that alternating words were more difficult than non-alternating words for both groups. Finally, the interaction between Group and Alternation was also significant ($F_1(1)=22{,}154$, $p<.001$, $\eta_p^2=.19$), showing that alternating words were more difficult for SLI children than for TD children.[40] Recall that if SLI children would store all plurals as unanalysed wholes, they should perform equally well on both alternating and non-alternating words. In contrast, these results show that alternating words are relatively hard for SLI children, which is unexpected under a Dual Mechanism account.

---

[40] However, this interaction is not found when the score correct is calculated as a percentage of all responses (p=.127), which excludes the large number of unmarked responses. This means that result should be interpreted with caution.

In sum, both groups of children have more difficulty producing alternating words than non-alternating words, as found in the first experiment. This result is unexpected if SLI children store all plurals as wholes. Furthermore, SLI children do not show larger lexical frequency effects (supporting a 'lexical strategy') than TD children.

### Non-words

The overall results are summarised below (comparable to Table 23 for Experiment I).

*Table 31: Results for non-words in numbers (%), Exp. II.*

| | TD | SLI |
|---|---|---|
| Non-alternation | 270 (83.6%) | 148 (52.9%) |
| Alternation | 11 (3.4%) | 15 (5.4%) |
| Bare stem | 19 (5.9%) | 63 (22.5%) |
| Stem change | 9 (2.8%) | 33 (11.8%) |
| S-plural | 13 (4.0%) | 18 (6.4%) |
| Missing | 1 (0.3%) | 3 (1.1%) |

The performance on non-words shows that intervocalic obstruents in non-words were most often realised as voiceless by both groups (e.g. [slɑtən]), matching the value of the singular. Moreover, most children never produce any voicing alternations. However, voicing alternations (e.g. [slɑdən]) were produced by both groups. For the control group, alternations were produced in 3.4% of cases (11/324), or 3.9% of inflected *-en* plurals. Note that this is very similar to the 3.5% of alternations for non-words produced by the 5-year-olds in Experiment I. SLI children produced alternations for non-words more often, i.e. in 5.4% of cases (15/280), which is 9.2% of inflected *-en* plurals. In the TD group, alternations were produced by 6 out of 27 children (22%), whereas 7 out of 24 SLI children produced alternations for non-words (29%). Recall that in the first experiment, around a third of children in each age group produced alternations for non-words.

The independent variable Group (TD, SLI) was entered in an ANOVA, which showed that there was a main effect on the number of non-alternating forms ($F_1(1)=18{,}752$, $p<.001$, $\eta_p^2=.277$). This means that typically developing children produced more non-alternating forms than SLI children. Table 31 shows that this difference was mainly due to an increased number of bare stems for SLI children, as was found for words. To further assess the differences between words and non-words, a separate analysis was performed on the number of bare stem errors (which reflects whether children could form a plural). For this analysis, Stimulus (word, non-word) and Group (TD, SLI) were entered as independent variables in a Univariate ANOVA, with the percentage of bare stems as dependent variable. Results only show a main effect of Group ($F_1(1)=11{,}234$, $p<.001$, $\eta_p^2 = .10$, ($F_2(1)=87{,}036$, $p<.001$, $\eta_p^2 = .64$), i.e. SLI children produce more bare stems overall. The effect of Stimulus was only significant in an item analysis ($F_2(1)=6{,}397$, $p=.015$, $\eta_p^2 = .11$), which shows that bare stems tend to be produced for non-words. Finally, there was no interaction, although an item analysis approached significance ($F_2(1)=3{,}176$, $p=.081$, $\eta_p^2 = .060$). This means that both groups tend to produce more null markings for non-words, even though this tendency seems greater for SLI children. This result is surprising given that SLI children are expected to be poorer at non-word inflection.

Furthermore, SLI children produced more stem change errors, half of which (17/33) involved an incorrect singular (e.g. [bɛp] ~ [bɛpə] in response to /bɛmp/). The same substitution errors that were found for words were also found for non-word plurals (e.g. [klat] ~ [klakə]), as well as simplification errors (e.g. [tɪnt] ~ [tɪtə] for dint /dɪnt/) and place changes (e.g. [bɛmp] ~ [bɛntən], [slat] ~ [slɑbən]).[41] The group of TD children produced fewer stem changes for non-words, almost half of which (4/9) involved an incorrect repetition of the singular, e.g. [bok] ~ [bokən] in response to /bop/). The remaining errors involved glide formation (e.g. [knot] ~ [knojə]) and simplification (i.e. [dɪnt] ~ [dɪnə]). An analysis of the singular realisations (see de Bree 2007) revealed that errors predominantly consisted of consonant substitutions (e.g. [flɑnt] as [flɑmp]) and omissions (e.g. [flɑnt] as [flɑt]). Furthermore, SLI children produced more incorrect realisations of singulars than the control group. Moreover, even though all children produce more incorrect singulars for non-words than for words, this effect is stronger for SLI children. This suggests that SLI children have more difficulty forming novel lexical representations.

In the first experiment, five (typically developing) children were found to produce overgeneralisations for words or non-words in a specific

---

[41] The latter case could reflect substitution by an existing word (i.e. *slab* 'bib'), just as [bop] ~ [bomə] 'trees' and [kɪmp] ~ [kɪpə] 'chickens'.

environment, although none produced alternations consistently within a certain context. Recall that in the present study, there are only two non-words in each phonological environment. Still, results show that there was only one TD child who was consistent within a certain environment, i.e. Jade produced both [jɪdə], [slɑdə] and *[pɛdə] 'caps'. In the SLI group, Yorens produced both [flɑndə] and [dɪndə], and Jelle produced both [flɑndə], [dɪndə], *[olifɑndə] 'elephants' and [tɛndə] 'tents'. It thus seems that there is some evidence for 'rule-based' behaviour in both groups, which is not predicted if SLI children only rely on explicit rule learning. Let us now turn to the distribution of alternations for each group.

Even though the number of non-word alternations is comparable across groups, the distribution of voicing alternations is very different, see Table 32 below (comparable to Table 25).

*Table 32: Distribution of non-word alternations in numbers (%), Exp. II.*

| | **TD** | | **SLI** | |
|---|---|---|---|---|
| | t ~ d | p ~ b | t ~ d | p ~ b |
| V_ | 4 (7% ) | 0 | 0.0% | 0.0% |
| V:_ | 2 (4%) | 0 | 2 (4.3%) | 4 (8.7%) |
| N_ | 4 (7%) | 1 (2%) | 5 (10.9%) | 4 (8.5%) |
| Total | 10 (6%) | 0.6% (1) | 7 (5.0) | 8 (5.7%) |

Results for the group of typically developing (TD) children resemble those of the children in the first experiment, in that they produce more alternations for stems with /t/ than /p/. Also, both groups produce the highest number of alternations for /d/ after nasals, which replicates results from the first experiment. Recall that such /d/ alternations are expected to occur in this environment, as most alternating plurals are found after nasals (e.g. /hɑndən/ 'hands'). Surprisingly, SLI children produced slightly more /b/ than /d/ alternations. Moreover, SLI children produced more alternations in 'lexical gap' environments, both after long vowels and nasals. Recall that evidence for such 'lexical gap' alternations was only found in three cases of *p~b* alternations after long vowels by the youngest children in the first experiment. In the group SLI children on the other hand, these unexpected alternations are even produced after nasals. The lexical gap alternations are shown in Table 33 below.

*Table 33: Lexical gap alternations by TD and SLI children (Exp. II).*

| **TD** | | **SLI** | |
|---|---|---|---|
| Hanna (F) | [kɪmp] ~ [kɪmbən] | Yorens (M) | [bɛmp] ~ [bɛmbən] |
| | | Mark (M) | [bɛmp] ~ [bɛmbən] |
| | | | [dep] ~ [debən] |
| | | Luuk (M) | [bɛmp] ~ [bɛmbən] |
| | | | [dep] ~ [debən] |
| | | Mees (M) | [bop] ~ [bobən] |
| | | Abe (M) [42] | [dep] ~ [debən] |
| | | | [kɪmp] ~ [kɪmbən] |
| | | | [ap] ~ [abən] 'monkeys' |
| | | | [lɑmp] ~ [lɑmbən] 'lamps' |

If we take these unexpected alternations to reflect an early rule of intervocalic and postnasal voicing, this would mean that the behaviour of SLI children provides more evidence for such phonological rules or constraints than the behaviour of typically developing children, both of the same age and younger (i.e. Exp. I). Recall that most overgeneralisations of /b/ were also produced by the youngest children in the first experiment. Such alternations were argued to reflect phonetic naturalness, which would be in accordance with an account using phonetically grounded phonological constraints. SLI children produce even more unexpected alternations with /b/, which suggests that such constraints play a more important role. Even though SLI children may have more difficulty with morphological or morpho-syntactic patterns, it is possible that they are more sensitive to 'early' or natural phonological constraints than typically developing children, which means they are less sensitive to lexical patterns. However, this also suggests that the evidence for rules or constraints emerging from the current experiments seems to be of a different type than the categorical rules in traditional phonological models, as they appear to be phonetically driven. Moreover, as we have seen, children do not apply a voicing rule (if they have one) in a consistent fashion. On the one hand, the evidence for rule-like behaviour is clearest when alternations with /b/ are produced, as there is less support for

[42] Note that Abe might have been influenced by his own name, which contains a /b/ after a long vowel (i.e. /abə/)

such realisations on the basis of Dutch. On the other hand, productions of /b/ are also phonetically likely. Furthermore, they were typically judged to be voiced by only a third of raters. This suggests that there may have been an (additional) effect of phonetic implementation, which is stronger for the SLI children.

To further investigate whether there is a relation between children's word knowledge and the production of alternations for non-words, children were divided into a group which produced alternations for non-words and a group which did not. The factors Group (TD, SLI) and AltsNonwords (yes, no) were entered in a Univariate ANOVA with score correct on alternating words as dependent variable. This analysis shows main effects of Group ($F_1(1)=22{,}511$, $p<.001$, $\eta_p^2=.32$), as SLI children performed worse on alternating plurals. Secondly, there was an effect of AltsNonwords ($F_1(1)=14{,}266$, $p<.001$, $\eta_p^2=.23$), which means that children who produced alternations for non-words performed better on existing plurals. Finally, there was an interaction between the two factors ($F_1(1)=4{,}409$, $p=.041$, $\eta_p^2=.09$), which indicates that this effect was greater for SLI children. The relation between voicing in words and non-words is shown for the two groups separately, in Figure 8 and
Figure 9 below, comparable to Figure 5 and Figure 6 above (page 162 and 166).

*Figure 8: Correct voicing in words and voicing in non-words TD children (Exp. II).*

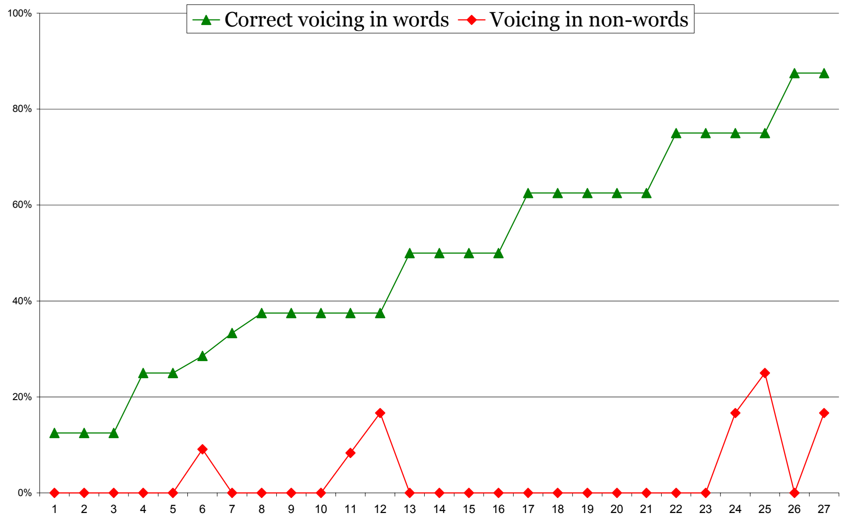
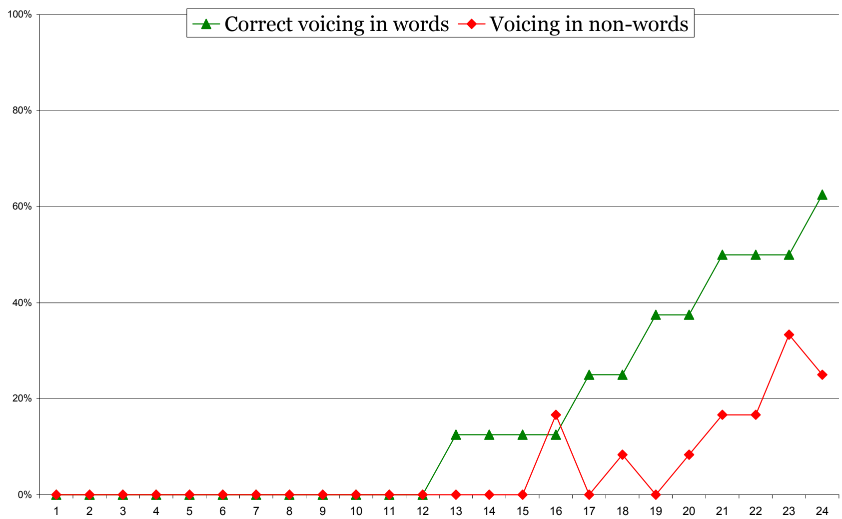



*Figure 9: Correct voicing in words and voicing in non-words SLI children (Exp. II).*



The first figure shows that the relationship between alternations in words and non-words is not very strong for typically developing 5-year-olds, which is confirmed by the lack of a significant correlation ($\rho$=.225, p=.131). Results for the SLI children show that children's alternations for non-words co-occurred with voicing in words, which is confirmed by a significant correlation ($\rho$=.779, p<.001). However, it is difficult to determine whether this result means that SLI children's alternations are based on their knowledge of existing words. After all, 'correct' voicing in words could have also been due to the overgeneralisation of voicing. If we look at words with /b/ separately, the only three children who produced [krɑbən] 'crabs' were those children who produced lexical gap alternations for /b/. If lexical factors influenced children's behaviour, it is clear that SLI children are not very sensitive to the distribution of voicing alternations. This could point towards an effect of poor generalisation, as children are less sensitive to effects of phonological environment. Furthermore, SLI children might have a more uniform treatment of words and non-words, which again reflects greater effects of phonological constraints.

In sum, results for non-words show that children predominantly produce non-alternating forms. As was found in the first experiment, the results support a general strategy of Paradigm Uniformity, as the rate of alternations is even lower than that expected on the basis of type frequency. However, voicing alternations for non-words are produced by both groups, and SLI children produce more unexpected 'lexical gap' alternations than the control group. Finally, both groups tend to produce more bare stems for non-words than for words, even though this tendency seems greater for SLI children.

To determine how the voicing contrast was realised, results of the acoustic analysis will be presented in the next section, before we turn to a discussion of the results.

### Acoustic Measurements

As in the first experiment, the auditory analysis was complemented by an acoustic analysis of the stimuli, to determine whether the perceived voicing contrast could be measured acoustically. The second goal of these measurements was to compare results to those of Experiment I, as well as compare the results of both the TD and the SLI group. This analysis only considered /t/ and /d/ for acoustic analysis, as there were too few items with *p~b* alternations. A total of 155 tokens of /t/ and 96 tokens of /d/ were measured in terms of closure and burst duration (in milliseconds), matched as closely as possible for items and child (i.e. contrasting tokens were taken from the same child and for items with the same rhyme as much as possible). Results are presented in Table 34.

*Table 34: Mean closure and burst duration for medial /t/ and /d/ in ms (sd), Exp. II.*

| | **Control** | | **SLI** | |
|---|---|---|---|---|
| | /t/ | /d/ | /t/ | /d/ |
| Closure duration | 104 (30) | 58 (16) | 124 (26) | 47 (11) |
| Burst duration | 42 (20) | 18 (11) | 30 (12) | 15 (6) |

The results in Table 34 show that /t/ is realised with a longer closure and burst duration than /d/, which is in line with the findings of Experiment I and previous findings by Kuijpers (1993ab). Results of a Multivariate ANOVA with Closure Duration and Burst Duration as dependent variables and Group (TD, SLI) and Auditory analysis (voiceless and voiced) as independent variables shows that both groups produce a reliable voicing contrast in terms of closure and burst duration. The effect of Auditory analysis is significant for both closure ($F(1, 300)=336.9$, $p<.001$) and burst ($F(1, 300)=125.5$, $p<.001$) duration. There is no effect of Group on closure duration ($p=.09$), although it approaches significance for burst duration ($p=.053$). However, an interaction was found between Group and auditory analysis for closure ($F(2,300)=5.9$, $p=.003$), but not for burst duration ($p=0.19$). This interaction is caused by the fact that SLI children show a longer mean closure duration for /t/.[43] The closure duration of the SLI group thus resembles those of younger typically developing children (see also Catts & Jensen 1983, Kuijpers 1993b).

### *5.2.4.7 Discussion*

The second Wug-test experiment involved the elicitation of plurals from words and non-words from a group of typically developing and language-impaired 5-year-olds. The results indicate that both groups of children are

[43] As was found in Zamuner et al. (2006b), the data do not support the idea that 'devoiced' realisations of underlying voiced segments (e.g. [t] in *[bɛten] 'beds') are more voiced than voiceless stops in correct realisations (e.g. [t] in [pɛten] 'caps').

able to use morphological rules (or generalisations) to inflect existing and novel words, even though SLI children used more bare stems. Importantly, SLI children showed a high ‘regularisation rate’ (i.e. errors such as *[betən] ‘beds’), and produced plurals for the majority of non-words. For both groups of children, alternating words (e.g. [bɛdən] ‘beds’) were more difficult than non-alternating words (e.g. [pɛtən] ‘caps’). Thus, it does not seem to be the case that SLI children differ from typically developing children by storing inflected words as unanalysed wholes. This is because they did not perform equally well (or better) on alternating (or ‘irregular’) forms. The poorer performance on alternating words does not seem to be due to an articulatory difficulty of producing voiced stops, as acoustic measurements established that both groups are able to make a reliable voicing contrast, matching the perceptual analysis.[44] Furthermore, as was found in Experiment I, there was a correlation between children’s performance on words and the frequency of the plural form, with typically developing children showing a higher sensitivity to the frequency of existing (alternating) plurals in Dutch. This result is not expected if TD children treat all plurals (both alternating and non-alternating) as regular.

The results further show that both groups of children produced voicing alternations for non-words, which is not expected if SLI children only rely on explicit rule learning (e.g. add *-en* to form a plural). In both groups, most of the alternations were produced for /t/ after nasals (e.g. [flɑndən]), which is expected given the distribution of alternations in Dutch. However, SLI children were found to be less sensitive to the distribution of alternations, as they produced more alternations in ‘lexical gap’ environments (e.g. [dep] ~ [debən]) than the control group. Recall that such alternations were also found for the youngest children in the first Wug-test, supporting the *phonological generalisation* hypothesis. As was found in the first experiment, there is evidence that children’s overgeneralisations are rooted in their knowledge of alternating words. Hence, children seem to base their overgeneralisations on the alternating words in their lexicon, which provides a schema for word formation (albeit an unproductive one). If this is the case, unexpected alternations could be caused by the fact that knowledge of the distribution of alternating patterns is not fully learned yet. Such an

[44] Items that were left out of analysis due to interrater disagreement are potentially relevant, as it could be the case that SLI children are less consistent in producing the contrast. However, most of the disagreement was actually due to the quality of the sound files used.

account would predict that SLI children are least sensitive to the distributional patterns, which is why they produce more unexpected alternations. Hence, both younger typically developing children and SLI children's behaviour may be influenced by phonological constraints leading to intervocalic or postnasal voicing. Results show that SLI children produce alternations for both coronal and labial obstruents after long vowels and nasals, showing that constraints target natural classes. However, as argued before, such constraints are probabilistic rather than categorical, as children did not behave consistently within environments (i.e. it was not the case that children who produced [debən] also produced [bobən]).

SLI children produced more stem changes, including cluster simplifications or substitutions for both singulars and plurals (e.g. *[hɑnə] for [hɑndən]). This shows that SLI children are aware of the underlying /nd/ cluster, as was argued in Chapter 4 for younger typically developing children (i.e. /nt/ clusters are not affected). This result also ties in with the finding that SLI children are more impaired in speech production, which might have an adverse effect on the subsequent acquisition of morphology. Hence, de Bree (2007) found that the same group of SLI children show impaired expressive phonology at a younger age (3 and 4), reflected by a lower *ppc* (percentage phoneme correct) score in picture naming and non-word repetition. Furthermore, 5-year-old SLI children showed poorer performance than age-matched controls on a rhyme oddity task, tapping phonemic awareness. Finally, perception studies with the same group of 5-year-olds showed that SLI children are worse at detecting phonemic mispronunciations (van Alphen et al 2004).

To conclude, it seems that we need to reformulate the notion of a linguistic rule to account for these data, taking the form of a frequency-based generalisation. As discussed in Chapter 2, these data could also be accounted for by models that allow rules or constraints to be probabilistic (e.g. Boersma & Hayes 2001, Albright & Hayes 2003). Crucially, in order to abstract generalisations, all children would have to be able to analyse (stored) inflected words. SLI children are then poorer at forming associations among these words and extracting generalisations from them. Thus, SLI children show poor performance on 'irregular' alternating plurals, poorer plural marking and unexpected alternations for non-words, which may be related to slow lexical learning and later and less efficient generalisation (Marchman & Bates 1994, Marchman, Wulfeck & Weismer 1999, Beckman & Edwards 2000, Montgomery 2002, Bates 2004). Even though vocabulary scores (i.e.

N-CDI scores) were not available for the group of SLI children, there is evidence from the literature that SLI children show slower lexical learning. For instance, a Dutch study comparing 136 8-year-old SLI children with 147 6-year-old controls revealed lexical-semantic differences, such as a smaller vocabulary size for SLI children (van Weerdenburg, Verhoeven & van Balkom 2006). Such an account would then predict that the effect of (phonetically motivated) phonological constraints is stronger in the absence of a well-developed lexicon. It thus seems that both probabilistic phonological models (e.g. Boersma & Hayes 2001) and usage-based models of language (Bybee 1985) are capable of explaining the aspects of children's behaviour discussed in this section.

Before we turn to a general discussion of these results, children's Wug-test behaviour will be compared to that of adult controls in the next section.

## 5.2.5 Elicitation Experiment III: Adults

### *5.2.5.1 Introduction*

In the previous experiments, children's behaviour was compared to the input data based on the ambient language. However, it is also interesting to compare these results to a group of adult controls. As discussed in previous chapters, adult Wug-test data have been gathered for Spanish diphthongisation (e.g. in [kɔntamɔs] ~ [kwɛnto] 'we/I count'), which revealed that not all possible phonological generalisations appear to be made by speakers (Kernan & Blount 1966, Bybee & Pardo 1981, Albright et al. 2001). Hence, diphthongised stems are more likely to be listed in the mental lexicon rather than derived from phonological rules. Furthermore, previous results show that Dutch speakers rely on analogy to existing words when they create past tense forms for novel verbs (Ernestus & Baayen 2003). In this study, subjects created past tense forms with *-de* for 24% of pseudo-verbs, in line with the probability that an 'underlying' final obstruent is voiced (see Chapter 3). However, adults have not been tested on plural formation with respect to voicing alternations.[45] Furthermore, previous results indicate that Dutch adults may overgeneralise both the *-en* and *-s* suffix (e.g. Baayen et al.

[45] Note that van Wijk (2007) found an overgeneralisation of voicing for one of her non-word stimuli for adults (i.e. *pleikoop* ~ *pleikoben*, which is a lexical gap alternation).

2002, van Wijk 2007), although for the present study only the -*en* plural is expected.

Before we turn to the current Wug-test experiment, let us first consider whether overgeneralisation errors occur in adult spontaneous speech. Recall that Dutch children make two types of errors, i.e. *[bɛtən] 'beds' and *[pɛdən] 'caps', even though the latter is rare in corpus data. Even though adults' errors have not been systematically studied, observational data suggest that both types of errors occur.[46] First of all, adults are sometimes unsure about the plural, leading to 'regularisation' errors such as **vingerhoeten* 'thimbles' and **vagebonten* 'vagabonds'.[47] Such errors were also observed for more common words, e.g. **oogpotloten* 'eyeliner', **spinneweppen* 'cobwebs' and **panten* 'buildings'. Overgeneralisations of voicing also occur, e.g. **linden* 'ribbons' and **kunstgebidden* 'dentures', showing that adults may occasionally produce both types of errors.

In the current study, adults were tested on the same set of non-words that was used for the two previous studies.[48] Importantly, adults are expected to be sensitive to the lexical distribution of voicing alternations, which predicts that 'lexical gap' alternations should not occur. In the previous experiments, children were shown to produce such unexpected alternations, which could reflect a strategy of phonological generalisation. By studying adult responses, it can be shown whether or not such alternations are unique for children. The first study was designed to compare adults and children directly, by eliciting a verbal response. A second study will be discussed in section 5.2.7, involving a written response.

[46] I am grateful to Judith van Wijk, Elise de Bree and Brigit van der Pas, who assisted me in my search for spontaneous voicing errors for adults. These errors were mainly produced on television shows and in spontaneous conversation between people that were unknown to the author, and did not result from any prompting or familiarity with the current research topic.
[47] This Paradigm Uniformity effect may also override orthography; the word *autoped* 'scooter' takes an -*s* plural but was erroneously inflected as **autopetten*. For some low frequent words both variants are considered correct (e.g. *boute* or *boude bewering* 'a bold statement').
[48] I am grateful to Eva Damen who tested subjects and assisted in designing the experiment. Results of the first two experiments are presented in her BA thesis.

### 5.2.6 Experiment IIIA: Verbal response

#### *5.2.6.1 Subjects*

A total of 42 subjects (33 female, 9 male) participated in Experiment IIIA, who were mostly students at Utrecht University (average age 20 years) and were paid for their participation.

#### *5.2.6.2 Materials*

The same set of 24 non-words was used for this experiment as in Experiment I (see Appendix I).[49] A set of 42 filler items was used, which were mono- or bi-syllabic non-words ending in a different final sound (e.g. *bree*, *peloem*, *bargel*, *flang*), see Appendix III for the full list of items and sentences. The items were presented in two lists, containing opposite orders.

#### *5.2.6.3 Procedure*

Subjects were seated in a soundproof booth in the phonetic lab at the Research Institute for Language and Speech (UiL-OTS) at Utrecht University. Recorded versions of the non-words were entered in a self-timed PowerPoint show, preceding sentences with blanks. Subjects were asked to produce plural and diminutive forms of non-words. After hearing the non-word, the subject was instructed to first repeat the stimulus in the first sentence (see example 1), after which the plural was produced in the second sentence (see example 2).

SOUND: [dɑp]

(1) *Ik zag een ________ op straat liggen.*
'I saw a ________ on the street'.

(2) *Verderop lagen nog een paar ________.*
'Further down there were a few other' ________.

[49] One subject noted that the item [jɪt] is in fact a low-frequent existing word *jid* 'jew'.

For half of the sentences, subjects were asked to provide a diminutive form (for filler items). This was done to divert attention from the plural alternation. Fillers were chosen in such a way that diminutive formation was not straightforward (sometimes two forms are possible, such as *peloempje* and *peloemetje*). Participants were instructed to read the entire sentence aloud. The experiment was not self-paced, i.e. the sound and the two sentences appeared within a time period of 8 seconds, which promoted an automatic and rapid response. Note that upon hearing a non-word, subjects did not know whether they would be required to produce a plural form or a diminutive. Only after reading the second sentence aloud, subjects could determine the required response. A pilot study showed that subjects had no difficulty completing the task. For a list of sentences see Appendix III. The items were presented in pseudo-randomised orders in two different lists. Responses were recorded using a DAT-recorder Aiwa HD S100 and a Sony microphone ECM MS957 and transferred into *Praat* for further analysis.

#### *5.2.6.4 Results*

As in the previous experiments, non-words were scored as either 'Non-alternating' or 'Alternating' (corresponding to non-word plurals with a voiceless or voiced medial obstruent) or belonging to one of the other four response categories (bare stems, stem changes, -*s* plurals and missing responses).

Results for Experiment IIIA show a striking resemblance to the results of the previous experiments. Again, most non-words were inflected without voicing alternations, i.e. 71% of the total. However, around a third of adults (33% or 14/42) produced overgeneralisations of voicing for non-words (e.g. [kɛt] ~ [kɛdən]), in a total of 2.3% of cases (23/1008) or 3.1% of inflected -*en* plurals (23/736). Recall that this percentage was 3.1% for the children in Experiment I (see section 5.2.3.4), and 3.4% for the group of 5-year-old controls from Experiment II (see section 5.2.4.6). The results are shown in Table 35 below, together with results from Experiment I.

*Table 35: Non-word alternations in numbers (%), Exp. I and Exp. III.*

| | **Children** | **Adults** |
|---|---|---|
| Non-alternation | 1055 (75.8%) | 713 (70.7%) |
| Alternation | 43 (3.1%) | 23 (2.3%) |
| s-plural | 23 (1.7%) | 243 (24.1%) |
| Bare stem | 116 (8.3%) | 1 (0.1%) |
| Stem change | 9 (0.6%) | 3 (0.3%) |
| Missing | 146 (10.5%) | 25 (2.5%) |

These results show that children produced slightly more voicing alternations than adults, although overall numbers are too low to investigate this difference statistically. To compare the results for adults to those of the children from Experiment I, the independent variable Group (3-year-olds, 5-year-olds, 7-year-olds, adults) was entered in a Univariate ANOVA, with the number of non-alternating forms as dependent variable. Results show a main effect of Group ($F_1(3)=5,275$, $p=.002$, $\eta_p^2=.14$), $F_2(3)=49,792$, $p=.000$). A post-hoc analysis showed that adults differed from the 7-year-olds (Tukey, $p=.006$), as the latter group produced more non-alternating forms (i.e. 96%). This difference was due to the large number of -*s* plurals produced by the adults.[50] Table 35 shows that there were few other responses.[51]

With respect to the distribution of voicing alternations, adults show a different pattern. Recall that six items (i.e. one per phonological environment) were chosen to resemble mono-morphemic items (e.g. [xlop] ~ [xlobə] due to globe 'globe'). The results show that in contrast to children, adults seem to have been influenced by this manipulation. The number of alternations was higher for four out of the six analogous items (4.4% or 11/252) than for the seven non-analogous items (1.6% or 12/756).[52] Many

---

[50] Recall that 5-year-olds produced –*s* plurals for 5% of the non-words in wug-test I. In contrast, adults produced -*s* plurals in 24% of cases. Some adults almost exclusively used -*s* plurals, e.g. one subject produced –*s* for 20 out of 23 non-words.

[51] Errors included an existing word substitution ([flɑnt] ~ [flɑŋkən] 'flanks') and one doubly marked plural ([dɪnt] ~ [dɪntsən]), as well as a stem change ([knot] ~ [knoɣən].

[52] A total of 33 additional subjects took part in a pilot experiment, which was carried out to test the procedure and determine whether adults would discover the aim of the experiment. In this study, a subset of 12 test-items was used, excluding the items that are analogous to existing simplex forms (e.g. /xlop/, which resembles *globe* 'globe') and the items that most resemble existing words (e.g. /jɔnt/, which resembles /hɔnd/ 'dog'). The procedure was identical to the experiments described above, and 42 filler items were used. Because of the subset of items used, the pilot study could be used to determine whether non-word alternations would be produced in

alternations were also produced for the item /jɔnt/, probably due to its high similarity to the alternating noun /hɔnd/ 'dog'. Before we discuss this issue further, let us turn to the distribution of alternations for each of the phonological environments (comparable to Table 25 on page 158).

*Table 36: Distribution of alternations for non-words in numbers (%), Exp. III.*

| | **t ~ d** | **p ~ b** |
|---|---|---|
| V_ | 0 | 8 (4.8%) |
| V:_ | 3 (1.8%) | 5 (3.0%) |
| N_ | 7 (4.2%) | 0 |
| Total | 10 (2.0%) | 13 (2.6%) |

This table shows that adults tend to produce t~d alternations after nasals, which is the environment in which these alternations are most frequent. However, (some) adults also produced voicing alternations in the 'lexical gap' environment of /b/ after long vowels (five subjects).[53] As in Experiment I, no alternations for /b/ after nasals were observed. Only two out of the five unexpected 'lexical gap' cases resulted in an existing mono-morphemic word (i.e. [xlop] ~ [xlobən], similar to *globe* 'globe'). Hence, the effect cannot solely be accounted for by the fact that these alternations resulted in existing words. As was argued for children, adult Wug-test behaviour can be accounted for when neighbours such as *deed ~ deden* 'did' or *neef ~ neven* 'nephews' influence subjects' behaviour on non-words with long vowels, predicting *p~b* alternations after short and long vowels but not nasals. When results are compared to the input based on the CELEX type frequency, adults also overextend voicing for /b/ after short vowels. This means that even though the occurrence of such alternations is unexpected, it does not reflect a child-specific effect. Still, adults produced slightly more /b/ alternations than /d/ alternations, although the difference is not significant. Finally, adults were not consistent in a certain phonological environment. This is

the absence of these 'triggering' non-words. Results of this pilot study showed that only three non-word alternations were produced by adults (0.8% or 3/396). This thus provides support for the claim that adults were sensitive to the fact that some non-words resembled existing simplex forms. Note that -*s* plurals were produced in 23% of cases (92/396).

53 Lexical gap alternations were produced twice for /xlop/, twice for /bop/ and once for /tap/.

illustrated in (40) below for the items /dɑp/ and /zwɑp/, yielding seven different combinations (excluding diminutive forms).

*(40)*

| | | | |
|---|---|---|---|
| [dɑpən], [zwɑpən] | (n = 14) | [dɑbən], [zwɑpən] | (n = 1) |
| [dɑps], [zwɑbən] | (n = 8) | [dɑbən], [zwɑbən] | (n = 1) |
| [dɑpən], [zwɑps] | (n = 2) | [dɑps], [zwɑpən] | (n = 1) |
| [dɑpən], [zwɑbən] | (n = 2) | | |

To conclude, adults' behaviour in a Wug-test was similar to that of children, producing mainly non-alternating forms. The main difference with children is their increased use of -*s* plurals, which might be due to an influence of English on their speech. Adults also produced voicing alternations for non-words, but they were more influenced by the existence of analogous mono-morphemic words (such as *zwabber* 'mop' or *globe* 'globe'). This 'word' effect was absent for children as these words are of low frequency. Importantly, adults were also found to produce alternations in a 'lexical gap' environment (i.e. /b/ after long vowels). Such unexpected alternations were also produced by children in the first two experiments, providing some support for the claim that children's behaviour in a Wug-test reflects phonological rather than analogical overgeneralisations. However, the fact that adults show the same behaviour may weaken this claim.

The final experiment was carried out to determine whether the same results could be obtained by eliciting a written response. Also, the effect of word class was assessed by repeating the experiment with novel verbs instead of nouns.

## 5.2.7 Experiment IIIB: Written response

### *5.2.7.1 Subjects*

A group of 64 university students (50 female and 14 male, average age 20 years) participated in Experiment IIIB-1 (elicitation of noun plurals) and were paid for their participation. A different group of 47 (43 female, 4 male, average age 20 years) university students participated in Experiment IIIB-2 (elicitation of verb plurals).

### *5.2.7.2 Materials*

Materials were the same as in the previous experiment (section 5.2.6.2). Due to the set-up of the experiment, items were always presented in the same order.

### *5.2.7.3 Procedure*

Subjects were tested while seated in a classroom at the end of a class. The non-word items were played to the students using the multimedia system available in the classroom, after which they wrote down an inflected form on the sheet in front of them. Each sentence was numbered, and a PowerPoint show automatically played each recorded item, showing only the item number (i.e. there was no visual presentation of the singular). After 3 seconds, the next non-word was played. The subjects were asked to either write down a noun plural or a diminutive, according to the sentence. Hence, there were also sentences that prompted a diminutive (see examples below).

SOUND: [dɑp]

(1) *Gisteren zag ik ineens twee* _________.
'Yesterday I suddenly saw two' _________.

SOUND: [pelum]

(2) *Mijn moeder heeft een klein* _________.
'My mother has a small' _________.

A total of 66 sentences were used for the set of 24 non-words and 42 filler items. The non-word test items were always inflected as plurals, the remaining 42 items were divided into plurals (9) or diminutives (33), such that half the sentences required a plural (24+9=33) and the other half required a diminutive (33). Subjects were asked to respond quickly and to write down the whole word rather than the suffix. In the second experiment, sentences elicited either the verbal plural or infinitive form or the past participle (see examples below).

SOUND: [dɑp]

(1) *Ik ga morgenochtend* _________.
'Tomorrow morning I am going to' _________.

SOUND: [pelum]

(2) *Eigenlijk heb ik nog nooit* _________.
'Actually I have never' _________.

There were two different versions, in which non-words were paired with different sentences to control for possible effects of the carrier sentence.

### *5.2.7.4 Results*

The results of Experiment IIIB are very similar to those of Experiment IIIA, although the number of voicing alternations (or plurals spelled with 'b' or 'd') is slightly lower, i.e. 2.1% of novel nouns (32/1536). Again, alternations were produced by around a third of subjects, i.e. 27% (17/64). The number of voicing alternations is even lower when subjects had to inflect novel verbs, i.e. 1.6% of cases (18/1128). Also, alternations were produced by fewer subjects, i.e. 15% (7/47). The results are summarised below.

*Table 37: Non-word alternations in numbers (%), Exp. III (written).*

| | **Nouns (n=64)** | **Verbs (n=47)** |
|---|---|---|
| Non-alternation | 1231 (80.1%) | 1028 (91.1%) |
| Alternation | 32 (2.1%) | 13 (1.2%) |
| s-plural | 182 (11.8%) | n.a. |
| Bare stem | 8 (0.5%) | 15 (1.3%) |
| Stem change | 30 (2.0%) | 40 (3.5%) |
| Missing | 53 (3.5%) | 32 (2.8%) |

A comparison shows that alternations are more common when nouns are elicited, which is in line with the finding that voicing alternations are most common for nouns (see table 2 in Chapter 4). Lexical gap alternations were also produced, for both nouns (three times *boben*, once *globen*, *deben*, *taben* and *bemben*) and verbs (twice *boben*, once *globen* and *tomben*). Note that most adults did not realise that the experiment was about voicing alternations, and were often not aware of the alternation.[54] Other responses included spelling errors (e.g. /xɔp/ - *gobjes*, /tap/ - *taabs*, /klat/ - *klaads*),

[54] For instance, when the experimenter asked a subject (after testing) why he had answered *jonten* and not *jonden,* he answered: "but then it becomes a different word!"

but these were not considered alternations. There were some inconsistent spellings possibly reflecting vowel alternations (e.g. /bop/ ~ *boopen* or /slɑt/ ~ *slaten*), stem changes (e.g. /bɛmp/ - *bebben*, /tɛp/ ~ *tedden*, /xɪnt/ ~ *ginderen*), as well as real word substitutions (e.g. /zwɑp/ ~ *zwakken* 'weak', /klat/ ~ *kwaden* 'evils').[55] Some errors revealed an influence of subjects' familiarity with English plurals (e.g. /fet/ ~ *faiths, /*tɛp/ ~ *taps*, /zwɑp/ ~ *swaps*, /bop/ ~ *boobs*). This suggests that the high number of *-s* plurals was (partly) due the influence of English. Finally, there are some differences between the two response modes, e.g. there were fewer *-s* plurals for nouns when responses were written down.[56]

### 5.2.8 Discussion

Results for Experiment III indicate that adults produce fewer overgeneralisations than children, even though the difference was not statistically significant. Furthermore, even adults occasionally produce 'lexical gap' alternations after long vowels (but not nasals).[57] Moreover, adults seem to be influenced by analogy to mono-morphemic items such as *globe* 'globe', which may point to the existence of product-oriented schemas (as predicted by usage-based models, e.g. Bybee 2001). Hence, even though voicing alternations are not expected to occur for /xlop/ on the basis of Dutch, the existence of mono-morphemic /xlobə/ 'globe' may have promoted the alternation.

The wug-test results suggest that adults also produced more /b/ alternations than expected on the basis of alternating nouns in Dutch, as was found the previous experiments. This might partly reflect the fact that intervocalic obstruents are phonetically slightly voiced. Such an account would tie in with findings by Ernestus (2000:169), who shows that

---

[55] There were also some doubly marked plurals (e.g. /jɪt/ – *jittens*, /dɑp/ – *dabbers*) and cases of zero marking, which is allowed when items are interpreted as mass nouns (i.e. *Bijna niemand heeft meer dan drie boop* 'almost no one has more than three *boop*').

[56] To test for effects of spelling, 19 additional subjects were tested in a separate experiment, in which the written form of the non-word was provided (spelled with *t* or *p*) as well as the auditory form. Hence, the non-word was provided in the margin of the answer sheet (e.g. *gloop*). This manipulation resulted in only 4.2% of *-s* plurals (19/456), produced by 4 out of 19 subjects and no alternations. This again suggests an influence of English (i.e. when the non-word is only heard, the speaker may be more likely to assign a loanword status to it).

[57] However, as noted in Chapter 4, the gap after long vowels is not a true lexical gap for adults, since there are low-frequent adjectives like *xenofoob* [p] ~ *xenofobe* [b] 'xenophobe'.

intervocalic obstruents tend to be realised and perceived as voiced in fast speech.[58] There is some evidence that speakers have access to such phonetic constraints (Fleishhacker 2005, Zuraw 2005), and such an account would explain why /b/ is more affected than /d/. However, it is unclear why adults would be more sensitive to phonetic conditioning than children, as they are arguably more sensitive to the distribution of alternations in Dutch. Apart from differences in the size of the lexicon, adults may also differ from children in that have access to an orthographic representation (see Keuleers et al. 2007).

The most striking difference with the child data is that adults produced more *-s* plurals for non-words. This may have been due to an influence of English, as subjects may have treated some non-words as English loanwords. In that sense, *-s* may have a 'default' status, which is supported by the observation that the use of the *-s* suffix is increasing in Dutch. As Keuleers et al. (2007) point out, the *-s* plural allows for maximal stem conservation (i.e. it does not affect prosodic structure), and is not ambiguous (i.e. it is only a plural suffix whereas *-en* is also used in the verbal domain). Finally, note that these results are not in line with a Dual Mechanism account (e.g. Pinker 1999), which would not predict the attested variability. Rather, current results support previous findings for Dutch pluralisation (see Baayen et al. 2001, van Wijk 2007 and Keuleers et al. 2007).

The Tilburg Memory-Based Learner or TiMBL (Daelemans et al. 2004) has been used to successfully model similar data on the basis of CELEX (see Keuleers et al. 2007 and Keuleers & Daelemans in progress for recent results concerning Dutch plural suffix selection). This analogical model contains a *learning* component (which is memory-based) and a *performance* component (which is similarity-based). To determine its success on the present data, a simulation study of the wug-test results was carried out (see Keuleers & Daelemans in progress).[59] Here, all possible responses were taken into account, i.e. *-s* plurals, unvoiced or voiced responses. In the memory-based learner, a lexical entry such as /pɛt/ can be assigned the feature values /p/, /ɛ/, and /t/, with the number of

[58] Neijt et al. (2006) found that adults are more tolerant towards incorrect voicing (**olifanden*) than incorrect devoicing (**liefte*), irrespective of morphological complexity.
[59] I am grateful to Emmanuel Keuleers from the University of Antwerp for running a simulation study on the wug-test data and discussing the data with me.

mismatching features determining the distance between its nearest neighbours. This distance can then be modified using a *modified value difference metric* (Cost & Salzberg 1993), which allows for the computation of a more fine-grained, *graded* similarity. Also, the model weighs features according to their information gain, or *gain ratio* (Quinlan 1993). In the current study, analogy was defined on the basis of the full feature set of the nearest neighbours of the non-word stimuli (at the 11 nearest distances). Appendix B shows the results of two simulations, the second of which also took into account the final grapheme of the stimulus. This is because the recorded singular non-words were read aloud on the basis of stimuli spelled with final 't' or 'p'. It is possible that the auditory form co-activated orthographic information (which would favour the voiceless candidate based on probability). Also, recall that incomplete neutralisation effects have been found for Dutch (Warner et al. 2004, Warner et al. 2006). For the present stimuli, the signal is likely to have contained phonetic detail corresponding to a voiceless final sound. If information about the final grapheme is taken into account, the model does not predict any voiced responses, whereas the model without final grapheme predicts only one voiced response (i.e. for *flant*). In Appendix B, feature weights are shown as well as a list of neighbours for some of the non-words, showing that neighbours with the highest token frequencies (e.g. *hand* 'hand') are not necessarily taken into account, as the model only uses type frequency. This list also shows that non-words with identical rhymes may have different sets of neighbours. Results show that the memory-based learner matches the response that the majority of subjects gave, for both children (for 23 out of 24 non-words) and adults (for 22 out of 24 non-words). Hence, the model succeeded in modelling plural formation in Dutch, in line with earlier findings by Daelemans et al. (2007). This result can be taken as an indication that an analogical learning algorithm based on nearest neighbours is viable. However, note that bare responses in the child data cannot be reliably modelled, as they do not occur in adult Dutch. Hence, results for the child data should be interpreted with caution. Moreover, the type of intra-word variation found in children's realisations (e.g. voiced and voiceless realisations of the same word in a single session) is possibly problematic for the MBL, although task effects are likely to play a role.

## 5.3 General discussion

In this chapter, elicitation experiments with children and adults were described, which were carried out to address the question of how and when voicing alternations are acquired. First, the general results for existing words will be discussed, after which we turn to the hypotheses formulated earlier. The elicitation experiments showed that alternating plurals (e.g. [bɛdən] 'beds') are not fully acquired by the age of 7, as even the oldest children produce these plurals correctly in under 60% of cases. Results further show that children's performance on alternating plurals correlates with the frequency of the plural form, indicating that they may be stored. Furthermore, a higher percentage of 'devoicing' errors for alternating words (e.g. *[hɔntən] 'dogs') was found than in corpus data. Recall that around a third of medial obstruents was devoiced in the CLPF database (both in mono-morphemic and bi-morphemic words), see Chapter 4. The present studies show an error percentage of over 40% for all age groups (with the exception of SLI children, whose error rate is over 50%). This is in line with earlier findings that show higher rates of overgeneralisations for elicitation experiments.

Let us now turn to the hypotheses as formulated in section 5.2.2. The *first* hypothesis was that children are guided by *Paradigm Uniformity*, which would predict that no alternations occur (i.e. *[pɛdən] 'caps' and [kɛt] ~ [kɛtən]). Similarly, theories of lexicon optimisation (Prince & Smolensky 1993) would not predict the occurrence of alternating forms (see Chapter 2). However, such forms were attested for all groups of children, ranging from 3-5% for words and 2-5% for non-words. Moreover, the youngest children did not produce a greater number of non-alternations (reflecting Paradigm Uniformity) than older children. The weaker version of this prediction is borne out, as children produce fewer alternations than expected on the basis of their input (around a third of all words with /t/ alternate, whereas children produced *t~d* alternations in around 5% of cases). On the other hand, the fact that adults produced fewer non-word alternations (around 1–2% of cases) indicates that children may be guided by phonological factors. Alternatively, children's non-word alternations indicate that the number of alternating plurals is relatively high in 5-year-olds' lexicons, which provides a basis for overgeneralisations. In a similar vein, MacWhinney & Skousen (1994) have pointed out that errors such as "yesterday it **snew*" for *snowed* would be rare for any adult speaker, but have been observed to occur for children in the early stages of lexical learning (i.e. when the irregular form 'knew' enters the lexicon).[60]

The general lack of productivity of voicing alternations can be accounted for in analogical or usage-based models, given the fact that the alternating pattern has a lower type frequency and a higher token frequency than the non-alternating pattern. Hence, the outcome of the present experiments mirrors results obtained for Spanish diphthongisation, which was found to be lexically specific (see e.g., Bybee & Pardo 1981, Albright, Andrade & Hayes 2001). As shown in Chapter 4, singulars of alternating pairs are not only relatively frequent, they are also more often realised as diminutives. Furthermore, there are a few highly frequent alternating plurals that dominate the child's input (e.g. *handen* 'hands'). This pattern may reflect the fact that the singular and plural form have weaker connections in case of alternations. The difference in frequency between the singular and plural as well as the non-transparent nature of the alternation may contribute to the relative 'autonomy' of the plural. The corpus data discussed in Chapter 4 further showed that children do not receive much evidence for alternating noun pairs in the input. Hence, it may be difficult to extract the alternating pattern on the basis of morphologically related forms, which is aggravated by the fact that plurals are relatively independent from singulars (due to their high frequency).

When alternating forms were produced for non-words, they did not occur randomly (i.e. providing evidence against the *second* hypothesis of *random selection*), as an effect of type of obstruent and rhyme was found.

The *third* hypothesis stated that children are guided by *phonological generalisations*, predicting that constraints such as intervocalic or postnasal voicing would influence children's behaviour. Such constraints can also be phonetically grounded (e.g. Stampe 1973, Kager 1999a, Hayes 1999). Importantly, errors such as *[pɛdən] 'caps' would be predicted to occur for the youngest children. However, such errors occurred more often for 5-year-olds than for 3-year-olds, which means that they do not necessarily reflect an early effect. This finding ties in with the fact that no such 'voicing' errors were found in early corpus data (see Chapter 4). The occurrence of alternations in lexical gap environments (e.g. *[abən] 'monkeys' and [tap] ~

[60] The observation that the occurrence of alternations is related to mastery of alternations in words may also be compatible with constraint-based models, under the assumption that overgeneralisations are 'triggered' by alternations in the input.

[tabən]) provided further evidence for phonological generalisations. These errors occurred both for the youngest children and SLI children, which suggests that phonological generalisations play a role when children are least sensitive to lexical distributions. Interestingly, children with SLI were found to produce many more lexical gap alternations than typically developing children, even extending /b/ to the nasal environment (e.g. [kɪmp] ~ [kɪmbən]). This suggests that these children are especially poor at generalising over their lexicon. However, subjects did not show consistent (rule-like) behaviour within a certain phonological environment (e.g. for the items [jɪt] and [mɪt]), which shows that phonological constraints are probabilistic rather than deterministic. Hence, results could be handled by some recent models such as the one proposed by Boersma & Hayes (2001).

The *fourth* hypothesis predicts an influence of *analogical generalisations,* which means that children are sensitive to lexical statistics. Regarding this hypothesis, the results lead to a number of possible conclusions. First, the hypothesis correctly predicts that errors (e.g. *[pɛdən] 'caps') and alternations for non-words (e.g. [kɛdən]) occur for stems with final /t/ rather than /p/, mirroring the likelihood of voicing based on the lexicon (see Chapter 4). Alternations tend to occur in the phonological environment for which there is most evidence in the input. Hence, children produce mostly *t~d* alternations after nasals (e.g. [flɑnt] ~ [flɑndən]), reflecting frequency rather than natural classes. Also, the rate of voicing alternations for non-words produced by children in a Wug-test showed a strong correlation with the CELEX probability of voicing, based on monosyllabic words with identical rhymes. This shows that children are sensitive to the probability of voicing in Dutch, as found by Ernestus & Baayen (2003) for adults. Results further indicated that both type and token frequency play a role, which is expressed in recent connectionist models of morphological processing (Moscoso del Prado Martín, Kostic & Baayen 2004). For instance, Moscoso del Prado Martín, Ernestus & Baayen (2004) trained a neural network model only on tokens, which closely simulated subjects' responses on novel Dutch past tense inflection. Finally, it is unclear to what extent factors such as stress and word class played a role. For instance, a simulation study showed that Dutch word class (N, V, A) is predictable on the basis of segmental information (Durieux & Gillis 2000). Hence, it is possible that children's expectations about word class may have interfered with plural formation.

Second, overgeneralisations and alternations for non-words did not occur more often for 3-year-olds than for 5-year-olds, contrary to expectations of phonological models. Rather, the occurrence of non-word alternations was shown to increase with word knowledge (i.e. performance on existing alternating plurals). In sum, the results indicate that the lexicon is more important than phonological constraints in accounting for children's overgeneralisations in acquisition, although multiple factors are likely to play a role.

The low productivity of the alternating pattern combined with the high token frequency of alternating forms suggests that alternating plurals in Dutch are likely to be stored in the lexicon rather than derived by rule. Such a view is in line with usage-based models of language, in which there is storage of all forms (i.e. alternating and non-alternating pairs), while lexical connections provide internal structure. Thus, as was found for Spanish diphthongisation, the voicing alternation may reflect a lexically specific pattern rather than a rule that is independent of the particular forms it describes (Bybee 2001:103). Moreover, children's performance on existing plurals correlated with the frequency of the plural form, lending further support for this claim. Recall that voicing alternations in Dutch are traditionally assumed to be regular, with an abstract underlying form (e.g. /bɛd/ 'bed') which undergoes final devoicing. There is as yet no evidence that young children restructure the underlying form /bɛt/ on the basis of an early rule of final devoicing. However, the Wug-tests discussed in this chapter have only considered children's formation of plurals on the basis of a novel singulars. Deriving singulars from newly heard plurals seems a straightforward task, given the fact that children may have knowledge about final devoicing. Such knowledge could be based on phonotactic knowledge gained in infancy, or even result from innate constraints. To test whether children have knowledge of alternations in relation to final devoicing, 'reverse' Wug-test experiments were performed, in which singulars were elicited from novel plurals. These experiments will be discussed in the next chapter.

# Chapter 6 Comprehension Experiments

Dutch children's productive knowledge of alternations was tested in a 'standard' Wug-test, described in the previous chapter. In a second series of experiments, children's comprehensive knowledge of voicing alternations was assessed. In a 'reverse' Wug-test, children were asked to respond to novel alternating plurals. In the first set of experiments (6.2), children were asked to form singulars of novel plurals (e.g. [kɛtən] or [kɛdən]). If the child has internalised knowledge of final devoicing in combination with alternating forms, this should lead to the same singular form in both cases (i.e. [kɛt]). Three separate experiments are discussed, involving a joint study with 2- to 4-year-old children (6.2.1), Experiment IV with 3- to 6-year-old children (6.2.2) and Experiment V with adults (6.2.3). In section 6.3, two additional comprehension experiments are discussed. In an earlier joint study using a 'picture selection' experiment (6.3.1), children were presented with novel alternating plurals and given a choice between two pictures. Finally, Experiment VI was an 'acceptability' experiment (6.3.2), in which a novel singular was offered to children (e.g. [kɛt]), who were then asked to decide whether a novel alternating plural form was appropriate, which could take the form of voicing alternation (e.g. [kɛdən]) or a place alternation (e.g. [kɛpən]), which does not occur in Dutch. A general discussion of the results will be provided in section 6.4.

## 6.1 Introduction

The current chapter focuses on comprehension experiments to investigate Dutch children's knowledge of voicing alternations. First, the phonological perspective on the acquisition of alternations will be discussed, focusing on recent constraint-based models (in which the link between phonotactics and alternations is made explicit). As discussed in Chapter 2, a final devoicing rule is traditionally considered to neutralise abstract underlyingly voiced obstruents (e.g. /d/ in /bɛd/ 'bed') in word- or syllable final position (Trommelen & Zonneveld 1979, Booij 1995). Dutch phonotactics (e.g. "voiced obstruents do not occur in final position") could be said to enforce the phonological process of final devoicing, which in turn produces a morpho-phonological alternation. Generally, children are thought to first acquire knowledge of contrast and phonotactics (which does not require knowledge of how specific lexical items vary), before they acquire morpho-phonological alternations (Hayes 2004). Phonotactics are learned on the basis of positive evidence from the input, without access to morphology and with a bias for the most restrictive grammar. As mentioned in Chapter 2, phonotactic knowledge is predicted to aid the subsequent acquisition of morpho-phonological alternations (Prince & Tesar 1999, Hayes 2004, Prince & Tesar 2004, Tesar & Prince 2004). This is a logical consequence of the fact that phonology is conspirational (i.e. phonotactics and alternations are derived from a single set of constraints). For instance, 'soft' phonotactic knowledge that geminates such as [dd] are not allowed at the word edge might lead the English learning child to discover a categorical constraint that filters out *[nidd] for *needed* (e.g. Albright & Hayes 2003).

Previous studies further show that phonotactic and distributional information can be used for segmenting continuous speech (Brent & Cartwright 1996, McQueen 1998) and detecting allophonic variation (Jusczyk, Hohne & Bauman 1999, Peperkamp & Dupoux 2002, White et al. 2006, Peperkamp, Rozenn Le Calvez & Dupoux in press). Furthermore, experimental research suggests that it is easier for English speakers to learn an artificial language with an alternation that reflects English phonotactics than one with an alternation that does not (Pater 2002, Pater & Tessier 2003; 2005). However, it is not clear whether such results (partly) reflect the relative salience of contrasts or the 'naturalness' of alternations (Beckman

2003). For instance, phonetic naturalness has been shown to play a role when the acquisition of alternations was studied using an artificial language paradigm, although results showed that unnatural patterns are learnable (see, e.g., Peperkamp, Skoruppa & Dupoux 2006 and references therein). Note that when the role of phonotactics is central, children's implicit knowledge of final devoicing is important. In a 'reverse' Wug-test, knowledge of voicing phonotactics in relation to alternations should lead children to relate a novel alternating plural (i.e. /kɛdən/) to a neutralised singular (i.e. /kɛt/). Recall that final devoicing has also been claimed to be a 'natural process' (see Chapter 2), in that it is phonetically natural and widespread across languages. For instance, as noted in Chapter 2, final devoicing errors occur for children learning English (e.g. Stampe 1973). In constraint-based theories such as Optimality Theory (see Chapter 2), emergent or innate markedness constraints such as *VOICED-CODA have been proposed to account for final devoicing (e.g. Kager 1999a), which would initially outrank faithfulness constraints (such as IDENT-IO(voice)) by a universal ranking property of the initial state: M >> F. Such a view entails that production rather than perception is important for knowledge of voicing neutralisation, as 'natural' production constraints rather than language-specific phonotactic knowledge might drive children's behaviour.

It was argued in Chapter 3 that voicing patterns show considerable variation even across word boundaries (e.g. [hɛbɪk] 'have I'). In terms of child-directed input, the clearest evidence for neutralisation of the voicing contrast is thus found in utterance-final position or words in isolation. As argued by Zamuner (2006), the fact that learners are not misled by the production of voicing in casual speech suggests that learners must be sensitive to probabilistic patterns in the input, to avoid setting up a system in which final voiced obstruents are permitted. Hence, a (statistical) learner may identify a pattern in which voiceless obstruents are frequent in utterance-final position, from which broader generalisations may emerge. It is unclear to what extent children are sensitive to these patterns, although previous research has shown that infants can use powerful statistical learning mechanisms (e.g. Saffran et al. 1996). On the other hand, if children possess innate knowledge, the probabilities with which voiceless and voiced obstruents occur do not play a deciding role. In OT, if no voiced obstruents occur in the input, the initial ranking of M>>F does not need to be altered, and final devoicing is maintained. In contrast, if learners encounter positive evidence for voiced codas, this would trigger a ranking in which a constraint such as *VOICED-CODA is demoted. As argued in Chapter 2, only recent models using probabilistic constraints (e.g. Boersma & Hayes 2001) can deal with noisy data. In sum, it is difficult to determine to what extent phonotactic learning is relevant, as both phonetic and phonotactic (or frequency-related) factors lead to final devoicing.

Before we turn to predictions for the comprehension experiments, let us consider what knowledge Dutch infants may already possess. In general, previous research has found that the perceptual system of children becomes language-specific in the second half of the first year (Werker & Tees 1984), and there is evidence for phonotactic knowledge at around nine months (e.g. Mattys et al. 1999). However, Zamuner (2006) has shown that 9- and 11-month-old Dutch-learning infants showed no preference for legal versus illegal voicing phonotactics in word-final position. To investigate whether this was due to the salience of contrasts in final position, discrimination tests were carried out. The results showed that 10-month-old infants did not discriminate between voicing and place of articulation (POA) contrast in word-final position, although they were sensitive to these contrasts in word-initial position. However, 16-month-olds were able to discriminate between POA contrasts in final position, but they were still not able to discriminate between word-final voiced and voiceless stops.[1] This pattern of results suggests that sensitivity to the phonotactics of the voicing contrast in Dutch appears late. Based on this study, Zamuner has argued that previous results on phonotactic knowledge in infants cannot be extended to all contrasts or all types of phonotactic patterns. Moreover, it suggests that positional neutralisation is hard to learn. This ties in with proposals by Steriade (1997) and others, who argue that phonological contrasts are neutralised in environments with poor perceptual cues. This approach follows earlier work by Ohala (1983) and Westbury & Keating (1986), who relate neutralisation to articulatory difficulty. Irrespective of Zamuner's findings, it is possible that Dutch children are guided by phonotactic knowledge, which could be acquired later.

Let us now turn to the 'reverse' Wug test experiment and hypotheses according to the models discussed in Chapter 2. The aim of the experiments described in the next section was to further investigate children's knowledge of final devoicing in relation to alternations. As existing words were also

[1] In contrast, Broersma (2005) shows that adult speakers of Dutch perceive the word-final voicing contrast in English.

elicited (plural-to-singular), results could also be compared to the results for the previous (singular-to-plural) elicitation task.

## 6.2 Reverse Wug-tests

### 6.2.1 Hypotheses

The experiments described in Chapter 5 showed that children's knowledge of voicing alternations in existing words is not robust, as they frequently produce voicing errors (e.g. *[hɔnten] 'dogs'). Furthermore, children were shown to be reluctant to extend the alternating pattern to non-words, producing alternations in only 4% of cases. However, it is possible that young children have knowledge of voicing *neutralisation* in Dutch, even though they did not consistently produce voicing *alternations*. Irrespective of whether the child has stored /bɛd/ or /bɛt/, knowledge of final devoicing should lead the child to infer the correct singular upon hearing a novel plural, provided that the child can decompose a complex form into a stem and an affix and detect the medial voicing contrast (see Chapter 3).

In the previous chapters we have discussed the general prediction that children are initially guided by *Paradigm Uniformity*. Previous studies using an artificial language paradigm provide further evidence for this claim. For instance, Tessier (2006) taught 4-year-olds a novel plural suffix (i.e. /dəl/), which was added to a CVC(C) non-word. Results suggest that when children repaired the novel consonant clusters they prefer Output-to-Output faithfulness (or Paradigm Uniformity) over both Markedness and Input-to-Output faithfulness. In the current study, a preference for Paradigm Uniformity could potentially yield two types of responses. In case of non-alternating plurals (e.g. /kɛtən/), the child is expected to produce uniform [kɛt], which is in accordance with final devoicing. However, in case of alternating plurals (e.g. /kɛdən/, a preference for Paradigm Uniformity (or Output-to-Output Faithfulness) would lead to the uniform singular [kɛd], violating Final devoicing (or Markedness). Such a response would indicate that O-O faithfulness is initially ranked high (as proposed in Hayes 2004). Note that plural-dominant words (such as [tɑndən] 'teeth') never seem to be realised as *[tɑnd] in the singular, showing that Paradigm Uniformity is overruled by final devoicing.

Secondly, constraint-based phonological models predict that if the constraint of final devoicing is in place (which could be universal or based on the phonotactics), children should be able to produce novel singulars from plurals with intervocalic voiced obstruents (e.g. /kɛdən/ ~ [kɛt]). Hence, both 'alternating' plurals (e.g. /kɛdən/) and 'non-alternating' plurals (e.g. /kɛtən/) should lead to the singular [kɛt]. In this scenario, Markedness (e.g. a constraint against final voicing) would outrank both I-O and O-O Faithfulness. Furthermore, if children are guided by phonological knowledge, they should be equally good at deriving singulars from words and non-words.

A third hypothesis is based on analogical models (see Chapter 2), according to which knowledge of alternations is lexically based rather than driven by phonological knowledge. For instance, usage-based models predict that abstract product-oriented *schemas* emerge from relations among words (Bybee 2001). Knowledge of phonotactics is based on emergent generalisations, but there is no explicit link between phonotactics and morpho-phonological alternations. Under this view, knowledge of voicing alternations is influenced by the (type) frequency of the pattern. As argued before, the alternating pattern is expected to be less productive than the non-alternating pattern due to its lower type frequency and higher token frequency (i.e. it is tied to a small number of highly frequent words such as *handen* 'hands'). Hence, children might be reluctant to extend the alternating pattern to non-words, which should also affect children's behaviour in a plural-to-singular task. This means that there could be a difference between 'alternating' plurals (e.g. [kɛdən]) and 'non-alternating' plurals (e.g. [kɛtən]), in that children are more likely to accept the latter as novel plurals. Furthermore, this hypothesis predicts that there will be a difference between words and non-words, since the alternating pattern is mainly based on existing word pairs.

### 6.2.2 Previous study

Zamuner, Kerkhoff & Fikkert (2006a, 2006b) devised an experiment that required children to neutralise voicing in word-final position.[2] A Reverse

[2] This study was carried out together with the Picture Naming Task and Word Imitation Task investigating production of mono-morphemic vs. bi-morphemic words reported in Chapter 5. Children were tested at the KindertaalLab at the Radboud University Nijmegen. Acoustic analyses were carried out by the present author. Results were also presented at IASCL 2005 as Zamuner, Kerkhoff, Fikkert and Westrek (2005).

Wug Task (see also van Wijk 2007) was carried out to test whether children have internalised knowledge of voicing neutralisation in relation to alternations. In this task, children were first presented with the plural of a non-word (e.g. /kɛtən/) and then asked to form a singular. Knowledge of final devoicing and its application in morpho-phonological contexts should lead children to be equally good at positing singulars from plurals with medial /t/ and /d/ (e.g. /kɛtən/ and /kɛdən/). However, if children have not (yet) learned how these domains interact, they should have more difficulty positing singulars from plurals with medial /d/. In this study, four groups of children participated (2-year-olds, 2;6-year-olds, 3;6-year-olds, and 4-year-olds). Subjects were 8 children between 2;3-2;4 (M=2;4), 18 children between 2;5-2;8 (M=2;7), 18 children between 3;6-3;8 (M=3;7), and 8 children between 4;3-5;1 (M=4;8). The set of plural non-words consisted of a subset of the non-word items used for the production experiments, restricted to the eight items with /t/ after short and long vowels (see Chapter 5, Appendix I). Hence, children heard four plurals with /t/ and four plurals with /d/, counterbalanced across subjects.

Results showed that novel singulars were produced more often when the non-word plurals had /t/ versus /d/. Hence, the mean number of correct singulars from plural non-words with /t/ was higher (1.82) than the number of correct singulars for plurals with /d/ (0.70). Clearly, for voicing to play a role, children as young as 3;6 must have correctly perceived the voicing contrast, extending the results found for children aged 4;6 (Kuijpers 1993ab) to a younger age group.[3] Overall, children's performance (i.e. producing singulars) was quite poor, ranging from 11% accurate for 2;6-year-olds to 25% accurate for 4-year-olds. The majority of children's errors were a plural response rather than a singular response. For example, children often responded using [kɛtən] or [kɛdən], commonly preceded by *een* [ən] 'a' or *één* [en] 'one'. In contrast, accuracy on filler items (existing *-s* plurals such as *kikkers* 'frogs') ranged from 80 to 100% for all age groups. Finally, a small correlation was found between children's productive vocabularies at age 2;6 and their ability to perform the Reverse Wug Test (i.e. to produce correct singulars).

Recall from Chapter 3 that a covert or subphonemic initial voicing contrast has been reported for child language in English (Macken & Barton

---

[3] Recall that Kuijpers (1993ab) showed that the youngest group of children (at 4;6) were able to correctly identify medial voiced and voiceless obstruents in non-words (e.g. *paddo* vs. *patto*).

1979), and that incomplete final neutralisation has been found for adult speakers of Dutch (e.g. Warner et al. 2004). To determine whether children had an underlying contrast that could not be perceived by transcribers, an acoustic analysis of children's singular responses was carried out (see Zamuner et al. 2006). However, this did not yield significant differences between the closure duration of [t] from plurals with /t/ (119 ms) and [t] from plurals with /d/ (135 ms). This means that there is no evidence for a covert contrast for neutralised segments.[4] Note that this result could be taken to support a phonological analysis, although the absence of incomplete neutralisation in child language could have a variety of causes (e.g. the fact that acoustic measurements are unreliable). Overall, the results suggest that children did not have a robust knowledge of voicing neutralisation in connection with alternations, as even 4-year-olds were poor at producing a singular with final /t/ from a plural with /d/ (e.g. /slɑdən/). At the same time, children never produced word-final /d/, indicating that the phonotactic pattern was never violated in favour of Paradigm Uniformity (e.g. resulting in /slɑdən/ ~ slɑ[d]). This could be interpreted as evidence for phonotactic learning, or the presence of a phonological devoicing rule. However, this outcome is also expected on the basis of the universal, phonetically motivated preference for final voiceless obstruents. One the one hand, one could argue that many such preferences do not occur categorically in early productions, pointing towards a phonological process of final devoicing. On the other hand, intervocalic voicing is also a candidate for a universal preference, but the data offer much weaker support for a phonological process. In sum, Zamuner et al. did not find unambiguous evidence that the knowledge of final devoicing aids the acquisition of alternations. Contrary to expectations, children were reluctant to accept plurals with /d/ as the source of back-formations, matching the general difficulty they have producing these plurals. Before we return to the hypotheses, results for a more detailed study using a reverse Wug-test are reported in the next subsection.

---

[4] Note that if children had maintained a covert contrast, the length difference would have been in the opposite direction. In other words, if the final [t] of neutralized /d/'s had been more 'd-like', durations should have been shorter compared to the /t/ condition.

### 6.2.3 Experiment IV: children

The corpus study reported in Chapter 4 revealed that children are most likely to be exposed to alternating plurals with /d/ after nasals (e.g. /hɔndən/ 'dogs') in the input. Likewise, the production experiments described in Chapter 5 showed that children performed best on existing plurals with /nd/. Furthermore, non-word alternations were mostly produced in this environment (e.g. [flɑnt] ~ [flɑndən]). It thus seems that knowledge of the alternating pattern is strongest for this context, while it is weakest for plurals with /b/ (both in terms of existing plurals and non-word alternations). In the previous experiment (Zamuner et al. 2005; 2006), the non-word items were limited to /t/ after short and long vowels. This selection was made to reduce the duration of the experiment, given young children's limited attention span. Furthermore, acoustic measurements in the nasal context were deemed less reliable, because it is more difficult to measure the onset of voicing after a nasal. The reverse Wug-test reported in this section involved stimuli in all environments used for the classic Wug-test reported in Chapter 5, to determine whether the effect of alternation holds for all rhymes (short vowel V_, long vowel VV_ and nasal N_, for both T and P). Furthermore, the previous experiment involved only novel plurals. The only existing words were *-s* plural fillers (e.g. *kikkers* 'frogs'), to avoid training of the test procedure. However, as overall accuracy was low in the Zamuner et al. study (even for novel plurals with /t/), the inclusion of the existing real word *–en* plurals in the test items might enhance children's performance. Therefore, in the current study both novel and existing word *-en* plurals were used. Furthermore, a systematic investigation of the difference between words and non-words was not possible in the Zamuner et al. experiment. Hence, the current study was aimed at comparing children's performance on words and non-words directly, by using both existing and novel *-en* plurals.

#### *6.2.3.1 Subjects*

A total of 30 children (21 girls, 9 boys) participated in the experiment.[5] One additional girl did not complete the experiment and was excluded from the

---

[5] Ten children participated in a pilot experiment, in which all 24 non-word plurals were used (Appendix I). The experiment was considered too long, as three children did not complete the pilot test, upon which half of the non-word items were removed. The results of the nine children

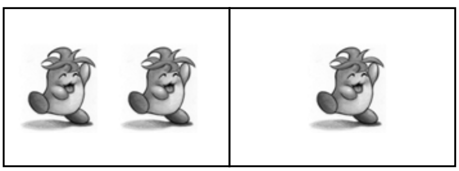
analysis. The mean age was 4;8, with ages ranging from 3;1 to 6;0 (this wider age range could potentially reveal age effects). Children were tested in a quiet room in the school or at home.

*6.2.3.2 Materials*

A total of 41 stimuli were presented to each child, including 3 practice items (*bomen* 'trees', *beren* 'bears', nonce *jikken*), 16 existing words (8 non-alternating and 8 alternating), 12 non-word plurals in all phonological environments (V_ , VV_ and N_ for T and P), and 10 fillers. The set of existing words was near-identical to the set used in the previous production experiments (e.g. /pɛtən/ 'caps' vs. /bɛdən/ 'beds'). The non-word plurals were taken from the same set of non-words used for the previous experiments (e.g. /kɛtən/ and /kɛdən/). The set of fillers included -*s* plurals (e.g. *varkens* 'pigs', *appels* 'apples') and irregular -*en* plurals (i.e. *eieren* 'eggs', *koeien* 'cows'), see Appendix IV for the full list of stimuli. The items were presented in two different orders.

*6.2.3.3 Procedure*

Children first saw two identical pictures while the recorded non-word plural was played, followed by a single picture (presented using PowerPoint, see the example below).



Recordings featured a female speaker who read non-word plurals aloud (written in their plural form, e.g. *slatten*, *deben*).[6]

In case of existing words, both the plural and the singular were elicited, to relate the two forms. Children were always asked to name the

who completed the pilot test were included in the overall results, without taking into account the 12 additional items.

[6] The recordings were made in the same way as described in the previous chapter (5.2.3.3) by the same speaker.

plural, and no recorded version was available. When a voicing error was produced for the plural (e.g. *bɛ[t]ən 'beds', *pɛ[d]ən 'caps'), the child was not corrected. When a child first produced a singular, the plural was elicited again before the next screen was shown. For existing words, singulars were only counted as correct if the pair was correct. Hence, when pairs such as *bɛ[t]ən ~ bɛ[t] 'beds', *pɛ[d]ən ~ pɛ[t] 'caps' or *[pɛts] ~ [pɛt] 'caps' were produced, the response was considered incorrect. When children failed to respond or provided a different word (e.g. *voeten* 'feet' ~ *teen* 'toe', *slapen* 'to sleep' for *bedden* 'beds', *hanen* 'roosters' for *kippen* 'chickens'), the response was coded as 'missing'. Diminutive plurals were also scored as 'missing' (e.g. [bɛtjəs] ~ [bɛt(jə)] 'beds'), but pairs for which only a diminutive singular was produced were considered correct (e.g. [bɛdən] ~ [bɛcə]), due to voicing being present in the plural. A plural response with a stem change was scored separately (e.g. [hɔnə] 'dogs').

Non-words were paired with colourful pictures of fantasy animals, identical to the ones used in the previous production experiments (Chapter 5). Voicing of the non-words was counterbalanced across subjects, such that each subject was tested on six plurals with /t/ or /d/ and six plurals with /p/ or /b/ (i.e. an individual child was presented with [flɑntən] or [flɑndən] but not both). There were always two items per environment, e.g. one child was presented with [kɛdən] and [slɑtən] (list 1), whereas another child was presented with [kɛtən] and [slɑdən] (list 2). For non-words, the experimenter introduced each plural "*Dit zijn twee* ____" ("These are two ____") upon which the pre-recorded plural was played. The experimenter then asked "*Nu is er nog maar een, wat is dit?*" ("Now there's only one, what's this?"), eliciting the singular. In case children indicated that they had forgotten the plural it was played back, and children were encouraged to provide a singular. For non-words, Correct responses were singulars (e.g. [slɑt], [slɑtjə]), while Incorrect responses included plural responses (e.g. [slɑdən], [slɑtən] [slɑtjəs]) and incorrect singular responses (e.g. [slɑ]). When children failed to give a response or responded with a diminutive (e.g. [slɑtjə]) or a different word (e.g. 'Pokemon'), the response was coded as missing. A plural response with a stem change was scored separately (e.g. /kɪmpən/ ~ [kɪpən], /xɔbən/ ~ [xɔdən], /tabən/ ~ [tapən]). When two responses were provided, the second one was considered in case of a self-correction (e.g. [kɛdən], [kɛt]).

### 6.2.3.4 Results

The results will be discussed separately for existing words and non-words.

**Words**

The results for words are shown in Table 38, where incorrect plurals refer to voicing errors (e.g. *[hutən] ‘hats’ or *[olifɑndən] ‘elephants’).

*Table 38: Results for word pairs in numbers (%), Exp. I.*

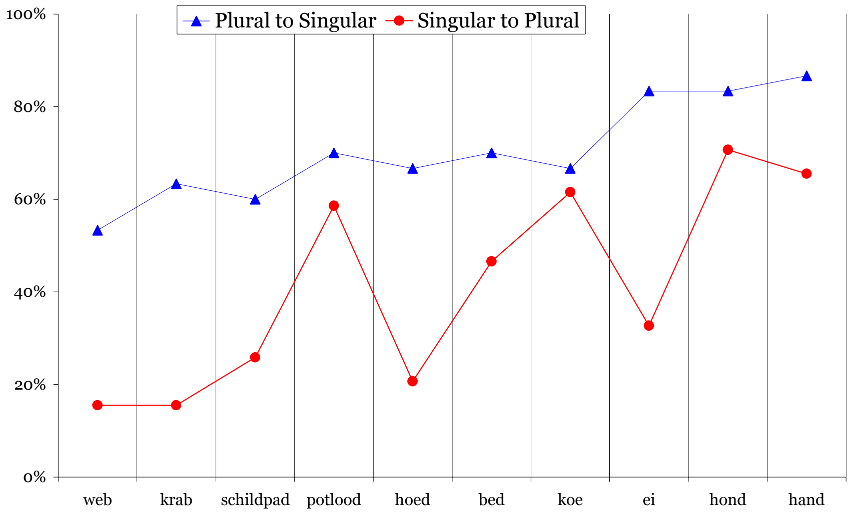
| | Non-alternating | Alternating | Total |
|---|---|---|---|
| Correct | 205 (85.4%) | 166 (69.2%) | 371 (77.3%) |
| Incorrect plural | 16 (6.7%) | 37 (15.4%) | 53 (11.0%) |
| Plural response | 4 (1.7%) | 3 (1.3%) | 7 (1.5%) |
| Other | 15 (6.3%) | 34 (14.2%) | 49 (10.2%) |

These data show that children can successfully derive singulars from existing plurals (in around 77% of cases), although pairs were not always correct.[7] For instance, children produced some incorrect plurals for non-alternating words, especially for *[boden] ‘boats’ and *[pɛdən] ‘caps’, which is presumably related to these items’ relatively low plural frequency.

Generally, children had more difficulty with alternating pairs, replicating the result for the previous experiments, in which plurals were elicited rather than singulars. A independent samples t-test with the proportion of correct responses as dependent variable showed that children scored significantly better on non-alternating words than on alternating words ($t(14)=3{,}070$, $p=.008$), mirroring results for the singular-to-plural elicitation. In Figure 10, performance on existing plurals (including the irregulars *koe* /ku/ ~ /kujən/ ‘cows’ and *ei* /ei/ ~ /eijərən/ ‘eggs’) is shown, with words ordered according to the CELEX frequency of the plural form. Children’s performance on these same words in the singular-to-plural task is added (see Figure 1, Chapter 5, note that these are two different groups of children).

[7] The lexical specificity of the alternating pattern is illustrated by Sietske (4;0), who produces *honden* ‘dogs’ and later remarks “*het lijken wel *zeehonten*” ‘they look like seals’ in response to a non-word picture.

*Figure 10: Performance correct on alternating and irregular plurals.*




This figure shows that performance on the plural-to-singular task reflected the frequency of plurals, as was found earlier in the singular-to-plural task of Chapter 5. Hence, performance ranged from 55% for *webben* 'webs' to 87% for *handen* 'hands', which was also among the most frequent plurals in the input (Chapter 4). As was found for the elicitation experiments described in Chapter 5, there was a significant correlation between children's performance on existing words and the CELEX frequency of the plural form ($\rho$=.750, p<.001), whereas there was no such correlation between performance and frequency of the singular (p=.154). This effect was due to performance on alternating words, which shows a strong correlation with frequency ($\rho$=.879, p=.004). However, the effect only approached significance for non-alternating words (p=.055).

Interestingly, the performance on alternating plurals (69%) was much better than in the previous study, in which the overall performance rate was around 40 to 50% for typically developing 5-year-olds. This effect was not found for non-alternating plurals, as performance was around 80 to 88% for 5-year-olds in the previous elicitation experiment, which is similar to the 86% found in the current study. Finally, children also scored better on the irregular plural *eieren* 'eggs' when the plural was elicited first. This possibly reflects an advantage for whole word retrieval for irregular plurals.

The results in Table 38 show that there were very few plural responses for existing words (e.g. [hɑndən] ~ [hɑndən] 'hands'), and there was no difference between alternating and non-alternating words.[8] Finally, children produced more diminutive plurals (coded as Other), especially for alternating words, e.g. *bedjes* 'beds-dim' (10 vs. 2). As this tendency was not found in the previous elicitation experiments, it is possible that when children are asked to first produce a plural, they are more likely to produce a diminutive (possibly revealing an avoidance strategy). Note that the observation that diminutives are more likely to occur for alternating words was also found in the corpus data (Chapter 4).

### Non-words

The results for non-words are shown in Table 39, containing the number of correct and incorrect responses for non-alternations (plurals with medial /t/ or /p/, e.g. /kɛtən/) and alternations (plurals with medial /d/ or /b/, e.g. /kɛdən/).

*Table 39: Results for non-words in numbers (%), Exp. I.*

| | **Non-alternating** | **Alternating** | **Total** |
|---|---|---|---|
| Correct | 95 (52.8%) | 56 (31.1%) | 151 (41.9%) |
| Plural response | 71 (39.4%) | 101 (56.1%) | 172 (47.8%) |
| Incorrect singular | 1 (0.6%) | 12 (6.7%) | 13 (3.6%) |
| Other | 13 (7.2%) | 11 (6.1%) | 24 (6.7%) |

These results show that children are able to derive non-word singulars from plurals in around 42% of cases, which is much lower than the overall success rate for words (77%). Crucially, children performed worse at deriving singulars from alternating plurals, mirroring the difficulty they have producing alternating plurals. A paired samples t-test revealed that the mean number of correct responses was higher for non-alternating non-word plurals than for alternating non-word plurals ($t(29)=4,860$, $p<.001$). This result thus confirms earlier findings by Zamuner et al. (2006), who also found that children tend to repeat the plural rather than produce a singular

[8] There was one stem change error, i.e. [hɔnə] for /hɔndən/ by Emilie at 4;6.

for alternating non-words.[9] Overall, children's performance on non-words was more accurate than that of children in the previous experiment (i.e. 42% accurate vs. 25% for 4-year-olds in the previous experiment). This is not likely to have been due to an age effect, as *mean* ages were identical. Rather, results suggest an effect of existing alternating plurals, which boosted children's performance on non-words (recall that the only existing words in the previous experiment were *-s* plurals). As in the previous study, the majority of errors were plural responses (e.g. [kɛdən] ~ [kɛdən]).

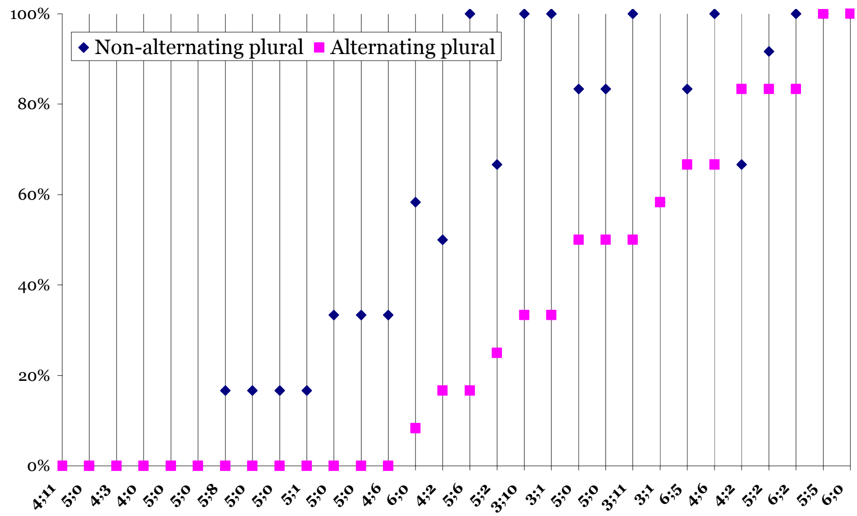
Children sometimes produced incorrect singulars of the type [kladən] ~ [kla] (resulting in incorrect affix stripping), which had also occurred sporadically in the Zamuner et al. study. It is possible that such errors resulted from analogy with a few existing pairs such as *la ~ laden* 'drawers' (here, the older singular *lade* co-exists with the singular *la*) or analogy to verbal pairs such as *ree ~ reden* 'drove'. Also, some errors may have been due to word effects (e.g. *fee* 'fairy' as a singular of the novel plural *feden*). This error type occurred more often in the current experiment, due to the alternating plurals in the nasal environment. Hence, an incorrect singular only occurred once for a non-word plural with /t/ (i.e. [jɔntən] ~ [jɔn]), whereas there were 16 cases for non-word plurals with /d/ or /b/ (including 9 for /b/ after nasals ([bɛmbən] ~ [bɛm]) and 3 for /d/ after nasals (e.g. [flɑndən] ~ [flɑn]). It is harder to account for these incorrect singulars, as there are no such pairs in Dutch, except for verbal pairs such as *kon ~ konden* 'could'. However, recall that errors such as [hɑnt] ~ [hɑnə] 'hands' occurred in corpus data (see Chapter 4) and in young children's elicited productions (see Chapter 5), reflecting the difficulty of the /nd/ cluster. Likewise, a final /nt/ or /mp/ cluster can also be considered articulatorily difficult. It is also possible that the difficulty is partly rooted in perception, which means that children may have misheard non-word plurals (i.e. /bɛmbən/ as /bɛmə/). There is some support for this in children's productions, e.g. /flɑndən/ was repeated as /flɑdən/ (by Sietske at 4;0). Finally, children never seemed to produce singulars with final voiced [d] or [b], although there may have been subtle differences.[10]

---

[9] Also note that a plural response for a voiced non-word sometimes resulted in devoicing, i.e. *globen ~ glopen* (Janno 4;11), *kladen ~ klaten* (Dieuwertje 5;2) and *tebben ~ teppen* (Didy 5;0).
[10] Also note that a response without final [n] (e.g. [flɑndə]) was considered a plural response, although it is possible that children intended this response as a singular. However, the production of final [n] is optional in Dutch, and the difference (i.e. between [flandən] and [flɑndə]) was mostly not clearly audible to the transcriber.

The performance of all children on both types of plural non-words is shown in Figure 11, with ages indicated.[11] Responses are ranked according to children's performance on alternating plurals rather than age, to indicate the relationship between the two types of plurals.

*Figure 11: Correct non-word singulars per child (%).*




This figure shows that six children were not able to produce singulars for any non-word plural (regardless of alternation), and seven children were only able to produce singulars for non-alternation plurals. Even though children were generally better at producing singulars for non-alternating plurals, one child (Stan, 4;2) produced more correct answers for alternating plurals (five vs. four items correct) and two children scored equally well on both types of plurals (note that the symbols overlap in Figure 11). Furthermore, the relation between performance and age is not very strong, as was found in the previous singular-to-plural elicitation experiments discussed in Chapter 5 (e.g. some 3-year-olds outperform 5-year-olds).

Let us now turn to the distribution of correct responses for non-words according to context, shown in Table 40. Recall that there are two

[11] In this figure all results were considered, including those of the 7 children who participated in the pilot experiment (who saw 24 non-word items each).

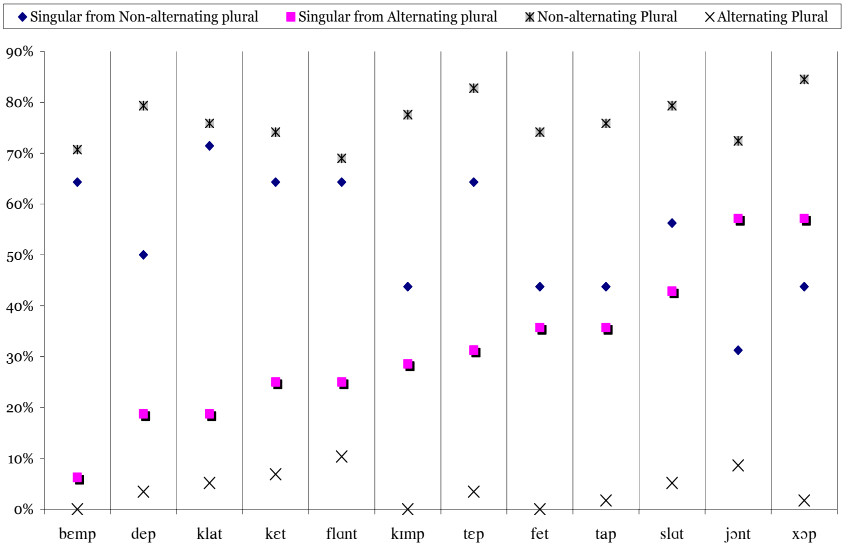
'lexical gap' environments, for which no alternations occur in Dutch (see Chapter 4), indicated by shading.

*Table 40: Distribution of correct non-word singulars in numbers (%), Exp.I.*

| | **Non-alternating** | | **Alternating** | |
|---|---|---|---|---|
| | t ~ t | p ~ p | d ~ t | b ~ p |
| V_ | 18 (60.0%) | 16 (53.3%) | 10 (33.3%) | 13 (43.3%) |
| V:_ | 17 (56.7%) | 14 (46.7%) | 8 (26.7%) | 8 (26.7%) |
| N_ | 14 (46.7%) | 16 (53.3%) | 12 (40.0%) | 5 (16.7%) |
| Total | 49 (54.4%) | 46 (51.1%) | 30 (33.3%) | 26 (28.9%) |

An ANOVA was carried out to determine whether there was an effect of phonological environment on performance correct, with independent factors Alternation (non-alternation, alternation), Final Obstruent (P, T) and Rhyme (V_, V:_, N_). Although there was a main effect of Alternation ($F(1)=11{,}292$, $p=.006$, $\eta_p^2=.49$), there were no other main effects or interactions. The absence of a significant effect of rhyme is probably due to the fact that numbers are low. Still, performance on /b/ plurals after nasals (e.g. [bɛmbən] ~ [bɛmp]) seems especially poor (17%), which might be related to the fact that it is phonotactically marked or that perceptual cues for voicing are weak in this environment (i.e. there is only one release in /mb/). Children's performance per non-word is shown in Figure 12 below, together with the proportion of *-en* plurals and alternations produced in the singular-to-plural Wug-test.

*Figure 12: Correct non-word singulars (Exp. I) compared with production of non-word plurals (Wug-test Ch. 5) per item (%).*




This figure confirms that producing plurals is easier than deducing singulars, even when only the production of singulars from non-alternating plurals is considered. Hence, correct *-en* plurals were produced for 79% of non-words (including alternations), whereas singulars were produced in 46% of cases (including incorrect singulars).[12] The figure also shows that not all 'lexical gap' items behave in the same way, e.g. performance on /bɛmbən/ was poorer than on /kɪmbən/. In general, performance on non-words which share the same rhyme was variable. For instance, the item /jɔnt/ was more often correct when offered as an alternating plural (i.e. /jɔndən/), while /flɑnt/ was more often correct when offered as a non-alternating plural (i.e. /flɑntən/). The high performance on /xɔbən/ is harder to explain, as plurals with /b/ are rare in children's input, except for the highly frequent verbal pair [hɛp] ~ [hɛbən] 'have-inf/plur' (but note that performance on /tɛbən/ is worse). In Table 41, children's performance is shown alongside with the CELEX probability of voicing based on noun plurals (with a frequency above zero)

[12] Cf. van Wijk (2007), who found that Dutch 3 yr-olds produced a plural for non-words with -en or -s in 60% of cases, whereas singulars were produced in only 28% of cases (but performance was high for existing words, i.e. 94%). Singular errors were also observed (e.g. *latopen* - *lato*).

with the same rhymes as the test-items and their average token frequency (see Chapter 5 and 6). For instance, there are no noun plurals with the rhyme /ɛmbən/ and one noun plural ending in /ɛmpən/, leading to a probability of voicing of 0 (0/1).

*Table 41: Correct non-word singulars from alternating plurals and the probability of voicing in plurals (%).*

| Rhyme | Correct singular | Probability types | Avg. token freq. pl. |
|---|---|---|---|
| /ɛmbən/ | 6.3 | 0 (0/1) | 0 |
| /ɪmbən/ | 28.6 | 0 (0/0) | 0 |
| /ebən/ | 18.8 | 0 (0/5) | 0 |
| /ɛbən/ | 31.3 | 0 (0/5) | 0 |
| /abən/ | 35.7 | 0 (0/6) | 0 |
| /ɔbən/ | 57.1 | 0 (0/11) | 0 |
| /ɛdən/ | 25.0 | 14.3 (1/7) | 12 |
| /edən/ | 35.7 | 25.0 (1/4) | 1 |
| /ɑdən/ | 42.9 | 27.3 (3/11) | 2 |
| /adən/ | 18.8 | 46.7 (7/15) | 10 |
| /ɑndən/ | 25.0 | 65.0 (13/20) | 53 |
| /ɔndən/ | 57.1 | 80.0 (8/10) | 14 |

This figure illustrates that the relationship between the probability of voicing and children's behaviour on non-word plurals is not strong (p>.1). Items ending in /ɑndən/ and /ɔndən/ are most likely to alternate (corresponding to the corpus data in Chapter 4). The probability of voicing was found to affect children's production of non-word alternations, which occurred predominantly /t/ after nasals (Chapter 5). The current experiment does not provide conclusive evidence for such an effect, as children scored well on the item ending in /ɔndən/ but not /ɑndən/. However, recall that the plural /hɑnden/ 'hands' has a very high token frequency and is among the first plurals in children's productions (see Chapter 5). This might mean that the pattern is less easily extracted in case of non-word plurals ending in /ɑndən/. As was found for existing words, tasks in which singulars rather than plurals are produced do not yield the same effects, i.e. the influence of highly frequent plurals might be greater in the plural-to-singular task. Hence, there is no apparent relationship between children's performance on the reverse Wug-test (e.g. /flɑndən/ ~ [flɑnt]) and the number of alternations produced in the singular-to-plural Wug-test (e.g. /flɑnt/ ~ [flɑndən]), and no correlation was found (see also Figure 12).

The production experiments (Chapter 5) showed that the alternating pattern was extended most often for /d/ after nasals, mirroring the distribution of alternations in the input. This pattern was not very strong in the reverse Wug-test, showing that different factors play a role. This suggests that children are aided by their knowledge of final devoicing, even though it does not lead them to treat all plurals in the same way.

In sum, results for non-words show that knowledge of the alternating pattern is not directly affected by rhyme context (i.e. the quality of the preceding vowel or nasal). However, errors such as /bɛmbən/ ~ /bɛm/ may reflect the fact that nasal clusters are less salient, as well as an effect of phonotactic likelihood, i.e. the 'lexical gap' in Dutch.

#### *6.2.3.5 Discussion*

Experiment IV elicited existing plural ~ singular word pairs and singulars for novel plurals. The results for existing words indicate that children are able to produce plural ~ singular pairs, even though alternating plurals were more difficult. This replicates the results for the previous elicitation experiments, in which children performed better on non-alternating words. However, children performed much better on alternating plurals than in the previous (singular-to-plural) experiment (see Chapter 5). This suggests that children perform worse on alternating plurals when the experimental task induces them to derive plurals from singulars. Standard elicitation (i.e. eliciting singulars before plurals) may induce a task effect, which leads children to produce more 'Paradigm Uniformity errors' (especially for plurals with low frequency, e.g. [hut] ~ *[hutən] 'hats'). This task may thus underestimate children's knowledge of alternating plurals, as eliciting plurals directly boosts their performance. Such an effect is reported in Kuczaj (1978), who found that English speaking children produced more overgeneralisation errors than found in spontaneous speech when they were supplied with the verb stem (as in 'I will drink my milk. I already ________ my milk.'). Furthermore, the finding suggests that existing alternating plurals are more easily accessed as whole forms (e.g. /hɑndən/ 'hands'). Hence, children might rely more on whole word storage when plurals are elicited directly (or in spontaneous speech). The results for words in the current task are indeed closer to younger children's use of alternating plurals in spontaneous speech in the CLPF corpus (see Chapter 4), which revealed an overall devoicing rate of 29% (e.g. *[hɑntən] 'hands'), while singular-to-plural elicitation experiments revealed a rate of around 40%. In the current plural-to-singular experiment, children produced 'devoicing' errors in only 16% of cases (e.g. *[bɛtən] 'beds'), compared to 7% 'voicing' errors for non-alternating words (e.g. *[pɛdən] 'caps'). Note that this 'overgeneralisation' rate is slightly higher than the 5% rate found for typically developing 5-year-olds in the previous elicitation experiments (see Chapter 5). This difference could be due to the presence of alternating non-words (such as /kɛdən/), which may have 'primed' the use of alternations. This suggestion is similar to MacWhinney's (1978) notion of analogy (see section 5.1), which is invoked when plurals such as *scarves* 'prime' the production of the novel pair *narf ~ narves*. The current result thus suggests that the opposite pattern may also occur, as novel plurals such as /kɛdən/ may prime overgeneralisations such as *[pɛdən] in words. Let us now turn to the more specific hypotheses discussed in section 6.2.1, taking into account the results for non-words.

Results of the reverse (plural-to-singular) Wug-test showed that *Paradigm Uniformity* was not the only factor guiding children's behaviour, as they did not produce pairs such as [hɑndən] ~ [hɑnd] 'hand' or [kɛdən] ~ [kɛd]. Children's reluctance to produce alternating plural ~ singular pairs might have been due to a preference for Paradigm Uniformity, as stripping off the *-en* suffix in case of novel alternating plurals exposes the voiced obstruent in a position in which it should not occur. The fact that children never violated the phonotactic pattern can be interpreted as evidence for phonotactic learning or the presence of an (innate or language-specific) rule or constraint against final voicing, supporting phonological models. However, it should be noted that this outcome is also expected on the basis of the universal, phonetically motivated preference for final voiceless obstruents. Furthermore, children had more difficulty deriving singulars from novel plurals in case of a stem alternation. Children were generally poor at deriving singulars from plural non-words, but they were worse at positing singulars when this would require knowledge of voicing neutralisation. This is unexpected if the (phonotactic) pattern of final devoicing had been successfully related to morpho-phonological alternations. Hence, the current results do not provide direct evidence that the knowledge of final devoicing aids the acquisition of alternations. This is because, contrary to expectations, children were reluctant to accept plurals with /d/ for back-formations, matching the general difficulty they have producing these plurals. If the (phonotactic) pattern of final devoicing had been successfully related to morpho-phonological alternations, children should not have more problems with novel alternating plurals than novel non-alternating plurals. The current results are therefore compatible with a strategy of Paradigm Uniformity, which predicts that children prefer non-alternating paradigms. Results are also in line with analogical models, which predict that children's knowledge of voicing alternations is lexically based. Hence, as was found for production, children have some difficulty extending the alternating pattern to novel words. This ties in with the second prediction of analogical models, as children performed worse on non-words (e.g. /kɛdən/ ~ [kɛt]) than on existing words (e.g. /bɛdən/ ~ [bɛt] 'bed'). Hence, incorrect plural responses (e.g. [kɛdən]) were almost exclusively produced for non-words (i.e. in around half of the cases overall), whereas this error type was rare for words.

In the next section, results of a control experiment with adults will be discussed, which was carried out to determine whether adults are able to correctly derive novel singulars from alternating non-words. This would allow a comparison between children's development and the adult end-state.

## 6.2.4 Experiment V: adults

A second reverse Wug-test was carried out, to test adults' productive knowledge of morpho-phonological alternations and final devoicing.

### *6.2.4.1 Subjects*

A total of 20 adult control subjects (16 female, 4 male) participated in the control experiment (average age 27 years). These were primarily university students or employees.

### *6.2.4.2 Materials*

The full set of 24 non-words was used for this experiment (see Appendix I). A set of 26 non-word fillers was used, consisting of novel plurals with stems ending in a different final obstruent, taking *-en* or *-s* (e.g. *blaas, herken, kroelen*). Additionally, a set of 50 existing word fillers was added, which contained 20 plurals with medial voiced obstruents (e.g. *ronden* 'rounds', *kinderen* 'children', *hoeden* 'hats', *leden* 'members', *bladen* 'trays'). The remaining 30 words were regular and irregular *-en* plurals (e.g. *kralen* 'beads', *koeien* 'cows', *vaten* 'barrels') and *-s* plurals (e.g. *leemtes* 'gaps'), see Appendix V for the full list of items. The words and non-words were mixed and two lists were created to control for order effects.

#### *6.2.4.3 Procedures*

Subjects were seated in a soundproof booth in the phonetic lab at the Research Institute for Language and Speech (UiL-OTS) at Utrecht University or in a quiet room. Subjects were asked to produce the singular of a real or non-word plural. The procedure was the same as that for the child-test (see 6.1.3.3), but no pictures were shown. Subjects were instructed to respond as quickly as possible to ensure an automatic response.

#### *6.2.4.4 Results*

The results for non-words are shown in Table 42, containing the number of correct and incorrect responses for non-alternating (e.g. /kɛdən/) and alternating plurals (e.g. /kɛtən/).

*Table 42: Results for non-words in numbers (%), Exp. II.*

| | **Non-alternating** | **Alternating** | **Total** |
|---|---|---|---|
| "Correct" | 237 (98.8%) | 222 (92.5%) | 459 (95.6%) |
| Plural response | 0 (0.0%) | 3 (1.3%) | 3 (0.6%) |
| Incorrect singular | 3 (1.3%) | 15 (6.3%) | 18 (3.8%) |

Results show that adults are generally capable of producing singulars for both types of non-word plurals, and there are hardly any plural responses.[13] Adults show a much better performance than the children in Experiment I, who produced singulars in only 31% of alternating plurals. However, adults also performed slightly better on non-alternating items, as they produced incorrect singulars (e.g. /kɪmbən/ ~ [kɪm]) more often for alternating plurals (as was found in Experiment I). A paired samples t-test revealed that the mean number of correct responses was higher for non-alternating non-word

[13] The only three cases were produced for the non-word /midən/, which corresponds to the simplex word *midden* 'middle'.

plurals than for alternating non-word plurals (t(23)=2,901, p=.008). Errors occurred mostly for plurals in the nasal environment (e.g. /jɔndən/ ~ [jɔn]), although there were two errors after long vowels (e.g. /kladen/ ~ [kla]) and even two errors after short vowels (e.g. /slɑdən/ ~ [slɑ]). Apparently, adults also have a tendency to be influenced by verbal pairs such as *kon* ~ *konden* 'could' (in 3.8% of cases vs. 3.6% of cases for children), showing that this behaviour is not child-specific.[14]

In the analysis above, "correct" responses included singulars in which adults left out final *-n*. This error occurred five times for a non-alternating plural (i.e. /fetən/ ~ [fetə]), possibly reflecting analogy to the existing singular *vete* 'feud'). Most of these errors occurred for alternating plurals (15% of cases), induced by non-word items intended to resemble mono-morphemic forms, showing a word effect (e.g. /tɔmbən/ ~ /tɔmbə/, due to the existing singular *tombe* 'tomb'), but also for other items (e.g. /dɪndən/ ~ [dɪndə]). As such pairs exist (e.g. *ronde* ~ *ronden* 'rounds') these responses were considered correct. Correct responses also included singulars with vowel alternations (e.g. /kladən/ ~ [klɑt]), which were produced for both types of plurals, in 10% of non-alternating items (23/240) and 9% of alternating items (22/240). Adults occasionally extended the non-productive vowel alternations to rhymes that would never undergo such alternations in Dutch (e.g. /depən/ ~ [dɛp]). Finally, items in 'lexical gap' environments were not more often incorrect, showing that adults' knowledge of voicing neutralisation is robust and not influenced by the probability of voicing in the lexicon.

### *6.2.4.5 Discussion*

The results of experiment V show that adult controls do not share children's reluctance to produce singulars from non-word plurals. Recall that the previous experiments showed that adults were not more likely than children to extend the pattern to novel plurals (Chapter 5), which suggests that the pattern is unproductive due to its lower type frequency. In the current experiment, singulars were readily produced for alternating plurals

[14] Errors for words also occurred, indicating that errors may be due to task effects (or uncertainty with respect to the correct singular), such as /stedən/ ~ *[stet] for [stɑt] 'city', /badən/ ~ */bat/ for /bɑt/ 'bath', /kledən/ ~ *[kledə] for [klet] 'rug', /bɔndən/ ~ *[bɔndə] for [bɔnt] 'union' and /kralen/ ~ *[kra] for [kral] 'bead'.

(although they were slightly more often incorrect). This shows that adults have productive knowledge of voicing alternations and final devoicing, as they extend the alternating pattern to non-words. Note that knowledge of alternating pairs combined with the general preference for voiceless final obstruents would suffice for the pattern to show productivity in a reverse Wug-test. The strength of the preference for both children and adults may indicate the presence of phonological knowledge of final devoicing. The obvious difference between adults and children is likely to be due to the fact that adults have internalised knowledge of final devoicing in relation to morpho-phonological alternations. This implies that the Dutch voicing alternation is not yet fully acquired at age six, as has been found for morphologically and lexically conditioned alternations (see Chapter 2). Implicit knowledge of alternations may be due to adults' increased experience with alternating pairs in the lexicon (i.e. there is a larger number of types in the adult lexicon), and may further have been augmented by adults' knowledge of orthography. Recall that knowledge of spelling renders the alternation more transparent (e.g. *bed* ~ *bedden* 'beds'). Finally, adults' behaviour was variable, as they produced a variety of different (correct and incorrect) singular types. For instance, vowel alternations (e.g. lengthening as in [bɑt] ~ [badən] 'baths') are traditionally argued to be lexically governed alternations, which are not expected to show productivity. However, the current experiment showed that adults extended this pattern to novel plurals (e.g. /klatən/ ~ [klɑt]). This finding illustrates the difficulty of treating morpho-phonological alternations as being either fully productive or unproductive (or regular vs. irregular).

Previous studies have investigated English-speaking adults' knowledge of the 'unnatural' /k/~/s/ alternation in derivationally related pairs such as *electric* ~ *electricity*. In a 'backformation' test, Pierrehumbert (2006) has observed a strong preference for non-alternations, although there was also evidence for productivity (i.e. forms with final /k/ were produced in response to derived novel forms with /s/). This productivity could not be explained by faithfulness or frequency effects (i.e. /s/ is more common word-finally), which leads Pierrehumbert to conclude that "knowledge of the alternation must be a generalisation over morphologically related word pairs". However, Ohala & Ohala (1987) provide evidence that speakers treat derivationally related pairs such as *extreme*/*extremity* in the same way as phonological isolates such as *pope*/*papal*. This shows that such knowledge may not be internalised for all types of alternation.

### 6.2.5 Discussion Reverse Wug-tests

The results of reverse Wug-test experiments show that children (aged 2;3 to 6;0) were generally poor at producing singulars from novel plurals, indicating that the task is more demanding than the standard Wug-test. Hence, children's performance on this task was worse than performance on the classic Wug-test (Chapter 5), in which children inflected the large majority of non-words (i.e. 81% of cases, including voicing errors and *-s* plurals). This difference seems to be due to the nature of the task, as inflecting singulars is arguably less demanding and more natural than 'affix stripping'. The results further showed that children's productive knowledge of the phonologically conditioned and 'natural' voicing alternation in plurals is not adult-like. Hence, children were worse at deriving singulars from novel alternating plurals such as /kɛdən/ (31% correct) than non-alternating plurals such as /kɛtən/ (53% correct). It is possible that the difference between non-alternating and alternating plurals could be due to the fact that children were more likely to regard the alternating non-words as mono-morphemic. Under this account, when children were presented with an alternating non-word (e.g. [slɑdən]), they would be more likely to perceive this word as a singular mono-morphemic form. The corpus results discussed in Chapter 4 provide some support for this possibility, as there were more mono-morphemic nouns in CELEX with medial /d/ than /t/ (see Table 5, Chapter 4). However, a similar count for nouns child-directed speech (from the van Kampen corpus) reveals that there is an equal number of mono-morphemic nouns with medial /d/ and /t/. Furthermore, mono-morphemic nouns with /t/ (e.g. *water* 'water') were more frequent. Thus, input frequencies do not support the idea that words with /d/ are more likely to be mono-morphemic. Moreover, children showed adequate knowledge of the singular - plural distinction for existing words. To conclude, it seems that many children had difficulty extending the alternating pattern to novel words, despite their knowledge of final devoicing. This suggests that children's knowledge of alternations is tied to lexical items rather than derived from their knowledge of final devoicing. In conclusion, there is no direct evidence that the Dutch voicing alternation is acquired on the basis of early knowledge of the word-final voicing neutralisation.

## 6.3 Acceptability experiments

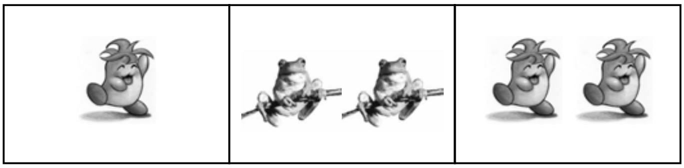
### 6.3.1 Introduction

The previous experiments showed that children have difficulty producing singulars from novel plurals. To investigate children's knowledge of singulars and plurals in a task that did not require production, Zamuner et al. (2006) carried out a 'picture selection' study.[15] If children's difficulty in producing novel singulars reflects the difficulty of the non-word production task, it was predicted that children would show better knowledge of singulars and plurals in comprehension. The previous experiments also found that children were more likely to produce plural responses for alternating non-words (e.g. /kɛdən/ ~ /kɛdən/). Therefore, the second goal of this study was to determine whether alternating plurals were also more difficult in a comprehension task (comparing [kɛtən] ~ [kɛt] to [kɛdən] ~ [kɛt]). If the distinction between non-alternating and alternating forms is also found in comprehension, this would suggest that it is difficult for children to determine the relationship between non-identical members of a paradigm. If the distinction is not found in comprehension, this would suggest that children can identify non-identical variants of a stem, but that their knowledge is not adult-like or tied to specific lexical items. Alternatively, this would indicate that there are task demands in production. Lastly, different ages were tested to determine whether there were developmental differences in children's comprehension of singulars and plurals, and in children's comprehension of alternating and non-alternating forms.

First, the results of the 'picture selection' study by Zamuner et al. (2006) will be discussed in more detail, before we turn to the other acceptability experiment. Two groups of 2;6- and 3;6-year-old children participated in the experiment, who were presented with both 'singular-to-plural and 'plural-to-singular' trials. Subjects were 18 children with a mean

[15] This work was carried out in collaboration with Tania Zamuner and Paula Fikkert, see also Zamuner, Kerkhoff & Fikkert (2006a). This research was supported by NWO grant 'Changing Lexical Representations in the Mental Lexicon' awarded to Paula Fikkert, by NWO subgrant 'The Early Acquisition of Phonotactics' awarded to Paula Fikkert and René Kager, and by NWO grant 'The Nature of Representations in Developmental Speech Perception and Production' awarded to Tania S. Zamuner. Thanks to Marleen van der Avoird and Ellen Westrek for assistance.

age of 2;7 and 15 children with a mean age of 3;7.[16] Children were tested on four non-words with /t/ and four non-words with /d/ (as before, the set of plural non-words was restricted to eight items with /t/ after short and long vowels, see Appendix I). Children were tested on either [kɛtən] ~ [kɛt] or [kɛdən] ~ [kɛt], but not both. The method was similar as described before, using fantasy animals in a PowerPoint show. In the singular-to-plural trials, the non-word was first presented in singular, ("Dit is een kɛt", 'This is a kɛt'). Upon the question ("Kun je de kɛtən/kɛdən vinden?" 'Can you find the kɛtən/kɛdən?"), children could choose between the correct plural, the singular or a filler picture, see the example below. Here, fillers were used as a control to ensure that children were not just randomly choosing pictures.



Results showed that children never chose the filler pictures, and that children understood the difference between singulars and plurals (i.e. children performed above chance on both types of trials).[17] Hence, children were able to correctly identify singulars and plurals, showing that children's difficulty with the singular-plural distinction is to be limited to production tasks that require children to posit novel singulars. Furthermore, results showed that children performed better on non-alternating non-words (i.e. with /t/) than alternating non-words (i.e. with /d/) on both types of trials (14 subjects). For instance, results for the singular-to-plural trials showed that the mean number correct for the youngest children was 0.77 for /t/ and 0.55 for /d/, while this score was 1.33 for /t/ vs. 1.00 for /d/ on the plural-to-singular trials. Note that there was no effect of trial type and no effect of age. Interestingly, although children performed above chance at picking the correct singulars or plurals, the non-alternating plurals (e.g. [kɛtən] ~ [kɛt]) were comprehended better than [kɛdən] ~ [kɛt]. This suggests that non-words with a surface alternation were more difficult to process than non-

[16] Children were first tested on the production experiments described in Zamuner et al. (2006a).
[17] Note that children were better at the singular-to-plural trials versus the plural to singular trials. However, this is primarily because when children were asked to find a singular, they sometimes would point to the plural picture and point to each individual picture.

words without a surface alternation, replicating the asymmetry found in production for comprehension. In the final comprehension experiment (Experiment VI), children were confronted with alternating items only, to determine whether they prefer voicing alternations over another type of alternation (i.e. a place alternation). Moreover, perception and production will be compared directly in this experiment, i.e. with the same group of children. This is important, because patterns often occur first in perception. Experiment VI will be described in the next section, before we turn to a general discussion of the results.

### 6.3.2 Experiment VI: children

In the production experiments described in the previous chapter, children did not always provide a plural form. When children were confronted with a plural (e.g. “are they [bɛdən]”)?, they would sometimes respond negatively, whereas they would accept a non-alternating form. This confirms the general pattern of results in which children are reluctant to accept ‘novel’ alternating forms, preferring non-alternations. The current ‘acceptability’ experiment was aimed at finding out whether children would accept a novel alternating form if they are not provided with a non-alternating alternative. Hence, it is possible that even the youngest children have enough knowledge of voicing alternations to prefer a voicing alternation over a place alternation (which does not occur in Dutch), even though they might not produce voicing alternations. To investigate this, children were tested on both words and non-words. The task was set up to ask children to judge a plural form, rather than to select the one they prefer. In case of a non-word (e.g. /kɛt/), the two alternatives were always a voicing alternation (e.g. /kɛdən/) or a place alternation (e.g. /kɛpən/). Note that the latter alternations do not occur in Dutch, although there is evidence for (other types of) place alternations cross-linguistically (see Lombardi 2001). For existing words, both the correct and incorrect alternatives were provided (e.g. /bɛdən/ and */bɛtən/ ‘beds’). Filler words and non-words for which both alternatives were incorrect were also added (e.g. */krokodɪmən/ and */krokodɪkən/ for /krokodɪlən/ ‘crocodiles’). Words were added to make the task less demanding, enabling children to respond to known words was thought to increase their confidence and interest in the game.

#### 6.3.2.1 Subjects

A total of 10 children were tested (ages ranging from 3;1 to 4;0), but data of 5 of them could not be included because they either failed to complete the perception task or accepted both answers in case of fillers, which were designed as a control. The results of the 5 remaining children (2 boys, 3 girls) who successfully completed both the production task and the perception task in a pilot session are discussed (ages 3;1, 3;2, 3;7, 3;0).

#### 6.3.2.2 Materials

A set of 12 non-words (taken from the set of 24 stimuli) was divided over the production and perception task. The production task included 6 non-words, 6 alternating words and 6 non-alternating words, two practice items and 6 fillers. In the perception task the same set of words and non-words was used, together with two practice items (one word, one non-word), three non-word fillers and two word fillers (see Appendix VI for a list of stimuli).

*Table 43: Example materials for the acceptability experiment.*

| Word | Non-word | Filler word | Filler non-word |
|---|---|---|---|
| /kɪp/ 'chicken' | /tɛp/ | /krokodɪl/ 'croc' | /flʌk/ |
| /kɪpən/? | */tɛtən/? | */krokodɪmən/? | */flʌzən/ |
| */kɪbən/? | /tɛbən/? | */krokodɪkən/? | */flʌmən/ |
| /hɑnd/ 'hand' | /dɪnt/ | /ber/ 'bear' | /baf/ |
| */hɑntən/? | /dɪndən/? | */beten/? | */banən/? |
| /hɑndən/? | */dɪmpən/? | */benən/? | */balən/? |

The materials were divided over two lists with opposite orders, i.e. the order of presentation of the two options was varied both within lists and across lists. This was so as to control for the general risk that children would always consider the second option as correct.

#### 6.3.2.3 Procedure

Children were all tested in their home environment. After the production test (for procedure see Chapter 5), children were given the opportunity to take a break, after which the second task would start. Children were told that the experimenter was not sure about the correct plural, and the child was asked to respond yes or no (e.g. “This is a /kɛt/”. “Now there are two, are they /kɛpən/?”. The filler items were added to provide a case in which both options were incorrect (see Table 43), to test whether children’s responses were reliable. Hence, these fillers would balance children’s general tendency to say ‘yes’. It was also hoped that such fillers would make children feel more comfortable about rejecting both options in case they wanted to do so. In the production task the answers were scored as ‘correct’ or ‘incorrect’ for words, and ‘alternation’, ‘non-alternation’ or ‘other’ for non-words. In the perception task, the answers were scored as ‘correct’ (i.e. in case of non-alternating words, accepting [pɛtən] ‘caps’ and rejecting *[pɛdən] or in case of alternating words, accepting [bɛdən] ‘beds’ and rejecting *[bɛtən]. Answers could also be ‘incorrect’ (accepting the wrong plural), ‘both’ (accepting both options), ‘neither’ (accepting neither option) or ‘missing’ (in case no answer was provided). In case of non-words, answers were scored as ‘voice’ (accepting [kɛdən] and rejecting [kɛpən]), ‘place’ (accepting [kɛpən] and rejecting [kɛdən]), ‘both’ (accepting both [kɛdən] and [kɛpən]), or ‘neither’ (accepting neither [kɛdən] nor [kɛpən]) or ‘missing’ (no response).

*6.3.2.4 Results*

The results will be discussed separately for words and non-words. First, the results for the *production* task are shown in Table 44.

**Words**

*Table 44: Results for words in the production task in numbers (%).*

| | **Non-alternating** | **Alternating** |
|---|---|---|
| correct | 19 (63.3%) | 10 (33.3%) |
| incorrect | 2 (6.7%) | 13 (43.3%) |
| missing | 9 (30.0%) | 7 (23.3%) |

The results show that children were better at non-alternating words, in line with the previous production experiments described in Chapter 5. The results of the acceptability experiments are shown in Table 45, for both non-alternating and alternating words.

*Table 45: Results for words in the acceptability task in numbers (%).*

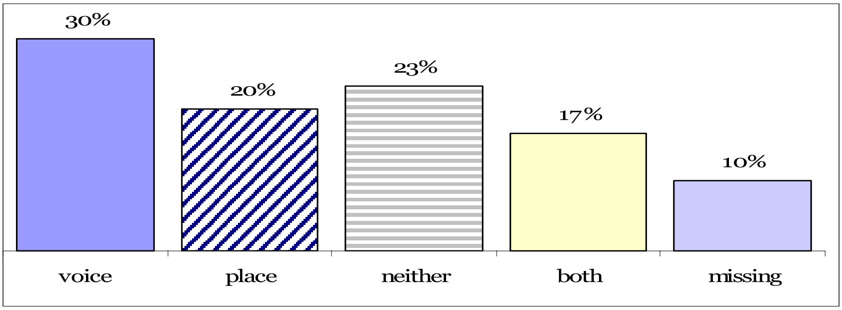
| | **Non-alternating** | **Alternating** |
|---|---|---|
| correct | 14 (46.7%) | 4 (13.3%) |
| incorrect | 2 (6.7%) | 6 (20.0%) |
| both | 9 (30.0%) | 9 (30.0%) |
| neither | 0 | 0 |
| missing | 5 (16.7%) | 11 (36.7%) |

These data show that 3-year-olds are not very good at this task, as they frequently accept both options for existing words. This could have been due to the order in which the words were provided, as children who were unsure of the correct plural would often change their mind on presentation of the second option. As was found before, children performed better on non-alternating words. Children rarely accepted an incorrect alternating plural (e.g. *[pɛdən]), whereas they were more likely to accept an incorrect non-alternating word (e.g. *[bɛtən]). In most cases, however, children preferred not to respond at all when they were unsure. Interestingly, they never rejected both options, a response that did occur for the non-words. When results are compared to the production task, it seems that they had more difficulty with the alternating words in the acceptability task than in the production task. This suggests that the acceptability task confuses children when they are unsure of the correct plural. In most cases there was correspondence between words in production and comprehension (e.g. producing [olifɑntən] and rejecting *[olifɑndən]). However, there was also considerable variation, i.e. children might produce a word incorrectly but still accept the correct form (e.g. Noa produces *[hutə] and accepts /hudən/ 'hats'). The results for non-words are discussed below.

### Non-words

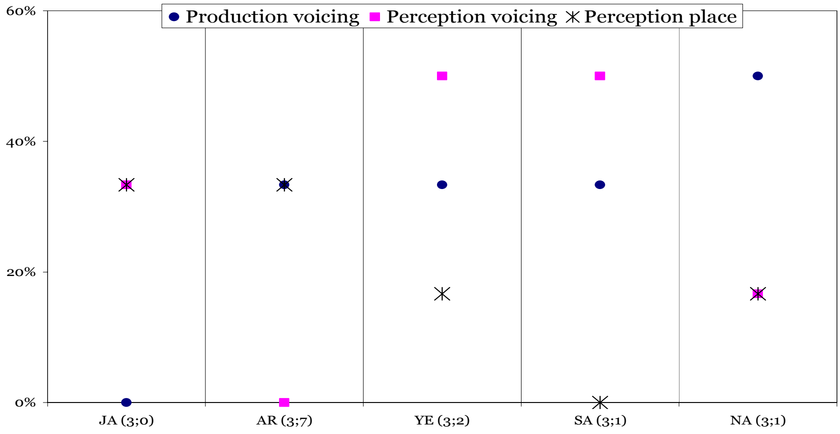
The production of non-words yielded only non-alternating responses (26/30), with 4 missing responses. Turning to the acceptance of non-words, overall results are shown in Figure 13. Recall that children never heard 'correct' non-alternating non-words but only 'correct' voicing alternations and 'incorrect' place alternations.

*Figure 13: Responses for non-words in the acceptability task.*




This outcome shows that children accepted both voicing alternations (9/30) and place alternations (6/30). Even though children were slightly more likely to accept a voicing alternation, this difference was not significant ($X^2$ = n.s.). This result is unexpected, as a place alternation does not occur in Dutch. However, children may have confused the identity of the final consonant of the singular, as many similar-sounding items (e.g. /kɛt/, /tɛp/) occurred in the experiment. Generally, the task was found to be somewhat unreliable, as children were apt to respond 'no' or 'yes' to both options. To compare children's comprehension and production of alternations, Figure 14 shows the production of voicing alternations in words and the acceptance of both types of alternations.

*Figure 14: Production of voicing alternations and acceptability of voicing and place alternations (%).*



This figure shows that there is a considerable amount of variation in response patterns. The first child (JA) clearly has very little knowledge of voicing alternations; he does not produce them and accepts them as readily as place alternations. Note that given his young age, this could also have been due to an inability to perceive the medial voicing contrast. The second child (AR) shows the unexpected pattern, accepting place alternations while rejecting voice alternations. This could have been due to a very strong preference for voiceless medial obstruents, although he does produce voicing alternations in existing words. Children YE and SA show the expected pattern, producing and accepting voicing alternations but not place alternations. Finally, child NA has difficulty accepting voicing alternations for novel words over place alternations, even though she produces the highest number of existing word alternations.

#### *6.3.2.5 Discussion*

The results of Experiment VI show that children are more likely to accept plurals like *[bɛtən] than *[pɛdən], in line with errors in production. The results for non-words showed that if children are given a choice between a voicing alternation (e.g. [kɛt] ~ [kɛdən]) and a place alternation (e.g. [kɛt] ~ [kɛpən]), they do not consistently prefer the former (which is the only 'legal' alternation). However, results should be interpreted with caution, as only few children succeeded in completing the task. Also, children sometimes accepted both or neither of the options, reflecting a possible response bias (e.g. Fritzley & Lee 2003). This experiment further showed that children may produce existing alternating plurals even though they perform poorly on non-word plurals, as was found in the previous experiments. Taken together, the results underscore children's reluctance to accept voicing alternations in novel words, showing that the alternation might be tied to specific lexical items.

### 6.3.3 General discussion

Previous studies have shown that when morpho-phonological alternations are extended to novel forms, type frequency of the pattern is the most important predictor of productivity (e.g. Bybee & Pardo 1981, Bybee 2001). Apparently, low type frequency not only restricts productivity when plurals are derived from novel singulars (see Chapter 5), it also restricts productivity in the reverse task of deriving a singular from a plural. Both the reverse Wug-test discussed in Zamuner et al. (2006) and Experiment VI showed that children were significantly worse at producing novel singulars in case of stem alternations. These results suggests that a strategy of *Paradigm Uniformity* guides children, resisting novel alternations. Furthermore, the result is in line with predictions of analogical models, according to which the alternating pattern may be tied to the lexical items that exhibit it. In this view, productivity for novel forms is limited until children attain adult-like knowledge of the alternation, in the form of a productive lexical schema. This view is further strengthened by the fact that children treat words and non-words differently. Hence, children do not seem to have a robust knowledge of voicing neutralisation in connection to alternations, as changing the identity of plural /d/ to singular /t/ was difficult for children as old as 5-years-of-age. Hence, the current investigation does not provide evidence for the claim that the early acquisition of phonotactics patterns may aid the acquisition of morphological alternations (cf. Hayes 2004). In other words, it does not support the view that the 'restructuring' of underlying representations (e.g. /bɛd/) from an earlier underlying form /bɛt/ (upon learning alternating plurals), is aided by the previously acquired rule of final devoicing.

The reverse Wug-test results also showed that children may be influenced by the likelihood that a novel form resembles an existing plural or singular. Previous findings by Anisfeld & Gordon (1968) show that older children (mean age 10;5) accepted non-words as English past tense forms more readily if the final consonant shared features with t/d. With respect to the acquisition of the English suffix alternation, acquisition research has shown that children often use base forms, especially when the stem already ends in *-d* (e.g. *need*) (e.g. Berko 1958, Derwing & Baker 1980). This phenomenon is called *affix-checking* by MacWhinney (1978), as children apparently think the form already contains an affix. Such effects are also found when children produce backformation errors, which also occurred in the current experiments. Hence, singulars such as **vark* (for *varken* 'pig') and **koei* (for *koe* 'cow') were produced for the filler items *varkens* 'pigs' and *koeien* 'cows'. This type of error is well attested in the literature, both for Dutch and other languages (e.g. Pinker & Prince 1988 report **bok* for *box*).[18]

[18] Den Os & Harder also found that Dutch words already ending in *-ən* (e.g. *toren* 'tower' and nonce *poden*) were considered to be plurals (causing a delay of *–s*). Similarly, van Wijk (2007)

Presumably, such errors result from the fact that the singular already resembles a plural form (i.e. *varken* 'pig'), reflecting product-oriented effects (Bybee & Slobin 1982). Although this error occurred sporadically in the singular-to-plural elicitation study, it occurred more often for plural-to-singular elicitation, showing task effects.

The effect of alternation appears to be greater in production tasks, because Zamuner et al. (2006) found that 2;6- and 3;6-year-old children were able to correctly identify both alternating singular and plural non-words in a comprehension task. However, children show the same reluctance to accept novel alternating forms in comprehension, as they score significantly worse on alternating plurals. Furthermore, the acceptability task showed that young children do not even consistently prefer a voicing alternation over a place alternation (which does not occur in Dutch). Hence, 3-year-olds are generally unwilling to accept novel voicing alternations, even though they may produce some existing alternating plurals. This provides further evidence for the claim that young children appear to treat such plurals as unanalysed wholes.

Children's problems with alternating forms in both production and comprehension might stem from the difficulty of relating the non-identical forms. This suggests that *Paradigm Uniformity* or stem-to-stem faithfulness guides children, because words that do not alternate have an advantage in both comprehension and production. However, note that this only applies when the members of a pair have /t/, as Paradigm Uniformity does not lead children to voice the singular in case of a plural with /d/ (e.g. /kɛdən/ ~ *[kɛd]). This means that Paradigm Uniformity is not an overriding preference for young children, as final devoicing may overrule it.

Both the production and comprehension experiments showed that there is considerable variation across children. As discussed in Chapter 2, recent proposals relate variable outcomes in morpho-phonology to conditional probabilities defined on word pairs (e.g. Skousen 1989, Albright & Hayes 2003, Baayen 2003, Daelemans & van den Bosch 2005, Ernestus & Baayen 2006). Pierrehumbert (2006) notes that morphological models such as those of Skousen (1989) and Baayen (2003) are particularly relevant because morphological derivation and back-formation is treated in the same way. A speaker may form analogies on the basis of a singular or a plural

found that bisyllabic non-words with penultimate stress (e.g. *pibos*) showed a high percentage of zero marking for 4 and 5 yr-olds.

form, which may show different rates of productivity. Variable outcomes have also been found in the domain of phonotactics. For instance, speakers are sensitive to *degrees* of phonotactic well-formedness, which is not easily captured by categorical rules or constraints (Zuraw 2000, Boersma & Hayes 2001, Frisch, Pierrehumbert & Broe 2004, Pater 2004). Hence, it seems that probabilistic models are needed to handle both static phonotactics and alternations.

As discussed in the previous chapter, the relative ease with which children learn phonotactics and morpho-phonological alternations across languages will differ according to such factors as their relative frequency and predictability (Bittner, Dressler & Kilani-Schoch 2003). Recall that corpus data showed that there are around twice as many non-alternating pairs than alternating pairs in the input, taking into account nouns, verbs and adjectives (see Chapter 4). Assuming that these corpora are representative, Dutch learners do not hear many alternating forms in the input, even though alternating singulars are associated with a high token frequency. In other words, children's attention is not drawn to the patterns that cue morpho-phonological alternations. Furthermore, as suggested in the previous chapter, the acquisition of alternations may be influenced by orthography and formal teaching, as children's lexical representations may change as soon as they learn to read and write (see, e.g. Michaels 1980, Goswami & Bryant 1990, Muneaux & Ziegler 2004). When children learn to spell, the alternating pattern can be seen more explicitly; the final sound of *hond* 'dog' is spelled with a 'd' but produced as [t]. This ties in with the observation that Dutch voicing alternations are acquired late. In a recent study by Gillis & Ravid (2006), Dutch-learning children between 6;0 and 12;0 were tested on their ability to spell pairs of nouns presented in the singular, with words containing final 't' or 'd' (e.g. *agent* 'officer' vs. *arend* 'eagle'). Even though children could infer the voicing by forming the plural (*agenten* and *arenden*), they were more likely to spell the final consonants based on how the words were produced, rather than their underlying morphology. Gillis & Ravid also studied how Hebrew-learning children spell similar morpho-phonological patterns, and found that these children rely more on morphology. Gillis & Ravid argue that this reflects the phonology and morphology of the respective languages. This type of cross-linguistic investigation of the acquisition of alternations is important, to determine to what extent language-specific influences play a role. For Dutch, knowledge of voicing alternations may be partly based on orthography (see also Dinnsen &

Charles-Luce 1984; Warner et al. 2004), although this issue awaits further research. Finally, Experiment V provided evidence that adults are capable of relating their knowledge of final devoicing to novel alternating plurals, showing implicit knowledge of voicing alternations. The difference between adults and children might be related to experience with particular lexical items and their knowledge of orthography. Furthermore, differences in metalinguistic abilities and general cognitive and maturational factors might play a role (see also Jones 1991).

In the next chapter, a summary of the previous chapters and a final conclusion will be provided.

# Chapter 7 Concluding Summary

*How children become competent users of a natural language*
*is not a logical problem but an empirical problem.*
*(Tomasello 2003:328)*

In this chapter the main findings of this dissertation are summarised, giving centre stage to the question of how and when Dutch voicing alternations are acquired. In Dutch, final neutralisation of the voice contrast leads to alternations in singular - plural pairs such as [bɛt] ~ [bɛdən] 'beds' or [hɑnt] ~ [hɑndən] 'hands'. Knowledge of voicing alternations can only be acquired on the basis of the alternating plural form (or possibly other derived forms), which needs to be related to the neutralised singular. The starting point of the current research concerned the observation that voicing alternations have been a dominant topic in phonological theory, while research into their acquisition is sparse. As noted by Bernhardt & Stemberger (1998:636), researchers studying morphological development tend to ignore or control for morpho-phonological alternations.

In Chapter 2, it was shown that morpho-phonological alternations can be roughly divided into three types, i.e. phonologically, morphologically, and lexically conditioned alternations, a subdivision that is related to the notions of regularity, productivity, and transparency. Generally, it is thought that regular, productive, and transparent alternations appear early, while unproductive, irregular, or opaque alternations may continue to develop into adulthood (e.g. Bernhardt & Stemberger 1998, Pierrehumbert 2003). Traditional generative approaches separate grammatical rules or constraints from the lexicon (e.g. Chomsky & Halle 1968, Prince & Smolensky 1993). Under this view, voicing alternations involve an abstract underlying representation (e.g. /bɛd/), which is converted to a surface phonetic form by a phonological rule of final devoicing, a regular process not restricted to a particular morphological domain. In acquisition, the surface form [bɛt] should be restructured once the child acquires the rule of final devoicing and has discovered that the (equally) 'natural' rule of intervocalic voicing does not belong to the grammar of Dutch. Natural processes constraining early child phonology have been proposed by Stampe (1973), which include both final devoicing and intervocalic voicing. Such processes are not only found in child language but are also common cross-linguistically. This relationship is also expressed in more recent constraint-based theories such as Optimality Theory (OT), in which such constraints are considered innate or emergent.

An important prediction of rule- or constraint-based models is that children are likely to 'overgeneralise' phonological generalisations such as intervocalic voicing, leading to errors such as *[pɛdən] 'caps'. A second (opposing) prediction is that constraints demanding 'paradigm-uniformity' (i.e. driving non-alternation) are ranked high in children's early grammars (e.g. Hayes 2004). Such constraints would lead to the opposite type of error, i.e. *[bɛtən] 'beds'. During acquisition, children need to arrive at the relevant rule selection or constraint ranking and adjust their lexical representations. An important assumption that underlies this view is that the processes involved are productive. In alternative models such as *Dual Mechanism* models (e.g. Pinker 1999), the distinction between regular ('default') and irregular forms is all important, as all irregular patterns are stored in memory. However, it is not easy to draw a line between regular and irregular processes, and the Dutch voicing alternation seems to position itself somewhere along this continuum (resembling to some extent 'mixed' alternations such as English *wife* ~ *wives*). Gradient effects have received much attention in analogical models of language, which do not distinguish between the grammar and the lexicon.

Analogical models include usage-based models (e.g. Bybee 2001, Tomasello 2003), in which rules or constraints are replaced by generalisations based on surface forms in the lexicon (e.g. both [bɛt] and [bɛdən] are stored and related by lexical connections). In such models, frequency of use predicts subjects' behaviour in a Wug-test, which is predicted to be probabilistic rather than rule-based. The most important prediction is that type frequency is the main determinant of productivity. If morpho-phonological alternations are associated with low type frequency and high token frequency, they are more likely to be treated as irregular forms, and acquired late. Crucially, knowledge of voicing phonotactics is not a necessary condition for learning morpho-phonological alternations. Under this view, children construct abstract schemas on the basis of words in the lexicon, while analogy may lead to overgeneralisations (e.g. *[pɛdən] 'caps'). However, errors of the opposite type (e.g. *[bɛtən] 'beds') are predicted to be dominant in Dutch, as the non-alternating pattern is more frequent. Recent versions of rule- or constraint-based theory (e.g. Boersma 2000, Boersma &

Hayes 2001) are closer to analogical models in that they allow for probabilistic effects. Furthermore, Albright & Hayes (2003) propose a 'rule-based' model that incorporates inductive learning of multiple rules, which may cover subregularities.

In Chapter 3, the Dutch voicing contrast (between voicing lead and short lag VOT) was discussed, which is generally acquired around the age of 3. Based on data from both production and perception, Kager et al. (to appear) and van der Feest (2004; 2007) argue that children's initial (devoicing) errors reflect the unmarked value, while children need to discover that the feature [voice] is active in Dutch. In order to acquire voicing alternations however, children also need to deal with the contextual variation in Dutch voicing patterns. Final devoicing is not only absent in some constructions (e.g. *heb ik* [hɛbɪk] 'have I'), the realisation of both final devoicing and assimilation is speaker- and context-dependent (e.g. Ernestus 2000). There is evidence that Dutch speakers interpret neutralised segments as underlyingly voiced or voiceless by making use of lexical patterns, which also affect adults' behaviour in a Wug-test (Ernestus & Baayen 2003; 2004). Hence, there is evidence that incomplete neutralisation plays a role when adults hear [t] in [bɛt], which activates the plural form [bɛdən]. Rather than constituting evidence for abstract underlying forms (e.g. Lahiri et al. 1990; 1992), Ernestus & Baayen (2007; to appear) argue that this finding reflects intraparadigmatic effects. Hence, the production and perception of a neutralised singular such as [bɛt] is influenced by lexical analogy to the co-activated full form [bɛdən]. These findings tie in with a more descriptive grammar of Dutch, in which probabilistic phonological or phonetic patterns are used by speakers.

In Chapter 4, new corpus data were presented on the basis of CELEX and CHILDES, to further determine the distribution of voicing alternations. Corpus data showed that alternating plurals often co-occur with vowel alternations or the irregular suffix *-eren,* while they are also associated with high token frequency. This mirrors the general interaction found between frequency and regularity, suggesting that alternating plurals are not fully regular. This pattern was also found in child-directed speech, both in Zamuner (2006) who studied the 'van de Weijer' corpus (van de Weijer 1998) and in the current study, which used the 'van Kampen' corpus (van Kampen 1997). The latter corpus contains mother-child interactions for Sarah (aged 1;6-5;2). There are roughly twice as many non-alternating than alternating *pairs* in the input, taking into account all word classes. However, frequent verbs and adjectives often undergo glide formation (e.g. *goed* [xut] ~ *goede* [xujə] 'good'). With respect to nouns, around 25% to 30% of types alternate, and the input to Sarah contained 16 non-alternating vs. 12 alternating noun *pairs*. Furthermore, highly frequent alternating plurals such as *handen* 'hands' and *kinderen* 'children' dominate early lexicons, which means they are likely to be stored as wholes. Moreover, children's input contains many diminutive forms (especially in case of 'alternating' singulars, e.g. *handje* 'hand-dim'), which do not provide evidence for the voicing alternation. Finally, it was argued that the relatively high frequency of the singular stem in relation to the plural form could affect the ability to form generalisations on the basis of an alternating singular - plural pair.

Child corpus data show that the most frequent alternating singular - plural pairs (e.g. *hand* ~ *handen* 'hands') are produced before the age of 3. For instance, Nina produces [hɑndə] 'hands' at 2;0 and [hɑnt] 'hand' at 2;2 (Zonneveld 2004:14). Children's errors in spontaneous speech were also studied in the van Kampen corpus and the CLPF corpus (Fikkert 1994, Levelt 1994). Results from the CLPF corpus (containing data from 12 children aged 1;0 to 2;11) showed that very young children distinguish between underlying /d/ and /t/. For instance, /d/ (but not /t/) after nasals was often deleted (e.g. [hɔnə] for /hɔndən/ 'dogs'), suggesting that children are aware of the voiced medial obstruent, contrasting with the voiceless one. However, this finding may also reflect misperception, as /d/ is difficult to perceive in this environment. Furthermore, children devoiced medial obstruents in around a third of noun plurals (e.g. *[bɛtən] 'beds'). This is the same overall percentage as found by van der Feest (2007) on the basis of the complete corpus, showing that mono- and bi-morphemic words were treated the same. Given that children also produced devoicing errors for mono-morphemic words, errors were argued to reflect *both* morphological *and* phonological factors, related to young children's difficulty to produce voicing.

Importantly, despite the anecdotal evidence that errors such as *[pɛdən] 'caps' are common, such errors only occurred sporadically in spontaneous speech data (i.e. only two out of twelve children in the CLPF corpus produced such errors: [habə] for /sxapən/ 'sheep' at 2;0 and [xodə] for /xrotə/ 'big' at 2;7). This finding indicates that early corpus data do not provide evidence for an early process of intervocalic voicing.

Finally, corpus data showed that the distribution of voicing alternations in Dutch is influenced by phonological context, e.g. /t/ is most likely to be voiced after nasals. On the other hand, there are very few *p~b* alternations in Dutch (except for the highly frequent pair *heb* ~ *hebben* 'have'), while there is a 'lexical gap' for /b/ after long vowels and nasals. Such frequency data are relevant to the formulation of specific predictions for the Wug-test studies discussed in Chapter 5 and 6. Hence, the set of non-words used for the elicitation experiments mirrored the phonological contexts that were used for the corpus counts.

In Chapter 5, experiments were described in which both existing words and non-word plurals were elicited, to investigate children's (productive) knowledge of voicing alternations. Children in several age groups were tested (Experiment I), as well as a group of 5-year-old typically developing children and SLI children (Experiment II) and adults (Experiment III). Results for existing words showed that even 7-year-olds are not fully competent, producing only 58% of alternating plurals correctly. A separate production experiment (testing children from 3;6 - 3;8) showed that intervocalic /d/ was produced more accurately in bi-morphemic words than mono-morphemic words (Zamuner, Kerkhoff & Fikkert 2006a). Hence, older children's difficulty with producing voicing alternations is not due to their inability to produce the medial voicing contrast. Rather, children's performance on alternating plurals was found to correlate with the frequency of the plural form. Non-word plurals were also elicited, to investigate children's productive knowledge of the alternating pattern. Importantly, when children are confronted with novel words such as /kɛt/, they need to infer the lexical representation from a neutralised singular. With respect to the possible outcome (i.e. [kɛtən] or [kɛdən]), four general hypotheses were formulated in Chapter 5: *Paradigm Uniformity*, *random selection*, *phonological generalisations* or *analogical generalisations*.

If children are guided by *Paradigm Uniformity*, they should produce non-alternations only (e.g. [kɛt] ~ [kɛtən]), or at least in larger numbers than expected on the basis of lexical frequency. Importantly, if such behaviour is due to a phonological constraint driving uniformity, non-alternations and errors such as *[bɛtən] 'beds' should occur more often for younger than for older children. Usage-based models would also predict that the alternating pattern is not very productive, because of its relatively low type frequency and high token frequency, combined with a reduced transparency (see section 2.3.2). However, such models predict that differences in productivity are based on a developing lexicon rather than rules or constraints in the grammar. Results of Experiment I (59 children aged 2;9 to 7;8) showed that errors such as *[bɛtən] 'beds' were produced in equal numbers by children in three age groups, in around 40% of cases. Such errors therefore occur more often in the elicitation task than in spontaneous speech, but no age effect was observed. In Experiment II (27 TD children and 24 SLI children with a mean age of 5;1 and 5;2 respectively), SLI children were also found to produce errors such as *[bɛtən] in more than 50% of cases, indicating that they are able to produce novel plurals. The results further showed that SLI children do not rely exclusively on lexically-based strategies and 'explicit rule learning' (c.f. Gopnik & Crago 1991, Goad & Rebellati 1994), because they distinguished between alternating and non-alternating words. Alternating words were more difficult than non-alternating words, as more voicing (e.g. *[bɛtən] 'beds') and stem change errors (e.g. *[hɑnə] for/ hɑndən/ 'hands') were produced for the former. Generally, as was found for Experiment I, performance on alternating words was affected by the frequency of the plural form. However, SLI children were found to be *less* sensitive to lexical frequency, contrary to expectations of certain 'generative' models such as the *Words and Rules* model.

Results also showed that both typically developing and SLI children produced overgeneralisations of the type *[pɛdən] 'caps' and alternations for non-words (e.g. [kɛt] ~ [kɛdən]) in 3-5% of cases. Results from Experiment III showed that adults produced fewer alternations for non-words (ranging from 2% for nouns and 1% for verbs). First, this shows that *Paradigm Uniformity* is not the only factor that guides children's word formation, even though more non-alternations were produced than expected on the basis of lexical frequency. Second, alternations for non-words were not produced randomly (i.e. by *random selection*), as there was an effect of type of final obstruent (/p/ or /t/) and rhyme (short vowel, long vowel or nasal) on the number of alternations produced.

To investigate whether children are guided by *phonological* or *analogical generalisations*, more specific predictions were formulated. Both phonological and analogical models might predict that alternations for non-words will predominantly be found for items such as [flɑnt], as postnasal voicing for /t/ is not only phonetically natural and typologically common, it is also supported in the target language. Likewise, voicing is expected after long rather than short vowels, which is also supported by the distribution of alternations in Dutch. However, for other environments models make different predictions.

First, if *phonological generalisations* play a role, voicing was predicted to occur in 'lexical gap' environments such as postnasal /p/. This is because natural classes and markedness are expected to play a role, rather than generalisations over words. In this view, young children are predicted to produce alternations such as [dep] ~ [debə] and [kɪmp] ~ [kɪmbə], which are unexpected on the basis of Dutch. Results for Experiment I showed that such alternations indeed occurred for the youngest children, although they were few in number (i.e. three in total). This suggests that young children may generalise from phonological knowledge when their lexicon is still relatively underdeveloped. The related prediction that long vowels and nasals are expected to pattern together was not borne out however, as /b/ alternations after nasals did not occur. Interestingly, Experiment II showed that children with SLI produced many more 'lexical gap' alternations than typically developing children, even extending /b/ to the nasal environment (showing a sensitivity to a natural class). This suggests that these children are especially poor at generalising over their lexicon, which may be partly due to unstructured lexical representations (see also Kerkhoff & de Bree 2005 and de Bree 2007). Adults were also found to produce lexical gap alternations in Experiment III, which further weakens their status as evidence for phonological generalisations. Generally, children were not found to be consistent within a particular phonological environment or for items with identical rhymes (e.g. producing both [jɪt] ~ [jɪdə] and [mɪt] ~ [mɪtə]), and their behaviour was not stable over a two-week period.

A second prediction of *phonological models* was that (phonetically grounded) phonological generalisations could favour voicing for /b/ rather than /d/, even though there is more evidence for /d/ in the ambient language. This prediction was not borne out, as children in all age groups produced more voicing alternations for /d/ than for /b/ (4.5% vs. 1.7%). However, alternations with /b/ were produced more often than expected on the basis of Dutch. This could be explained by the relative phonological strength of /b/ over /d/, both universally (e.g. Ohala 1983:195), and for Dutch (Zonneveld 1983:306), which is also phonetically grounded (e.g. van Alphen 2004). Surprisingly, this effect seemed stronger for adults, who produced slightly more alternations for /b/ than /d/ (although numbers were low).

A third prediction of phonological models was that children's overgeneralisations of voicing in words or alternations for non-words should decrease with age. This was not the case, as 3- and 5-year-olds produced the same number of alternations for non-words, and 5-year-olds produced more overgeneralisations of voicing for words (e.g. *[pɛdən] 'caps'). However, adults produced fewer voicing alternations than children, which means that 5-year-olds might be influenced by phonological generalisations.

Alternatively, as suggested by MacWhinney & Skousen (1994) and others, this outcome is in line with the observation that irregular forms (such as **knew*) are relatively frequent in smaller lexicons, and may form a basis for analogy. For instance, errors such as 'yesterday it **snew*' for *snowed* would be rare for adults but has been observed for children. This ties in with an important prediction of *analogical models*, which states that knowledge of existing plurals should affect overgeneralisations of voicing. Under this view, alternations are acquired on the basis of words stored in the lexicon, which predicts that acquisition is influenced by lexical frequency. In other words, the lexicon rather than age affects children's rate of overgeneralisations. Results showed that children's overgeneralisations (e.g. *[pɛdən] 'caps', *[olifɑndən] 'elephants') correlated with the number of correctly realised alternating plurals (e.g. [bɛdən] 'beds', [hɑndən] 'hands'), suggesting that overgeneralisations are based on a semi-productive *schema* for word formation. This is in line with the observation that most overgeneralisation errors were produced for the word with the lowest plural frequency (i.e. *[pɛdən] 'caps'). Such a schema is likely to be developed around age 5, after children have presumably built up a 'critical mass' of singular - plural pairs. Hence, this ties in with recent proposals according to which phonological and morphological generalisations are made on the basis of a 'critical mass' of lexical items (e.g., Marchman & Bates 1994, Bates & Goodman 1997, Beckman & Edwards 2000). Alternatively, it is possible that children's overgeneralisations are based on intervocalic voicing, resulting in 'correct' plurals (e.g. [bɛdən] 'beds' on the basis of /bɛt/) and alternations for non-words (e.g. [kɛdən]). However, this account is less plausible, given that children's performance on alternating plurals correlated with lexical frequency. Hence, children performed better when the alternating plural form was more frequent.

When children do extend the alternating pattern to non-words (e.g. [kɛdən] or [flɑndən], they were found to be influenced by the probability of voicing based on existing words. This replicates previous results for Dutch past tense inflection by adults, who were shown to be sensitive to lexical probabilities (Ernestus & Baayen 2001; 2003). The present study shows that there is a high correlation between children's rate of alternations for non-words and the probability of voicing based on nouns with identical rhymes. Moreover, a preliminary simulation study showed that a memory-based learner (TiMBL) based on CELEX closely matches subjects' behaviour (see Keuleers et al. 2007, Keuleers & Daelemans in progress).

In Chapter 6, 'reverse' Wug-test experiments were carried out, to test children's productive knowledge of final neutralisation and alternations. Results for elicitation of words showed that children were able to correctly derive singulars from plurals for both alternating and non-alternating words. Furthermore, children were better at producing alternating plurals in a plural-to-singular elicitation task than in a standard singular-to-plural elicitation task (described in Chapter 5). Results of the plural-to-singular elicitation task suggest that both alternating plurals and irregular *-eren* plurals are more easily accessed as whole forms (e.g. /hɑndən/ 'hands' and /eijərən/ 'eggs'). This finding is important from a methodological perspective, as it shows that singular-to-plural elicitation may underestimate children's knowledge of existing alternating plurals (or morpho-phonological alternations in general).

Children were also asked to produce singulars from *novel* plurals (see also Zamuner, Kerkhoff, & Fikkert 2006b). If children have productive knowledge of voicing phonotactics or final devoicing, they should be equally good at producing singulars from non-alternating (e.g. [kɛtən] ~ [kɛt]) and alternating plurals (e.g. [kɛdən] ~ [kɛt]). According to constraint-based phonological theories, early phonotactic knowledge is predicted to aid the subsequent acquisition of alternations (e.g. Hayes 2004). However, results showed that children were worse at deriving singulars for alternating plurals, showing a clear preference for non-alternations. This result is surprising under the assumption that the (phonotactic) pattern of final devoicing had been successfully related to morpho-phonological alternations. Hence, there is no direct evidence from this experiment that the knowledge of final devoicing aids the acquisition of alternations, as suggested by Hayes (2004). In contrast, adults showed implicit knowledge of the alternation as they were able to derive singulars for both types of plurals. The fact that children did not produce final voiced obstruents (e.g. *[kɛd]) shows that final devoicing may overrule the preference for *Paradigm Uniformity*. It was argued that this result may be based on a phonological rule or constraint of final devoicing, given the strength of the preference. However, from a phonological perspective, it is unclear why the data did not provide evidence for such a phonological rule or constraint of intervocalic voicing.

The comprehension experiments in Chapter 6 showed that 3-year-olds are generally reluctant to accept novel voicing alternations, even though they may produce some (frequent) alternating plurals. Moreover, children accepted novel place alternations (e.g. [kɛt] ~ [kɛpən]) as well as voicing alternations (e.g. [kɛt] ~ [kɛdən]), even though the former does not occur in Dutch. In conclusion, comprehension experiments confirm that the voicing alternation is acquired late, as children around age six are not yet able to relate the pattern productively to novel words.

In conclusion, the experimental results indicate that the lexicon plays a deciding role in the acquisition of voicing alternations. The Dutch voicing alternation is acquired late rather than early, which suggests that it does not pattern with regular or 'automatic' alternations. Even though there is some evidence for phonological generalisations in the productions of young children (around age 4), lexical effects seem dominant. Hence, voicing alternations for non-words are affected by the likelihood of voicing based on existing words in Dutch, as predicted by analogical models (Ernestus & Baayen 2003). Finally, even if voicing neutralisation is acquired by Dutch children (reflecting both language specific phonotactics and phonetic naturalness), there is no evidence that this knowledge is subsequently applied to the acquisition of alternations.

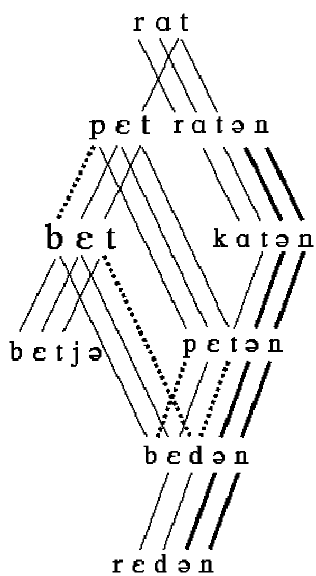
The findings presented in this dissertation are not easily captured in one theoretical framework. A suitable model for the acquisition of morpho-phonology would need to be able to handle at least the following main findings:

1. Dutch voicing alternations are associated with low type frequency and high token frequency, which is reflected in child-directed speech (Chapter 4).
2. There is little evidence for a general effect of intervocalic voicing in children's natural speech (Chapter 4).
3. Children's performance on existing alternating plurals is poor, as even 7-year-olds produce errors such as *[bɛtən] 'beds' (Chapter 4 and 5).
4. Children's performance on existing alternating (but not non-alternating) plurals in an elicitation task is influenced by the frequency of the plural form (Chapter 5).
5. The effect of 'Paradigm Uniformity' is equally strong at all ages, i.e. for 3- to 7-year-olds (Chapter 5 and 6).
6. Overgeneralisations of voicing for words and alternations for non-words are produced by 5-yr-olds rather than 3-yr-olds (Chapter 5).

7. Alternations for non-words are determined by the probability of voicing based on Dutch noun plurals rather than natural classes (Chapter 5).
8. Evidence for alternations in unexpected 'lexical gap' contexts is scarce and not limited to (TD) children, as both children with SLI and adults may produce such alternations (Chapter 5).
9. Alternations for non-words are produced variably within a phonological environment, even when items have identical rhymes (Chapter 5).
10. Children do not easily extend the alternating pattern to non-words in either a classic Wug-test or a reversed Wug-test, even though they might produce alternations for existing words (Chapter 5 and 6).
11. There is no direct evidence that children acquire voicing alternations on the basis of previously acquired knowledge of final devoicing (Chapter 6)

Taken together, these findings suggest a model in which generalisations arise over inflected word forms. There are several candidate analogical models, including Skousen (1992), Daelemans et al. (1999/2004), Bybee (2001) and Baayen (2003). These models all relate variable outcomes in morpho-phonology to conditional probabilities on word pairs (see Pierrehumbert 2006). Recent models incorporating probabilistic rules or constraints such as Boersma (1998), Boersma & Hayes (2001) or Albright & Hayes (2003) can also capture variable outcomes. These models are like analogical models, as a strict division between grammar and lexicon is not necessarily upheld. For instance, in the model proposed by Boersma (1998) the relation between underlying form and meaning is determined by lexical constraints. However, the present results show that children treat words and non-words differently, which is not easily accounted for by rule-or constraint-based models. Furthermore, these approaches do not take product-oriented effects into account, and do not incorporate meaning or lexical frequency in a structural way. In contrast, usage-based models predict that frequency of use has a deciding influence on subjects' behaviour, which is probabilistic rather than rule-based. The late acquisition of the alternation combined with children's reluctance to extend the pattern to non-words is more compatible with usage-based models such as the one proposed by Bybee (e.g. 2001), described in Chapter 2. The low productivity of the alternating pattern mirrors results obtained for diphthongisation in Spanish, which was found to be lexically specific and not often extended to novel words (see e.g., Bybee & Pardo 1981, Albright, Andrade & Hayes 2001). Children's knowledge of voicing alternations could be represented in a network model, in which lexical connections yield word-internal morphological structure (see also Chapter 2, section 2.3). In such a model, there are separate lexical representations for different inflectional forms, such as singulars and plurals (see also Blevins 2003). An example is shown in Figure 15 below, where font size reflects lexical strength, while a dotted line represents the weakest connections (in case of shared features) and thicker lines represent stronger connections (e.g. for the plural suffix). Note that there is no principled distinction between shared features for alternating or non-alternating segments.

*Figure 15: A schema for Dutch voicing alternations.*




This figure illustrates that singular and plural forms may have weaker connections in case of voicing alternations. In this type of model, the interplay between phonology and morphology can be viewed as phonetically grounded phonological constraints interacting with the paradigmatic organisation of the lexicon (in line with Bybee 2001, Pierrehumbert 2001, Ernestus & Baayen to appear; 2007). Phonological generalisations such as final devoicing (the 'automatic' phonological rules of generative theory) are seen as 'a set of neuromotor production schemas' that are highly practised or 'automised' (Bybee 2001:64). Here, the regularity, productivity and transparency of morpho-phonological processes are considered as gradient notions. Hence, principles of morpho-phonological alternation are seen as the most abstract level of generalisation, arising over relations among words (Pierrehumbert 2003). Furthermore, diachronic aspects of morpho-phonological alternations are considered to be important. For instance, final devoicing can be considered a phonetically-based sound change (e.g. changing the Middle Dutch singular *bedde* 'bed' after the schwa was lost), which resulted in alternations due to the fact that plurals (e.g. *bedden* 'beds') were resistant to paradigmatic levelling. According to Bybee, there is a "universal and almost inevitable diachronic trajectory of morphologization of phonetic processes" (Bybee 2001:68).

Usage-based models of the type mentioned above capture product-oriented effects, based on paradigmatic relations between surface forms. This mirrors a trend in recent OT models, which propose Output-to-Output correspondence constraints that directly relate surface forms to each other (e.g. McCarthy 1995;1999, Benua 1995;1997, Kager 1999ab, cf. Kiparsky 2000; to appear, who argues against such powerful Faithfulness constraints). For instance, Kager (1999a:418-420) has proposed to list both allomorphs /bɛd/ and /bɛt/ for Dutch, with affix selection through constraint interaction. As discussed in Chapter 2, O-O constraints may ultimately obviate the need for abstract underlying representations such as /bɛd/ 'bed' altogether (see, e.g., Burzio 1996, Steriade 1997). Separate storage of both the singular and the inflected plural form (e.g. [bɛt] and [bɛdən] 'beds') ties in with psycholinguistic findings for Dutch, which have been used to argue against the idea that lexical storage is maximally economical (e.g. Baayen et al. 1997; 2003) and that representations are abstract (e.g. Ernestus & Baayen 2007).

Despite the fact that voicing alternations are arguably a prime case for abstract underlying forms in traditional generative theory, the current findings are more in line with a lexical analysis of voicing alternations. Generalisations can be viewed as emergent patterns over the lexicon rather than symbolic rules, as children were found to be sensitive to the lexical distribution of voicing alternations rather than natural classes. Hence, children's generalisations may initially be surface-true and item-based, gradually becoming more and more abstract, with varying degrees of productivity (Bybee 2001, Tomasello 2005, Goldberg 2005). In this view, Dutch children initially start with a small number of unanalysed plural forms (such as the frequent plural *tanden* 'teeth' or *kinderen* 'children'). Even when subsequent plurals are stored in combination with their singular, this does not lead them to extract a pattern that is easily extended to novel forms, both in the production of plurals and in back-formations. As a larger number of alternating *pairs* are learned and stored in the lexicon, a semi-productive schema may be formed, which is abstract to the extent that it may be used in word formation. Phonetically grounded phonological constraints may play a role in the early stage of lexical development, but lexical factors are dominant, reaching a peak at around age 5. Even though productivity of such a schema for novel *singulars* will remain low (given the results for adults), the child would eventually attain adult-like performance on novel *plurals*, a process that might not be completed by age 7. Furthermore, knowledge of spelling conventions (acquired at around age 6) might increase children's knowledge of the voicing alternation. Hence, knowledge of Dutch spelling (e.g. *bed* ~ *bedden*) would render the pattern more robust due to its increased transparency. Finally, the existence of an abstract schema does not mean that underlying representations are abstract, as even the adult lexicon may contain both surface forms (i.e. /bɛt/ and /bɛdən/).

The current study shows that the Dutch voicing alternation is acquired late rather than early, suggesting that it patterns with irregular forms. This challenges the categorical distinction between such notions as regular vs. irregular, or the strict division between phonologically, morphologically and lexically conditioned alternations. Furthermore, the current findings illustrate that it is difficult to distinguish between rules or constraints on the one hand and analogy on the other. This is because more recent models allow for multiple generalisations over stored 'regular' word forms (e.g. Albright & Hayes 2001). However, the problem of distinguishing between rules and analogy is not particular to the current investigation into the Dutch voicing alternation. Hence, a general question is raised whether models' predictions are explicit enough to be empirically testable.

Future research could focus on cross-linguistic investigations, to determine to what extent language-specific influences play a role. For instance, the voicing alternation is not reflected in the spelling in languages such as Catalan, and voicing alternations are more widespread in languages such as Polish. In this way, the role of the input and the relative frequency of voicing alternations can be further investigated. Future studies could also focus on the role of orthography in changing children's lexical representations. Longitudinal studies or experiments with older children are also indicated, to find the point at which children's performance becomes adult-like. Finally, it is important to compare results obtained for the Dutch voicing alternation to the acquisition of other types of alternations. For instance, the difference between the Dutch voicing alternation and the 'early' voicing alternation in the English plural may reflect the fact that the former occurs in the stem rather than the suffix.

In general, future studies should examine the nature of the relationship between children's developing lexicons and the acquisition of morpho-phonological alternations in more detail. Hence, the precise nature of children's productive lexicons could be determined, as well as the relationship between lexicon size and overgeneralisations. The role of word class and paradigmatically related forms (such as diminutives) should also be explored further. Predictions of various models could also be tested by computer simulations based on actual child lexicons. Such research may provide more insight into the principles and mechanisms underlying the acquisition of morpho-phonology.

## Bibliography


Aguilar-Mediavilla, E. M., Sanz-Torrent, M. & Serra-Raventós, M. (2002). A comparative study of the phonology of pre-school children with specific language impairment (SLI), language delay (LD) and normal acquisition. *Clinical Linguistics & Phonetics*, 16: (8) 573-596.

Albright, A. (2002). Islands of reliability for regular morphology: evidence from Italian. *Language*, 78: 684-709.

Albright, A., Andrade, A. & Hayes, B. P. (2001). Segmental Environments of Spanish Diphthongization. In: A. Albright and T. Cho (eds.). *UCLA Working Papers in Linguistics 7 (Papers in Phonology 5)*. UCLA, 117-151.

Albright, A. & Hayes, B. P. (2003). Rules vs. analogy in English past tenses: a computational / experimental study. *Cognition*, 90: 119-161.

Alegre, M. & Gordon, P. (1999). Frequency effects and the representational status of regular inflections. Journal of Memory and Language. *Journal of Memory and Language*, 40: 41-61.

Anderson, J. R. (1974). *The Organization of Phonology*. New York: Academic Press.

Anderson, S. R. (1981). Why Phonology Isn't Natural. *Linguistic Inquiry*, 12: 493–539.

—. (1982). Where's morphology? *Linguistic Inquiry*, 13: 571-612.

Anisfeld, M. & Gordon, M. (1968). On the Psychophonological Structure of English Inflectional Rules. *Journal of Verbal Learning and Verbal Behavior*, 7: 973-979.

Anttila, A. (1997). Deriving Variation from Grammar. In: F. Hinskens et al. (eds.). *Variation, Change, and Phonological Theory*. Amsterdam: John Benjamins, 35-68.

—. (2002a). Morphologically conditioned phonological alternations. *Natural Language & Linguistic Theory*, 20: 1-42 [ROA 425].

—. (2002b). Variation and Phonological Theory. In: J. K. Chambers et al. (eds.). *The Handbook of Language Variation and Change*. Oxford: Blackwell, 206-243.

Anttila, A. & Cho, Y.-m. Y. (1998). Variation and Change in Optimality Theory. *Lingua*, 104: (Special issue on Optimality Theory) 31–56.

Archangeli, D. (1988). Aspects of underspecification theory. *Phonology*, 5: 183-207.

Archangeli, D. & Pulleyblank, D. (1994). *Grounded Phonology*. Cambridge, MA: MIT Press.

Aronoff, M. (1976). *Word Formation in Generative Grammar*. Cambridge, Mass.: MIT Press.

Aronoff, M. & Anshen, F. (1998). Morphology and the Lexicon: Lexicalization and Productivity. In: A. Spencer and A. Zwicky (eds.). *The Handbook of Morphology*. 237-247.

Avery, P. & Idsardi, W. J. (2001). Laryngeal dimensions, completion and enhancement. In: T. A. Hall (ed.) *Distinctive Feature Theory*. Berlin: Mouton/de Gruyter, 41-70.

Baayen, H. (1989). *A corpus-based approach to morphological productivity: Statistical analysis and psycholinguistic interpretation*. PhD Dissertation, Vrije Universiteit Amsterdam.

—. (1991). Quantitative aspects of morphological productivity. *Yearbook of Morphology 1991*, 109-149.

—. (2003). Probabilistic approaches to morphology. In: R. Bod et al. (eds.). *Probability theory in linguistics*. Cambridge (MA): The MIT Press, 229–287.

Baayen, H., Dijkstra, T. & Schreuder, R. (1997). Singulars and Plurals in Dutch: Evidence for a parallel Dual-Route Model. *Journal of Memory and Language*, 37: 94-117.

Baayen, H., McQueen, J. M., Dijkstra, T. & Schreuder, R. (2003). Frequency effects in regular inflectional morphology: revisiting Dutch plurals. In: H. Baayen and R. Schreuder (eds.). *Morphological structure in language processing*. Berlin: Mouton de Gruyter, 355–390.

Baayen, H., Piepenbrock, R. & Gulikers, L. (1995). *The CELEX lexical database (CD-ROM)*. Philadelphia, PA. Linguistic Data Consortium, University of Pennsylvania.

Baayen, H., Schreuder, R., de Jong, N. H. & Krott, A. (2002). Dutch inflection: The rules that prove the exception. In: S. G. Nooteboom (ed.) *Storage and computation in the language faculty*. Dordrecht: Kluwer Academic Publishers

Barlow, M. & Kemmer, S. (eds.). (2000). *Usage-based Models of Language*. Stanford, CA: CSLI Publications.

Bassano, D., Laaha, S., Maillochon, I. & Dressler, W. U. (2004). Early acquisition of verb grammar and lexical development: evidence from periphrastic constructions in French and Austrian German. *First Language*, 24: (1) 33-70.

Bates, E. (2004). Explaining and interpreting deficits in language development across clinical groups: Where do we go from here? *Brain and Language*, 88: 248-253.

Bates, E. & Goodman, J. C. (1997). On the Inseparability of Grammar and the Lexicon: Evidence from Acquisition, Aphasia and Real-time Processing. *Language and Cognitive Processes*, 12: (5/6) 507 - 584.

Bates, E. & MacWhinney, B. (1987). Competition, variation, and language learning. In: B. MacWhinney (ed.) *Mechanisms of language acquisition*. New Jersey: Erlbaum, 157-193.

Baudouin de Courtenay, J. (1895). *An Attempt at a Theory of Phonetic Alternations*. Bloomington: Indiana University Press, 1975. (Trans. E. Stankiewicz).

Baumann, M. (1995). *The production of syllables in connected speech*. Nijmegen: Benda drukkers. PhD Dissertation, University of Nijmegen.

Beckman, J. N. & Pierrehumbert, J. B. (2003). Interpreting 'phonetic interpretation' over the lexicon. *Papers in Laboratory Phonology VI*. Cambridge, UK: Cambridge University Press, 13-37.

Beckman, M. E. (2003). Input Representations (Inside the Mind and Out). In: G. Garding and M. Tsujimura (eds.). *WCCFL 22 Proceedings*. Somerville, MA: Cascadilla Press, 70-94.

Beckman, M. E. & Edwards, J. (2000). The ontogeny of phonological categories and the primacy of lexical learning in linguistic development. *Child Development*, 71: 240-249.

Beers, M. (1995). *The Phonology of Normally Developing and Language-Impaired Children*. Dordrecht: ICG Printing. PhD Dissertation, University of Amsterdam.

Benua, L. (1995). Identity effects in morphological truncation. In: J. N. Beckman (ed.) *Papers in Optimality Theory, University of Massachusetts Occasional Papers in Linguistics, 18*. Amherst MA: GLSA

—. (1997). *Transderivational Identity: Phonological Relations between Words*. PhD Dissertation, University of Massachusetts Amherst.

—. (2000). *Phonological Relations Between Words*. New York: Garland Publishing Inc.

Berendsen, E. (1983). Final devoicing, assimilation, and subject clitics in Dutch. In: H. Bennis and W. U. S. van Lessen Kloeke (eds.). *Linguistics in the Netherlands 1993*. Dordrecht: Foris Publications, 21-30.

—. (1986). *The Phonology of Cliticization*. Dordrecht: Foris Publications.

Berent, I., Pinker, S. & Shimron, J. (2002). The nature of regularity and irregularity: evidence from Hebrew nominal inflection. *Journal of psycholinguistic research*, 31: (5) 459-502.

Berko, J. (1958). The Child's Learning of English Morphology. *Word*, 14: 150-177.

Bernhardt, B. H. & Stemberger, J. P. (1998). *Handbook of Phonological Development. From the Perspective of Constraint-Based Nonlinear Phonology*. San Diego: Academic Press.

Bertram, R., Schreuder, R. & Baayen, H. (2000). The balance of storage and computation in morphological processing: the role of word formation type, affixal homonymy, and productivity. *Journal of Experimental Psychology: Learning, Memory and Cognition*, 26: (2) 489-511.

Bishop, D. (1994). Grammatical errors in specific language impairment: Competence or performance limitations. *Applied Psycholinguistics*, 15: 507-550.

Bittner, D., Dressler, W. U. & Kilani-Schoch, M. (2003). *Development of Verb Inflection in First Language Acquisition: A Cross-Linguistic Perspective*. Berlin: Mouton de Gruyter.

Blevins, J. P. (2003). Stems and paradigms. *Language*, 79: (737-767).

Bodor, P. & Barcza, V. (2007). Acquisition of diminutives in Hungarian. In: I. Savickiene and W. U. Dressler (eds.). *The Acquisition of Diminutives: A cross-linguistic perspective*. 231–262.

Boersma, P. (1997). How we learn variation, optionality, and probability. In: R. J. J. H. van Son (ed.) *Proceedings of the Institute of Phonetic Sciences 21*. Amsterdam: University of Amsterdam, Institute of Phonetic Sciences, 43–58.

—. (1998). *Functional Phonology: formalizing the interaction between articulatory and perceptual drives*. The Hague: Holland Academic Graphics. PhD Dissertation, University of Amsterdam.

—. (1999). Phonology-semantics interaction in OT, and its acquisition. In: R. Kirchner et al. (eds.). *Papers in Experimental and Theoretical Linguistics*. Edmonton: University of Alberta, 24-35.

Boersma, P. & Hayes, B. P. (2001). Empirical tests of the gradual learning algorithm. *Linguistic Inquiry*, 32: (1) 45-86.

Boersma, P. & Levelt, C. C. (2000). Gradual constraint-ranking learning algorithm predicts acquisition order. In: E. V. Clark (ed.) *Proceedings of the 30th Child Language Research Forum*. Stanford, Calif.: CSLI Publications, 229–237.

Boersma, P. & Weenink, D. (1996). *Praat: doing phonetics by computer*. University of Amsterdam. Report of the Institute of Phonetic Sciences, University of Amsterdam, 132.

Bond, Z. & Wilson, H. (1980). Acquisition of the voicing contrast by language-delayed and normal-speaking children. *Journal of Speech and Hearing Research*, 23: 152-161.

Booij, G. E. (1977). *Dutch Morphology. A study of word formation in generative grammar*. Lisse: The Peter de Ridder Press. Studies in Generative Grammar.

—. (1981). *Generatieve fonologie van het Nederlands*. Utrecht/Antwerpen: Spectrum.

—. (1985). Lexical phonology, final devoicing, and subject pronouns in Dutch. In: H. Bennis and F. Beukema (eds.). *Linguistics in the Netherlands*. Dordrecht: Foris Publications

—. (1995). *The phonology of Dutch*. Oxford: Clarendon.

—. (1997). Non-derivational Phonology Meets Lexical Phonology. In: I. Roca (ed.) *Derivations and Constraints in Phonology*. Oxford: University Press, 261-288.

—. (2002). *The Morphology of Dutch*. Oxford: Oxford University Press.

—. (2004). Review of J. L. Bybee, Phonology and language use. *Studies in Language*, 28: (1) 225-237.

Booij, G. E. & Rubach, J. (1987). Postcyclic versus Postlexical Rules in Lexical Phonology. *Linguistic Inquiry*, 18: (1) 1-44.

Bortolini, U. & Leonard, L. B. (2000). Phonology and Children with Specific Language Impairment: Status of Structural Constraints in two Languages. *Journal of Communication Disorders*, 33: 131-150.

Bortolini, U., Leonard, L. B. & Caselli, M. C. (1998). Specific Language Impairment in Italian and English: Evaluating Alternative Accounts of Grammatical Deficits. *Language and Cognitive Processes*, 13: (1) 1-20.

Braine, M. D. S. (1974). On what might constitute learnable phonology. *Language*, 50: 270-299.

Breiman, L., Friedman, J., Olshen, R. & Stone, C. (1984). *Classification and Regression Trees*. Belmont, California: Wadsworth International Group.

Brent, M. R. & Cartwright, T. A. (1996). Distributional regularity and phonotactic constraints are useful for segmentation. *Cognition*, 61: 93-125.

Brink, D. (1975). Voice assimilation in Dutch: some refinements. *Acta Linguistica Hafniensia*, 16: 11-19.

Brockhaus, W. (1995). *Final Devoicing in the Phonology of German*. Tübingen: Max Niemeyer Verlag. Linguistische Arbeiten, vol. 337.

Broersma, M. (2005). Perception of familiar contrasts in unfamiliar positions. *Journal of the Acoustical Society of America*, 117: 3890-3901.

Broselow, E., Chen, S. & Wang, C. (1998). The emergence of the unmarked in second language acquisition. *Studies in Second Language Acquisition*, 20: 261-280.

Brown, R. (1973). *A first language: the early stages*. Cambridge, Massachusetts: Harvard University Press.

Burnage, G. (1990). *CELEX, a guide for users*. Nijmegen: CELEX Centre for Lexical Information.

Burzio, L. (1996). Surface constraints versus underlying representation. In: J. Durand and B. Laks (eds.). *Current trends in phonology: Models and methods*. Paris/Salford: CNRS/University of Salford Publications, 97-122.

—. (1998). Multiple correspondence. *Lingua*, 103: 79–109.

—. (2000). Cycles, Non-Derived Environment Blocking, and Correspondence. In J. Dekkers, F. van der Leeuw, and J. van de Weijer (Eds.), Optimality Theory: Phonology, Syntax and Acquisition, 47-87. Oxford: Oxford University Press.

—. (2002). Surface-to-surface morphology: When your representations turn into constraints. In: P. Boucher (ed.) *Many Morphologies*. Somerville, MA: Cascadilla Press, 142-177. [ROA-341].

Bybee, J. L. (1985). *Morphology: A study of the relation between meaning and form*. Amsterdam: John Benjamins Publishing Company.

—. (1995). Regular Morphology and the Lexicon. *Language and Cognitive Processes*, 10: (5) 425-455.

—. (1998). Usage-based Phonology. In: M. Darnell (ed.) *Functionalism and Formalism in Linguistics*. Amsterdam / Philadelphia: Benjamins, 211-243.

—. (2000). Lexicalization of sound change and alternating environments. In: M. Broe and J. B. Pierrehumbert (eds.). *Papers in Laboratory Phonology V, Acquisition and the Lexicon*. Cambridge: Cambridge University Press, 250-268.

—. (2001). *Phonology and language use*. Cambridge: Cambridge University Press.

—. (2002a). Phonological evidence for exemplar storage of multiword sequences. *Studies in Second Language Acquisition*, 24: 215–221.

—. (2002b). Word frequency and context of use in the lexical diffusion of phonetically conditioned sound change. *Language Variation and Change*, 14: 261–290.

Bybee, J. L. & Hopper, P. (eds.). (2001). *Frequency and the Emergence of Linguistic Structure. Typological Studies in Language*. Amsterdam/Philadelphia: John Benjamins.

Bybee, J. L. & Moder, C. L. (1983). Morphological classes as natural categories. *Language*, 59: (2) 251-270.

Bybee, J. L. & Pardo, E. (1981). On lexical and morphological conditioning of alternations: a nonce-probe experiment with Spanish verbs. *Linguistics*, 9: 937-968.

Bybee, J. L. & Slobin, D. I. (1982). Rules and schemas in the development and use of the English past tense. *Language*, 58: (2) 265-289.

Cameron-Faulkner, T. & Carstairs-McCarthy, A. (2000). Stem Alternants as Morphological Signata: Evidence from Blur Avoidance in Polish Nouns? *Natural Language and Linguistic Theory*, 18: 813–835.

Catts, H. W. & Jensen, P. J. (1983). Speech timing of phonologically disordered children: voicing contrast of initial and final stop consonants. *Journal of Speech and Hearing Research*, 26: 501-510.

Cazden, C. B. (1968). The acquisition of noun and verb inflections. *Child Development*, 39: 433-448.

Charles-Luce, J. (1985). Word-final devoicing in German: effects of phonetic and sentential contexts. *Journal of Phonetics*, 15: 309-324.

—. (1993). The effects of semantic context on voicing neutralisation. *Phonetica*, 50: 28-43.

Charles-Luce, J. & Dinnsen, D. A. (1987). A reanalysis of Catalan devoicing. *Journal of Phonetics*, 15: 187-190.

Chen, M. (1970). Vowel lenght variation as a function of the voicing of the consonant environment. *Phonetica*, 22: 129-159.

Chevrot, J.-P. & Fayol, M. (2001). Acquisition of French Liaison and Related Child Errors. In: A. Almgren et al. (eds.). *Research on Child Language Acquisition*. Cascadilla Press, 760-774.

Cho, T. & Keating, P. A. (2001). Articulatory and acoustic studies on domain-initial strengthening in Korean [2001/4]. *Journal of Phonetics*, 29: (2) 155-190.

Chomsky, N. (1965). *Aspects of the Theory of Syntax*. Cambridge, Massachusetts: MIT Press.

—. (1976). A Review of B. F. Skinner's Verbal Behavior. In: L. A. Jakobovits and S. M. Murray (eds.). *Readings in the Psychology of Language*. Prentice-Hall, 142-143.

—. (1980). *Rules and Representations*. New York: Columbia University Press.

—. (1995). *The Minimalist Program*. Cambridge, MA: The MIT Press.

Chomsky, N. & Halle, M. (1968). *The sound pattern of English*. New York: Harper and Row.

Clahsen, H. (1999). Lexical entries and rules of language: A multidisciplinary study of German inflection. *Behavioral and Brain Sciences*, 22: 991-1060.

Clahsen, H., Aveledo, F. & Roca, I. (2002). The Development of Regular and Irregular Verb Inflection in Spanish. *Journal of Child Language*, 29: (3) 591-622.

Clahsen, H., Marcus, G., Bartke, S. & Wiese, R. (1995). Compounding and inflection in German child language. In: G. E. Booij and J. v. Marle (eds.). *Yearbook of Morphology 1995*. 115-142.

Clahsen, H., Rothweiler, M., Woest, A. & Marcus, G. F. (1992). Regular and irregular inflection in the acquisition of German noun plurals. *Cognition*, 45: 225-255.

Clements, G. N. (1985). The geometry of phonological features. *Phonology Yearbook*, 2: 225-252. [Reprinted as chapter 211 in Goldsmith (1999)].

Cole, J. S. (1995). The cycle in phonology. In: J. A. Goldsmith (ed.), The handbook of phonological theory, 70-113. Cambridge, MA: Blackwell.

Cole, J. S. & Hualde, J. I. (1998). The object of lexical acquisition: A UR-free model. Proceedings of the Chicago Linguistics Society, 34, 447-458.
Conti-Ramsden, G. (2003). Processing and linguistic markers in young children with specific language impairment. *Journal of Speech, Language, and Hearing Research*, 46: 1029-1037.
Conti-Ramsden, G. & Windfuhr, K. (2002). Productivity with word order and morphology: a comparative look at children with SLI and children with normal language abilities. *International Journal of Language and Communication Disorders*, 37: (1) 17-30.
Corbett, G. & Fraser, N. (1993). Network Morphology: a DATR account of Russian nominal inflection. *Journal of Linguistics*, 29: 113-142.
Corbett, G., Hippisley, A., Brown, D. & Marriot, P. (2001). Frequency, regularity and the paradigm: a perspective from Russian on a complex relation. In: J. Bybee and P. Hopper (eds.). *Frequency and the emergence of linguistic structure*. Amsterdam/Philadelphia: John Benjamins, 201-226.
Cost, S. & Salzberg, S. (1993). A weighted nearest neighbor algorithm for learning with symbolic features. *Machine Learning*, 10: 57-78.
Croft, W. (1991). *Syntactic categories and grammatical relations: The cognitive organization of information*. Chicago: University of Chicago Press.
—. (2001). *Radical Construction Grammar*. Oxford: Oxford University Press.
Croft, W. & Cruse, D. A. (2004). *Cognitive Linguistics*. Cambridge: Cambridge University Press.
Curtin, S. L. (2002). *Representational Richness in Phonological Development*. PhD Dissertation, University of Southern California.
Dabrowska, E. (2001). Learning a morphological system without a default: the Polish genitive. *Journal of Child Language*, 28: 545-574.
—. (2006). Low-level schemas or general rules? The role of diminutives in the acquisition of Polish case inflections. *Language Sciences*, 28: 120-135.
Daelemans, W., Gillis, S., Durieux, G. & van den Bosch, A. (1993). Learnability and Markedness in Data-Driven Acquisition of Stress. In: T. M. Ellison and J. M. Scobbie (eds.). *Computational Phonology. Edinburgh Working Papers in Cognitive Science 8*. 157-178.
Daelemans, W. & van den Bosch, A. (2005). *Memory-Based Language Processing*. Cambridge: Cambridge University Press. Studies in Natural Language Processing.
Daelemans, W., Zavrel, J., van der Sloot, K. & van den Bosch, A. (2004). TiMBL: Tilburg Memory-Based Learner version 5.1. Reference Guide, ILK Technical Report – ILK 04-02: Tilburg University / University of Antwerp.
Damhuis, R., de Glopper, K., Boers, M. & Kienstra, M. (1992). Woordenlijst voor 4- tot 6-jarigen. Een streeflijst voor kleuters: Projectbureau OVB Rotterdam.
Davis, K. (1995). Phonetic and Phonological contrasts in the acquisition of voicing: voice onset time production in Hindi and English. *Journal of Child Language*, 22: (2) 275-305.

de Bree, E. (2007). *Dyslexia and Phonology: A study of the phonological abilities of Dutch children at-risk of dyslexia*. Utrecht: LOT.
de Haas, W. & Trommelen, M. (1993). *Morfologisch Handboek van het Nederlands. Een overzicht van de woordvorming*. 's Gravenhage: SDU Uitgeverij.
de Houwer, A. (1990). *The acquisition of two languages from birth: A case study*. Cambridge: Cambridge University Press.
de Jong, J. (1999). *Specific language impairment in Dutch: inflectional morphology and argument structure*. Enschede: Print Partners Ipskamp. Groningen Dissertations in Linguistics 28.
de Jong, K. J. (1998). Stress-related Variation in the Articulation of Coda Alveolar Stops: Flapping Revisited. *Journal of Phonetics*, 26: 283 - 310.
de Lacy, P. (2004). Maximal words and the Maori Passive. In: J. J. McCarthy (ed.) *Optimality Theory in phonology: a reader*. Blackwell, 495-512.
de Saussure, F. (1915/1969). *Cours de Linguistique Générale. 3rd edition*. Paris: Payot.
de Schutter, G. & Taeldeman, J. (1986). Assimilatie van stem in de zuidelijke Nederlandse dialekten. In: M. Devos and e. al. (eds.). *Vruchten van zijn akker. Opstellen van (oud)medewerkers en oud-studenten voor Prof. V.F. Vanacker.*: Rijksuniversiteit Gent. Updated digital library version 2002, 91-133.
de Villiers, J. & de Villiers, P. (1973). A cross-sectional study of the acquisition of grammatical morphemes in child speech. *Journal of Psycholinguistic Research*, 2: 267-278.
de Vleeshauwer, M. (1986). *Grammaticale analyse bij drie-jarigen: een toepassing van LARSP in de verwerking van Verhulst-Schlichting*. Leuven: Katholieke Universiteit Leuven.
den Os, E. & Harder, R. (1987). De verwerving van de regels voor meervouds- en verkleinwoordsvorming in het Nederlands. *De nieuwe taalgids*, 80: (3) 240-245.
Derwing, B. L. & Baker, W. J. (1980). Rule learning and the English inflections (with special emphasis on the plural). In: G. D. Prideaux et al. (eds.). *Experimental linguistics: Integration of theory and applications*. Ghent: E.Stroy-Scientia, 247-272.
Derwing, B. L. & Skousen, R. (1994). Productivity and the English past tense: Testing Skousen's Analogy Model. In: S. D. Lima et al. (eds.). *The Reality of Linguistic Rules*. Amsterdam: John Benjamins, 193–218.
Devine, A. M. & Stephens, L. (1978). The Greek Appositives: Toward a Linguistically Adequate Definition of Caesura and Bridge. *Classical Philology*, 73: (4) 314-328.
Dinnsen, D. A. (1996). Context-sensitive underspecification and the acquisition of phonemic contrast. *Journal of Child Language*, 23: 57-79.
Dinnsen, D. A. & Barlow, J. (1998). On the characterization of a chain shift in normal and delayed phonological acquisition. *Journal of Child Language*, 25: 61–94.
Dinnsen, D. A. & Charles-Luce, J. (1984). Phonological neutralization, phonetic implementation, and individual differences. *Journal of Phonetics*, 12: 49-60.

Dinnsen, D. A. & Garcia-Zamora, M. (1971). The three degrees of vowel length in German. *Papers in Linguistics*, 4: 111-126.

Dollaghan, C. A. et al. (1999). Maternal education and measures of early speech and language. *Journal of Speech Language and Hearing Research*, 42: (1432-1444).

Donegan, P. J. & Stampe, D. (1979). The study of Natural Phonology. In: D. A. Dinnsen (ed.) *Current Approaches to Phonological Theory*. Bloomington: Indiana University Press, 126-173.

Downing, L. J., Hall, T. A. & Raffelsiefen, R. (eds.). (2005). *Paradigms in Phonological Theory*: Oxford University Press.

Dresher, B. E. (1999). Child Phonology, Learnability, and Phonological Theory. In: W. Ritchie and T. K. Bhatia (eds.). *Child Phonology, Learnability, and Phonological Theory. Handbook of Child Language Acquisition*. San Diego: Academic Press, 299-346.

Durieux, G. & Gillis, S. (2000). Predicting grammatical classes from phonological cues: An empirical test. In: B. Höhle and J. Weissenborn (eds.). *Approaches to bootstrapping: phonological, syntactic and neurophysiological aspects of early language acquisition*. Amsterdam: Benjamins, 189-232.

Eddington, D. (2000). Analogy and the dual-route model of morphology. *Lingua*, 110: 281-298.

—. (2001). Surface analogy and spelling rules in English vowel alternations. *Ms.*

Edwards, J., Beckman, J. N. & Fletcher, J. (1991). The Articulatory Kinematics of Final Lengthening. *Journal of the Acoustical Society of America*, 89: 369-382.

Edwards, J., Fourakis, M., Beckman, M. E. & Fox, R. A. (1999). Characterizing knowledge deficits in phonological disorders. *Journal of Speech, Language and Hearing Research*, 42: 169-186.

Edwards, M. L. & Shriberg, L. (1983). *Phonology: Applications in Communicative Disorders*. San Diego, CA: College-Hill Press.

Elman, J. L. et al. (1996). *Rethinking innateness: A connectionist perspective on development*. MIT Press.

Ernestus, M. (1997). The Voicedness of Intervocalic Word-Final Stops in Dutch. In: J. Coerts and H. de Hoop (eds.). *Linguistics in the Netherlands*. Amsterdam: John Benjamins, 73-84.

—. (2000). *Voice assimilation and Segment Reduction in casual Dutch*. PhD Dissertation, Free University of Amsterdam/Holland Institute of Generative Linguistics.

—. (2003). The role of phonology and phonetics in Dutch voice assimilation. In: J. M. van de Weijer et al. (eds.). *The Phonological Spectrum*. Amsterdam / Philadelphia: John Benjamins Publishing Company, 119-144.

Ernestus, M. & Baayen, H. (2001). Choosing between the Dutch past-tense suffixes –te and –de. In: T. van der Wouden and H. de Hoop (eds.). *Linguistics in the Netherlands 2001*. Amsterdam, 77-87.

—. (2003). Predicting the unpredictable: Interpreting neutralized segments in Dutch. *Language*, 79: 5-38.

—. (2004). Analogical effects in regular past tense production in Dutch. *Linguistics*, 42: (5) 873–903.

—. (2006). The functionality of incomplete neutralisation in Dutch: The case of past-tense formation. In: L. M. Goldstein et al. (eds.). *Papers in Laboratory Phonology VIII*. Berlin: Mouton de Gruyter, 27-49.

—. (2007). Paradigmatic effects in auditory word recognition: The case of alternating voice in Dutch. *Language and Cognitive Processes*, 22: 1-24.

—. (to appear). Intraparadigmatic effects on the perception of voice. In: J. M. van de Weijer and E. J. van der Torre (eds.). *Voicing in Dutch*.

Ernestus, M., Lahey, M., Verhees, F. & Baayen, H. (2004). Lexical frequency and voice assimilation in complex words in Dutch. *The Journal of the Acoustical Society of America*, 115: (5) 2610.

Féry, C. (2003). Final Devoicing and the stratification of the lexicon in German. In: J. M. van de Weijer et al. (eds.). *The Phonological Spectrum. Volume I: Segmental Structure*. Amsterdam / Philadelphia: John Benjamins, 145-169.

Fikkert, P. (1994). *On the acquisition of prosodic structure*. The Hague: Holland Academic Graphics. PhD Dissertation.

—. (1995). State-of-the-Article: Acquisition of Phonology. *Glot International*, 1: (8) 3-8.

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

Fikkert, P. & Levelt, C. C. (2002). Putting Place into Place. Paper presented at *GLOW Phonology Workshop*, Utrecht.

—. (to appear). How does place fall into place? The lexicon and emergent constraints in the developing phonological grammar. In: P. Avery et al. (eds.). *Contrast in phonology: Perception and Acquisition*. Berlin: Mouton

Fikkert, P., Levelt, C. C. & Zamuner, T. S. (2005). Perception through production? *Ms., Radboud University of Nijmegen & Leiden University*.

Fleishhacker, H. (2005). *Similarity in Phonology: Evidence from Reduplication and Loan Adaptation*. PhD Dissertation, UCLA, Los Angeles.

Forrest, K. & Morrisette, M. L. (1999). Feature analysis of segmental errors in children with phonological disorders. *Journal of Speech and Hearing Research*, 42: 187-194.

Fourakis, M. & Iverson, G. K. (1984). On the 'incomplete neutralization' of German final obstruents. *Phonetica*, 41: 140-149.

Frisch, S., Pierrehumbert, J. B. & Broe, M. (2004). Similarity Avoidance and the OCP. *Natural Language & Linguistic Theory*, 22: (1) 179 - 228.

Fritzley, V. H. & Lee, K. (2003). Do Young Children Always Say Yes to Yes-No Questions? A Metadevelopmental Study of the Affirmation Bias. *Child Development*, 74: (5) 1297–1313.

Gafos, A. (2006). Dynamics in grammar: Comments on Ladd and Ernestus & Baayen. In: L. Goldstein et al. (eds.). *Varieties of phonological competence: Papers in Laboratory Phonology 8*. New York: Mouton de Gruyter

Gekiere, P. (1983). De verwerving van morfonologische verschijnselen door jongeren van drie tot vijftien jaar. (Focus: Sint-Eloois-Vijve). *Ms., Ghent University.*

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

Gillis, S. & Ravid, D. (2006). Typological effects on spelling development: a crosslinguistic study of Hebrew and Dutch. *Journal of child language*, 33: 621-659.

Gnanadesikan, A. (2004). Markedness and faithfulness constraints in child phonology. In: R. Kager et al. (eds.). *Constraints in Phonological Acquisition.* Cambridge: Cambridge University Press, 73-108.

Goad, H. & Rebellati, C. (1994). Pluralization in familial language impairment. *McGill Working Papers in Linguistics*, 10: 24-40.

Goldberg, A. E. (1995). *Constructions. A Construction Grammar Approach to Argument Structure.* Chicago, IL: The University of Chicago Press.

—. (2003). Constructions: a new theoretical approach to language. *Trends in Cognitive Sciences*, 7: (5) 219-224.

Goldinger, S. D. (1997). Words and Voices: Perception and production in an episodic lexicon. In: K. Johnson, J. W. Mullennix (eds.), Talker variability in speech processing, 33-65. San Diego: Academic.

Goldsmith, J. A. (1993). Phonology as an intelligent system. In: D. J. Napoli and J. Kegl (eds.). *Bridges between psychology and linguistics: A Swarthmore Festschrift for Lila Gleitman.* Hillsdale, NY: Lawrence Erlbaum Associates, 247-267.

Gopnik, M. (1990). Feature blindness: A case study. *Language Acquisition*, 1.

Gopnik, M. & Crago, M. (1991). Familial aggregation of a developmental language disorder. *Cognition*, 39: 1-50.

Gopnik, M. & Goad, H. (1997). What underlies inflectional error patterns in genetic dysphasia? *Journal of Neurolinguistics*, 10: (2/3) 109-137.

Gordon, P. & Alegre, M. (1999). Is There a Dual System for Regular Inflections? *Brain and Language*, 68: 212–217.

Goswami, U. & Bryant, P. E. (1990). *Phonological skills and learning to read.* Hillsdale, NJ: Erlbaum.

Grijzenhout, J. (2000). *Voicing and devoicing in English, German, and Dutch; evidence for domain-specific identity constraints.* SFB 282 working paper no. 116. 'Two papers on constraint domains.' HHU Düsseldorf 1.

Grijzenhout, J. & Joppen-Hellwig, S. (2002). The lack of onsets in German child phonology. In: I. Lasser (ed.) *The Process of Language Acquisition (Proceedings of GALA 1999).* Frankfurt am Main: Peter Lang Verlag, 319-339.

Grijzenhout, J. & Krämer, M. (2000). Final devoicing and voicing assimilation in Dutch derivation and cliticization. In: B. Stiebels and D. Wuenderlich (eds.). *Lexicon in Focus.* Berlin: Akademie Verlag, 55-82.

Grunwell, P. (1987). *Clinical phonology.* London: Croom Helm.

Gussenhoven, C. & Bremmer, R. H. (1983). Voiced fricatives in Dutch: Sources and present-day usage. *North-Western European Language Evolution*, 2: 55-71.

Gussenhoven, C. & Jacobs, H. (1998). *Understanding Phonology*. London: Arnold. Understanding Language Series.

Hahn, U. & Nakisa, R. C. (2000). German inflection: single route or dual route? *Cognitive Psychology*, 41: (4) 313-360.

Hale, K. (1973). Deep-surface canonical disparities in relation to analysis and change: an Australian example. In: T. Sebeok (ed.) *Current Trends in Linguistics, Vol. II: Diachronic, Areal, and Typological Linguistics*. The Hague: Mouton, 401-458.

Hale, M. & Reiss, C. (1998). Formal and Empirical arguments concerning phonological acquisition. *Linguistic Inquiry*, 29: 656-683.

Hall, T. A. (1992). *Syllable Structure and Syllable-Related Processes in German*. Tübingen: Niemeyer. Linguistische Arbeiten 276.

Halle, M. (1973). Prolegomena to a theory of word formation. *Linguistic Inquiry*, 4: (3-16).

Halle, M. & Mohanan, K. P. (1985). Segmental Phonology of Modern English. *Linguistic Inquiry*, 16: 57-116.

Hare, M., Elman, J. L. & Daugherty, K. G. (1995). Default Generalisation in Connectionist Networks. *Language and Cognitive Processes*, 10: (6) 601-630.

Harms, R. (1973). *Some non-rules in English*. Indiana University Linguistics Club.

Harris, J. (1985). Spanish diphthongization and stress: a paradox resolved. *Phonology Yearbook*, 2: 31-45.

Hay, J. (2000). *Causes and Consequences of Word Structure*. PhD Dissertation, Northwestern University.

—. (2001). Lexical Frequency in Morphology: Is Everything Relative? *Linguistics*, 39: (6) 1041-1070.

Hay, J. B. & Baayen, R. H. (2005). Shifting paradigms: gradient structure in morphology [2005/7]. *Trends in Cognitive Sciences*, 9: (7) 342-348.

Hayes, B. P. (1989). Compensatory lengthening in moraic phonology. *Linguistic Inquiry*, 20: 253-306.

—. (1995). On What to Teach the Undergraduates: Some Changing Orthodoxies in Phonological Theory. In: L. Ik-Whan (ed.) *Linguistics in the Morning Calm 3*. Seoul: Hanshin, 59-77.

—. (1999). Phonetically-Driven Phonology: The Role of Optimality Theory and Inductive Grounding. In: M. Darnell et al. (eds.). *Functionalism and Formalism in Linguistics*. Amsterdam: John Benjamins, 243-285.

—. (2004). Phonological Acquisition in Optimality Theory: The Early Stages. In: R. Kager et al. (eds.). *Fixing Priorities: Constraints in Phonological Acquisition*. Cambridge: Cambridge University Press, 158-203.

Hayes, B. P., Kirchner, R. & Steriade, D. (eds.). (2004). *Phonetically-Based Phonology*: Cambridge University Press.

Hayes, B. P. & Stivers, T. (1996). A Phonetic Account of Postnasal Voicing. *Ms*.

Heemskerk, J. & Zonneveld, W. (2000). *Uitspraakwoordenboek*. Utrecht: het Spectrum.

Hoff-Ginsberg, E. (1991). Mother-child conversation in different social classes and communicative settings. *Child Development*, 62: 782-796.
Hooper [Bybee], J. B. (1976). *An introduction to natural generative phonology*. New York: Academic Press.
Hornstein, N. & Lightfoot, D. (1981). *Explanation in linguistics: the logical problem of language acquisition*. London; New York: Longman.
Hume, E. & Johnson, K. (2001). A model of the interplay of speech perception and phonology. In: E. Hume and K. Johnson (eds.). *The role of speech perception in phonology*. Academic Press, 3–26.
Ingram, D. (1981). *Procedures for the Phonological Analysis of Children's Language*. Baltimore: University Park Press.
—. (1989). *First Language Acquisition: Method, Description, Explanation*. Cambridge: Cambridge University Press.
Inkelas, S. (1990). *Prosodic Constituency in the Lexicon*. New York: Garland.
—. (1994). The consequences of optimization for underspecification. *NELS*, 25: 287-302.
Inkelas, S., Orgun, O. & Zoll, C. (1997). The implications of lexical exceptions for the nature of grammar. In: I. Roca (ed.) *Derivations and constraints in phonology*. Oxford: Clarendon, 393-418.
Itô, J. & Mester, A. (1997). Sympathy theory and German truncations. In: V. Miglio and B. Moreen (eds.). *University of Maryland Working Papers in Linguistics 5*. 117-139. [ROA-211-0897].
—. (1999). The Phonological Lexicon. In: N. Tsujimura (ed.) *The Handbook of Japanese Linguistics*. Oxford: Blackwell, 62-100.
—. (2003). *Japanese Morphophonemics. Markedness and Word Structure*. Cambridge, Mass: MIT Press.
Itô, J., Mester, A. & Padgett, J. (1995). NC: Licensing and Underspecification in Optimality Theory. *Linguistic Inquiry*, 26: (4) 571-613.
Iverson, G. K. (1982). Voice alternations in Lac Simon Algonquin. *Journal of Linguistics*, 19: 161-164.
Iverson, G. K. & Salmons, J. C. (1995). Aspiration and laryngeal representation in Germanic. *Phonology*, 12: 369-396.
—. (2003). Legacy specification in the laryngeal phonology of Dutch. *Journal of Germanic Linguistics*, 15: 1-22.
Jackendoff, R. S. (1975). Semantic and Morphological Regularities in the Lexicon. *Language*, 51: 639-671.
Jaeger, J. J. (1984). Assessing the psychological status of the vowel shift rule. *Journal of Psycholinguistic Research*, 13: 13-36.
—. (1986). On the acquisition of abstract representations for English vowels. Phonology Yearbook 3, 71-97.
Jakobson, R. (1941). *Child Language, Aphasia and Phonological Universals*. The Hague: Mouton, 1968. (Trans. A. Keiler).

Jansen, W. (2004). *Laryngeal contrast and phonetic voicing: a laboratory phonology approach to English, Hungarian, and Dutch.* Groningen Dissertations in Linguistics 47.

Jassem, W. & Richter, L. (1989). Neutralization of voicing in Polish obstruents. *Journal of Phonetics*, 17: (317-325).

Jensen, J. T. (2004). *Principles of Generative Phonology. An introduction.* Amsterdam: John Benjamins. Current Issues in Linguistic Theory, vol. 250.

Joanisse, M. F. & Seidenberg, M. S. (1998). Specific Language Impairment in children: An impairment in grammar or processing? *Trends in Cognitive Sciences*, 2: (7) 240-246.

Johnson, K. (1996). Speech perception without speaker normalization. In: K. Johnson and J. W. Mullennix (eds.). *Talker Variability in Speech Processing*. San Diego: Academic Press, 145-166.

Jones, N. K. (1991). Development of morphophonemic segments in children's mental representations of words. *Applied Psycholinguistics*, 12: (217-239).

Jongman, A., Sereno, J. A., Raaijmakers, M. & Lahiri, A. (1992). The phonological representation of [voice] in speech perception. *Language and Speech*, 35: (1,2) 137-152.

Jun, S. A. (1993). *The phonetics and phonology of Korean prosody*. PhD Dissertation, Ohio State University.

—. (1995). Asymmetrical prosodic effects on the laryngeal gesture in Korean. In: B. Connell and A. Arvaniti (eds.). *Phonology and phonetic evidence: Papers in Laboratory Phonology IV.* Cambridge: Cambridge University Press, 235-253.

Jusczyk, P. W., Hohne, E. A. & Bauman, A. (1999). Infants' sensitivity to allophonic cues for word segmentation. *Perception & Psychophysics*, 61: 1465-1476.

Kager, R. (1996). On Affix Allomorphy and Syllable Counting. In: U. Kleinhenz (ed.) *Interfaces in Phonology. (Studia Grammatica, 41).* Berlin: Akademie Verlag., 155-171. [ROA-188].

—. (1999a). *Optimality Theory*. Cambridge: Cambridge University Press. Cambridge Textbooks in Linguistics.

—. (1999b). Surface Opacity of Metrical Structure in Optimality Theory. In: B. Hermans and M. Van Oostendorp (eds.). *The Derivational Residue in Phonological Optimality Theory*. Amsterdam/Philadelphia: John Benjamins Publishing Company, 207-245.

—. (to appear). Lexical irregularity and the typology of contrast. In: K. Hanson and S. Inkelas (eds.). *The Structure of Words*. MIT Press

Kager, R., Pater, J. & Zonneveld, W. (eds.). (2004). *Constraints in Phonological Acquisition.* Cambridge: Cambridge University Press.

Kager, R. et al. (to appear). Consequences of voicing acquisition for the representation of laryngeal features. In: E. J. van der Torre and J. M. van de Weijer (eds.). *Voicing in Dutch.* Amsterdam: Benjamins

Katamba, F. (1989). *An introduction to phonology*. Harlow: Longman.

Kazazis, K. (1969). Possible evidence for (near-)underlying forms in the speech of one child. *Chicago Linguistics Society*, 5: 382-388.

Keller, F. & Asudeh, A. (2002). Probabilistic Learning Algorithms and OT. *Linguistic Inquiry*, 33: (2) 225-244.

Kemps, R., Ernestus, M., Schreuder, R. & Baayen, H. (2005). Prosodic cues for morphological complexity: The case of Dutch plural nouns. *Memory & Cognition*, 33: (3) 430-446.

Kenstowicz, M. (1994). *Phonology in Generative Grammar*. Cambridge, Massachusetts: Blackwell Publishers. Blackwell Textbooks in Linguistics.

—. (1996). Base-identity constraints and uniform exponence: Alternatives to cyclicity. In: J. Durand and B. Laks (eds.), Current trends in phonology: Models and methods, 363-393. Paris X and University of Salford: University of Salford Publications.

Kenstowicz, M. & Kisseberth, C. W. (1977). *Topics in Phonological Theory*. New York: Academic Press.

—. (1979). *Generative Phonology: Description and Theory*. San Diego: Academic Press, Inc.

Kerkhoff, A. (2004). Acquisition of Voicing alternations. In: S. Baauw and N. J. van Kampen (eds.). *Proceedings of GALA 2003*. Utrecht: LOT, 269-280.

Kerkhoff, A. & de Bree, E. (2005). Acquisition of Morpho-phonology in Children with Specific Language Impairment and Typically Developing Children. In: A. Kerkhoff et al. (eds.). *UiL-OTS Yearbook 2004*. Utrecht: LOT, 37-51.

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

Keuleers, E. & Daelemans, W. (in progress). Memory-based learning models of inflectional morphology: A methodological case study.

Keuleers, E. et al. (2007). Dutch plural inflection: The exception that proves the analogy. *Cognitive Psychology*, 54: (4) 283-318.

Kiparsky, P. (1968). *How abstract is phonology?* Bloomington, Indiana: Indiana University Linguistics Club.

—. (1975). What are phonological theories about? In: D. Cohen and J. Wirth (eds.). *Testing linguistic hypothesis*. New York: Wiley, 187–209.

—. (1982a). *Explanation in Phonology*. Dordrecht: Foris Publications. Publications in Language Series.

—. (1982b). Lexical phonology and morphology. In: I. S. Yang (ed.) *Linguistics in the Morning Calm*. Seoul: Hanshin, 3-91.

—. (1996). Allomorphy and Morphophonology. In: R. Singh (ed.) *Trubetzkoy's Orphan, Proceedings of the Montréal Roundtable "Morphonology: Contemporary Responses"*. John Benjamins Publishing

—. (2000). Opacity and Cyclicity. *The Linguistic Review*, 17: 351-367.

—. (to appear). *Opacity and Paradigm Effects*. Stanford: CSLI.

Kisseberth, C. W. (1970). On the functional unity of phonological rules. *Linguistic Inquiry*, 1: 291-306.

Köpcke, K. M. (1988). Schemas in German plural formation. *Lingua*, 74: 303-335.

—. (1998). The acquisition of plural marking in English and German revisited: schemata versus rules. *Journal of Child Language*, 25: 293-319.

Kopkalli, H. (1993). *A Phonetic and Phonological Analysis of Final Devoicing in Turkish*. Unpublished Dissertation, The University of Michigan.

Koutsoudas, A., Sanders, G. & Noll, C. (1974). The application of phonological rules. *Language*, 50: 1–28.

Krott, A., Baayen, H. & Schreuder, R. (2001). Analogy in morphology: modeling the choice of linking morphemes in Dutch. *Linguistics*, 39: (1) 51-93.

Kuczaj, S. A. (1977). The acquisition of regular and irregular past tense forms. *Journal of Verbal Learning and Verbal Behavior*, 16: 589-600.

—. (1978). Children's judgements of grammatical and ungrammatical irregular past-tense verbs. *Child Development*, 49: 319-326.

Kuijpers, C. (1993a). *Temporal Coordination in Speech Development. A study on voicing contrast and assimilation of voice*. PhD Dissertation, University of Amsterdam.

—. (1993b). The voiced-voiceless distinction in speech development of young Dutch children. *Journal of Phonetics*, 21: 313-327.

Laaha, S. & Gillis, S. (2007). Typological perspectives on the acquisition of noun and verb morphology. *Antwerp Papers in Linguistics*, 112.

Laaha, S. et al. (2007). Weakly inflecting languages: French, Dutch, and German. In: S. Laaha and S. Gillis (eds.). *Typological perspectives on the acquisition of noun and verb morphology*. Antwerp Papers in Linguistics 112, 21-35.

Laalo, K. (2007). Diminutives in Finnish child-directed and child speech. In: I. Savickiene and W. U. Dressler (eds.). *The Acquisition of Diminutives: A cross-linguistic perspective*. 263–278.

Labov, W. (1972). *Socio-linguistic patterns*. Philadelphia: University of Pennsylvania Press.

Lachter, J. & Bever, T. G. (1988). The relation between linguistic structure and associative theories of language learning: A constructive critique of some connectionist learning models. *Cognition*, 28: 195-247.

Ladefoged, P. & Maddieson, I. (1996). *The Sounds of the World's Languages*. Oxford: Blackwell Publishers Ltd.

Lahiri, A., Jongman, A. & Sereno, J. (1990). The pronominal clitic [der] in Dutch: A theoretical and experimental approach. *Yearbook of Morphology*, 3: 115-127.

Lahiri, A. & Reetz, H. (2002). Underspecified recognition. In: C. Gussenhoven and N. Warner (eds.). *Laboratory phonology 7*. Berlin: Mouton de Gruyter

Langacker, R. (1987). *Foundations of Cognitive Grammar, Volume 1*. Stanford, CA: Stanford University Press.

—. (1988). A usage-based model. In: B. Rudzka-Ostyn (ed.) *Topics in Cognitive Linguistics*. Amsterdam: Benjamins, 127-161.

—. (1990). *Concept, Image, and Symbol: the cognitive basis of grammar*. New York: Mouton de Gruyter.

—. (1991). *Foundations of Cognitive Grammar, Volume 2*. Stanford: Stanford University Press.

—. (2000). *Grammar and conceptualization*. Berlin, New York: de Gruyter.

Lasky, R. E., Syrdal-Lasky, A. & Klein, R. E. (1975). VOT discrimination by four to six and a half month old infants from Spanish environments. *Journal of Experimental Psychology*, 20: 215-225.

Leonard, L. B. (1998). *Children with Specific Language Impairment*. Boston: MIT Press.

Leonard, L. B. et al. (1985). Homonymy and the voiced-voiceless distinction in the speech of children with specific language impairment. *Journal of Speech and Hearing Research*, 28: 215-224.

Leonard, L. B., Eyer, J. A., Bedore, L. M. & Grela, B. G. (1997). Three Accounts of the Grammatical Morpheme Difficulties of English-Speaking Children With Specific Language Impairment. *Journal of Speech, Language and Hearing Research*, 40: 741-753.

Leonard, L. B., Salameh, E. K. & Hansson, K. (2001). Noun phrase morphology in Swedish-speaking children with specific language impairment. *Applied Psycholinguistics*, 22: 619-639.

Levelt, C. C. (1994). *On the acquisition of place*. PhD Dissertation, Leiden University, HIL Dissertation Series 8.

Levelt, W. J. M., Roelofs, A. & Meyer, A. S. (1999). A theory of lexical access in speech production. *Behavioral and Brain Sciences*, 22: (1) 1-38.

Lieber, R. (1981). *On the Organization of the Lexicon*. Bloomington: Indiana University Linguistics Club. PhD Dissertation, MIT.

Lisker, L. & Abramson, A. (1964). A cross-linguistic study of voicing in initial stops: Acoustical measurements. *Word*, 20: 384–422.

Lohuis-Weber, H. & Zonneveld, W. (1996). Phonological acquisition and Dutch word prosody. *Language Acquisition*, 5: 245-283.

Lombardi, L. (1991). *Laryngeal Features and Laryngeal Neutralization*. PhD Dissertation, University of Massachusetts, Amherst. Published in 1994 by Garland, New York.

—. (1995). Laryngeal Neutralization and Syllable Wellformedness. *Natural Language & Linguistic Theory*, 13: 39–74.

—. (1997). Restrictions on direction of voicing assimilation: an OT account. *Maryland Working Papers in Linguistics*, 4: 89-115.

—. (1999). Positional Faithfulness and Voicing Assimilation in Optimality Theory. *Natural Language and Linguistic Theory*, 17: 276-302.

—. (2001). Why Place and Voice are different: constraint-specific alternations in Optimality Theory. In: L. Lombardi (ed.) *Segmental Phonology in Optimality Theory: Constraints and Representations*. Cambridge: Cambridge University Press, 13-45.

Macken, M. A. (1980). The child's lexical representation: the 'puzzle–puddle–pickle' evidence. *Journal of Linguistics*, 16: 1–17.

Macken, M. A. & Barton, D. (1978). The acquisition of the voicing contrast in English: a study of voice onset time in word-initial stop consonants. *Journal of Child Language*, 7: 41-74.

—. (1979). A longitudinal study of the acquisition of the voicing contrast in American-English word-initial stops, as measured by voice onset time. *Journal of Child Language*, 7: 41–74.

—. (1980). The acquisition of the voicing contrast in Spanish: A phonetic and phonological study of word-initial stop consonants. *Journal of Child Language*, 7: 433 – 458.

MacWhinney, B. (1978). *The acquisition of morphophonology*. Monographs of the Society for Research in Child Development, serial 174, vol. 34, nos. 1-2.

—. (1991/2000). *The CHILDES Project: Tools for Analyzing Talk*. Hillsdale, NJ: Lawrence Erlbaum Associates.

MacWhinney, B., Leinbach, J., Tarban, R. & McDonald, J. (1989). Language Learning: Cues or rules? *Journal of Memory and Language*, 28: 255-288.

MacWhinney, B. & Skousen, R. (1994). Productivity and the English past tense: Testing Skousen's Analogy Model. In: S. D. Lima et al. (eds.). *The Reality of Linguistic Rules*. Amsterdam: Benjamins, 193–218.

Manaster Ramer, A. (1996). A letter from an incompletely neutral phonologist. *Journal of Phonetics*, 24: (4) 477-489.

Maratsos, M. (2000). More overregularizations after all. *Journal of Child Language*, 28: 32-54.

Marchman, V. & Bates, E. (1994). Continuity in lexical and morphological development: a test of the critical mass hypothesis. *Journal of Child Language*, 21: (2) 339-366.

Marchman, V., Mulroy, N., Fishman, E. & Schwartz, J. (1997). *The relationship between vocabulary growth and the productive use of the English plural suffix*. Poster presented at Biennial meeting of Society for Research in Child Development (SRCD), Washington, DC.

Marchman, V., Wulfeck, B. & Weismer, S. E. (1999). Morphological Productivity in Children With Normal Language and SLI: A Study of the English Past Tense. *Journal of Speech, Language, and Hearing Research*, 42: (206-219).

Marcus, G. F., Brinkmann, U., Clahsen, H. & Wiese, R. (1995). German inflection: the exception that proves the rule. *Cognitive Psychology*, 29: 189-256.

Marcus, G. F. et al. (1992). Overregularization in language acquisition. Monographs of the Society for Research in Child Development, Serial no. 228, Vol. 57.

Marshall, C. R. (2004). *The morpho-phonological interface in Specific Language Impairment*. PhD Dissertation, University College London.

Marslen-Wilson, W. D., Komisarjevsky-Tyler, L., Waksler, R. & Older, L. (1994). Morphology and meaning in the English mental lexicon. *Psychological Review*, 101: (1) 3–33.

Mascaró, J. (1987). Underlying voicing recoverability of finally devoiced obstruents in Catalan. *Journal of Phonetics*, 15: 183-186.

—. (1995). A reduction and spreading theory of voicing and other sound effects. *Catalan Working Papers in Linguistics*, 4: 267-328.

Mattys, S. L., Jusczyk, P. W., Luce, P. A. & Morgan, J. L. (1999). Phonotactic and prosodic effects on word segmentation in infants. *Cognitive Psychology*, 38: 465-494.

Maye, J., Werker, J. & Gerken, L. A. (2002). Infant sensitivity to distributional information can affect phonetic discrimination. *Cognition*, 82: 1001–1111.

McCarthy, J. J. (1998). Morpheme structure constraints and paradigm occultation. In: C. Gruber et al. (eds.). *CLS 34, vol. II: The Panels*. Chicago: Chicago Linguistic Society

—. (2005). Optimal paradigms. In: L. Downing et al. (eds.). *Paradigms in Phonological Theory*. Oxford: Oxford University Press, 170-210.

McCarthy, J. J. & Prince, A. (1993). Generalized alignment. In: G. E. Booij and J. Van Marle (eds.). *Yearbook of Morphology*. Dordrecht: Kluwer Academic Publishers, 79-154.

—. (1999). Prosodic Morphology. In: J. A. Goldsmith (ed.) *Phonological Theory, The Essential Readings*. Oxford: Blackwell, 238–288.

McMahon, A. (2000). *Change, chance, and optimality*. Oxford: Oxford University Press.

McQueen, J. M. (1998). Segmentation of continuous speech using phonotactics. *Journal of Memory and Language*, 39: 21 – 46.

McQueen, J. M. & Cutler, A. (1998). Morphology in word recognition. In: A. Spencer and A. Zwicky (eds.). *Handbook of Morphology*. 406-427.

McQueen, J. M., Dijkstra, T. & Schreuder, R. (2003). Dutch inflectional morphology in spoken-and written-word recognition. In: H. Baayen and R. Schreuder (eds.). *Language Processing*.

Menert, L. (1994). *Experiments on voice assimilation in Dutch: Prosodic structures and tempo*. PhD Dissertation, University of Utrecht.

Menn, L. (1971). Phonotactic rules in beginning speech: a study in the development of English discourse. *Lingua*, 26: 225-251.

—. (1980). Phonological Theory and Child Phonology. In: J. Yeni-Komshian et al. (eds.). *Child Phonology: Vol. 1, Production*. New York: Academic Press

Menn, L. & Matthei, E. (1992). The "Two-Lexicon" Account of Child Phonology: Looking Back, Looking Ahead. In: C. A. Ferguson et al. (eds.). *Phonological Development: Models, Research, Implications*. York Press, Inc., 211-247.

Mester, A. (1994). The Quantitative Trochee in Latin. *Natural Language & Linguistic Theory*, 12: 1-61.

Mester, A. & Ito, J. (1989). Feature predictability and underspecification: Palatal prosody in Japanese mimetics. *Language*, 65: 258–293.

Michaels, D. (1980). Spelling and the Phonology of Tense Vowels. *Language and Speech*, 23: 379-392.

Minnebo, I. (1983). Het verwerven van het meervouds- en diminutiefsysteem en morfeemstruktuurkondities bij 3- tot 15-jarigen. Een onderzoek in het Gentse. *Ms., Ghent University*.

Mohanan, K. P. (1982). *Lexical Phonology*. MIT. PhD Dissertation.

—. (1986). *The Theory of Lexical Phonology*. Dordrecht: Reidel.

Montgomery, J. W. (2002). Examining the nature of lexical processing in children with Specific Language Impairment: Temporal processing or processing capacity deficit? *Applied Psycholinguistics*, 23: 447–470.

Moscoso del Prado Martín, F., Ernestus, M. & Baayen, R. H. (2004). Do type and token effects reflect different mechanisms? Connectionist modeling of Dutch past-tense formation and final devoicing. *Brain and Language*, 90: 287-298.

Moscoso del Prado Martín, F., Kostic, A. & Baayen, H. (2004). Putting the bits together: an information theoretical perspective on morphological processing. *Cognition*, 94: (1) 1-18.

Moskowitz, B. A. (1973). Cognitive development and the acquisition of language. In: T. E. Moore (ed.), Cognitive development and the acquisition of language, 223-260. New York: Academic Press.

Muneaux, M. & Ziegler, J. C. (2004). Locus of orthographic effects on spoken word recognition: Novel insights from the neighbour generation task. *Language and Cognitive Processes*, 19: (5) 641-660.

Myers, S. (2002). Gaps in factorial typology: The case of voicing in consonant clusters. [ROA 509]. *Ms*.

Myers, S. & Hansen, B. (to appear). The origin of vowel length neutralization in final position: Evidence from Finnish speakers. *Natural Language and Linguistic Theory*.

Neijt, A., van de Hulsbeek, K. & Schreuder, R. (2006). Over liefte en olifanden. Sonoriteit van consonanten in het midden van woorden. Paper presented at *TiN-dag 2006*, Utrecht.

Nespor, M. & Vogel, I. (1986). *Prosodic phonology*. Dordrecht: Foris Publications.

Neu, H. (1980). Ranking of constraints on /t,d/ deletion in American English: A statistical analysis. In: W. Labov (ed.) *Locating language in time and space*. New York: Academic Press, 37-54.

Niemi, J. (1999). Production of Grammatical Number in Specific Language Impairment: An Elicitation Experiment on Finnish. *Brain and language*, 68: (1-2) 262-267.

Oetting, J. B. & Horohov, J. E. (1997). Past-tense marking by children with and without specific language impairment. *Journal of Speech, Language, and Hearing Research*: 62-74.

Oetting, J. B. & Rice, M. L. (1993). Plural Acquisition in Children with Specific Language Impairment. *Journal of Speech and Hearing Research*, 36: 1236-1248.

Ohala, J. J. (1983). The origin of sound patterns in vocal tract constraints. In: P. F. MacNeilage (ed.) *The Production of Speech*. New York: Springer, 189-216.

Ohala, J. J. & Ohala, M. (1986). Testing hypotheses regarding the psychological manifestation of morpheme structure constraints. In: J. J. Ohala, J. J. Jaeger (eds.), Experimental phonology, 239-52. San Diego: Academic Press.

—. (1987). Psycholinguistic probes of native speakers' phonological knowledge. In: W. U. e. a. Dressler (ed.) *Phonologica 1984*. Cambridge University Press, 227-233.

Orsolini, M., Fanari, R. & Bowles, H. (1998). Acquiring regular and irregular inflection in a language with verb classes. *Language and Cognitive Processes*, 13: 425-464.

Ota, M. (1999). *Phonological theory and the acquisition of prosodic structure: evidence from child Japanese.* PhD Dissertation, Georgetown University.

—. (2003). *The Development of Prosodic Structure in Early Words. Continuity, divergence and change.* John Benjamins. Language Acquisition and Language Disorders 34.

Owen, A. J., Dromi, E. & Leonard, L. B. (2001). The phonology-morphology interface in the speech of Hebrew-speaking children with Specific Language Impairment. *Journal of Communication Disorders*, 34: 323-337.

Paradis, C. & Prunet, J.-F. (1991). *The special status of coronals: Internal and external evidence.* San Diego: Academic Press.

Pater, J. (1999). Austronesian nasal substitution and other NC effects. In: R. Kager et al. (eds.). *The Prosody-Morphology Interface.* Cambridge: Cambridge University Press, 310-343.

—. (2000). Unpublished course handout from Ling 751, University of Massachusetts, Amherst.

—. (2002). Does phonotactic knowledge aid rule learning? Paper presented at *Boston University Applied Linguistics Speaker Series.*

—. (2004). Exceptions and Optimality Theory: Typology and Learnability. Paper presented at *Conference on Redefining Elicitation: Novel Data in Phonological Theory*, New York University.

Pater, J. & Coetzee, A. (2005). Lexically Specific Constraints: Gradience, Learnability, and Perception. *Proceedings of the 3rd Seoul International Conference on Phonology.* Seoul, 85-119.

Pater, J. & Tessier, A. M. (2003). Phonotactic Knowledge and the Acquisition of Alternations. In: M. J. Solé et al. (eds.). *Proceedings of the 15th International Congress on Phonetic Sciences.* 1777-1180.

—. (2005). Phonotactics and alternations: Testing the connection with artificial language learning. In Kathryn Flack and Shigeto Kawahara (eds.),. In: K. Flack and S. Kawahara (eds.). *University of Massachusetts Occasional Papers in Linguistics.* 1-16.

Pater, J. & Werle, A. (2003). Direction of Assimilation in Child Consonant Harmony. *Canadian Journal of Linguistics*, 48: (3/4) 385-408.

Paul, H. (1886). *Prinzipien der Sprachgeschichte.* Halle: Newmeyer.

Peperkamp, S. & Dupoux, E. (2002). Coping with Phonological Variation in Early Lexical Acquisition. In: I. Lasser (ed.) *The Process of Language Acquisition.* Berlin: Peter Lang Verlag, 359–385.

Peperkamp, S., Rozenn Le Calvez, J.-P. N. & Dupoux, E. (in press). The acquisition of allophonic rules: statistical learning with linguistic constraints. To appear in: Cognition.

Peperkamp, S., Skoruppa, K. & Dupoux, E. (2006). The Role of Phonetic Naturalness in Phonological Rule Acquisition. In: D. Bamman et al. (eds.). *Proceedings of* *the 30th Annual Boston University Conference on Language Development*. Somerville, MA: Cascadilla Press

Piaget, J. (1924). *Judgment and reasoning in the child.* London: Routledge & Kegan Paul, 1928.

Pierrehumbert, J. B. (2001). Exemplar dynamics: Word frequency, lenition and contrast. In: J. Bybee and P. Hopper (eds.). *Frequency and the emergence of linguistic structure*. Amsterdam / Philadelphia: John Benjamins

—. (2006). The Statistical Basis of an Unnatural Alternation. In: L. Goldstein et al. (eds.). *Laboratory Phonology VIII. Varieties of Phonological Competence.* Berlin: Mouton de Gruyter, 81-107.

Pinker, S. (1984). *Language Learnability and Language Development*. Cambridge, MA / London, England: Harvard University Press.

—. (1986). Productivity and conservatism in language acquisition. In: W. Demopoulos and A. Marras (eds.). *Language learning and concept acquisition*. Norwood. NJ: Ablex

—. (1989). *Learnability and cognition: the acquisition of argument structure.* Cambridge: MIT Press.

—. (1998). Words and Rules. *Lingua*, 106: 219-242.

—. (1999). *Words and Rules: the Ingredients of Language*. New York: HarperCollins Publishers Inc.

Pinker, S. & Prince, A. (1988). On language and connectionism: Analysis of a parallel distributed processing model of language acquisition. *Cognition*, 28: (1-2) 73-193.

Piroth, H. G. & Janker, P. M. (2004). Speaker-dependent differences in voicing and devoicing of German obstruents. *Journal of Phonetics*, 32: 81–109.

Plank, F. (2000). Morphological re-activation and phonological alternations: Evidence for voiceless restructuring in German. In: A. Lahiri (ed.) *Analogy, Levelling, Markedness: Principles of change in phonology and morphology*. Berlin & New York: Mouton de Gruyter, 171-191.

Plunkett, K. & Marchman, V. (1991). U-shaped learning and frequency effects in a multi-layered perceptron: implications for child language acquisition. *Cognition*, 38: 43-102.

—. (1993). From rote learning to system building: Acquiring verb morphology in children and connectionist nets. *Cognition*, 48: 21-69.

Plunkett, K. & Nakisa, R. C. (1997). A connectionist model of the Arabic plural system. *Language and Cognitive Processes*, 12: (5/6) 807-836.

Port, R. F. (1981). Linguistic timing factors in combination. *Journal of the Acoustical Society of America*, 69: (1) 262-274.

—. (1996). The discreteness of phonetic elements and formal linguistics: response to A. Manaster Ramer. *Journal of Phonetics*, 24: (4) 491-511.

Port, R. F. & Crawford, P. (1989). Incomplete Neutralization and pragmatics in German. *Journal of Phonetics*, 17: 257-282.

Port, R. F. & Leary, A. P. (2005). Against Formal Phonology. *Language*, 81: (4) 927-964.

Port, R. F., Mitleb, F. & O'Dell, M. L. (1981). Neutralization of obstruent voicing in German is incomplete. *Journal of the Acoustical Society of America*, 70: (S10).

Port, R. F. & O'Dell, M. L. (1985). Neutralization of syllable-final voicing in German. *Journal of Phonetics*, 13: 455-471.

Prasada, S. & Pinker, S. (1993). Generalisation of regular and irregular morphological patterns. *Language and Cognitive Processes*, 8: (1) 1-56.

Prince, A. & Smolensky, P. (1993). *Optimality Theory: Constraint Interaction in Generative Grammar*. Technical report, Rutgers University.

—. (2004). *Optimality theory: constraint interaction in generative grammar*. Malden, MA: Blackwell.

Prince, A. & Tesar, B. (1999). *Learning phonotactic distributions*. Technical Report RuCCS-TR-54, Rutgers University, New Brunswick [ROA-353].

—. (2004). Learning Phonotactic Distributions. In: R. Kager et al. (eds.). *Constraints in Phonological Acquisition*. Cambridge: Cambridge University Press

Pulleyblank, D. & Turkel, W. J. (2000). Learning Phonology: Genetic Algorithms and Yoruba Tongue-Root Harmony. In: J. Dekkers et al. (eds.). *Optimality Theory. Phonology, Syntax and Acquisition*. Oxford: Oxford University Press, 554-591.

Pye, S. (1986). Word-final devoicing of obstruents in Russian. *Cambridge Papers in Phonetics and Experimental Linguistics*, 5: 1-10.

Quinlan, J. R. (1993). *C4.5: Programs for machine learning*. San Mateo: MorganKaufmann.

Rice, M. L. & Oetting, J. B. (1993). Morphological Deficits of Children with SLI: Evaluation of Number, Marking, and Agreement. *Journal of Speech and Hearing Research*, 36: 1249-1257.

Rimac, R. & Smith, B. L. (1984). Acoustic characteristics of flap productions by American-English speaking children and adults: Implications concerning the development of speech motor control. *Journal of Phonetics*, 12: 387-396.

Roca, I. & Johnson, W. (1999). *A course in phonology*. Oxford, UK / Malden, Mass: Blackwell Publishers.

Roelofs, A. (1996). Serial order in planning the production of successive morphemes of a word. *Journal of Memory and Language*: (35) 854–876.

Rubach, J. (1989). Final Devoicing and Cyclic Syllabification in German. *Linguistic Inquiry*, 21: (1) 79-94.

Rubach, J. & Booij, G. E. (2001). Allomorphy in Optimality Theory: Polish Iotation. *Language*, 77: 26-60.

Rumelhart, D. E. & McClelland, J. L. (1986). *Parallel Distributed Processing: Explorations in the Microstructure of Cognition*. Cambridge, MA: MIT Press.

—. (1987). Learning the past tenses of English verbs. In: B. MacWhinney (ed.) *Mechanisms of language acquisition*. Hillsdale, NJ: Erlbaum

Ruys, E. & Trommelen, M. (2003). Constraints on post-lexical processes in Dutch. *Linguistics in the Netherlands 2003*. John Benjamins, 141-153.

Saffran, J. R., Aslin, R. N. & Newport, E. L. (1996). Statistical learning by 8-month-old infants. *Science*, 274: 1926–1928.

Sanders, E. K. (1972). When are speech sounds learned? *Journal of Speech and Hearing Disorders*, 37: 55-63.

Schaerlaekens, A. M. (1980). *De taalontwikkeling van het kind. Een oriëntatie in het Nederlandstalig onderzoek*. Groningen: Wolters-Noordhoff bv.

Schaerlaekens, A. M. & Gillis, S. (1987). *De Taalverwerving van het Kind*. Groningen: Wolters Noordhoff.

Scobbie, J. M., Gibbon, F., Hardcastle, W. J. & Fletcher, P. (2000). Covert contrast as a stage in the acquisition of phonetics and phonology. In: M. Broe and J. B. Pierrehumbert (eds.). *Papers in Laboratory Phonology V: Language Acquisition and the Lexicon*. Cambridge: Cambridge University Press, 194-207.

Seidenberg, M. S. & MacDonald, M. C. (1999). A probabilistic constraints approach to language acquisition and processing [1999/0]. *Cognitive Science*, 23: (4) 569-588.

Sereno, J. & Jongman, A. (1997). Processing of English inflectional morphology. *Memory and Cognition*, 25: 425-437.

Siegel, D. (1977). *Topics in English morphology*. Doctoral Dissertation, MIT, Cambridge, MA.

Simons, E. (2005). The acquisition of the English voice contrast by native speakers of vernacular Belgian Dutch. In: M. Tzakosta et al. (eds.). *Developmental Paths in Phonological Acquisition. Special issue of Leiden Papers in Linguistics 2.1*. 93-115.

Skousen, R. (1975). *Substantive Evidence in Phonology*. The Hague: Mouton.

—. (1989). *Analogical Modeling of Language*. Dordrecht: Kluwer.

—. (1992). *Analogy and Structure*. Dordrecht: Kluwer Academic Publishers.

Slis, I. H. (1985). *The voiced-voiceless distinction and assimilation of voice in Dutch*. PhD Dissertation, University of Nijmegen.

—. (1986). Assimilation of voice in Dutch as a function of stress, word boundaries and sex of speaker and listener. *Journal of Phonetics*, 14: 311-426.

Slis, I. H. & Cohen, A. (1969). On the complex regulating the voiced-voiceless distinction. *Language and Speech*, 12: 137-155.

Slis, I. H. & van Heugten, M. (1989). Voiced-voiceless distinction in Dutch fricatives. In: H. Bennis and A. van Kemenade (eds.). *Linguistics in the Netherlands*. Dordrecht: Foris, 123-132.

Slobin, D. I. (1985). Crosslinguistic evidence for the language-making capacity. In: D. I. Slobin (ed.) *The crosslinguistic study of language acquisition*. Hillsdale, N.J: Lawrence Erlbaum

Slowiaczek, L. & Dinnsen, D. A. (1985). On the neutralizing status of Polish word-final devoicing. *Journal of Phonetics*, 13: 325-341.

Slowiaczek, L. & Szymanska, H. J. (1989). Perception of word-final devoicing in Polish. *Journal of Phonetics*, 17: 205-212.

Smith, N. (1973). *The Acquisition of Phonology: A Case Study*. Cambridge: Cambridge University Press.

Smits, R., Warner, N., McQueen, J. M. & Cutler, A. (2003). Unfolding of phonetic information over time: A database of Dutch diphone perception. *Journal of the Acoustical Society of America*, 113: 563-574.

Smolensky, P. (1996). On the comprehension/production dilemma in child language. *Linguistic Inquiry*, 27: 720-731.

Snijders, J. T., Tellegen, P. J. & Laros, J. A. (1988). *Snijders-Oomen niet-verbale intelligentie test. Verantwoording en handleiding*. Groningen: Wolters-Noordhoff.

Snow, C. E., Smith, N. S. & Hoefnagel-Höhle, M. (1980). The acquisition of some Dutch morphological rules. *Journal of Child Language*, 7: 539-553.

Snow, D. (1997). Children's acquisition of speech timing in English: A comparative study of voice onset time and final syllable vowel lengthening. *Journal of Child Language*, 24: 35–56.

Sommerstein, A. (1974). On Phonotactically Motivated Rules. *Journal of Linguistics*, 10: 71–94.

Souman, A. & Gillis, S. (2007). A longitudinal study of the acquisition of diminutives in Dutch. In: I. Savickiene and W. U. Dressler (eds.). *The Acquisition of Diminutives: A cross-linguistic perspective*. 183–206.

Spencer, A. (1991). *Morphological Theory. An Introduction to Word Structure in Generative Grammar*. Cambridge: Blackwell.

Stampe, D. (1973). *A Dissertation on Natural Phonology*. (Published by Garland 1979). PhD Dissertation, University of Chicago.

Steriade, D. (1995). Underspecification and markedness. In: J. A. Goldsmith (ed.) *The handbook of phonological theory*. Cambridge: Blackwell, 114-174.

—. (1997). Phonetics in Phonology: The Case of Laryngeal Neutralization. *Ms*. UCLA, Los Angeles, Calif.

—. (1999). Lexical Conservatism in French Adjectival Liaison. In: B. Bullock et al. (eds.). *Formal Perspectives in Romance Linguistics*. John Benjamins, 243-270.

—. (2000). Paradigm Uniformity and the Phonetics-Phonology Boundary. In: M. Broe and J. B. Pierrehumbert (eds.). *Papers in Laboratory Phonology V, Acquisition and the Lexicon*. Cambridge University Press, 313-334.

Stoel-Gammon, C. & Dunn, C. (1985). *Normal and Disordered Phonology in Children*. Austin, TX: Pro-Ed.

Strauss, S. (ed.) (1982). *U-shaped behavioral growth*. New York: Academic Press.

Taelman, H. (2004). *Syllable omissions and additions in Dutch child language. An inquiry into the function of rhythm and the link with innate grammar*. PhD Dissertation, University of Antwerpen.

Taft, M. (1979). Recognition of affixed words and the word frequency effect. *Memory & Cognition*, 7: (4) 263-272.

—. (1994). Interactive-activation as a framework for understanding morphological processing. *Language and Cognitive Processes*, 9: (271-294).

Taft, M. & Forster, K. I. (1976). Lexical storage and retrieval of polymorphemic and polysyllabic words. *Journal of Verbal Learning and Verbal Behavior*, 15: 607-620.

Tallal, P. et al. (1996). Language comprehension in language-learning impaired children improved with acoustically modified speech. *Science*, 271: (5245) 81-84.

Tesar, B. & Prince, A. (2004). Using Phonotactics to Learn Phonological Alternations. *The Proceedings of CLS 39, Vol. II: The Panels [ROA-620].*

Tesar, B. & Smolensky, P. (1998). Learnability in optimality theory. *Linguistic Inquiry*, 29: 229–268.

—. (2000). *Learnability in OT.* MIT Press.

Tessier, A. M. (2006). Testing for OO-Faithfulness in Artificial Phonological Acquisition. In: D. Bamman et al. (eds.). *BUCLD 30: Proceedings of the 30th annual Boston University Conference on Language Development.* Somerville, MA: Cascadilla Press, 619-630.

Tomasello, M. (1992). *First verbs: a case study of early grammatical development.* Cambridge: C.U.P.

—. (1998). *The New Psychology of Language, Volume 1: Cognitive and Functional Approaches to Language Structure.* Mahwah, NJ: Lawrence Erlbaum.

—. (2003). *Constructing a Language: A Usage-Based Theory of Language Acquisition.* Cambridge, MA/London: Harvard University Press.

—. (2005). Beyond formalities: The case of language acquisition. *The Linguistic Review*, 22: 183-197.

Trommelen, M. (1983). *The Syllable in Dutch, with Special Reference to Diminutive Formation.* Dordrecht: Foris. PhD Dissertation, Utrecht University.

Trommelen, M. & Zonneveld, W. (1979). *Inleiding in de Generatieve Fonologie.* Muiderberg: Coutinho.

Trubetzkoy, N. S. (1929). Sur la morphonologie. *Travaux du cercle linguistique de Prague*, 1: 85-88.

—. (1939). *Principles of Phonology.* Berkeley: University of California Press, 1969. (Trans. C. Baltaxe).

van Alphen, P. (2004). *Perceptual relevance of prevoicing in Dutch.* PhD Dissertation, University of Nijmegen (MPI Series in Psycholinguistics 24).

van Alphen, P. et al. (2004). Early language development in children with a genetic risk for dyslexia. *Dyslexia*, 10: 265-288.

van de Hulsbeek, K. (2005). *De hoogde van de zeste rij poorden wordt verlaagd door zwaarten.* Unpublished Bachelor thesis, Radboud University Nijmegen.

van de Weijer, J. (1998). *Language Input for Word Discovery.* Wageningen: Ponsen & Looijen bv. PhD Dissertation (MPI Series in Psycholinguistics 9).

van der Feest, S. (2004). Acquiring voicing in Dutch: the role of function words. In: L. Cornips and J. Doetjes (eds.). *Linguistics in the Netherlands 2004.* Amsterdam: John Benjamins

—. (2007). *Building a phonological lexicon. The acquisition of the Dutch voicing contrast in perception and production.* PhD Dissertation, University of Nijmegen.

van der Lely, H. K. J. & Christian, V. (2000). Lexical word formation in children with grammatical SLI: a grammar-specific versus an input-processing deficit? *Cognition*: 33-63.

van der Lely, H. K. J. & Ullman, M. T. (2001). Past tense morphology in specifically language impaired and normally developing children. *Language and Cognitive Processes*, 16: (2/3) 177-217.

van Doorne, B. (1992). De verwerving van het meervouds- en diminutiefsysteem bij kinderen van drie tot achttien jaar en de invloed van sociale variabelen. *Ms., Ghent University.*

van Haeringen, C. B. (1947). *De meervoudsvorming in het Nederlands*. Amsterdam: N.V. Noord-Hollandsche Uitgevers Maatschappij.

—. (1955). Is tat juist, op tie manier. *Taal en Tongval*, 7: 49-57.

van Hout, R. (1989). Analyzing language variation by multiple regression: The case of (t) deletion in a Dutch dialect. In: R. van Hout (ed.) *New methods in dialectology*. Dordrecht, Holland: Foris

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

van Oostendorp, M. (2002). Uitzonderingen op final devoicing. Paper presented at *Morfologiedagen*, Meertens Instituut, Amsterdam.

van Valin, R. D. (1991). Functional linguistic theory and language acquisition. *First Language*, 11: 7-40.

van Weerdenburg, M., Verhoeven, L. & van Balkom, H. (2006). Towards a typology of specific language impairment. *Journal of Child Psychology and Psychiatry*, 47: (2) 176-189.

van Wijk, J. (2007). *The Acquisition of the Dutch Plural*. Utrecht: LOT.

Velten, H. V. (1943). The growth of phonemic and lexical patterns in infant language. *Language*, 19: 281-292.

Venneman, T. (1972). On the Theory of Syllabic Phonology. *Linguistische Berichte*, 18: 1-18.

Vervenne, M. (1983). Een morfonologisch onderzoek naar de verwerving van flexieverschijnselen bij kinderen tussen drie en vijftien jaar. *Ms., Ghent University.*

Wang, H. S. & Derwing, B. L. (1986). More on English vowel shift: The back vowel question. Phonology Yearbook 3, 99-116.

Warner, N., Good, E., Jongman, A. & Sereno, J. (2006). Orthographic versus morphological incomplete neutralization effects. *Journal of Phonetics*, 34: 285-293.

Warner, N., Jongman, A., Sereno, J. & Kemps, R. (2004). Incomplete neutralization and other sub-phonemic durational differences in production and perception: Evidence from Dutch. *Journal of Phonetics*, 32: 251-276.

Wedel, A. (2004). *Self-organization and the development of higher-order phonological patterns*. PhD Dissertation, University of California, Santa Cruz.

Werker, J. F. & Tees, R. C. (1984). Cross-language speech perception: Evidence for perceptual reorganization during the first year of life. *Infant Behavior and Development*, 7: 49–63.

Westbury, J. R. & Keating, P. A. (1986). On the naturalness of stop consonant voicing. *Journal of Linguistics*, 22: 145-166.

Wetzels, W. L. & Mascaró, J. (2001). The typology of voicing and devoicing. *Language*, 77: (2) 207-244.

Whalen, D. H. (1991). Infrequent words are longer in duration than frequent words. *Journal of the Acoustical Society of America*, 90: (4) 2311.

White, L., Peperkamp, S., Kirk, C. & Morgan, J. L. (2006). *A developmental trajectory for the statistical learning of phonological alternations*. Poster presented at the 31st Boston University Conference on Language Development, Boston, MA.

Whitehurst, G. J. et al. (1991). The Continuity of Babble and Speech in Children with Specific Expressive Language Delay. *Journal of Speech and Hearing Research*, 34: 1121-1129.

Wiese, R. (1996). *The Phonology of German*. Oxford: Clarendon Press.

Wijnen, F. & Elbers, L. (1993). Effort, production skill, and language learning. In: C. A. Ferguson et al. (eds.). *Phonological development*. Timonium, MD: York

Wijnen, F., Krikhaar, E. & den Os, E. (1994). The (non)realization of unstressed elements in children's utterances: evidence for a rhythmic constraint. *Journal of Child Language*, 21: (1) 59-83.

Wilson, B. & Peters, A. M. (1988). What are you cookin' on a hot? Movement Constraints in the Speech of a Three-Year-Old Blind Child. *Language*, 64: (2) 249-273.

Wissing, D. (1971). *Fonologie en morfologie van die simplekse selfstandig naamwoord in Afrikaans; 'n transformasioneel-generatieve beskrywing*. Utrecht: RUU. PhD Dissertation.

Wright, R. (2002). Factors of lexical competition in vowel articulation. In: J. Local et al. (eds.). *Papers in laboratory phonology VI*. Cambridge: Cambridge University Press

Wunderlich, D. (1996). Minimalist Morphology: the role of paradigms. *Yearbook of Morphology 1995*: 93-114.

Wunderlich, D. & Fabri, R. (1995). Minimalist morphology: An approach to inflection. *Zeitschrift für Sprachwissenschaft*, 14: (2) 236–294.

Xu, F. & Pinker, S. (1995). Weird past tense forms. *Journal of Child Language*, 22: 531-556.

Yang, C. D. (2003). *Knowledge and Learning in Natural Language*. Oxford University Press.

Yip, M. (2004). Phonological Markedness and Allomorph Selection in Zahao. *Language and Linguistics*, 5: (4) 969-1001.

Yu, A. C. L. (2004). Explaining final obstruent voicing in Lezgian: phonetics and history. *Language*, 80: (1) 73-97.

Zamuner, T. S. (2006). Sensitivity to Word-Final Phonotactics in 9- to 16-Month-Old Infants. *Infancy*, 10: (1) 77-95.

—. (2007). Phonotactics and morpho-phonological alternations: An analysis of the learner's input. Manuscript under review.

Zamuner, T. S., Gerken, L. A. & Hammond, M. (2005a). The acquisition of phonology based on input: A closer look at the relation of cross-linguistic and child language data. *Lingua*, 115: (10) 1403-1426.

Zamuner, T. S., Kerkhoff, A. & Fikkert, P. (2006a). Acquisition of voicing neutralization and alternations in Dutch. In: D. Bamman et al. (eds.). *Proceedings of the 30th Boston University Conference on Language Development*. Somerville, MA: Cascadilla Press, 701-712.

—. (2006b). Children's knowledge of how phonotactics and morphology interact. *Submitted to Language & Linguistic Processes*.

Zamuner, T. S., Kerkhoff, A., Fikkert, P. & Westrek, E. (2005b). Dutch children's acquisition of morpho-phonological alternations in plural formation. Poster presented at IASCL 2005.

Zink, I. & Lejaegere, M. (2002). *N-CDIs: Lijsten voor Communicatieve Ontwikkeling. Aanpassing en hernormering van de MacArthur CDIs van Fenson et al. []*. Leusden: Acco.

Zonneveld, W. (1978). *A Formal Theory of Exceptions in Generative Phonology*. Dordrecht: Foris. PhD Dissertation, Utrecht University.

—. (1983). Lexical and Phonological Properties of Dutch Voicing Assimilation. In: M. van Broeke et al. (eds.). *Sound and Structure, Studies for Antonie Cohen*. Dordrecht: Foris, 297-312.

—. (2004). De verwerving van een morfologisch proces: Nederlandse meervoudsvorming. *Nederlandse Taalkunde*, 9: 1-28.

—. (to appear). Issues in Dutch devoicing - and the role of Local Conjunction. In: E. J. van der Torre and J. M. van de Weijer (eds.). *Voicing in Dutch*.

Zuraw, K. (2005). The Role of Phonetic Knowledge in Phonological Patterning: Corpus and Survey Evidence from Tagalog. *Ms*. UCLA, Los Angeles.

Zuraw, K. R. (2000). *Patterned Exceptions in Phonology*. PhD Dissertation, UCLA.

Zwitserlood, P., Schriefers, H., Lahiri, A. & van Donselaar, W. (1993). The role of syllables in the perception of spoken Dutch. *Journal of Experimental Psychology: Learning, Memory and Cognition*, 19: 260-271.

**Appendix I. Stimuli Exp. I: age groups**

**Non-words Exp. I**

| | **T** | **P** |
|---|---|---|
| V_ | [kɛt] | [tɛp] |
| | [slɑt] | [xɔp] |
| | [jɪt] | [dap] |
| | [mɪt] | [zwɑp] |
| V:_ | [klat] | [tap] |
| | [fet] | [dep] |
| | [knot] | [bop] |
| | [zot] | [xlop] |
| N_ | [flɑnt] | [bɛmp] |
| | [jɔnt] | [kɪmp] |
| | [dɪnt] | [fɔmp] |
| | [xɪnt] | [tɔmp] |

**Complex words Exp. I**

| **Word** | | | **CELEX** | | **AoA*** | **SL†** |
|---|---|---|---|---|---|---|
| | | | freq sg. | freq pl. | | |
| **Non-alternating** | | | | | | |
| *pet* | [pεt] | ‘cap’ | 16 | 2 | - | 0.80 |
| *voet* | [vut] | ‘foot’ | 96 | 129 | 3.9 | 0.34 |
| *tent* | [tεnt] | ‘tent’ | 20 | 7 | 6.2 | 1.00 |
| *olifant* | [olifɑnt] | ‘elephant’ | 6 | 4 | - | 0.86 |
| *kip* | [kɪp] | ‘chicken’ | 19 | 14 | - | 0.46 |
| *aap* | [ap] | ‘monkey’ | 12 | 9 | - | 0.41 |
| *schaap* | [sxap] | ‘sheep’ | 11 | 15 | - | 0.57 |
| *lamp* | [lɑmp] | ‘lamp’ | 21 | 10 | 4.8 | 0.54 |
| **Alternating** | | | | | | |
| *bed* | [bεt] | ‘bed’ | 284 | 12 | - | 0.21 |
| *schildpad* | [sxilpɑt] | ‘turtle’ | 4 | 2 | - | ‡ |
| *hoed* | [hut] | ‘hat’ | 31 | 4 | 5.6 | 0.69 |
| *potlood* | [pɔtlot] | ‘pencil’ | 10 | 2 | - | 0.39 |
| *hand* | [hɑnt] | ‘hand’ | 645 | 377 | 3.9 | 0.30 |
| *hond* | [hɔnt] | ‘dog’ | 107 | 53 | 4.0 | 0.18 |
| *(spinne)web* | [wεp] | ‘web’ | 3 | 0 | - | 1.35 |
| *krab* | [krɑp] | ‘crab’ | 2 | 0 | 6.8 | - |

* Age of Acquisition norms reflect ratings of Dutch four- and five-letter words by 559 undergraduates from Ghent University, collected by Ghyselinck, de Moor & Brysbaert (2000).

† The *StreefLijst* (SL) ratings were taken from the *streeflijst voor 4- tot 6-jarigen* ‘target list vocabulary for 4- to 6-year-olds’ (Damhuis et al. 1992) and reflect judgments of 2585 words by 71 teachers in Dutch kindergarten schools. Teachers were asked to indicate at which point in time children should be able to produce a certain word (e.g. upon starting kindergarten). The scores ranged between 0 and 3; lower scores reflect an earlier point in time.

‡ This word was erroneously omitted from the AoA ratings although it occurs in the *Streeflijst* (Damhuis et al. 1992:39).

**Mono-morphemic words Exp. I**

| **/t/** | | | **/d/** | | |
|---|---|---|---|---|---|
| *spetter* | [spɛtər] | ‘spatter’ | *modder* | [mɔdər] | ‘mud’ |
| *appel* | [ɑpəl] | ‘apple’ | *Dribbel* | [drɪbəl] | ‘Spot’ |
| *peper* | [pepər] | ‘pepper’ | *ober* | [obər] | ‘waiter’ |
| *panter* | [pɑntər] | ‘panther’ | *vlinder* | [vlɪndər] | ‘butterfly’ |
| *rimpels* | [rɪmpəls] | ‘wrinkles’ | *ladder* | [lɑdər] | ‘ladder’ |

## Appendix II. Stimuli Exp. II : SLI and TD children

**Non-words Exp. II**

| | T | P |
|---|---|---|
| V_ | [slɑt] | [xɔp] |
| | [jɪt] | [dap] |
| V:_ | [klat] | [dep] |
| | [knot] | [bop] |
| N_ | [flɑnt] | [bɛmp] |
| | [dɪnt] | [kɪmp] |

**Words Exp. II**

| Non-alternating | | | Alternating | | |
|---|---|---|---|---|---|
| *pet* | [pɛtən] | 'cap' | *bed* | [bɛdən] | 'beds' |
| *voet* | [vutən] | 'foot' | *hoed* | [hudən] | 'hats' |
| *olifant* | [olifɑntən] | 'elephants' | *hand* | [hɑndən] | 'hands' |
| *tent* | [tɛntən] | ' tents' | *hond* | [hɔndən] | 'dogs' |
| *kip* | [kɪpən] | 'chickens' | *schildpad* | [sxilpɑdən] | 'turtles' |
| *schaap* | [sxapən] | 'sheep' | *potlood* | [pɔtlodən] | 'pencils' |
| *lamp* | [lɑmpən] | 'lamps' | *web* | [wɛbən] | 'webs' |
| | | | *krab* | [krɑbən] | 'crabs' |

## Appendix III. Stimuli Exp. III: adults

### List of experimental stimuli

1. *spuut*
2. *baaf*
3. [tɛp]
4. *gétel*
5. *kroe*
6. [slɑt]
7. *bree*
8. *praan*
9. [tɔmp]
10. *bárgel*
11. [klat]
12. *bládil*
13. [bop]
14. *mérkel*
15. [xɪnt]
16. *brin*
17. [zwɑp]
18. *bólek*
19. *bloor*
20. [zot]
21. *gamát*
22. *pruu*
23. [tap]
24. *kiroon*
25. *klak*
26. [flɑnt]
27. *flie*
28. *berel*
29. [xɔp]
30. *klon*
31. *zoján*
32. [bɛmp]
33. *dang*
34. *wádok*
35. [dep]
36. *bluik*
37. [jɪt]
38. *kla*
39. *flímpel*
40. [jɔnt]
41. *padee*
42. *slépa*
43. [dɑp]
44. *knui*
45. *peloem*
46. [fet]
47. *brégel*
48. *blaa*
49. [kɪmp]
50. *grok*
51. *padáng*
52. [kɛt]
53. *kalan*
54. [xlop]
55. *bralíng*
56. [knot]
57. *blins*
58. *flang*
59. [dɪnt]
60. *padir*
61. [fɔmp]
62. *miroos*
63. *slónkel*
64. [mɪt]
65. *káling*

### Experimental sentences

1. Mijn moeder heeft al jaren een klein ______.
2. Voor de zekerheid heb ik twee ______ aangeschaft.
3. Ik heb thuis wel drie ______.
4. We hebben alle ______ mee naar huis genomen.
5. Kinderen spelen vaak op een klein ______.
6. Gisteren zag ik ineens twee ______.
7. In die vallei stonden nog veel meer ______.
8. Zelfs een klein ______ vind ik eng.
9. De meeste ______ zijn van slechte kwaliteit.
10. In sommige grotten leven tientallen ______ bij elkaar.
11. Gisteren heb ik een paar ______ gekocht.
12. Je ziet ook wel eens een klein ______.
13. Bijna niemand heeft meer dan drie ______.
14. In China worden ______ beschouwd als delicatesse.
15. Meestal zie je twee ______ tegelijk.
16. Thuis heeft mijn oom een klein ______.
17. Ik zorg ervoor dat ik altijd twee ______ bij de hand heb.
18. In de dierentuin in Amersfoort is een klein ______ geboren.
19. Speciaal voor kinderen kun je ook een klein ______ kopen.
20. Astrologen hebben vorige week drie ______ ontdekt.
21. Deze onderzoeker heeft een klein ______ gevonden.
22. Mijn dochter heeft een grote verzameling ______.
23. Dat kind eet altijd wel drie ______ tegelijk.
24. Ik heb een klein ______ in de kast hangen.

25. Gisteren zag ik ook al een klein ______.
26. Twee ______ zijn eigenlijk al genoeg.
27. Mijn broertje van vier heeft een klein ______ gekregen.
28. In Artis leven nog twee ______.
29. Ik zou minstens vier ______ willen hebben.
30. In het bos staat een klein ______.
31. In Japan kun je nog veel ______ vinden.
32. Gisteren liet ik mijn zoontje twee _______ zien.
33. Op de veluwe zag ik laatst een klein ______.
34. Al jaren zijn ______ een delicatesse in Japan.
35. Gelukkig kreeg ik van mijn oom twee ______.
36. Mijn oma heeft veel ______ in de kast hangen.
37. Mijn vriendin had helaas al vier ______.
38. Ik wist niet dat ______ gevaarlijk kunnen zijn.
39. Ik zou ook al blij zijn met een klein ______.
40. Ik heb thuis zes verschillende ______.
41. Ik heb een klein ______ gekocht.
42. In het park staan een paar ______.
43. Laatst hebben we twee ______ gekregen.
44. Met een paar ______ kun je al een heerlijke maaltijd bereiden.
45. Mijn moeder heeft een klein ______.
46. De dansleraar bracht laatst een heleboel ______ mee.
47. In het echt heb ik wel eens een klein ______ gezien.
48. Het museum van oudheden te Leiden heeft twee ______ in bezit.
49. Dat bedrijf beschikt over enkele ______.
50. Ik heb alleen nog maar een heel klein ______ gezien.
51. Een klein ______ eten kan geen kwaad.
52. Deze week ga ik drie ______ kopen.
53. Ik ben vorig jaar gebeten door een klein ______.
54. Er lagen drie ______ op straat.
55. Een klein ______ kun je in de tuin tegenkomen.
56. Op kantoor worden veel ______ gebruikt.
57. Ik heb al drie ______ verspild.
58. Ik gebruikte laatst een klein ______ in de keuken.
59. Ik haal het liefst twee ______ tegelijk.
60. Mijn nicht kreeg een klein ______ van haar vriend.
61. Ik zag wel drie ______ in die winkel.
62. Mijn oom is ______ gaan sparen.
63. Mijn zus heeft een klein ______ gekregen.
64. Ik zag laatst een paar ______.
65. Dat is wel een heel klein ______!

## Appendix IV. Stimuli Exp. IV: children

### Non-word plurals

| | /t/ | /d/ | /p/ | /b/ |
|---|---|---|---|---|
| V_ | [ketən] | [kedən] | [tepən] | [tebən] |
| | [slɑtən] | [slɑdən] | [xɔpən] | [xɔbən] |
| V:_ | [klatən] | [kladən] | [depən] | [debən] |
| | [fetən] | [fedən] | [tapən] | [tabən] |
| N_ | [flɑntən] | [flɑndən] | [kɪmpən] | [kɪmbən] |
| | [jɔntən] | [jɔndən] | [bɛmpən] | [bɛmbən] |

### Word plurals

| Non-alternating | | | Alternating | | |
|---|---|---|---|---|---|
| *petten* | [pɛtən] | 'caps' | *bedden* | [bɛdən] | 'beds' |
| *voeten* | [vutən] | 'feet' | *hoeden* | [hudən] | 'hats' |
| *boten* | [botən] | 'boats' | *potloden* | [pɔtlodən] | 'pencils' |
| *olifanten* | [olifɑntən] | 'elephants' | *handen* | [hɑndən] | 'hands' |
| *tenten* | [tɛntən] | 'tents' | *honden* | [hɔndən] | 'dogs' |
| *kippen* | [kɪpən] | 'chickens' | *schildpadden* | [sxilpɑdən] | 'turtles' |
| *schapen* | [sxapən] | 'sheep' | *webben* | [wɛbən] | 'webs' |
| *lampen* | [lɑmpən] | 'lamps' | *krabben* | [krɑbən] | 'crabs' |

### Fillers

1. *ridders* 'knights'
2. *appels* 'apples'
3. *balonnen* 'baloons'
4. *eieren* 'eggs'
5. *varkens* 'pigs'
6. *paraplus* 'umbrellas'
7. *koeien* 'cows'
8. *vlinders* 'butterflies'
9. *kikkers* 'frogs'
10. *vissen* 'fish'

## Appendix V. Stimuli Exp. V: adults

### Non-word plurals Exp. V

| | /t/ | /d/ | /p/ | /b/ |
|---|---|---|---|---|
| V_ | [kɛtən] | [kɛdən] | [tɛpən] | [tɛbən] |
| | [slɑtən] | [slɑdən] | [xɔpən] | [xɔbən] |
| | [jɪtən] | [jɪdən] | [dapən] | [dabən] |
| | [mɪtən] | [mɪdən] | [zwɑpən] | [zwɑbən] |
| V:_ | [klatən] | [kladən] | [tapən] | [tabən] |
| | [fetən] | [fedən] | [depən] | [debən] |
| | [knotən] | [knodən] | [bopən] | [bobən] |
| | [zotən] | [zodən] | [xlopən] | [xlobən] |
| N_ | [flɑntən] | [flɑndən] | [bɛmpən] | [bɛmbən] |
| | [jɔntən] | [jɔndən] | [kɪmpən] | [kɪmbən] |
| | [dɪntən] | [dɪndən] | [fɔmpən] | [fɔmbən] |
| | [xɪntən] | [xɪndən] | [tɔmpən] | [tɔmbən] |

**Filler non-word plurals Exp. V**

1. *kimmen*
2. *blaas*
3. *lanken*
4. *kallen*
5. *taas*
6. *kroeien*
7. *fliejen*
8. *zaten*
9. *melen*
10. *huis*
11. *snaden*
12. *ronten*
13. *aten*
14. *varken*
15. *boeten*
16. *klajen*
17. *herken*
18. *gems*
19. *kleimen*
20. *klangen*
21. *hoes*
22. *kroelen*
23. *flieden*
24. *laken*
25. *klaas*
26. *blijden*

**Filler word plurals Exp. V**

1. *kikkers*
2. *fluiten*
3. *boeken*
4. *ballen*
5. *matten*
6. *zolen*
7. *bedden*
8. *bloemen*
9. *zonde-n*
10. *kralen*
11. *steden* (irr.)
12. *kinnen*
13. *baden* (irr.)
14. *kansen*
15. *vlinders*
16. *bladen* (irr.)
17. *koeien*
18. *bakkers*
19. *lade-n*
20. *petten*
21. *ronde-n*
22. *raten*
23. *padden*
24. *straten*
25. *trede-n*
26. *paden* (irr.)
27. *lepels*
28. *kleden*
29. *hoeden*
30. *stangen*
31. *leemtes*
32. *raden*
33. *apen*
34. *magen*
35. *leden* (irr.)
36. *blinde-n*
37. *zaden*
38. *bekers*
39. *kinderen*
40. *ruiten*
41. *beken*
42. *snedes*
43. *voeten*
44. *waarde-n*
45. *smeden* (irr.)
46. *bomen*
47. *palen*
48. *bonden*
49. *rokken*
50. *vaten* (irr.)

## Appendix VI. Stimuli Exp. VI: children

### Production

**Non-words**

| | T | P |
|---|---|---|
| V_ | [kɛt] | [tɛp] |
| | [jɪt] | [xɔp] |
| V:_ | [klat] | [tap] |
| | [fet] | [dep] |
| N_ | [flɑnt] | [bɛmp] |
| | [dɪnt] | [kɪmp] |

**Words**

| Non-alternating | | | Alternating | | |
|---|---|---|---|---|---|
| *pet* | [pɛtən] | ‘caps’ | *bed* | [bɛdən] | ‘beds’ |
| *voet* | [vutən] | ‘feet’ | *hoed* | [hudən] | ‘hats’ |
| *boot* | [botən] | ‘boats’ | *hand* | [hɑndən] | ‘hands’ |
| *olifant* | [olifɑntən] | ‘elephants’ | *hond* | [hɔndən] | ‘dogs’ |
| *tent* | [tɛntən] | ‘ tents’ | *schildpad* | [sxilpɑdən] | ‘turtles’ |
| *kip* | [kɪpən] | ‘chickens’ | *potlood* | [pɔtlodən] | ‘pencils’ |
| *pop* | [pɔpən] | ‘dolls’ | *aap* | [apən] | ‘monkey’ |
| *schaap* | [sxapən] | ‘sheep’ | *web* | [wɛbən] | ‘webs’ |
| *lamp* | [lɑmpən] | ‘lamps’ | *krab* | [krɑbən] | ‘crabs’ |

**Practice**

1. *beer* ‘bear’
2. *jik*

**Fillers**

1. *koe* ‘cow’
2. *kikker* ‘frog’
3. *vlinder* ‘butterfly’
4. *ei* ‘egg’
5. *appel* ‘apple’
6. *varken* ‘pig’
7. *krokodil* ‘crocodile’
8. *dribbel* ‘Spot’

**Perception**

**Non-words**

| | |
|---|---|
| jit | /jɪdən/ - /jɪpən/ |
| klaat | /klapən/ - /kladən/ |
| dint | /dɪndən/ - /dimpən/ |
| tep | /tɛtən/ - /tɛbən/ |
| deep | /detən/ - /debən/ |
| bemp | /bɛmbən/ - /bɛntən/ |

**Words (Non-alternating)**

| | |
|---|---|
| pet 'cap' | /pɛdən/ - /pɛtən/ |
| boot 'boat' | /botən/ - /bodən/ |
| olifant 'elephant' | /olifɑntən/ - /olifɑndən/ |
| kip 'chicken' | /kɪpən/ - /kɪbən/ |
| schaap 'sheep' | /sxabən/ - /sxapən/ |
| lamp 'lamp' | /lɑmbən/ - /lɑmpən/ |

**Words (Non-alternating)**

| | |
|---|---|
| bed 'bed' | /bɛdən/ - /bɛtən/ |
| hoed 'hat' | /hutən/ - /hudən/ |
| potlood 'pencil' | /potlotən/ - /potlodən/ |
| hand 'hand' | /hɑntən/ - /hɑndən/ |
| hond 'dog' | /hɔndən/ - /hɔntən/ |
| web 'web' | /wɛbən/ - /wɛpən/ |

**Practice**

| | |
|---|---|
| boom 'tree' | /botən/ - /bomən/ |
| jik | /jɪkən/ - /jɪtən/ |

**Fillers**

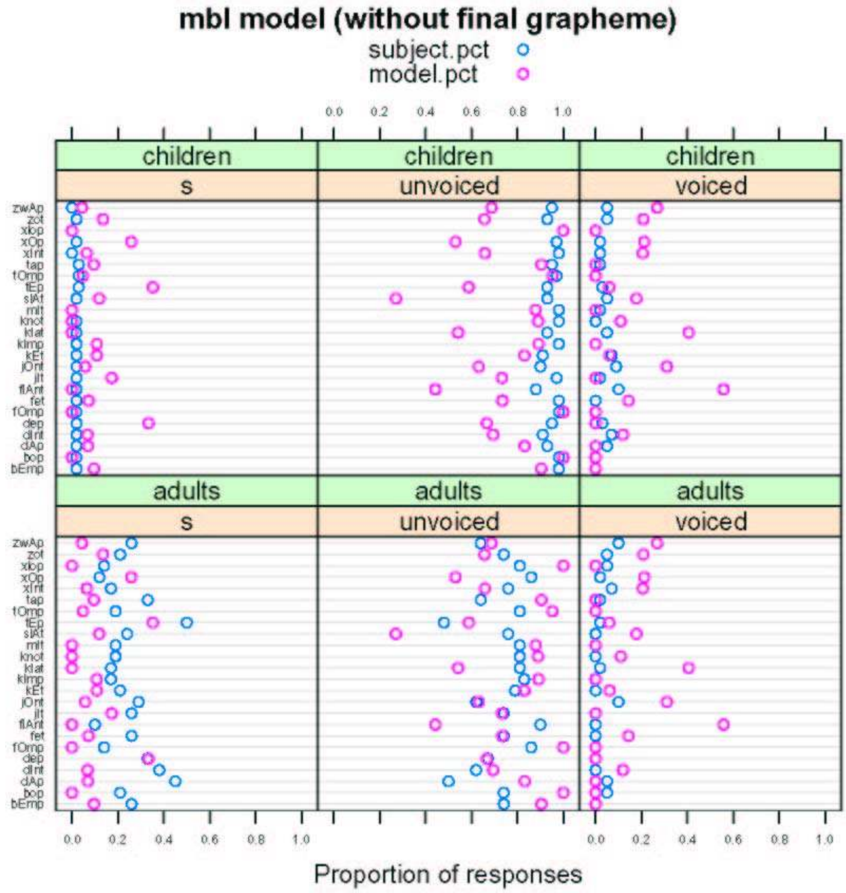
| | |
|---|---|
| krokodil 'crocodile' | /krokodɪmən/ - /krokodɪkən/ |
| beer 'bear' | /betən/ - /benən/ |
| | |
| fluk | /flʌmən/ - /flʌzən/ |
| baaf | /balən/ - /banən/ |
| kla | /klatən/ - /kladən/ |

**Appendix A. Corpus data selection**

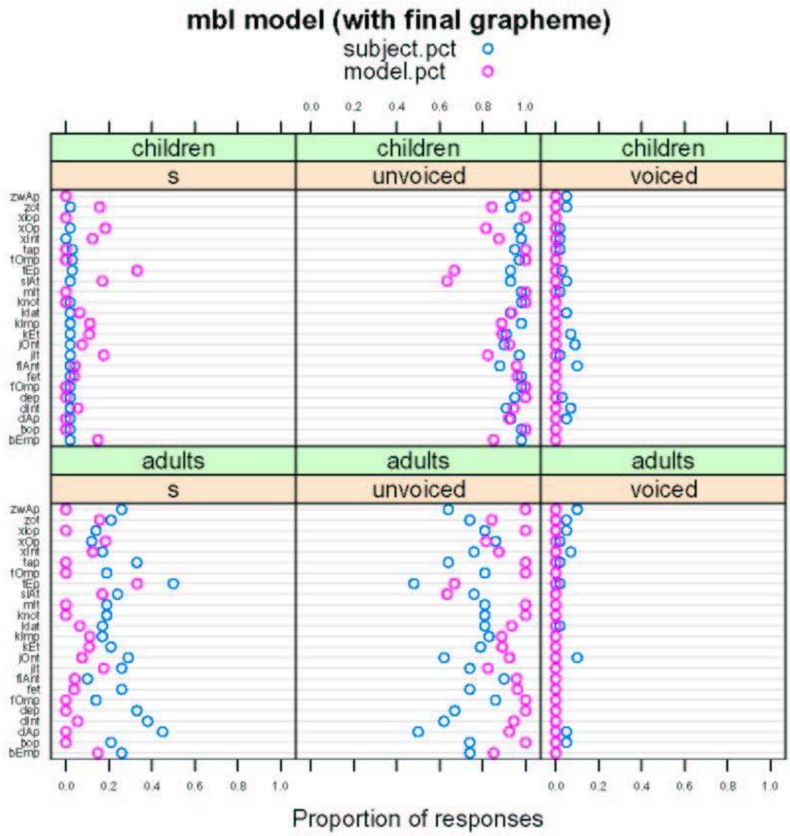
The CELEX frequency count chosen for this analysis represents the frequency over a total of one million words. The CELEX database contains 42,380,000 words, but for this measure (INLMln) the frequency is scaled down to a range of 1 to 1,000,000 by dividing the normal 42 million word frequency (INL) for each word by the number of words in the whole corpus, and then multiplying by 1,000,000. This manipulation makes the frequency count more transparent and comparable to other corpora based on one million words, such as the Brown corpus, see Burnage (1990).

The count included all word forms except personal names (e.g. *David*), place names (e.g. *Engeland* 'England') and numbers (e.g. *dertienduizend* 'thirteen thousand'). Stems of more than three syllables were excluded. All word forms were counted as separate types, as it is often an arbitrary decision whether compounds or particle verbs have a compositional meaning or not. Thus, there are many nouns with *bed* 'bed', including types of beds (e.g. *bloembed* 'flower bed') and nouns with non-compositional meanings (e.g. *pissebed* 'wood louse', *hunebed* 'megalithic grave'). This is especially common for verbs, e.g. *houden* 'hold' occurs in around 40 different complex verbs, such as *achterhouden* 'hold back' and *onderhouden* 'maintain'. For verbs, only 1st person present and past singulars (e.g. *vind* 'find', *vond* 'found') were included, as well as the infinitive and plural past tense. The comparative suffix was included (*groot ~ groter* 'bigger'), but the feminine suffix *-e* (e.g. *acrobaat ~ acrobate* 'acrobat-fem') and agentive suffix *-er* (e.g. *werker* 'worker') were not, as it is unclear to what extent speakers relate nouns like *kapper* 'hairdresser', *loodgieter* 'plumber' or *wethouder* 'city councillor' to a corresponding verb. Plurals that take the *-s* suffix (including diminutives) were not considered. These are almost exclusively English loanwords, e.g. *clubs* 'clubs', *trends* 'trends' (note that *-s* is especially common for /b/ stems). Voicing alternations in the suffix were not considered (e.g. 3rd person present /kɛnt/ vs. past tense [kɛndə] 'knew'), because here /t/ is not part of the stem /kɛn/ (compare vɪnt] ~ [vɪndən] 'find').

**Appendix B. Simulation memory-based learner TiMBL [1]**




[1] The model's memory contains all noun lemmas from CELEX with a stem and plural form above zero (i.e. attested in the INL corpus), representing ~19.000 forms. Each exemplar is represented by the last two syllables of the stem (coded as onset-nucleus-coda), and the stress pattern of the final two syllables (e.g. *portret* as (p,O,r,tr,E,t,False,True,t), *hand* as (-,-,-,h,A,nt,-,True,d). The distance between exemplars was calculated with the MVDM metrics (and overlap metrics for the stress patterns). The model creates fully specified new plural forms, instead of adding -s or -en to the stem. For the analogical set, all exemplars on the 11 nearest distances were taken, weighted with an inverse distance decay (i.e. candidates at a greater distance have less influence on production). See Daelemans et al. 2004 (TiMBL reference guide) and Keuleers & Daelemans (in progress) for more information.


mbl model (with final grapheme)
subject.pct
model.pct
children
s
unvoiced
voiced
adults
zwAp
zot
xlop
xOp
xlnt
tap
fOmp
fEp
slAt
mlt
knot
klat
klmp
kEt
jOnt
jlt
flAnt
fet
dep
dlnt
dAp
bop
bEmp
0.0
0.2
0.4
0.6
0.8
1.0
Proportion of responses

**Feature weights (information gain)**

| Feature | without final grapheme | with final grapheme |
|---|---|---|
| onset syllable 1 | 0.04 | 0.04 |
| nucleus syllable 1 | 0.05 | 0.05 |
| coda syllable 1 | 0.06 | 0.06 |
| onset syllable 2 | 0.13 | 0.13 |
| nucleus syllable 2 | 0.26 | 0.26 |
| coda syllable 2 | 0.32 | 0.32 |
| stress syllable 1 | 0.05 | 0.05 |
| stress syllable 2 | 0.12 | 0.12 |
| final grapheme | - | 0.37 |

**Neighbours for non-word stimuli**

| | | | |
|---|---|---|---|
| **kEt** (VT) | | **flAnt** (NT) | |
| 0.019405 | nEt-nEt@ | 0.012865 | krAnt-krAnt@ |
| 0.019748 | vEt-vEt@ | 0.015106 | plAnt-plAnt@ |
| 0.021927 | pEt-pEt@ | 0.016886 | klAnt-klAnt@ |
| 0.026979 | sEt-sEts | 0.021291 | strAnt-strAnd@ |
| 0.029718 | kut-kut@ | 0.026640 | lAnt-lAnd@ |
| 0.047017 | kIt-kIt@ | 0.032451 | brAnt-brAnd@ |
| 0.047856 | bEt-bEd@ | 0.032540 | stAnt-stAnd@ |
| 0.048969 | zMt-zMt@ | 0.032756 | bAnt-bAnd@ |
| 0.049396 | wEt-wEt@ | 0.034538 | rAnt-rAnd@ |
| 0.049467 | vut-vut@ | 0.038940 | wAnt-wAnd@ |
| 0.049524 | fMt-fMt@ | 0.040237 | flArt-flArd@ |
| **dInt** (NT) | | **dep** (V:P) | |
| 0.016758 | tInt-tInt@ | 0.016758 | tep-teps |
| 0.029831 | fInt-fInt@ | 0.019425 | zep-zep@ |
| 0.029899 | kInt-kInd@r@ | 0.029899 | kep-keps |
| 0.034088 | pInt-pInt@ | 0.029899 | kep-kep@ |
| 0.044584 | lInt-lInt | 0.036562 | dOp-dOp@ |
| 0.052274 | sMnt-sMnts | 0.047995 | tek-teks |
| 0.055939 | kwInt-kwInt@ | 0.049132 | den-den |
| 0.058501 | vrInt-vrInd@ | 0.049308 | rep-rep@ |
| 0.058852 | bInt-bInt@ | 0.050185 | zew-zew@ |
| 0.060479 | wInt-wInd@ | 0.053320 | tOp-tOp@ |
| 0.063776 | tEnt-tEnt@ | 0.053804 | dop-dop@ |
| **xInt** (NT) | | **kImp** (NP) | |
| 0.025473 | bInt-bInt@ | 0.005028 | kIst-kIst@ |
| 0.026250 | wInt-wInd@ | 0.015867 | kIk-kIks |
| 0.030312 | kwInt-kwInt@ | 0.025419 | mIst-mIst@ |
| 0.031904 | lInt-lInt@ | 0.030011 | nImf-nImf@ |
| 0.032447 | fInt-fInt@ | 0.034450 | nIxt-nIxt@ |
| 0.033364 | pInt-pInt@ | 0.035716 | dIN-dIN@ |
| 0.038629 | vrInt-vrInd@ | 0.035811 | kMw-kMw@ |
| 0.044533 | kInt-kInd@r@ | 0.036258 | mIk-mIk@ |
| 0.047017 | xEnt-xEnt@ | 0.037026 | dIrk-dIrk@ |
| 0.049863 | sprInt-sprInts | 0.037178 | lIst-lIst@ |
| 0.052240 | plInt-plInt@ | 0.037794 | pIk-pIk@ |

**List of Tables**

**List of Figures**

**Samenvatting in het Nederlands (Summary in Dutch)**

**1. Inleiding**

In deze dissertatie is onderzocht wanneer en hoe de Nederlandse stemalternantie wordt verworven. The term alternantie wordt gebruikt wanneer een enkel morfeem twee of meer vormen aanneemt, afhankelijk van de fonologische of morphologische context waarin het zich bevindt. In paren zoals *bed* ~ *bedden* wordt de laatste klank van het enkelvoud stemloos uitgesproken (als bɛ[t]), terwijl het meervoud een stemhebbende klank bevat (bɛ[d]ən). Het proces van finale neutralisatie of *final devoicing* leidt hier dus tot alternantie van de stam *bed* (in dit geval tussen /t/ en /d/). Ook labialen en fricatieven ondergaan dit proces, en naast meervouden zijn er ook alternerende paren voor werkwoorden (zoals *heb* ~ *hebben*) en adjectieven (zoals *lief* ~ *lieve*). Er bestaan ook niet-alternaterende paren zoals *pet* ~ *petten*, waaruit blijkt dat het Nederlands geen fonologisch proces kent dat een obstruent tussen vocalen stemhebbend maakt. Kennis over stemalternantie kan alleen verworven worden als het kind de meervoudsvorm of andere complexe vorm relateert aan de stam. Er is niet veel bekend over de verwerving van dit soort morfo-fonologische processen, hoewel het proces van *final devoicing* centraal staat in de fonologische theorievorming.

In deze studie is gebruik gemaakt van zowel corpus data als experimentele data om de verwerving van stemalternanties te onderzoeken. In experimenten werd kinderen gevraagd om meervouden te vormen van bestaande woorden (zoals *bed* en *pet*) en niet-bestaande woorden of non-woorden (zoals *ket* [kɛt]). In het geval van non-woorden (de zogenaamde 'Wug-test') kunnen kinderen in principe twee mogelijke meervouden vormen ([kɛtən] of [kɛdən]). Aan de hand van twee typen theoretische modellen (fonologische modellen en analogiemodellen) zijn hypotheses gevormd over (1) het soort fouten of 'overgeneralisaties' dat kinderen maken in woorden en (2) het voorkomen van stemalternantie in non-woorden.

In hoofdstuk 2 wordt de achtergrondliteratuur over morfo-fonologische alternanties besproken en worden de twee typen modellen nader uitgewerkt. Er worden ruwweg drie soorten alternantie onderscheiden: fonologisch, morfologisch of lexicaal geconditioneerd. Dit onderscheid heeft te maken met het al dan niet regelmatig, productief en transparent zijn van de alternantie, factoren die ook invloed hebben op het moment van verwerving. In traditionele generatieve modellen van fonologie worden grammaticale regels of ‘constraints’ gescheiden van het lexicon. In deze modellen wordt de stemalternantie als regelmatig beschouwd; de fonologische regel van *final devoicing* wordt toegepast op een *abstracte* representatie: /bɛd/ verandert in [bɛt]. Kinderen zouden de enkelvoudsvorm die ze horen (eindigend op [t]) dan ook moeten herstructureren tot de ‘correcte’ lexicale representatie (van /bɛt/ naar /bɛd/). Op deze manier is het lexicon zo economisch mogelijk georganiseerd, terwijl de grammatica de benodigde regel bevat om in alle gevallen de juiste oppervlaktevorm te realiseren. Binnen de generatieve fonologie wordt ook aangenomen dat het kind tijdens de taalverwerving beïnvloed wordt door ‘natuurlijke’ processen. Zo heeft Stampe (1973) al voorgesteld dat zowel *final devoicing* als *intervocalic voicing* (het stemhebbend worden van obstruenten tussen vocalen) natuurlijke processen zijn, die zowel in kindertaal als in talen van de wereld voorkomen. Meer recente modellen zoals de Optimaliteitstheorie (OT) gaan uit van soortgelijke fonologische beperkingen of ‘constraints’. Een belangrijke voorspelling van dit soort modellen is dat kinderen *fonologische* generalisaties maken zoals *intervocalic voicing*. In het geval van het Nederlandse kind zou dit proces leiden tot fouten van het type **pedden* *[pɛdən] voor *petten*, vooral bij de jongste kinderen. Er zijn aanwijzingen dat dit soort fouten ook daadwerkelijk door kinderen gemaakt worden, al is een systematisch onderzoek nooit eerder uitgevoerd. Een tweede voorspelling die voortkomt uit constraint-gebaseerde modellen is dat kinderen gevoelig zijn voor uniformiteit binnen een paradigma, ofwel *Paradigm Uniformity constraints* (Hayes 2004). Hierdoor zouden jonge kinderen juist niet-alternerende of ‘stamgetrouwe’ vormen produceren, waardoor fouten van het omgekeerde type ontstaan zoals **betten* *[bɛtən] voor *bedden*.

In het tweede type modellen, ook wel analogie- of ‘usage-based’ theorieën genoemd (Bybee 2001, Tomasello 2003), worden het lexicon en de grammatica niet van elkaar gescheiden. Noties als regelmatig of transparent worden beschouwd als gradueel, terwijl frequentie van gebruik bepaalt hoe productief een bepaald proces is. Volgens deze modellen kunnen complexe of gelede woorden zoals meervouden opgeslagen worden in het lexicon (bijvoorbeeld zowel [bɛt] als [bɛdən]). Generalisaties worden gevormd op basis van analogie met bestaande woorden in het lexicon, in plaats van door abstracte regels. Deze (inherent probabilistische) generalisaties kunnen leiden tot een bepaalde mate van productiviteit, maar alleen als een patroon voldoende (type-)frequent is. Fouten zoals *[bɛtən] en *[pɛdən] kunnen dus optreden, maar dit hangt af van de frequentie en transparantie van de morfo-fonologische alternantie. De hypotheses die voortkomen uit beide typen modellen worden in deze dissertatie getoetst.

In hoofdstuk 3 wordt een overzicht gegeven van eerdere studies op het gebied van het Nederlandse stemcontrast en de representatie van morfologisch complexe woorden. Het stemcontrast wordt door kinderen al rond driejarige leeftijd verworven, nadat ze in eerste instantie vooral stemloze obstruenten produceren (in initiële positie). De (fonotactische) kennis over het voorkomen van *final devoicing* is belangrijk voor de verwerving van de stemalternantie. Echter, samenstellingen zoals *handdoek* [hɑnduk] of constructies als *heb ik* [hɛbɪk]) laten zien dat final devoicing niet altijd optreedt. Ook zijn er aanwijzigingen dat neutralisatie van stem niet compleet is in het Nederlands, waardoor bijvoorbeeld de [t] van *pet* en de [t] van *bed* een verschillende duratie hebben (op sub-fonemisch niveau). Recente studies laten zien dat sprekers gebruik kunnen maken van deze incomplete neutralisatie. Ook wijzen eerdere studies uit dat volwassenen zich laten leiden door lexicale analogie bij het interpreteren van geneutraliseerde segmenten. Hoewel deze resultaten analogiemodellen ondersteunen, is het niet duidelijk of deze bevindingen ook voor taalverwerving gelden.

**2. Corpusdata en experimenten**

In hoofdstuk 4 worden nieuwe corpusdata gepresenteerd, verzameld op basis van CELEX (geschreven taal uit diverse bronnen) en CHILDES (interactie tussen moeder en kind). Allereerst blijkt uit corpusdata dat een stemalternantie vaak gepaard gaat met onregelmatigheden als klinkeralternantie (*stad ~ steden*) of het *-eren* suffix (*kind ~ kinderen*). Bovendien blijken meervouden met een stemalternantie een hoge tokenfrequentie te hebben, iets dat vaak optreedt bij onregelmatige vormen. De resultaten laten verder zien dat er ongeveer twee keer zoveel niet-alternerende vormen als alternerende vormen (naamwoorden, werkwoorden en bijvoegelijk naamwoorden) voorkomen in taal gericht tot een kind (de ‘input’). Voor alternerende meervouden van naamwoorden (zoals *bedden*) werd slechts een percentage van 25-30% gevonden. Het is ook belangrijk om te kijken hoeveel *paren* er in de input voorkomen. Zo bevatte de input van Sarah (1;6-5;2) in het ‘van Kampen’ corpus 16 niet-alternerende paren (zowel *olifant* als *olifanten*) en 12 alternerende paren (zowel *hond* als *honden*). Ook domineert een klein aantal hoogfrequente woorden (met name *handen* en *kinderen*) het vroege lexicon. Productiedata laten zien dat de meest frequente paren (zoals *hand ~ handen*) al correct geproduceerd worden voordat kinderen drie jaar oud zijn. Resultaten uit het CLPF corpus waarin producties van 12 kinderen van 1;0 tot 2;11 zijn opgenomen tonen aan dat zeer jonge kinderen 'onderliggende' /d/ and /t/ al verschillend behandelen. Een veel voorkomende fout is bijvoorbeeld deletie van /d/ (maar niet /t/) na een nasaal (zoals *[hɔnə] voor /hɔndən/). Dit kan echter ook komen doordat /d/ in deze context moeilijk te onderscheiden is. Ongeveer een derde van de meervoudsvormen die kinderen produceerden waren van het type *[bɛtən], waarin de mediale obstruent als stemloos gerealiseerd wordt. Aangezien kinderen dit soort fouten ook maken in mono-morfemische woorden (zoals *baby* of *ander*), spelen fonologische factoren blijkbaar een rol bij de verwerving van stemalternantie. Hoewel er veel anekdotische aanwijzingingen zijn voor fouten van het type *[pɛdən], wordt dit type fout slechts zeer sporadisch in spontane taaldata aangetroffen (zie ook van der Feest 2007). Op basis van corpusdata vinden we dus geen ondersteuning voor een vroeg proces van *intervocalic voicing* in kindertaal.

Uit corpusdata blijkt ook dat de distributie van stemalternanties beïnvloed wordt door de fonologische context waarin de finale obstruent voorkomt. Zo is de kans op een t~d alternantie het grootst na nasalen (zoals in *hand ~ handen*), terwijl p~b alternanties alleen na korte vocalen voorkomen (zoals in het hoogfrequente *heb ~ hebben*). Er is blijkbaar een 'toevallig' *lexical gap* voor p~b alternanties na lange vocalen en nasalen (met andere woorden, alternanties als *aap ~* aben* komen niet voor). Dit gegeven is belangrijk om voorspellingen van eerdergenoemde modellen te kunnen toetsen. Als kinderen immers gevoelig zijn voor de lexicale distributie van het alternantiepatroon, zouden overgeneralisaties in *lexical gap* contexten niet voor mogen komen. De non-woorden die gebruikt zijn in de elicitatie experimenten beschreven in hoofdstuk 5 en 6 bevatten dan ook eindklanken in verschillende fonologische contexten (bijv. *ket*, *flant* en *taap*), om te onderzoeken of kinderen zich bij meervoudsvorming laten leiden door lexicale patronen of juist door fonologisch gestuurde voorkeuren (zoals natuurlijke klassen).

In hoofdstuk 5 worden experimenten beschreven waarin kinderen zowel bestaande woorden als non-woorden in het meervoud moesten omzetten. De resultaten voor bestaande woorden (zoals *bedden*) laten zien dat 7-jarigen nog verassend veel fouten maken: in slecht 58% van de gevallen werden correcte alternerende meervouden gevormd. Kinderen blijven dus lang fouten maken als *[bɛtən], ook als ze al wel in staat zijn de mediale /d/

en /b/ uit te spreken (zie ook Zamuner, Kerkhoff & Fikkert 2006a). Kinderen presteerden beter naarmate de frequentie van de meervoudsvorm hoger was. Naast woorden werden ook non-woorden aangeboden, die gekoppeld werden aan een fantasie-figuurtje (bijv: "Dit is een *ket*", "Nu zijn er twee ..."?). In het geval van een non-woord moet de lexicale representatie worden afgeleid uit een enkelvoudsvorm die altijd op /t/ eindigt. Er zijn vier algemene hypothesen wat betreft de mogelijke meervoudsvorm (*ketten* [kɛtən] of *kedden* [kɛdən]): '*Paradigm Uniformity*', *willekeurige selectie*, *fonologische generalisatie* of *analogische generalisatie*.

Als kinderen geleid worden door *Paradigm Uniformity* zouden ze alleen of vooral niet-alternerende vormen (zoals [kɛt] ~ [kɛtən]) en fouten als *[bɛtən]) moeten produceren. De tweede hypothese betreft de logische mogelijkheid dat kinderen wel degelijk alternanties produceren, maar dat deze *willekeurig* optreden. Als echter aangetoond kan worden dat dit niet het geval is, zijn de laatste twee hypothesen van toepassing. De overgeneralisaties en alternanties voor non-woorden kunnen dan voortkomen uit *fonologische* of *analogische* generalisaties. Om tussen deze twee hypotheses te kiezen zijn meer specifieke voorspellingen nodig, die hieronder beschreven worden.

De resultaten van Experiment I (60 kinderen in de leeftijd van 2;9 tot 7;8) laten zien dat 'stamgetrouwe' meervouden zoals *[bɛtən] door alle kinderen in gelijke mate geproduceerd worden (in ongeveer 40% van de gevallen). Het is dus niet zo dat een voorkeur voor uniforme meervouden alleen voor de jongste kinderen geldt. In Experiment II zijn twee groepen 5-jarigen getest; een controlegroep van 27 kinderen en een groep van 24 kinderen met een specifieke taalstoornis (*Specific Language Impairment* of *SLI*). Het is interessant om deze laatste groep te testen, omdat in generatieve modellen wordt verondersteld dat kinderen met SLI niet in staat zijn regels te vormen. De resultaten laten echter zien dat SLI kinderen iets vaker fouten van het type *[bɛtən] maken (in meer dan 50% van de gevallen). Dit toont aan dat deze kinderen in staat zijn productieve meervouden te vormen, en dat ze net als de controlegroep meer moeite hebben met alternerende meervouden dan met niet-alternerende meervouden. SLI kinderen waren bovendien *minder* gevoelig voor de frequentie van meervouden, een uitkomst die onverwacht is als ze alle meervouden ongeanalyseerd opslaan.

De resultaten laten ook zien dat kinderen overgeneralisaties van stem voor woorden (*[pɛdən]) en alternanties voor non-woorden (zoals [kɛt] ~ [kɛdən]) produceerden in 3-5% van de gevallen, iets vaker dan volwassenen

(1-2%). Dit wijst erop dat *Paradigm Uniformity* niet de enige factor is die bij de woordvorming een rol speelt. Alternanties werden echter niet *willekeurig* geproduceerd; er is een effect van de finale obstruent (/p/ of /t/) en de klank die aan de finale obstruent voorafging (korte vocaal, lange vocaal of nasaal).
Om te onderzoeken of kinderen door *fonologische* of *analogische* generalisaties geleid worden bij het produceren van alternanties werden meer specifieke hypotheses opgesteld. In sommige gevallen leiden de twee modellen tot dezelfde voorspellingen. Zo zouden stemalternanties vaak moeten voorkomen voor items zoals *flant*, ofwel omdat alternanties voor /t/ na nasalen (zoals *hand* ~ *handen*) het meest frequent zijn in het Nederlands, ofwel omdat een fonologisch proces van *postnasal voicing* (voor /t/ en /p/ na nasalen) een rol speelt. Echter, als kinderen beïnvloed worden door dit soort fonologische generalisaties, zouden ze ook overgeneralisaties moeten produceren in een omgeving waar dit in het Nederlands niet voorkomt, zoals voor /p/ na nasalen (bijv. *kimp* [kɪmp] ~ *kimben* [kɪmbən]) en /p/ na lange vocalen (bijv. *taap* [tap] ~ *taben* [tabən]). In een analogiemodel worden zulke *lexical gap* generalisaties juist niet verwacht, omdat kinderen zich enkel laten leiden door de meervouden in hun lexicon.

De resultaten voor Experiment I laten zien dat alternanties in de onverwachte *lexical gap* contexten voorkomen, al is het zeer sporadisch (slechts drie keer, en alleen voor /p/ na lange vocalen). Het is dus mogelijk dat jonge kinderen zich laten leiden door (aangeboren) fonologische kennis. Er werden echter geen p~b alternanties na nasalen geproduceerd, terwijl deze klanken zich fonologisch gezien als lange vocalen zouden moeten gedragen. Bovendien produceerden volwassenen ook onverwachte alternanties zoals [tap] ~ [tabən]. Wel produceerden de SLI kinderen uit Experiment II meer onverwachte alternanties dan de controlegroep, wat kan wijzen op ongestructureerde lexicale representaties (zie de Bree 2007 en Kerkhoff & de Bree 2005). In het algemeen werden alternanties niet consequent in een bepaalde fonologische context geproduceerd, en was er zelfs variatie tussen items met hetzelfde rijm (zoals *zwap* en *dap*). Dit betekent dat een kind bijvoorbeeld zowel *zwap* ~ *zwappen* als *dap* ~ *dabben* produceerde.

Een tweede voorspelling waarin fonologische modellen zich onderscheiden van analogiemodellen is dat volgens de eerste /b/ alternanties vaker zouden moeten voorkomen dan /d/ alternanties, omdat /b/ zowel fonetisch als fonologisch minder gemarkeerd is. Analogiemodellen voorspellen echter het omgekeerde, omdat /d/ veel vaker in de input voorkomt (kinderen kennen vaak slechts enkele /b/ alternanties, zoals *heb* ~ *hebben*). De puur fonologisch gestuurde voorspelling wordt niet door de data ondersteund, aangezien kinderen in alle leeftijdsgroepen meer /d/ alternanties produceerden (4.5% vs. 1.7%).

Een derde en laatste voorspelling van fonologische modellen is dat overgeneralisaties afnemen naarmate kinderen ouder worden. Dit was niet het geval; 3-jarigen en 5-jarigen produceerden evenveel alternanties voor non-woorden, en overgeneralisaties voor woorden (**pedden* *[pɛdən], **olifanden* *[olifɑndən]) kwamen zelfs vaker voor bij 5-jarigen. Bovendien hield het aantal overgeneralisaties verband met het aantal correct gerealiseerde alternerende meervouden (*bedden* [bɛdən], *handen* [hɑndən]). Dit resultaat sluit aan bij voorspellingen van analogiemodellen, waarin overgeneralisaties zijn gebaseerd op de meervouden die het kind kent. Als het lexicon nog weinig woorden bevat zullen onregelmatige vormen zoals *bedden* relatief frequent zijn, waardoor ze een basis kunnen vormen voor overgeneralisatie. De resultaten laten ook zien dat alternanties voor non-woorden (zoals *kedden* [kɛdən] of *flanden* [flɑndən]) beïnvloed worden door de kans dat bestaande woorden met hetzelfde rijm alterneren. Kinderen laten zich dus leiden door de mate waarin het non-woord op een bestaand woord lijkt en niet door fonologische generalisaties of natuurlijke klassen. In een computationele simulatiestudie werd dit resultaat ook ondersteund; het gedrag van proefpersonen kwam overeen met dat van een 'memory-based learner' (TiMBL) op basis van nomina in het CELEX corpus (zie bijv. Keuleers et al. 2007).

In hoofdstuk 6 is een aantal experimenten beschreven waarin de productieve kennis van final devoicing en alternanties is onderzocht. Voor bestaande woorden werd kinderen gevraagd eerst het meervoud en daarna het enkelvoud te produceren (*bedden* ~ *bed*). De resultaten laten zien dat kinderen geen moeite hebben met het vormen van enkelvouden van bestaande woorden, zowel voor niet-alternerende als alternerende meervouden. Kinderen waren in de meervoud-naar-enkelvoud taak wel beter in staat alternerende meervouden te produceren dan in de enkelvoud-naar-meervoud taak van hoofdstuk 5, hoewel deze meervouden (*bedden*) nog steeds moeilijker waren dan niet-alternerende meervouden (*petten*).

Om te onderzoeken of kinderen in staat zijn het enkelvoud van niet bestaande meervouden te maken werden ook vervoegde non-woorden zoals *ketten* [kɛtən] of *kedden* [kɛdən] aangeboden (zie ook Zamuner, Kerkhoff, & Fikkert 2006b). Deze taak is tegenovergesteld aan de Wug-test, omdat kinderen nu de stam van het affix moeten scheiden (bijvoorbeeld: "Dit zijn twee *flanden*, "Nu is er één ..."). Als kinderen beschikken over een fonologische regel of constraint van *final devoicing* (mogelijk afgeleid uit de eerder opgedane 'fonotactische' kennis dat woorden nooit eindigen op een stemhebbende obstruent), zouden ze geen moeite moeten hebben het enkelvoud te produceren van vormen met mediale /d/ of /b/. Met andere woorden, als vroege fonologische kennis toegepast wordt bij het leren van alternanties (zoals voorspeld door Hayes 2004 en anderen), zouden kinderen zowel niet-alternerende paren (*ketten* [kɛtən] ~ *ket* [kɛt]) als alternerende paren moeten kunnen produceren (*kedden* [kɛdən] ~ *ket* [kɛt]). Aangezien het enkelvoud altijd in een stemloze obstruent eindigt, is er maar een enkelvoudsvorm mogelijk. De resultaten laten echter zien dat kinderen (in de leeftijd van 2;3 tot 6;0) meer moeite hebben met het produceren van het enkelvoud voor vormen met een mediale /d/ ([kɛdən] ~ [kɛt]) dan mediale /t/ ([kɛtən] ~ [kɛt]). Hoewel enkelvoudsvormen als *[kɛd] niet voorkwamen, hadden kinderen de neiging het meervoud te herhalen ("Dit is één *kedden*"). De 'omgekeerde' Wug-test experimenten laten dus dezelfde voorkeur voor niet-alternerende paren zien als de productie-experimenten. Volwassenen waren overigens wel in staat om enkelvouden voor beide typen meervouden te produceren. Deze resultaten wijzen erop dat de Nederlandse stemalternantie pas laat verworven wordt, ondanks het feit dat kinderen waarschijnlijk al vroeg over (fonetisch gestuurde) kennis van *final devoicing* beschikken.

In hoofdstuk 6 worden tenslotte 'acceptatie'-experimenten beschreven, waarbij kinderen nieuwe paren kregen aangeboden die ofwel een stemalternantie ([kɛt] ~ [kɛdən]) ofwel een alternantie in plaats van articulatie ([kɛt] ~ [kɛpən]) bevatten (bijvoorbeeld: "*Dit is een ket*". "*Zijn dit twee kedden*"? en "*Zijn dit twee keppen*"?), waarna het kind 'ja' of 'nee' kon zeggen. De resultaten lieten zien dat jonge kinderen (3;0-3;7) weliswaar iets vaker het eerste type alternantie accepteerden, maar dat ze ook vaak een plaatsalternantie accepteerden of beide vormen afwezen. De meeste kinderen produceerden wel stemalternanties voor bestaande woorden, terwijl ze de alternantie toch niet accepteerden in non-woorden.

**3. Samenvatting en conclusie**

De resultaten van bovenstaande experimenten wijzen op een grote invloed van het lexicon bij het verwerven van de Nederlandse stemalternantie. Samengevat zijn de belangrijkste bevindingen:

1. De Nederlandse stemalternantie heeft een lage typefrequentie en een hoge tokenfrequentie, ook in spraak gericht tot kinderen (H4).
2. Er is weinig evidentie voor een fonologisch proces van *intervocalic voicing* in spontane taaldata van jonge kinderen (H4).
3. Kinderen hebben moeite met bestaande alternerende meervouden, aangezien zelfs 7-jarigen nog fouten maken van het type **betten* *[bɛtən] (H4 en H5).
4. Het aantal correcte alternerende meervouden in een elicitatietaak correleert met de frequentie van de meervoudsvorm (H5).
5. Effecten van stamgetrouwheid of *Paradigm Uniformity* zijn even sterk op verschillende leeftijden (H5 en H6).
6. Overgeneralisaties van stem voor woorden en alternanties voor non-woorden worden vaker door 5-jarigen dan door 3-jarigen geproduceerd (H5).
7. Alternanties voor non-woorden worden voorspeld door de kans op stemalternantie voor Nederlandse naamwoorden en niet door natuurlijke klassen (H5).
8. Onverwachte *lexical gap* alternanties zijn schaars, en worden ook door kinderen met SLI en volwassenen geproduceerd (H5).
9. Alternanties voor non-woorden zijn variabel in een bepaalde fonologische context, zelfs voor items met hetzelfde rijm (H5).
10. Kinderen hebben moeite het patroon van stemalternantie toe te passen op non-woorden, zowel in een klassieke Wug-test als in een omgekeerde Wug-test, ook als ze wel bestaande alternerende paren produceren (H5 en H6).
11. Er is geen directe evidentie dat kinderen de Nederlandse stemalternantie verwerven op basis van eerder verworven kennis van *final devoicing* (H6).

Deze bevindingen wijzen op een model waarin probabilistische generalisaties worden gevormd op basis van woorden in het lexicon. Zowel analogiemodellen als de meest recente regel- of constraint-gebaseerde modellen bevatten de notie van variabele generalisaties. Het onderscheid tussen bestaande woorden en non-woorden en de rol van lexicale frequentie worden echter beter verklaard door analogiemodellen, zoals voorgesteld door Bybee (2001). In Bybee's network model komt morfologische structuur voort uit de lexicale verbanden tussen (complexe) woorden zoals enkelvouden en meervouden (zie Figuur 15 op pagina 260). Er is in dit model sprake van interactie tussen fonologie (fonetisch gefundeerde fonologische processen) en morfologie (de paradigmatische organisatie van het lexicon), maar niet in de vorm van regels die los staan van de vormen die ze beschrijven. Productieve schema's kunnen gevormd worden op basis van patronen in het lexicon, maar hierbij is type frequentie de belangrijkste voorspeller (woorden met een hoge tokenfrequentie zullen juist meer autonoom zijn). Ook zijn er in dit model geen abstracte 'onderliggende' representaties (zoals /bɛt/), omdat zowel [bɛt] als [bɛdən] in het lexicon worden opgenomen.

De stemalternantie die optreedt in paren als *bed ~ bedden* wordt door velen beschouwd als een van de belangrijkste argumenten voor het bestaan van abstracte representaties. Toch passen de huidige bevindingen met betrekking tot de verwerving van stemalternanties beter in een lexicale analyse van dit fenomeen. Het gegeven dat de Nederlandse stemalternantie relatief laat verworven wordt suggereert bovendien dat deze niet volledig regelmatig is. Binnen de voorgestelde analyse kunnen regelmatigheid, productiviteit en transparantie beter beschouwd worden als graduele noties, en is er geen strikte indeling mogelijk tussen fonologisch, morfologisch of lexicaal geconditioneerde alternanties.

**Curriculum Vitae**

Annemarie Kerkhoff was born on 4 November 1974 in Nijmegen (The Netherlands). After obtaining a VWO diploma from the 'Rijnlands Lyceum' in Oegstgeest, she completed a Propaedeutic degree in Dutch Language and Literature at the University of Groningen. She studied General Linguistics (Neurolinguistics) from 1996 to 2000 at the same university, and spent six months at the University of Newcastle upon Tyne (UK) as an Erasmus student. After completing her MA degree, she received a VSB grant and spent ten months at the University of Essex (UK), where she was enrolled in an MPhil programme in Psycholinguistics. From 2001 to 2006 she was employed as a PhD student at the Utrecht Institute of Linguistics OTS (funded by NWO). This thesis is the result of the research carried out during that period. Currently, she is employed as a postdoctoral researcher in the project 'Category formation in phonology and grammar: Distributional learning in children with and without a developmental language delay' at Utrecht University.