

## Representations of [Voice]: Evidence from Acquisition

René Kager<sup>a</sup>, Suzanne van der Feest<sup>b</sup>, Paula Fikkert<sup>b</sup>,  
Annemarie Kerkhoff<sup>a</sup> & Tania S. Zamuner<sup>c</sup>  
<sup>a</sup> *Utrecht University*, <sup>b</sup> *Radboud University Nijmegen*  
<sup>c</sup> *University of British Columbia*

We consider two theories of laryngeal representation, one using a single feature [voice] generalizing across prevoicing languages and aspiration languages, and the other using multiple features: [voice] for pre-voicing languages and [spread glottis] for aspiration languages. We derive predictions for children's early productions, and test these for three Germanic languages. Children acquiring Dutch, a prevoicing language, show de-voicing of stops, while available data from German, an aspiration language, show de-aspiration. Although the difference might simply reflect intrinsic properties of children's early production and perception systems, we argue that a representational account is in order, based on multiples features. The case is made for English, an aspiration language, based on the early productions of a single child. A laryngeal harmony pattern is found which spreads voicelessness from coda to onset, which is argued to involve activity of [spread glottis]. This is interpreted as evidence for a laryngeal representation involving multiple features.

### 1. Introduction

A great deal of recent research has addressed the representation of laryngeal features (Avery 1996, Avery & Idsardi 2001, Iverson & Salmons 1995, 2003, Jessen 1989, 1996, Jessen & Ringen 2002, Lombardi 1991, 1995, 1999, 2001, Salmons & Iverson, to appear, Steriade 1995, 1997, Van Rooy & Wissing 2001, Vaux 1998, Wetzels & Mascaró 2001). Debates in the literature focus on several major aspects of phonological theory, including monovalent versus bivalent features, gradient versus categorical processes, and phonetic detail in phonological representation. In discussions of these issues, evidence from a range of sources has been used, including phonetics, phonology, typology, and diachrony. Surprisingly, very little evidence from acquisition has been brought to bear on the issue of laryngeal representation.

Acquiring the laryngeal phonology of a language amounts to identifying the relevant contrasts, building up a representation of laryngeal features, and learning to produce these contrasts in an adult-like fashion. By studying children's developing language systems, we can gain insight into how laryngeal features are represented. Acquisition patterns provide a way to test claims about the representations of laryngeal features. This paper presents corpus analyses of the acquisition

of initial obstruents in Dutch and German, and the acquisition of initial and final obstruents in English. Children's productions of voiced and voiceless obstruents were analysed for realizations of laryngeal features. These analyses revealed a number of interesting error patterns.

First, productions of children acquiring Dutch, which is a so-called prevoicing language, differ from productions of children acquiring German and English, which are aspiration languages. Production errors of children learning Dutch tend to favour voiceless stops in initial position, whereas children learning German (Grijzenhout & Joppen-Hellwig 2002) and English (Menn 1971, Smith 1973) exhibit the opposite error pattern, producing more voiced stops. This confirms earlier results from acquisition studies of prevoicing languages such as Spanish and Hindi (Macken & Barton 1980b, Davis 1995) and of aspiration languages such as English and German (Macken & Barton 1980a).

Second, the frequency of voicing in children's targets that children attempt to produce does not reflect the error patterns related to voicing. For example, while production errors of young children acquiring Dutch show a trend toward initial voiceless stops (/b/ → [p] and /d/ → [t]), a statistical analysis of children's targets reveals rather the reverse trend: children attempt more voiced than voiceless word-initial stops. This finding implies that children's error patterns reflect factors other than frequency of targets in children's productions, such as articulatory factors or featural representation.

Lastly, children's voicing errors in English turn out to be conditioned by the laryngeal specification of segments occurring later in the word; more specifically, the devoicing of initial stops is triggered by a following voiceless obstruent. However, no 'harmonic' effect is found for the voicing of initial obstruents when followed by voiced obstruents.

These acquisition data were used to test different theories' claims of how laryngeal features should be represented in languages that display a two-way laryngeal contrast: either with a single binary feature [ $\pm$ voice] (Wetzels & Mascaró 2001), or multiple language-dependent features, specifically monovalent [voice] and [spread glottis] (Iverson & Salmons 1995).<sup>1</sup> Results from this study will be argued to support Iverson & Salmons' theory, in which aspiration languages (including English and German) use the feature [spread glottis], while prevoicing languages (including Dutch) use [voice].

Importantly, our study supports a multiple feature view, under which languages use one constant active feature to represent their laryngeal contrasts in all positions, initial and final. The harmony pattern observed in English acquisition data supports this view: interaction between initial and final obstruents implies a shared monovalent featural representation for voicing in these positions, which abstracts from specific phonetic realization. A purely phonetic account cannot readily account for this pattern.

This paper is organized as follows. Section 2 will discuss the major theories of laryngeal representation to be considered, stating predictions they make for error patterns in children's productions in English, Dutch and German. In section 3, we will discuss Dutch acquisition data, which we will compare with German data in

section 4. Then we will find that children acquiring Dutch, a prevoicing language, produce errors involving devoicing of stops that are voiced in the target word, whereas the available German data show rather the reverse pattern, in which consonants that are voiceless in target words are produced as ‘voiced’ or more accurately, as plain unaspirated stops. In comparing the acquisition data from Dutch, a prevoicing language, and German, an aspiration language, two major interpretations will be considered: a phonetic one, based on intrinsic properties of children’s early production and perception systems, and a phonological one, based on the Multiple Feature Hypothesis. In section 5, these hypotheses will be tested on English, using a corpus-based study of the early productions of a single child, whose laryngeal error patterns will be discussed in detail. We will argue that error patterns involve the activity of [spread glottis] in a laryngeal harmony pattern affecting only voiceless consonants in coda and onset of a word. This will be interpreted as evidence for a representation of laryngeal contrasts involving multiple features, [voice] (for prevoicing languages) and [spread glottis] (for aspiration languages). Finally, we will discuss consequences of our findings in section 6.

## 2. *Theories of laryngeal representation*

### 2.1 *One versus multiple features*

The phonetic realization of laryngeal contrasts<sup>2</sup> varies across languages. The main acoustic cue associated with voicing is voice onset time or VOT, which refers to the time between a segment’s release and the beginning of vocal cord vibration. There are a number of ways in which laryngeal contrasts are realized in languages (Cho & Ladefoged 1999). In the languages studied in this paper, a two-way contrast is employed. However, there are also languages that employ a six-way contrast, as for example Beja (Cushitic) and Igbo (Kwa) (Ladefoged 1973, cited in Iverson & Salmons 1995:382). For the purpose of this paper, however, it is important to note the VOT differences in stops<sup>3</sup> across Dutch, German and English. These values (based on Lisker & Abramson 1964, Braunschweiler 1997) are given in Table 1.

	Voicing Lead	Short Lag VOT	Long Lag VOT
Dutch	-4 ms: b, d	0-25 ms: p, t	
German		16 ms: b, d	51 ms: p, t
English		32 ms: b, d	59 ms: p, t

Table 1: *VOT in Dutch, German and English*

The laryngeal contrasts in these three languages can be divided into those that exhibit voicing lead (where voicing begins before the release), short lag VOT (where voicing begins at the time of the release or shortly afterwards), and long lag VOT (where there is a delay between the release and the beginning of voicing). In languages such as Dutch, the contrast in initial position is one between voicing lead and short lag VOT, while in aspiration languages such as German,

the initial contrast is one between short lag and long lag VOT (where long lag VOT results in voiceless aspirated stops). Similar to Dutch are languages such as French and Spanish. Similar to German are English and most other Germanic languages (except Dutch and Germanic languages such as Afrikaans, Frisian, and Yiddish).

The question then naturally arises as to whether the featural representations of prevoicing and aspiration languages are different. There are two primary views in the literature. The standard approach within current phonological theories assumes that a single feature captures the laryngeal contrasts of all languages with a binary contrast, generalizing across prevoicing languages and aspiration languages. This single feature is either a binary feature [ $\pm$ voice] (Steriade 1995, Wetzels & Mascarcó 2001), or monovalent [voice] (Mester & Ito 1989, Cho 1990, Lombardi 1995, 1996). For the purposes of this paper, we refer to these theoretical variants as the Single Feature Hypothesis. Laryngeal specifications for Dutch, German and English for both variants are given in Tables 2 and 3.

	Voicing Lead	Short Lag VOT	Long Lag VOT
Dutch	[+voice]	[-voice]	
German		[+voice]	[-voice]
English		[+voice]	[-voice]

Table 2: *Laryngeal feature representation for Dutch, German and English under the Single Feature Hypothesis, using a binary feature [ $\pm$ voice]*

	Voicing Lead	Short Lag VOT	Long Lag VOT
Dutch	[voice]	[ ]	
German		[voice]	[ ]
English		[voice]	[ ]

Table 3: *Laryngeal feature representation for Dutch, German and English under the Single Feature Hypothesis, using a monovalent feature [voice]*

Note that the laryngeal contrasts of Dutch, German and English are captured with the same distinction between [+voice] and [-voice] (or [+voice] and [ ]), but the acoustic correlates for these features differ for Dutch versus German and English.

A second approach to laryngeal features was advanced by Jessen (1989, 1996) and Iverson & Salmons (1995, 2003), who argue that laryngeal features are best represented with multiple monovalent features such as [voice] and [spread glottis]. In languages with a binary laryngeal contrast, only one of these (the active feature) is underlyingly specified. A language's selection of laryngeal feature can be diagnosed by its active phonological processes, and it tends to correlate with VOT properties of stops. This approach will be referred to as the Multiple Feature Hypothesis since it assumes two monovalent features, [voice] and [spread glottis].<sup>4</sup> According to Iverson & Salmons, prevoicing languages, such as Dutch, represent the laryngeal contrast by a monovalent feature [voice], such that voiced

stops are specified and voiceless stops are unspecified. Aspiration languages, such as German and English, select the active feature [spread glottis], such that aspirated stops (voiceless) are specified, and unaspirated stops (voiced or voiceless) lack specification, indicated by [ ]. The laryngeal specifications under this approach for Dutch, German and English are given in Table 4.

	Voicing Lead	Short Lag VOT	Long Lag VOT
Dutch	[voice]	[ ]	
German		[ ]	[spread glottis]
English		[ ]	[spread glottis]

Table 4: *Laryngeal feature representation for Dutch, German and English under the Multiple Feature Hypothesis, [voice] and [spread glottis]*

Under the Multiple Feature Hypothesis, the Dutch voicing contrast is expressed by representing pre-voiced stops with the feature [voice], while voiceless segments lack specification in their phonological representation. Aspiration languages such as German and English represent their laryngeal contrast with [spread glottis] on aspirated stops, and lack of specification on plain (unaspirated voiceless) stops. Note that both [voice] and [spread glottis] are abstract phonological features in the sense that their phonetic realizations vary and depend on the position in the word. For example, [spread glottis] is realized with maximal aspiration (i.e. fully abducted vocal folds) only in the onset of foot-initial syllables, while other positions have weaker implementations (Iverson & Salmons 1995:377).

Having sketched the Single Feature Hypothesis and the Multiple Feature Hypothesis, we are now in a position to turn to acquisition, which provides a testing ground for theories of laryngeal feature representation.

## 2.2 Acquisition of laryngeal contrasts: Previous studies

Previous studies on the acquisition of voicing have found developmental differences between prevoicing and aspiration languages. With respect to the time course of acquisition, it appears that laryngeal contrasts are acquired later in prevoicing languages than in aspiration languages (Macken & Barton 1980a,b, Davis 1995). While the Dutch contrast is acquired some time around the age of three (Kuipers 1993a,b, Beers 1995), the English contrast is acquired relatively early, by the age of two (Macken & Barton 1980a). Previous research (Davis 1995) has indicated a role of acoustic salience in these acquisition differences, where prevoicing (voicing lead) is argued to be less salient than aspiration (long lag VOT). This suggests that some of the acquisition differences seen in languages are to some extent attributable to ease of perception and, possibly, production. However, differences in acoustic salience across languages do not exclude the possibility that differences in acquisition are due to different feature representations across languages. This paper will explore the phonetic versus phonological accounts for the patterns seen in acquisition.

### 2.3 Further assumptions and predictions

The accuracy of children's productions can be taken to reflect children's phonological knowledge and representations. Children's production errors have been argued to reflect innate universal grammar (Jakobson 1941/1968 and others), given that children's production errors can often be characterized as neutralizing to the unmarked value. For example, children often delete final consonants and produce CV syllables, e.g., *taart* /tɑ:rt/ 'cake' is produced as [tɑ:] in Dutch (Fikkert 1994), *Tag* /tag/ 'day' as [dɑ:] in German (Kerstin 1;5, see below), and *tape* /teyp/ as [t<sup>h</sup>e:] in English (Seth 1;7, see below). These production errors can be interpreted as reflecting phonological knowledge, such as the knowledge that the universally preferred syllable shape is a CV syllable. It is a well-known observation (Jakobson 1941/1968) that criteria for markedness based on cross-linguistic evidence are supported by language acquisition, as children tend to produce the least marked properties before more marked ones (cf. Zamuner 2003, Zamuner, Gerken & Hammond 2005). Returning to the example of syllable structure, we note that children initially produce the least marked CV syllables, before producing more marked syllable shapes, such as those with final consonants (Fikkert 1994, Levelt et al. 2000).

Assuming Jakobson's hypothesis that children's initial errors reflect the unmarked values of phonological features, we can derive a number of predictions regarding children's error patterns, based on the Single Feature Hypothesis and the Multiple Feature Hypothesis. These predictions are given in Table 5.

	Representation	Unmarked	Error Type
Single Feature Hypothesis	[±voice]	[-voice]	[+voice] → [-voice]
	or		
	[voice]	[ ]	[voice] → [ ]
Multiple Feature Hypothesis	[voice]	[ ]	[voice] → [ ]
	(prevoicing languages)		
	[spread glottis]	[ ]	[spread glottis] → [ ]
	(aspiration languages)		

Table 5: Predictions of the acquisition of laryngeal features based on the Single Feature Hypothesis and Multiple Feature Hypothesis

Recall that the Single Feature Hypothesis makes use of a single binary feature of [±voice] or monovalent feature [voice]. The unmarked value for this theory is invariant across languages, because all languages utilize the same feature to represent laryngeal contrasts. With a binary feature, this unmarked value is [-voice]. If children's initial productions tend toward the unmarked value, this would predict that the direction of errors is cross-linguistically uniform, and should be independent of the language that children are acquiring. Accordingly, children learning Dutch, German or English are all predicted to make devoicing errors [+voice] → [-voice]. These errors would affect words starting with (marked) [+voice] con-

sonants, while words starting with (unmarked) [-voice] consonants should not be affected. (For a monovalent feature, predictions are essentially the same, although the specifications are slightly different.) In contrast, the Multiple Feature Hypothesis would predict differences between prevoicing and aspiration languages regarding the types of consonants that are affected. Languages with prevoicing should display devoicing errors [voice] → [ ] (omission of the feature [voice]) in words starting with (marked) voiced consonants, while words starting with (unmarked) [ ] voiceless consonants should not be affected, whereas children acquiring aspiration languages should produce de-aspiration errors [spread glottis] → [ ] (omission of the feature [spread glottis]) in words starting with (marked) aspirated consonants, while words starting with (unmarked) unaspirated consonants should not be affected. In Table 6, the predictions are spelled out in terms of phonetic symbols.

	Dutch	German	English
Single Feature Hypothesis	/b/ → [p]	/b/ → [p <sup>h</sup> ]	/b/ → [p <sup>h</sup> ]
Multiple Feature Hypothesis	/b/ → [p]	/p <sup>h</sup> / → [p]	/p <sup>h</sup> / → [p]

Table 6: Predictions of laryngeal errors for Dutch German and English, based on the Single Feature Hypothesis and Multiple Feature Hypothesis

In sum, using Jakobson's hypothesis that children's errors are changes in the direction of the unmarked, theories of laryngeal representation make different predictions about consonants which are prone to undergo errors in acquisition. Hence, the acquisition of Dutch versus German and English provides an excellent test case for these theories.

### 3. Dutch

To test the predictions of the Single and Multiple Feature hypotheses, we collected acquisition data from Dutch, German, and English. Different corpora from the CHILDES database were studied. We will start with a discussion of the Dutch data, which were taken from the CLPF database (Fikkert 1994, Levelt 1994). The data from 11 Dutch monolingual children whose ages range between 1;0 and 2;11 were studied; this involved approximately 20,000 utterances. Examples of voicing errors are below.

#### (1) Examples of laryngeal errors in Robin's utterances

- |    |        |          |      |           |
|----|--------|----------|------|-----------|
| a. | douche | 'shower' | tus  | (1;10.21) |
| b. | dier   | 'animal' | tiɪ  | (1;10.21) |
| c. | beer   | 'bear'   | pi   | (1;7.13)  |
| d. | bal    | 'ball'   | pɑl  | (1;7.13)  |
| e. | baby   | 'baby'   | pipi | (1;8.10)  |
| f. | thuis  | 'home'   | dœys | (1;5.10)  |

## (2) Examples of laryngeal errors in Tom's utterances

- |    |       |         |     |          |
|----|-------|---------|-----|----------|
| a. | boot  | 'boat'  | pɔ  | (1;5.0)  |
| b. | bal   | 'ball'  | pɑ  | (1;5.14) |
| c. | bed   | 'bed'   | pɛt | (1;5.28) |
| d. | doen  | 'do'    | tun | (2;1.14) |
| e. | paard | 'horse' | bat | (1;3.24) |

Only productions of initial stops /b/, /p/, /d/ and /t/ were considered. These stops had to be faithfully realized for place of articulation to be included in our analyses. Dutch lacks the voicing contrast in velars; hence we did not consider the velar stop /k/. Also, fricatives were not included, because in many Dutch regions the distinction between voiced and voiceless fricatives is disappearing (Slis & van Heugten 1989, Ernestus 2000, Van de Velde et al. 1996, Van de Velde & van Hout 2001).

Not all children were monitored for the same period of time: Figure 1 shows the ages of the different children in the database. Note that at the beginning and end of the age span, data were collected for only one or two children: Tom at 1;0, Leon at 2;9 and Noortje at 3;0.

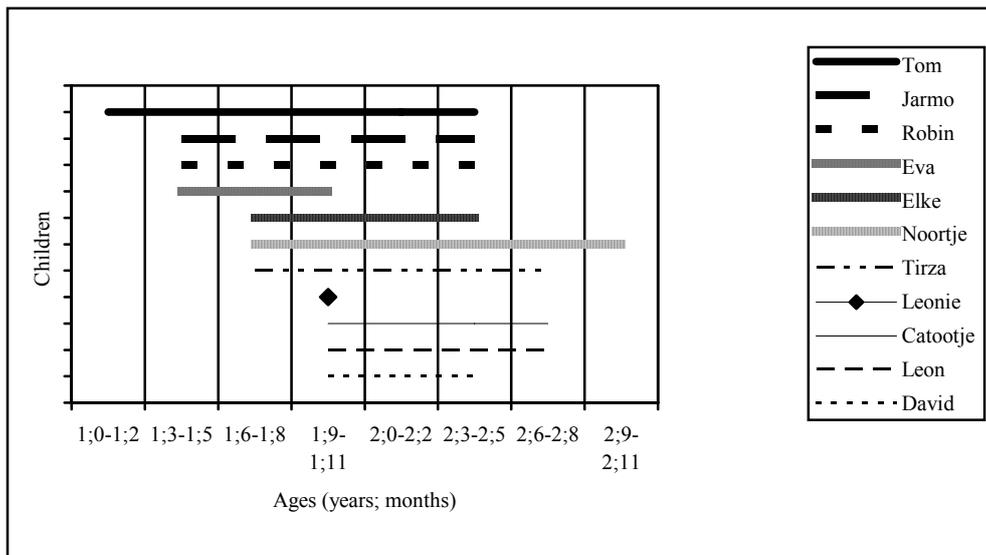


Figure 1: Breakdown of children's ages from the CLPF database

In Figure 2, the number of target words is given for each child. Note that the number of tokens is given here: for types, a similar pattern was found. In the remainder of the discussion of Dutch, we will only present results from token analyses. Figure 2 shows that all children attempted more voiced targets (/b/ and /d/-initial words, a total of 4871 targets) than voiceless targets (/p/ and /t/-initial words, a total of 2244 targets). On the basis of this, one might predict that children will be more accurate when producing voiced targets than when producing voiceless targets, but we will see that we find the opposite pattern: overall, children produced more voiceless than voiced stops.

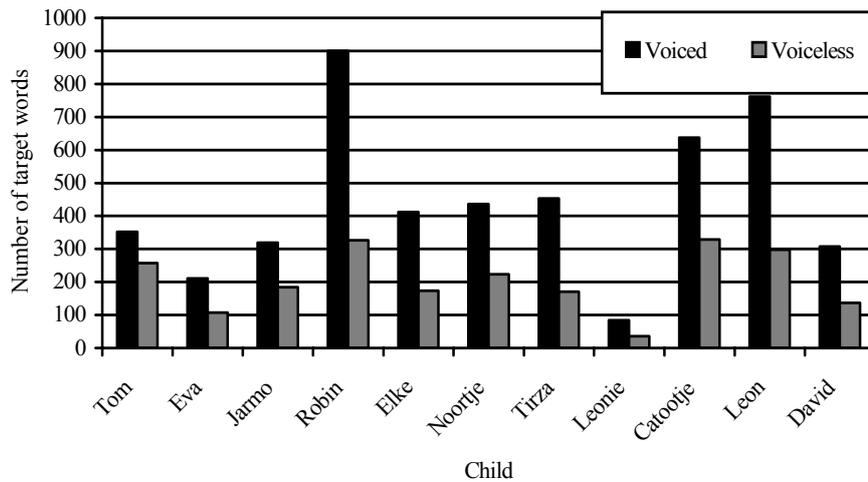


Figure 2: Number of target words (in tokens) per child

Figure 3 shows the error percentages of all children when producing word-initial stops. These percentages were determined by averaging the percentage error rate across children.

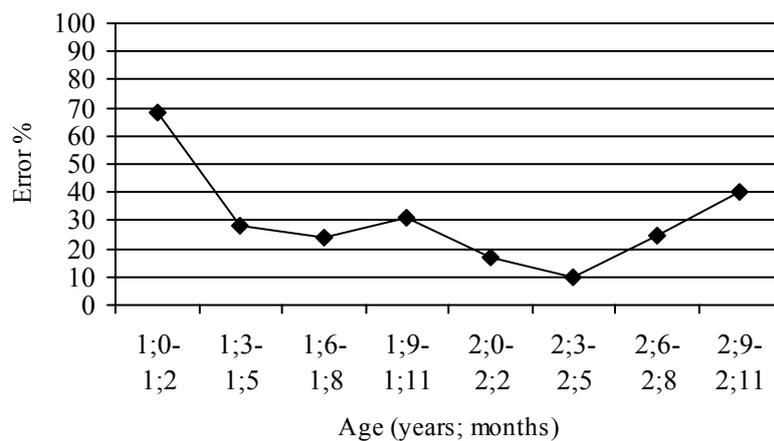


Figure 3: Error percentages of all children

In Figure 3, all errors children made in the voicing value of word initial segments are shown. Thus it collapses errors made both in voiced and voiceless segments. Children's productions become more faithful during the time they were studied. Apparently, the overall development curve is U-shaped, but this appearance is caused by the fact that at the age periods 1;0, 2;9 and 3;0 there are only data from one or two children. Noortje, who is the single child providing data at 3;0, was found to be late in her overall phonological development (Fikkert 1994). Hence, she causes a rise of the curve at this point. Her error rate for the production of initial stops remains quite high during the entire period in which she was studied.

When errors are broken down by place of articulation, we see that this factor does not play any crucial role in the production of the voice value. In Figure 4, the errors are broken down for labial-initial words (/b/ and /p/) versus alveolar-initial words (/d/ and /t/). There is no significant difference (*t-test*,  $p=0.11$ , two-tailed) between the error rates of these two places of articulation, hence, we find no evidence that voicing in labials is either more or less difficult than voicing in alveolars.

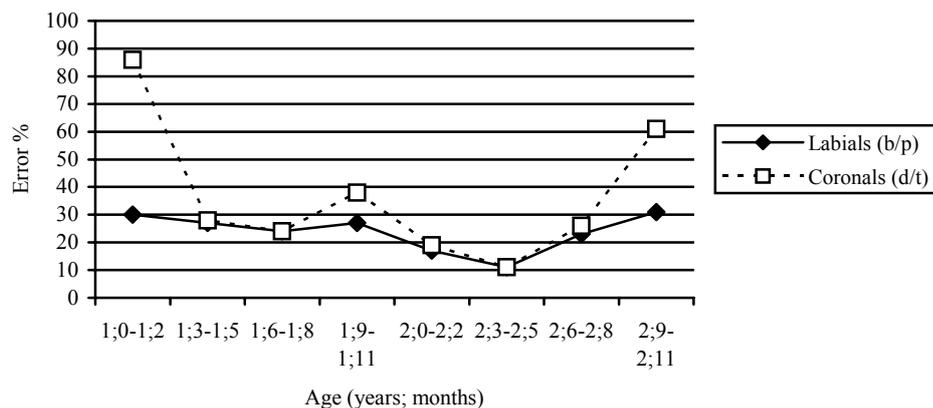
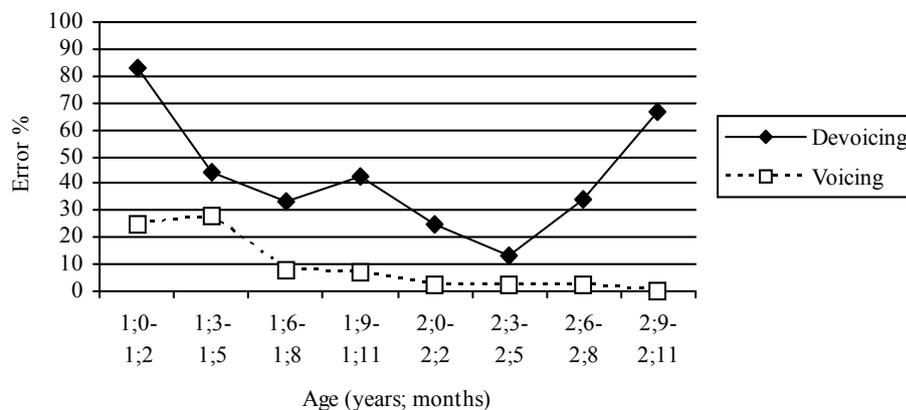


Figure 4: Percentage of errors in coronal and labials

In Figure 5, error rates are split for voicing errors (e.g., /p/ and /t/ produced as /b/ and /d/) and devoicing errors (e.g., /b/ and /d/ produced as /p/ and /t/). This clearly shows that there were more devoicing errors ( $M=42.75$ ,  $SD=22.6$ ) than voicing errors ( $M=9.25$ ,  $SD=11.01$ ). This difference is significant (*t-test*,  $p\leq 0.01$ , two-tailed), and holds for every stage. For all children, we see that devoicing errors persist well into the third year, while the rate of voicing errors drops to almost 0%.

We can also examine the extent to which Dutch children's initial 'voicing' production patterns reflect the distribution of voicing in the input (van der Feest 2004, 2007). For this analysis, we analyzed child-directed speech from the van de Weijer corpus (van de Weijer 1998). This corpus contains speech directed to a child between the ages of 2;6 and 2;9 (a selection of 18 days appears in the corpus). We conducted (type and token) counts of initial voiced and voiceless stops for different places of articulation. Results are summarized in Table 7.

Figure 5: *Percentage of voicing and devoicing errors*

	Labials		Alveolars			
	p	b	t		d	
types	151 (40.7%)	220 (59.3%)	104 (41.3%)		148 (58.7%)	
tokens	1492 (30.9%)	3342 (69.1%)	1481 (13.6%)		9389 (86.4%)	

Table 7: *Distribution of voicing in child-directed speech from van de Weijer corpus*

There is a preference for voiced stops in both type and token counts in child-directed speech. This means that the errors patterns seen in Dutch production data (voiceless stops are produced before voiced stops) cannot be accounted for by input frequencies.<sup>5</sup>

To summarize these data, we can say that overall, Dutch children acquire the voicing system quite late, having not yet completed it by the age of 2;6. We have seen that for the acquisition of the Dutch voicing contrast, there is no significant effect of place of articulation. Also, although more target words have voiced onsets, children make more errors in voiced than in voiceless initial segments, while their overall productions contain more voiceless than voiced segments. These findings support the featural specification [voice] for Dutch, assuming that unmarked voiceless segments are acquired before the marked voiced segments.

However, the acquisition data from Dutch are consistent with both the Single Feature Hypothesis and the Multiple Feature Hypothesis. Under the latter hypothesis, voiceless segments are assumed to be unspecified for the monovalent feature [voice], and hence, predicted to be acquired before specified voiced segments, while under the former hypothesis (assuming a single binary feature [ $\pm$ voice]), voiceless segments would also be predicted to be acquired first. Here, voiceless segments are specified as [-voice], and would be less marked than voiced segments, which are specified as [+voice]. Since predictions from these hypotheses are identical, Dutch acquisition data could, in principle, never produce

any crucial evidence deciding between these hypotheses. It is important, though, that the acquisition patterns cannot be explained on the basis of input frequency.

Still, the two approaches predict different orders of acquisition for the German segments. The Multiple Feature Hypothesis, which assumes that aspiration languages such as German represent the laryngeal contrast by [spread glottis], would predict that voiced segments are acquired first, since these are unspecified for [spread glottis], and hence unmarked. The Single Feature Hypothesis, on the other hand, assumes voiceless segments to be universally specified as [–voice], and for that reason would predict such (unmarked) segments to be acquired first. We now turn to a discussion of German to see which of the two approaches is supported by our acquisition data.

#### 4. *German*

German is an aspiration language, which differs from prevoicing languages such as Dutch in encoding its two-way laryngeal contrast by aspiration versus non-aspiration, at least in word onset position. As pointed out above, the Multiple Feature Hypothesis represents the German laryngeal contrast as one of [spread glottis] for aspirated /p/, versus [ ] for plain /b/ (Jessen 1996, Jessen & Ringen 2002). Accordingly, the prediction from this hypothesis is that the production errors of children acquiring German will be predominantly of the /ph/ → [b] (or ‘lenition’) type, matching a neutralization of the feature [spread glottis]. Phonetically, such errors would amount to a failure to realize aspiration on a stop that is lexically specified as [spread glottis].

German data were collected from the Nijmegen Database in CHILDES (MacWhinney 1999).<sup>6</sup> We considered data from the only child in the database for which sufficient phonetic transcription was available, Kerstin (aged 1;3–3;4). From this large longitudinal database (containing approximately 25,000 utterances) we selected Kerstin’s productions between ages 1;0 and 2;2, which allowed us to track her development with respect to laryngeal specifications.

Some characteristic examples of Kerstin’s de-aspiration errors are given below:

##### (3) Examples of ‘voicing’ errors in Kerstin’s utterances

- |    |       |         |  |
|----|-------|---------|--|
| a. | Papa  | ‘daddy’ | baba (1;5.7, 1;6.20, 1;7.24, 1;10.3, 1;11.20, 2;0.5) |
| b. | Puppe | ‘doll’  | bibbaa (1;3.22), bubbaa (1;3.22)                     |
| c. | Tag   | ‘day’   | daa (1;5.3)  |
| d. | Teddy | ‘Teddy’ | diddie (1;5.6), dide (1;6.13), didi (1;7.24, 1;8.22) |
| e. | Turm  | ‘tower’ | dum (2;3.1)  |

Note that the informal transcription indicated in the corpus of items such as *Papa* as ‘baba’ and *Tag* as ‘daa’ suggests pre-voicing, rather than just de-aspiration. This was presumably due to a language-specific bias on the part of the transcribers, who may have perceived unaspirated stops as lenis stops /b, d/. We will interpret the transcriptions conservatively as evidence for de-aspiration only. In comparison, only a very small number of errors in the opposite direction was found.

The single clear example is *Becher* (1;5.17) transcribed in the corpus as ‘peschel’, presumably phonetically [p<sup>h</sup>ɛʃəl]. The near complete absence of initial devoicing/aspiration in Kerstin’s utterances contrasts with the situation in English, as we will see in section 6.

Quantitative analysis confirms that Kerstin’s errors are almost exclusively of the ‘voicing’ type: an unaspirated realization of stops corresponding to aspirated stops in the adult language. See Figure 6.

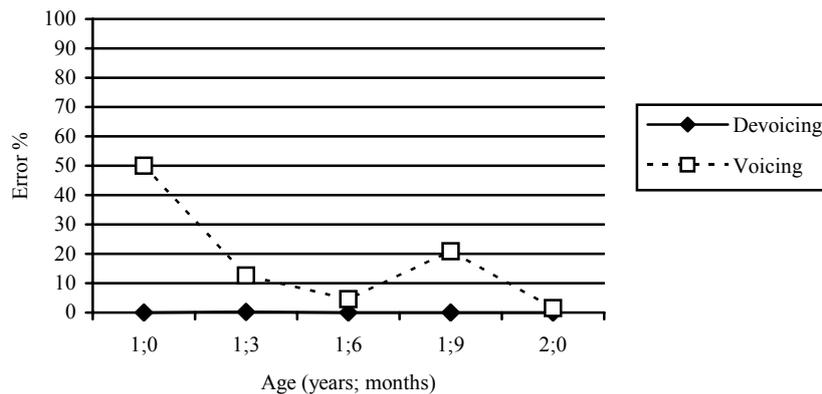


Figure 6: *Onset voicing versus devoicing in Kerstin's errors*

Errors of the ‘devoicing’ type were extremely rare, while ‘voicing’ errors are abundant. This matches earlier observations for German acquisition (Grijzenhout & Joppen-Hellwig 2002). The prediction of the Multiple Feature Hypothesis is thus borne out.

To what extent does Kerstin’s initial ‘voicing’ pattern reflect the statistics of the input? We carried out an analysis of child-directed speech in the same CHILDES corpus, based on utterances from caretakers present during recording sessions in the relevant periods (Kerstin’s age 1;0–1;12). We conducted (type and token) counts of initial voiced and voiceless stops for different places of articulation. Results are summarized in Table 8.

	Labials		Alveolars		Velars	
	p	b	t	d	k	g
types	32 (28.57%)	80 (71.43%)	52 (37.14%)	88 (62.86%)	94 (45.63%)	112 (54.37%)
tokens	121 (20.9%)	458 (79.1%)	157 (6.24%)	2361 (93.76%)	638 (48.55%)	676 (51.45%)

Table 8: *Distribution of voicing in child-directed speech from Kerstin's corpus*

There is a noticeable trend toward voiced initial stops (especially for labials) in child-directed speech during Kerstin’s second year.<sup>7</sup> Kerstin’s error pattern is thus

compatible with the input she received. But although Kerstin's input matches the direction of her errors, input statistics alone cannot account for the error pattern. This is because Kerstin produced virtually no errors of the devoicing type during her second year, a much stronger result than might be expected on the basis of the input pattern alone.

In sum, the error pattern of a German child between ages 1;0 and 2;3, showing abundant voicing errors in initial stops, but hardly any devoicing errors, is naturally accounted for by the Multiple Feature Hypothesis as omission of the active feature [spread glottis], but not by the Single Feature Hypothesis, while an input-based account offers only a partial explanation.

### **5. Interpretation and further predictions**

Summarizing so far, we found that the acquisition of the initial voicing contrast in Dutch is rather slow, and completed beyond the age of 2;6. Errors are predominantly of the 'devoicing' type. For German, the initial contrast is acquired earlier, and seems completed by the age of 2;0. Errors are overwhelmingly of the 'voicing' type (presumably, lenition or de-aspiration).

Two plausible interpretations of these findings suggest themselves, one phonological and the other phonetic. A strongly phonological account, which we have been assuming thus far, seeks a featural basis for the observed developmental differences. On the assumption that errors in early productions target the unmarked feature value, findings for Dutch and German would favour the Multiple Feature Hypothesis over the Single Feature Hypothesis, since only the former predicts differences in error patterns between the languages. The Multiple Feature Hypothesis models the Dutch voicing contrast on a monovalent feature [voice], and hence would correctly predict production errors of Dutch children to result in featurally unspecified stops, which are phonetically interpreted as 'voiceless'. The German laryngeal contrast, as opposed to Dutch, is based on a monovalent feature [spread glottis], which predicts German children's production errors to favour unspecified stops, phonetically realized as 'unaspirated' (that is, lenis and voiceless). The Single Feature Hypothesis, on the other hand, represents both languages by a single feature [voice], and hence would not predict any differences in the directionality of laryngeal errors between German and Dutch developmental patterns, as both languages would represent their contrasts by a single feature. (Note that, as indicated in Table 6, differences may occur between phonetic errors patterns in voicing and aspiration languages due to the language-particular implementation of the specifications [voice] and [ ].)

However, an alternative articulatory interpretation might be proposed, which would explain differences in error patterns between the languages, and hence would leave no room for testing the two representational hypotheses discussed above. According to what we will refer to as the 'Articulatory Effort Hypothesis', young children's initial preference for short lag VOT (that is, unaspirated voiceless stops) is due to lack of articulatory skills necessary to produce stops with either long lag VOT (aspiration) or short lead VOT (prevoicing). This would correctly predict that early German productions show a lack of aspiration, while early Dutch productions show a lack of prevoicing. To account for the developmental

differences between German and Dutch (i.e., age of acquisition of the contrast), the additional assumption would be needed that prevoicing is more difficult to produce than aspiration (Kewley-Port & Preston 1974, van Alphen, this volume). Alternatively, a perceptual account may be given following Davis (1995) and others: on the basis of greater perceptual salience of long lag VOT as compared to short lead VOT, children acquire the laryngeal contrast in aspiration languages earlier than in prevoicing languages.

The Articulatory Effort Hypothesis explains properties of children's errors in relation to the target language, but its validity need not rule out a role of featural representations in the explanation of error patterns. We are thus facing the following question: when considering children's laryngeal errors, how to distinguish effects of articulatory effort from effects of feature specifications? Here is an attempt to tease the two kinds of effects apart.

The Articulatory Effort Hypothesis would predict that errors correlate with the overall motoric complexity of a target. As is well-known, the articulatory effort required for the realization of a gesture may also depend on its position in an utterance. For example, it is much easier to maintain voicing in intervocalic contexts than in word-initial or final contexts. However, motoric effort should be independent of the presence of a target elsewhere in the utterance, specifically when they are not adjacent (for example, when two consonants are separated by a vowel), or when the targets are articulatorily diverse. Cross-linguistically, the articulatory gestures for laryngeal contrasts and the acoustic cues are quite varied. Cues include VOT, closure duration, duration of the preceding vowel (Keating 1984). Within a language, choice of laryngeal gesture may depend on a segment's position in the word, in the syllable, or on neighbouring segments. For example, English realizes laryngeal contrasts in onset mainly by VOT, and laryngeal contrasts in coda mainly by duration of the preceding vowel, closure duration, or glottalization. In sum, the Articulatory Effort Hypothesis would predict few interactions in error patterns between articulatorily heterogeneous positions, such as the onset and coda in English.

In contrast, a phonological account would predict contrastive specifications to appear in children's error patterns, which abstracts from fine-grained phonetic realization depending on position. For example, a phonological account would predict cases of 'laryngeal harmony' between onset and coda, in which only contrastive features would harmonize, not redundant ones.<sup>8</sup> (Cross-linguistic studies on laryngeal cooccurrence patterns include MacEachern 1997, Hanson 2001, Rose & Walker 2001.) It should be emphasized that by 'harmony' we generally refer to any kind of interaction between segments which produces identical contrastive feature specifications, without implying autosegmental spreading resulting in doubly-linked features. As Fikkert & Levelt (2002) argue, consonant harmony at early stages of development may be driven by a general requirement for stops to be featurally similar, regardless of whether similarity is achieved by spreading, by default, or by phonologically active features.

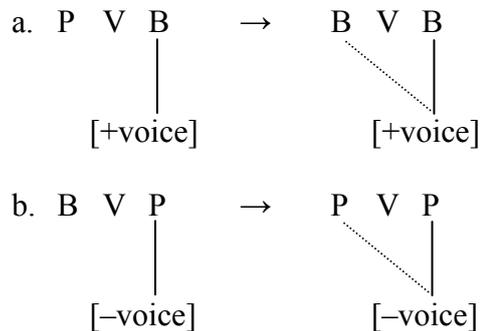
Radical underspecification of contrastive features would make an additional prediction that production errors reflect activity of the specified feature only, to the exclusion of the unspecified value. Under the Multiple Feature Hypothesis,

laryngeal features are monovalent, mainly to capture the observation that voiceless unaspirates are unmarked both in prevoicing and in aspiration languages. Languages differ as to which feature is specified: [voice] in prevoicing languages such as Dutch, and [spread glottis] in aspiration languages such as English. Underspecification thus creates predictions about harmony, since only unspecified segments should be the targets, assimilating non-locally to specified segments.

Different harmony effects would be predicted to occur, depending on the two featural approaches under comparison. Consonant harmony of place of articulation in children's early productions is typically anticipatory (Menn 1971, Smith 1973, Pater & Werle 2001, 2003, Fikkert & Levelt 2002), which leads us to expect a similar asymmetry for laryngeal harmony. For this reason, we will consider predictions for a hypothetical anticipatory harmony pattern, in which laryngeal errors in the onset anticipate the coda's specification.

Under a Single Feature Hypothesis with a binary feature [ $\pm$ voice], symmetrical error patterns within a language would be predicted, since both values are active, and potentially induce errors. This would predict a pattern with both devoicing and voicing in onsets, depending on whichever feature value is specified in the coda.

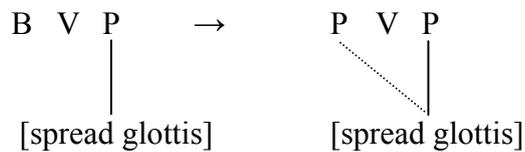
(4) Harmonies predicted by the Single Feature Hypothesis [ $\pm$ voice]



Note that a variant of the Single Feature Hypothesis based on a monovalent feature [voice] would only predict harmony of the type (4a), not (4b). Thus, if anticipatory laryngeal harmony were to be found in children's English, this could only be /PVB/ → [BVB]. This makes a strong prediction, which allows a rather straightforward test of these two versions of the Single Feature Hypothesis.

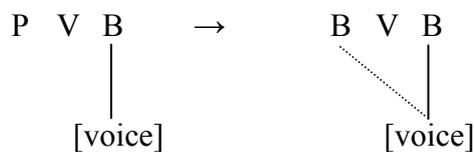
The Multiple Feature Hypothesis predicts patterns to be asymmetrical, and to correlate with a language's 'active' feature, either [voice] or [spread glottis]. That is, languages whose specified feature is [spread glottis] would be predicted to display only one kind of error: onset devoicing triggered by a voiceless coda /BVP/ → [PVP], but not onset voicing triggered by a voiced coda /PVB/ → [BVB], because voiced codas would be laryngeally unspecified [ ].

## (5) Harmony predicted by the Multiple Feature Hypothesis: aspiration languages



Under the Multiple Feature Hypothesis, prevoicing languages should only display harmonies involving voiced segments, /PVB/ → [BVB]:

## (6) Harmony predicted by the Multiple Feature Hypothesis: prevoicing languages



Note that the Multiple Feature Hypothesis and the monovalent version of the Single Feature Hypothesis make similar predictions for prevoicing languages. Predictions differ between the ‘monovalent’ frameworks, however, for aspiration languages. If the Multiple Feature Hypothesis is correct, English uses monovalent [spread glottis], and hence should display anticipatory harmony of the ‘devoicing’ type /BVP/ → [PVP], whereas if the Single Feature Hypothesis is correct, anticipatory harmony should be of the ‘voicing’ type, /PVB/ → [BVB].

In sum, to compare predictions made by the Single Feature Hypothesis and Multiple Feature Hypothesis, we must distinguish monovalent and binary variants of the latter. If children’s productions were to contain systematic patterns of voicing harmony, but not devoicing harmony, this pattern would be compatible with both the Multiple Feature Hypothesis and the monovalent version of the Single Feature Hypothesis, but it would form evidence against its binary version. Next, if harmony of the voicing and devoicing type were to systematically co-occur in a child’s productions, this would support the binary version of the Single Feature Hypothesis, but form evidence against both monovalent accounts. Finally, if we were to find that children’s production errors consistently display devoicing harmony, while lacking voicing harmony, this asymmetrical pattern would favour the Multiple Feature Hypothesis, but constitute evidence against the monovalent and binary variants of the Single Feature Hypothesis. The logical options are summarized in Table 9:

	Voicing harmony only	Both voicing and devoicing harmony	Devoicing harmony only
SFH, binary [±voice]	contra (circumstantial)	pro	contra (circumstantial)
SFH, monovalent [voice]	pro	contra	contra
MFH [voice] or [sg]	pro	contra	pro

Table 9: *Evidential status of hypothetical harmony patterns for the Single Feature Hypothesis (binary and monovalent) and Multiple Feature Hypothesis*

Let us now turn to a test case for the Multiple Feature Hypothesis against the Single Feature Hypothesis: English.

## 6. *English*

English acquisition data may serve as a test case, since this language precisely meets the conditions under which harmonic anticipations of laryngeal features might occur. First, English matches German (but not Dutch) in being an aspiration language. Hence, under the Multiple Feature Hypothesis [spread glottis] is specified, predicting this feature to be active in children's early phonologies. Indeed, English-learning children display de-aspiration errors (Menn 1971). Second, unlike German, English lacks syllable-final laryngeal neutralization, so that coda obstruents are specified contrastively. Since onsets and codas both license laryngeal specification, harmony effects become potentially visible. This meets the logical requirement which must be fulfilled for testing for positional interactions involving [spread glottis].<sup>9</sup>

### 6.1 *Earlier studies*

Earlier studies (such as Smith 1973) provide evidence for initial voicing and final devoicing in children's productions. Let us first turn to some data from Smith (1973). During the first half of his third year (ages 2;2–2;6), Amahl realized most of his initial stops as plain (voiceless unaspirated), by a general neutralization of initial laryngeal contrasts. In this period, initial neutralization affects voiced targets (for example, bell), as well as voiceless ones (for example, pen). (We adopt Smith's transcription.)

(7) Initial stops realized as voiceless unaspirated, irrespective of targets

(ages 2;2–2;6)

a. bell [b̥ɛ] (2;2)

b. pen [p̥ɛn] (2;2)

In the same period, Amahl also neutralized most word-final stops to voiceless. Since this has audible consequences only for voiced targets (e.g. [mɒb̥] for knob), the overall effect is one of final devoicing.

- (8) Final stops realized as voiceless (unaspirated), irrespective of targets  
(ages 2;2–2;5)
- a. knob [mɒb̥] (2;2)
  - b. stop [dɒp] (2;2)

Amahl's development during this period is shown in Figure 7. The error percentages were calculated as the proportion of target words of which the laryngeal realization, [p], [b̥], or [b], deviated from the target specification, /p/ or /b/.<sup>10</sup> For target /b/, for example, any realizations deviating from it, either [p] or [b̥], were considered as devoicing errors.

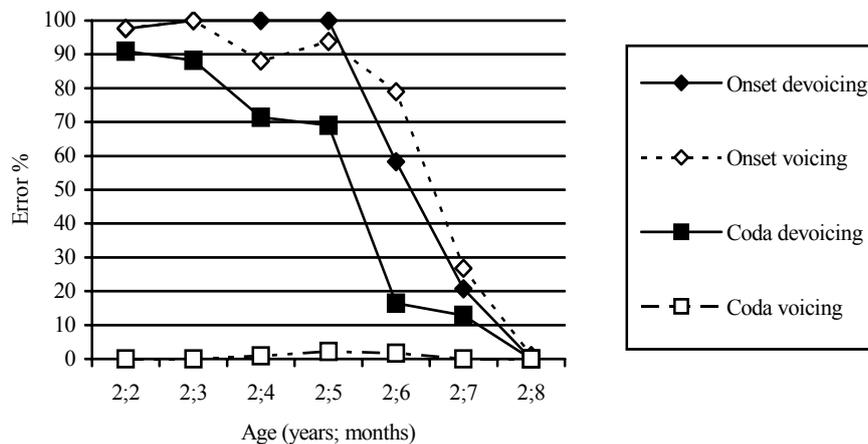


Figure 7: Amahl's development (2;2–2;8)

Note that the final contrast is acquired slightly earlier than the initial contrast, by about a month. The error rate of final devoicing drops sharply round age 2;5, while that of initial neutralization (as shown in the two topmost lines) follows at the distance of about a month.

The observation that Amahl's laryngeal contrast stabilizes slightly earlier in final than in initial position may come as a surprise, given that the child is unlikely to receive input with the laryngeal distinction directly realized on the final stop, whereas initial stops have robust VOT cues.<sup>11</sup> However, other cases are known of children acquiring English who mastered the laryngeal distinction in codas before it emerged in onsets (Clark & Bowerman 1986:55, fn. 5, Vihman & Ferguson 1987:383, Fey & Gandour 1982, but see Stoel-Gammon & Buder 1999). More generally, other consonant types, such as fricatives and liquids, are more likely to be first acquired in final position (Ferguson 1978, Stoel-Gammon 1985).

At best, we can offer speculative accounts of the coda-onset lag in Amahl's laryngeal development, as no acoustic data are available for verification. One ac-

count would attribute the lag to production factors rather than lexical representation; Amahl may have mastered control over vowel duration, the primary realization of laryngeal contrast in coda, before the gestural coordination between release and voice onset which is required for aspiration. Under this scenario, lexical representations in onset and coda are stable at an earlier stage, setting the stage for rapid across-the-board changes once the relevant gestures are mastered. Indeed, Amahl shows a rapid development of the laryngeal contrast in onset and coda, both of which are completed within approximately three months. Nevertheless, an explanation of the coda-onset lag based on developing lexical representations cannot be ruled out, because laryngeal error rates vary somewhat between individual lexical items, an observation which is difficult to explain under a production-only account. For example, during period 10-11 (at the age of 2;6) all three occurrences of *bread* had neutralized onsets, while all three occurrences of *Braj* (a name) were realized correctly. Hence, it is quite possible that Amahl's laryngeal contrast was lexically represented in final position before it emerged in initial position.

Relative strength of the laryngeal contrast in coda position will become a major factor in our central case study, to which we turn next.

### 6.2 *Seth: a case study*

The data in this section were taken from a large CHILDES database (Wilson & Peters 1988), containing approximately 12,500 utterances, with a total number of 39,000 words. All data were from a single monolingual child, named Seth, aged between 1;7–4;1, who was acquiring American English. Utterances in the database are matched with target words in plain orthography, and are phonetically transcribed at a level allowing for qualitative and quantitative analysis of voicing patterns. The original sound files were kindly made available to us in digitized form by Ann Peters and Brian MacWhinney for further transcription and acoustic analysis.

We monitored Seth's development between ages 1;7 and 2;5, the period during which major changes in the laryngeal contrast took place, and at the end of which Seth's productions of the contrast were largely indistinguishable from adults.

**6.2.1 *Initial devoicing.*** We first focused on word-initial position and collected all of Seth's productions of content word<sup>12</sup> targets containing an initial voiced or voiceless stop. This allowed us to search for factors which possibly influenced the proportion of two major types of error: initial 'voicings' (actually, de-aspirations resulting in plain stops) and initial 'devoicings' (actually, aspirations). The resulting dataset contained 227 types and 4354 tokens. Table 10 shows type and token distributions of initial voiced and voiceless target stops, for different places of articulation.

	Labials		Alveolars		Velars		Total
	p	b	t	d	k	g	
types	34 (37.78%)	56 (62.22%)	45 (63.38%)	26 (36.62%)	48 (72.73%)	18 (27.27%)	227
tokens	902 (58.95%)	628 (41.05%)	689 (42.98%)	914 (57.02%)	553 (45.29%)	668 (54.71%)	4354

Table 10: *Distribution of initial stop targets in Seth's productions (types & tokens)*

A representative set of examples of initial errors in Seth's early utterances are given below:

(9) Initial devoicing in Seth's utterances

Labials

- a. bark [paak] 1;8  
 b. boy [pay] 1;9  
 c. bike [payk] 1;10  
 d. backpack [pəkpək] 1;11

Alveolars

- e. Dabee [tabiy] 1;7  
 f. doughnut [townaa] 2;0  
 g. dog [tagiy] 2;2  
 h. drink [trunk] 2;5

Velars

- i. geese [kiys] 1;8  
 j. go [ko] 1;8  
 k. got it [kaaDit] 2;2  
 l. get [ket] 2;3

(10) Initial voicing in Seth's utterances

Labials

- a. penny [bəniy] 1;11  
 b. play [bwedjʌ] 1;11  
 c. peanut butter [bi bʌdʌ] 1;11  
 d. play [beydl] 2;1

Alveolars

- e. tape [deyp] 1;9  
 f. trunk [drʌŋk] 1;10  
 g. tan [dən] 2;0  
 h. tell [dəʌ] 2;0

Velars

- i. kiss it [giset] 1;8  
 j. kitchen [gisʌN] 1;9  
 k. cookie [gukiy] 1;10  
 l. cool [guw] 1;10

Although both types of initial errors are abundant, Seth makes more devoicing errors than voicing errors in initial position, as shown in Table 11. This presents token counts of voiced and voiceless targets (/B/ and /P/) and their realization (voiced [B] or voiceless [P]) over a succession of four three-month periods (1;7–

2;5), where the last column gives the results collapsed across periods. In the notation we use, B refers to voiced (labial, coronal, or velar) stops, and P to voiceless stops. Error percentages are indicated in cells for unfaithful realizations. For example, for the period 1;7–1;9, the corpus contains 415 targets with initial voiced stops, 372 of which were realized faithfully, and 43 of which (10.4%) were devoiced.

	Age 1;7–1;9		1;10–1;12		2;0–2;2		2;3–2;5		1;7–2;2	
	Realiz.	[B]	[P]	[B]	[P]	[B]	[P]	[B]	[P]	[B]
target /B/	372	43	273	17	690	31	775	9	2210	100
target /P/	24	537	16	627	9	551	1	379	50	2094
chi-square	$\chi^2 = 13.8$		$\chi^2 = 6.7$		$\chi^2 = 7.6$		$\chi^2 = 2.4$		$\chi^2 = 13.6$	
	$p \leq 0.001$		$p \leq 0.01$		$p \leq 0.01$		(n.s.)		$p \leq 0.001$	

Table 11: *Distribution of initial errors in Seth's productions*

Over the monitored period (1;7–2;5), both error types decreased continuously. To determine whether devoicing errors were more frequent than voicing errors, a series of chi-square tests was conducted. Results were significant for all periods except the period of 2;3–2;5.

Figure 8 shows the direction of initial errors in tokens (devoicing versus voicing) as it develops between the ages of 1;7 and 2;5.

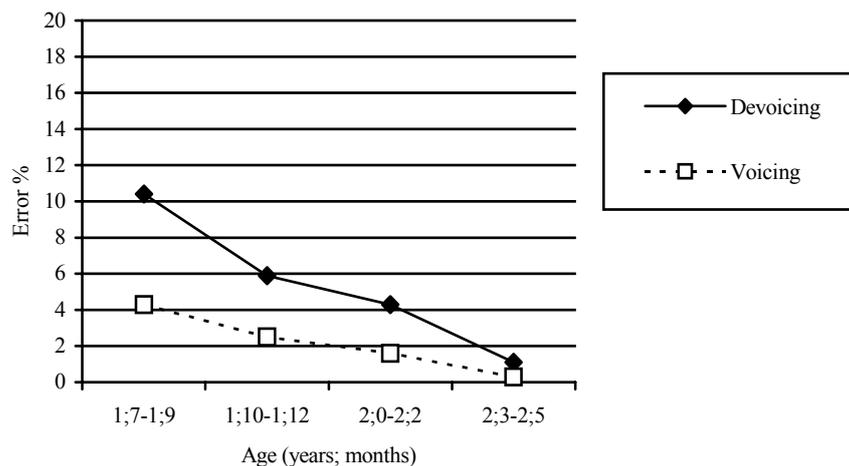


Figure 8: *Direction of initial errors in Seth's productions*

Note that devoicing errors occur about twice as frequently as voicing errors throughout Seth's development.

The dominance of devoicing errors extends to types, as Table 12 shows:<sup>13</sup>

	[B...]	[P...]
target /B.../	68	32
target /P.../	19	108
Chi-square	$\chi^2 = 9.3$ $p \leq 0.01$	

Table 12: *Distribution of voicing and devoicing errors in Seth's productions (types)*

The high proportion of devoicing errors in the early productions of an English learning child apparently clashes with our previous observations for German, where 'voicing' (actually de-aspirating, lenition) errors prevailed. It seems to run against the typological prediction made in section 2, according to which aspiration languages would display errors of the 'voicing' (de-aspiration) type, so that English would parallel German. Upon closer inspection, however, we see that Seth's initial devoicings are not simply neutralizations to the unmarked value.

When we differentiate Seth's initial devoicing errors according to their contexts in the word, it becomes clear that following consonants play a major conditioning role. Figure 9 shows that initial devoicing is much more frequent in targets in which a voiceless obstruent follows (e.g., *bark*, *drink*, *geese*) than in targets which have no following voiceless obstruent (e.g., *boy*, *dog*, *go*):

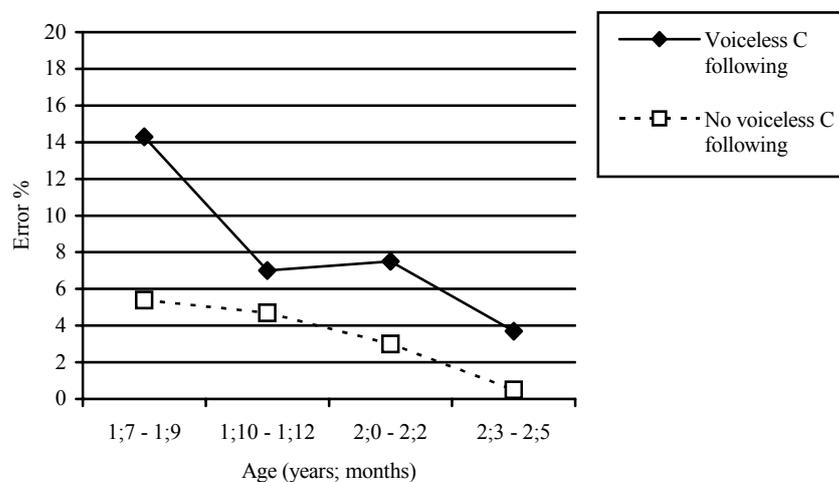


Figure 9: *Initial devoicing in Seth's productions: the role of following consonants*

We used a chi-square test to test the difference in error rates between target categories (words with a following voiced obstruent, a following voiceless obstruent, or no following obstruent), over the entire period (1;7–2;5), and found a strong effect ( $\chi^2 = 45.4$ ,  $p \leq 0.001$ ).

Next, to establish whether voiced obstruents differed from sonorants in their effects on initial devoicing, we broke down the category 'no voiceless consonant following'. Table 13 shows devoicing in three target types: (a) /B...P/ targets, which have a following voiceless obstruent (e.g. *back*, *diaper*, *grass*), (b) /B...B/

targets, which have a voiced obstruent (e.g. *baby*, *dog*, *give*), and (c) /B...R/ targets, which have a following sonorant or vowel ('R') (e.g. *ball*, *door*, *gone*). For each period, Seth's productions are broken down into faithful [B...] and unfaithful [P...] realizations of the target, and error rates are indicated in 'unfaithful' cells.

	Age 1;7-1;9		1;10-1;12		2;0-2;2		2;3-2;5		1;7-2;2	
Realiz.	[B]	[P]	[B]	[P]	[B]	[P]	[B]	[P]	[B]	[P]
/B...P/	198	33	132	10	197	16	156	6	683	65
/B...B/	77	6	65	5	256	7	370	1	768	19
/B...R/	97	4	76	2	237	8	249	2	659	16

Table 13: *Initial devoicing as a function of following consonants*

Throughout Seth's development, initial devoicing rate is highest for /B...P/ targets. Over the four periods, this error type reaches a much higher average (of 8.7%) than targets /B...B/ and /B...R/ (both 2.4%). Note that in all three categories, a gradual overall reduction of devoicing errors occurs.

The difference between /B...P/ targets and the other targets /B...B/ and /B...R/ turned out to be statistically significant, as Table 14 shows. This compares initial devoicing rates for three targets (/B...P/, /B...B/, /B...R/) for all four periods. Chi-square tests were conducted for the token distribution in Table 12, establishing that /B...P/ targets undergo devoicing significantly more often than targets /B...B/ and /B...R/, while any differences in devoicing rate between /B...B/ and /B...R/ targets are non-significant.

	1;7-1;9	1;10-1;12	2;0-2;2	2;3-2;5	1;7-2;2
/B...P/ versus /B...B/	$\chi^2 = 2.8$ (n.s.)	$\chi^2 = 0.001$ (n.s.)	$\chi^2 = 6.0$ ( $p \leq 0.025$ )	$\chi^2 = 10.3$ ( $p \leq 0.01$ )	$\chi^2 = 29.2$ ( $p \leq 0.001$ )
/B...P/ versus /B...R/	$\chi^2 = 7.6$ ( $p \leq 0.01$ )	$\chi^2 = 2.0$ (n.s.)	$\chi^2 = 4.1$ ( $p \leq 0.05$ )	$\chi^2 = 4.4$ ( $p \leq 0.05$ )	$\chi^2 = 26.4$ ( $p \leq 0.001$ )
/B...B/ versus /B...R/	$\chi^2 = 0.9$ (n.s.)	$\chi^2 = 1.7$ (n.s.)	$\chi^2 = 0.2$ (n.s.)	$\chi^2 = 0.9$ (n.s.)	$\chi^2 = 0.003$ (n.s.)

Table 14: *Initial devoicing as a function of following consonants*

In sum, devoicing in /B...P/ targets is significantly more frequent than for other targets across periods. It is more frequent than devoicing for /B...B/ targets in two out of four periods, and more frequent than devoicing for /B...R/ targets in three out of four periods. Also, /B...B/ and /B...R/ targets cannot be distinguished in terms of the initial devoicing rate.

We interpret these results as follows. In targets that begin with a voiced obstruent, a following ‘P’ (voiceless) segment triggers initial devoicing, as compared to a following ‘B’ (voiced) or ‘R’ (sonorant) segment, which behave as inactive with respect to initial devoicing. Initial devoicing in /B...P/ targets results in outputs with identical laryngeal features between the word onset (which undergoes it) and a following [spread glottis] obstruent (which triggers it). This is arguably a case of laryngeal harmony of the type that was predicted in section 5 (Table 9). Consequently, this finding supports predictions of the Multiple Feature Hypothesis: English, a language using [spread glottis] to represent its laryngeal contrast, should display activity of this feature in laryngeal harmony, if such harmony were to be found.

Note also that segment types predicted to be phonologically inactive by the Multiple Feature Hypothesis, ‘B’ (voiced obstruents) or ‘R’ (sonorants), are indeed inactive. Devoicing rates for target words with following ‘B’ or ‘R’ segments fall well below the rate observed for /B...P/ targets. Their shared behaviour is predicted by lack of specification for [spread glottis] under the Multiple Feature Hypothesis. Note that no other featural theory under consideration predicts a shared behaviour, since voiced obstruents will be marked [voice], while sonorants will not bear a distinctive laryngeal representation.

How to account for the fact that initial devoicing marginally occurs in the other targets /B...B/ and /B...R/? We attribute the initial devoicing rate for these targets, which amounts to 2.4% on average in the period 1;7–2;5, to the instability of early lexical representations. That is, the early lexicon contains incomplete featural information for lexical items, manifesting itself in variable productions, with both voiced and voiceless realizations. We suggest that during early production, lexically incomplete featural information is supplemented by three sources. First, context-free markedness effects (that is, omission of [spread glottis]) amount to context-free neutralization, which we observed as voicing in /P.../ targets. Second, copying of the active feature [spread glottis] amounts to laryngeal harmony in /B...P/ targets. Thirdly, a certain amount of random selection occurs. Initial devoicing in targets /B...B/ and /B...R/ occurs when random selection fills in incomplete features in early lexical representations. The fact that initial devoicing is facilitated by, but not categorically restricted to, /B...P/ targets, can thus be explained by an interplay of harmony effects and random specification.

To verify the amount of initial ‘devoicing’ (actually, aspiration) in Seth’s productions, we now turn to the results of phonetic analysis.

*6.2.2 Phonetic analysis.* In order to determine whether devoicing results in a full merger with target voiceless stops, we carried out narrow phonetic transcriptions and conducted acoustic measurements of VOT. First, all stop-initial items (between ages 1;7 and 1;9) were extracted from the digitized sound material. Next, 79 items were removed due to bad quality. For the remaining 605 items, narrow phonetic transcriptions were made by five transcribers (the current authors), and an acoustic analysis (VOT measurements) was conducted. Figure 10 shows mean VOT values (in msec) for Seth’s voiced and voiceless stops (ages 1;7–1;8). Seth’s values closely approximate the adult VOT values.

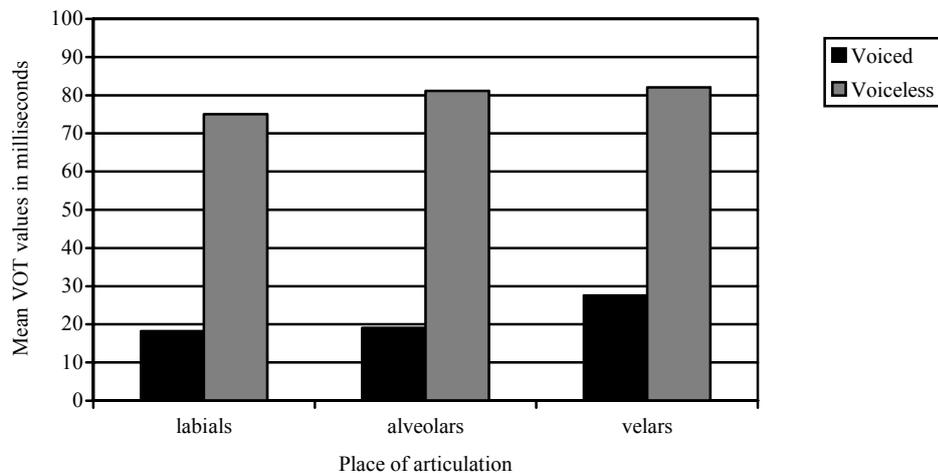
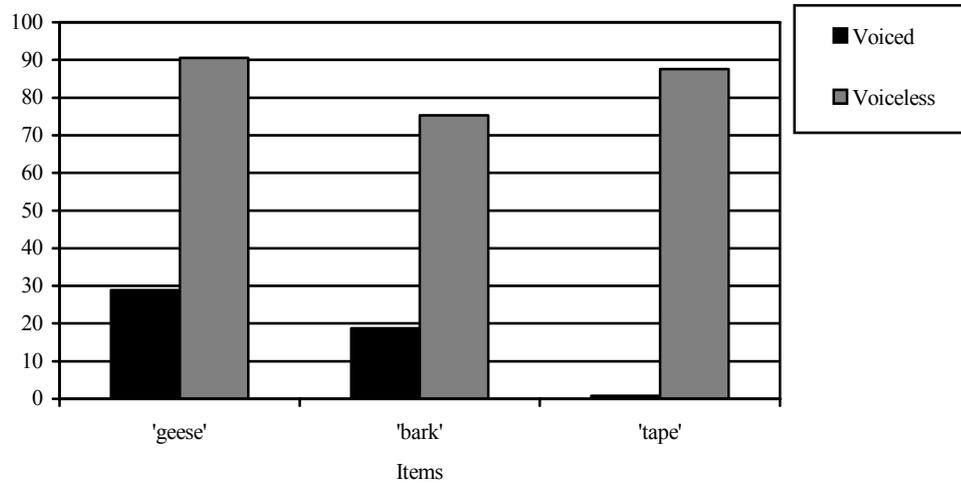


Figure 10: Mean VOT values in milliseconds for labials, alveolars and velars

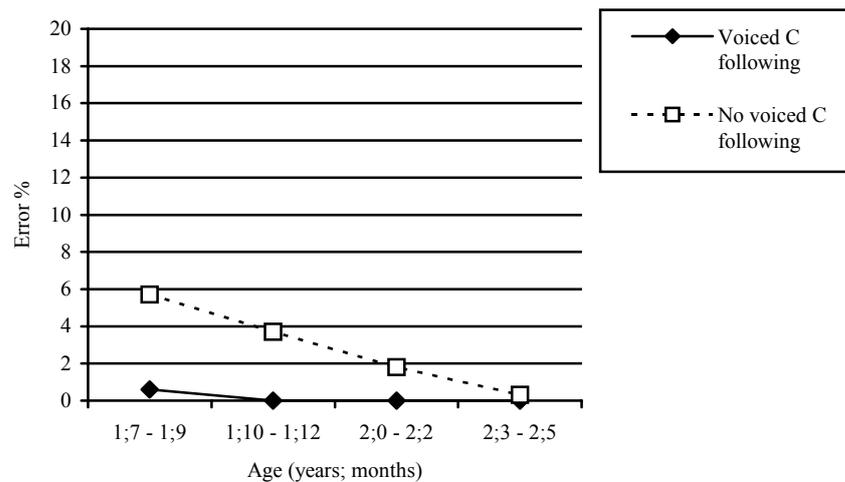
The following criteria were applied for determining an item's error status. A 'devoicing error' was defined as an item whose target has a voiced onset and which was realized with VOT > 30 msec (for labials and alveolars), or with VOT > 50 msec (for velars), and which was also categorized as 'voiceless' by a native listener (one of the current authors). A 'voicing error' was defined as an item whose target has a voiceless onset and which was realized with VOT < 30 msec (for labials and alveolars), or with VOT < 50 msec (for velars), and which was also categorized as 'voiced' by the native listener.

Analysis showed that Seth's devoicing errors resulted in initial stops (e.g. geese, bark) with mean VOT values which approximate mean VOT values for target voiceless stops (e.g. tape). See Figure 11 below. On the basis of these findings, we feel safe in assuming that devoicing errors are 'categorical', in the sense that devoiced target consonants are acoustically indistinguishable from faithfully realized voiceless consonants.

**6.2.3 Initial voicing.** We now turn to target words whose initial consonants are voiceless, and look into patterns of initial voicing, in order to find out whether voiced obstruents behave as phonologically inactive, as predicted by the Multiple Feature Hypothesis.

Figure 11: *Errors are categorical*

First, we are interested in the question whether Seth's productions show any effects of anticipatory voicing harmony, analogously to initial devoicing.

Figure 12: *Initial voicing: the role of following voiced consonants*

Surprisingly, hardly any voicing harmony occurs. The initial voicing rate for /P...B/ targets falls significantly below that of other targets /P...P/ and /P...R/ ( $\chi^2 = 11.7$ ,  $p \leq 0.001$ ). That is, the prediction from the Multiple Feature Hypothesis that voiced obstruents are phonologically inactive is confirmed. We momentarily put aside the question of what causes the harmony-avoiding pattern in /P...B/ targets, and break down the data for /P...P/ and /P...R/ targets, respectively.

Data are broken down for following consonants 'P', 'B' and 'R' in Table 15, which is the counterpart of Table 13 (for initial devoicing).

	Age 1;7-1;9		1;10-1;12		2;0-2;2		2;3-2;5		1;7-2;2	
	Real.: [B]	[P]	[B]	[P]	[B]	[P]	[B]	[P]	[B]	[P]
/P...P/	23	278	7	295	3	304	1	242	34	1119
/P...B/	1	155	0	214	0	66	0	29	1	464
/P...R/	0	104	9	118	6	181	0	108	15	511

Table 15: *Initial voicing in relation to following consonants (tokens)*

/P...P/ and /P...B/ targets display a gradual decrease of initial voicing errors during Seth's development. As noted above, the case of /P...B/ is most interesting since it shows no voicing errors except one in the first period. As compared to /P...B/, target /P...R/ shows more errors.

Differences between categories turn out to be actually much smaller than in the case of initial devoicing, when measured by chi-square tests. Table 16 compares initial voicing rates of targets /P...P/, /P...B/, and /P...R/.

	1;7-1;9	1;10-1;12	2;0-2;2	2;3-2;5	1;7-2;2
/P...P/ versus /P...B/	$\chi^2 = 10.1$ ( $p \leq 0.01$ )	$\chi^2 = 5.0$ ( $p \leq 0.025$ )	$\chi^2 = 0.7$ (n.s.)	$\chi^2 = 0.1$ (n.s.)	$\chi^2 = 11.7$ ( $p \leq 0.001$ )
/P...P/ versus /P...R/	$\chi^2 = 8.4$ ( $p \leq 0.01$ )	$\chi^2 = 5.7$ ( $p \leq 0.025$ )	$\chi^2 = 3.2$ (n.s.)	$\chi^2 = 0.4$ (n.s.)	$\chi^2 = 0.01$ (n.s.)
/P...B/ versus /P...R/	$\chi^2 = 0.7$ (n.s.)	$\chi^2 = 15.6$ ( $p \leq 0.001$ )	$\chi^2 = 2.2$ (n.s.)	(n.s.)	$\chi^2 = 10.8$ ( $p \leq 0.001$ )

Table 16: *Initial voicing in relation to following consonants (tokens)*

Phonological inactivity of final /B/ (voiced obstruents) becomes clear from Tables 15 and 16. Since the rate of initial voicing for /P...B/ targets falls far below that of other targets /P...P/ and /P...R/, we may safely conclude that final voiced obstruents fail to condition anticipatory voicing harmony. Non-occurrence of voicing harmony is straightforwardly predicted by the Multiple Feature Hypothesis on the assumption that English is a [spread glottis] language (see again Table 9).

Table 16 also shows significant differences during the first two periods between on the one hand, /P...P/ and /P...B/, and on the other hand, /P...P/ and /P...R/. Since in terms of its featural specification [spread glottis], /P/ forms no natural class with /B/, nor with /R/, both of which are unspecified, these findings are compatible with the Multiple Feature Hypothesis. Note, however, that the relative ease of initial voicing in /P...P/ targets is not predicted by this hypothesis, which has nothing to say about non-harmonic effects. If initial voicing amounts to context-free neutralization, the question is why final /P/ apparently facilitates it. Below we will offer an answer based on the maintenance of lexical contrast. But first we turn to another surprising result shown in Table 16.

The surprising finding is that /P...B/ and /P...R/, which ought to be indistinguishable by the non-specification of /B/ and /R/, nevertheless significantly differ

in period 1;10-1;12, where /P...B/ hardly shows any initial voicing, while /P...R/ does. Upon closer inspection, we find that the inhibitory effect in /P...B/ targets is caused by a single word, *please*, having 170 occurrences in this period, all of which fail to undergo initial voicing. Since *please* is a highly frequent item in Seth's corpus, the inactivity of its initial consonant may be attributed to its relatively stable lexical representation. The high rate of initial voicing in /P...R/ targets in the middle periods (1;10-2;2) is mostly due to alveolar stops: five cases of *turn*, all realized in rapid succession in a single recording session, and two cases of *tan*. Interestingly, the voiced realizations of these items were always preceded by nasals (for example, [sʌn dən] 'sun tan', and [n da' af] 'N turn off', where the nasal is apparently a realization of 'want'), suggesting a post-nasal voicing process in Seth's phonology. Since alveolar stops regularly alternate with flaps in American English, these may also have been early attempts at flapping. In sum, we suggest that initial voicing in /P...R/ targets is due to phonological factors applying across word boundaries, rather than phonological activity of sonorants following within the target.

In sum, there is no evidence for voicing harmony triggered by /P...B/, confirming the predictions of the Multiple Feature Hypothesis. Target /P...R/ shows no initial voicing except for a temporary unexplained increase during a single period. For /P...P/, some development is visible, with high levels of initial voicing during earlier periods, followed by a sharp decline. It thus seems that segment types /P/, /B/ and /R/ cannot be distinguished as to their effect on initial voicing. This finding is predicted by the central hypothesis that 'voicing' errors amount to a context-free neutralization, a delinking of [spread glottis] → [ ], not conditioned by segments elsewhere in the word.

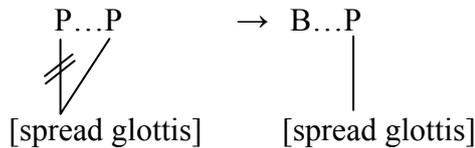
Finally, we turn to the unexplained low proportion of initial voicing errors in /P...B/ targets. As in German, we attribute initial voicing to context-free omission of [spread glottis], reflecting the instability of early lexical representations. Neutralization occurs at an average level of 2.9% in /P...P/ targets, whereas it is strongly inhibited in /P...B/ (average 0.2%). Our speculative account again starts from the hypothesis that initial voicing amounts to delinking of [spread glottis]. We observe that the key difference between the targets /P...P/ and /P...B/ is that the former contains two [spread glottis] elements, and the latter only one. We suggest that blocking of initial delinking in /P...B/ targets reflects an avoidance of wholesale deletion of [spread glottis].

- (11) Initial delinking resulting in complete loss of specified [spread glottis]

$$\begin{array}{ccc}
 \text{P...B} & \rightarrow & \text{B...B} \\
 \begin{array}{c} | \\ \hline | \end{array} & & \\
 \text{[spread glottis]} & & 
 \end{array}$$

The avoidance of initial delinking in /P...B/ and /P...R/ may thus be construed as a way of maintaining the laryngeal contrast by preserving the single occurrence of [spread glottis].

(12) Initial delinking resulting in partial loss of specified [spread glottis]



In contrast, /P...P/ targets have two occurrences, so that initial delinking still preserves one.

In sum, the observed asymmetry between onset devoicing and onset voicing in Seth's early word productions gives evidence from acquisition for the activity of [spread glottis] in English. Phonologically active coda obstruents are lexically specified for this feature, while voiced obstruents and sonorants are unspecified, hence phonologically inactive. The Single Feature Hypothesis predicts the observed asymmetry, while other models under consideration fail to account for it.

However, two important questions remain. First, the issue of directionality of harmony effects arises: is laryngeal harmony from coda to onset matched by harmony in the reverse direction, triggered by the onset, and effected in the coda? Second, a major issue arises as to whether the laryngeal harmony as witnessed in /B...P/ targets is due to lexical specification, or rather to surface realization. That is, we assumed that the influence of [spread glottis] is located at the level of lexical representation, but have not presented any evidence bearing on this. We will discuss both issues below.<sup>14</sup>

To answer these questions, we need to look at word onsets and codas separately. For this reason, we will now turn to a selection of Seth's productions, his monosyllables.

*6.2.4 Directionality of harmony: Seth's monosyllables.* Although the Multiple Feature Hypothesis makes no predictions<sup>15</sup> about the directionality of laryngeal harmony, we are nevertheless interested in directionality effects in Seth's data, for two reasons. First, directional asymmetries in laryngeal harmony might provide clues about early lexical representations, related to the relative strength of specification for onset and coda consonants. Second, directionality in laryngeal harmony would strengthen the similarity with other types of consonantal harmony in language acquisition and in speech production. A well-known asymmetry in directionality is found in consonant harmony in children's productions (Menn 1971, Smith 1973, Pater & Werle 2001, 2003, Fikkert & Levelt 2002) as well as in adult speech errors (Fromkin 1973, Shattuck-Huffnagel 1979, Stemberger 1991a,b) including anticipations of voicelessness. We will first address the issue of whether laryngeal harmony is unidirectional in Seth's productions, affecting onsets only, or bidirectional, affecting codas as well.

Before comparing contextual effects in onset and coda devoicing, we briefly address the relative stability of the laryngeal contrasts in onsets and codas in Seth's monosyllables. Figure 13 shows laryngeal error rates for final and initial stops, respectively, during four stages.

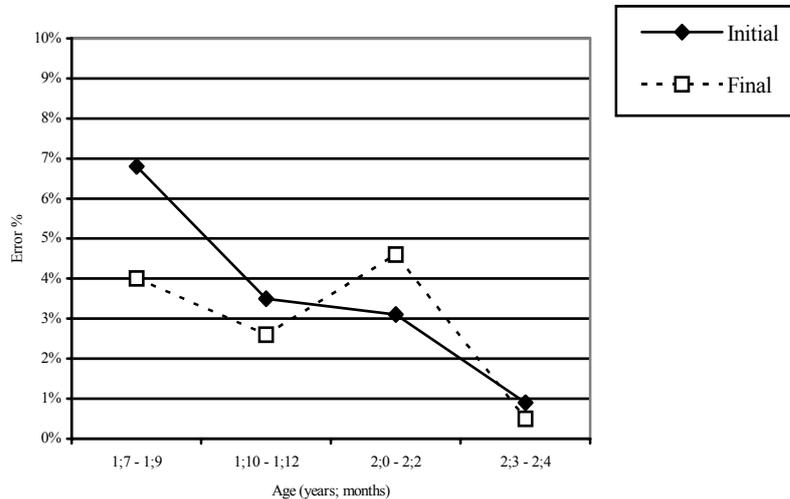


Figure 13: *Laryngeal error rates for initial and final obstruents*

During the earliest stage (1;7–1;9), when Seth's anticipatory devoicing pattern was at its peak, error rates for initial obstruents are slightly (although not significantly) above those for final ones. This suggests that the laryngeal contrast in final position is relatively stable as compared to initial position in early stages of phonological development. Speculatively, this would tie in with our earlier finding about Amahl's development (§6.2, Figure 7), who acquired a stable laryngeal contrast in coda slightly before it stabilized in onset.

The relative stability of the coda contrast naturally leads to the expectation that onset devoicing may be facilitated by voiceless codas, but not vice versa. This expectation stems from the assumption (see §6.2.1) that laryngeal harmony is related to the instability of early lexical representations in the sense that harmony amounts to the influence of a relatively stable lexical specification (typically, a coda) onto a less stable one (typically, an onset). To test the predicted lack of harmony in coda devoicing, we compared onset and coda devoicing in Seth's monosyllables, assessing the degree to which each is contextually conditioned by a voiceless obstruent. From the earlier mentioned database we extracted all targets (151 types, 3064 tokens) starting with a stop, which was voiced in about half of the cases (51.0% of types and 52.8% of tokens). We also extracted all targets (146 types, 2406 tokens) ending in a stop, which was voiced in about one third of the cases (32.9% of types, 28.9% of tokens).

To assess the relative strength of contextual factors in onset and coda devoicing, we first counted voiced and voiceless realizations of /B.../ monosyllables into three categories, whose target ended in a voiceless obstruent, voiced obstruent, or sonorant (including vowels). We will refer to these categories as /B...P/, /B...B/

and /B...R/ respectively. We then counted voiced and voiceless realizations<sup>16</sup> of /...B/ monosyllables, for the three target categories /P...B/, /B...B/, and /R...B/. Characteristic examples of coda devoicing in Seth's database are [bet] 'bed' (1;9) in the category /B...B/, and [fayNt] 'find' (2;2) in the category /P...B/. The results are shown in Tables 17 and 18.

	[B...]	[P...]
/B...P/	544	42
/B...B/	402	9
/B...R/	553	11
<i>chi-square</i>	$\chi^2 = 25.15, p \leq 0.001$	

Table 17: *Initial devoicing in Seth's monosyllables, relative to context*

	[B...]	[P...]
/B...P/	90	5
/B...B/	312	25
/B...R/	65	1
<i>chi-square</i>	$\chi^2 = 3.48, (n.s.)$	

Table 18: *Final devoicing in Seth's monosyllables, relative to context*

Again, a strong effect is found for the contextual conditioning of initial devoicing ( $\chi^2 = 25.15, p \leq 0.001$ ). This effect is due to the high proportion of initial devoicing in /B...P/ targets, as compared to other targets. Moreover, as compared to initial devoicing, final devoicing is not contextually conditioned. In spite of some minor differences between targets, there is no effect of onset type ( $\chi^2 = 3.48, n.s.$ ), supporting our hypothesis that final devoicing is context-free.

These results point to an interesting asymmetry in the directionality of laryngeal harmony. While onset devoicing anticipates the coda's voicelessness, coda devoicing shows no perseveration of onset voicelessness. That is, laryngeal harmony shares its directionality with other consonantal harmony processes found in acquisition and adult speech errors (see references at the beginning of this section).

Again we argue that the positional interactions in Seth's productions give evidence for an abstract representation of the laryngeal contrast, which is unified between the onset and the coda. As we saw earlier, the phonetic realization for laryngeal contrast in English strongly differ between onsets and codas. In onsets, VOT is the primary cue, while in codas, duration of the preceding vowel, closure duration, and (for some dialects) glottalization, are main cues. Since anticipation of the coda's voicelessness by the onset cannot be reduced to anticipation of the articulatory gestures involved, we have a case that the positional interaction occurs at a more abstract level: that of contrastive specification. This gives evidence from acquisition for representations of laryngeal contrasts involving monovalent

features. More precisely, activity of ‘voiceless’ obstruents (with non-activity of ‘voiced’ obstruents and sonorants) supports the specification of the feature [spread glottis], rather than [ $\pm$ voice].

6.2.5 *The level of harmony: lexical versus surface specification.* The final issue is whether initial devoicing in /B...P/ targets is due to lexical specification, or an effect of surface realization. We already saw some evidence from Seth’s early productions that laryngeal harmony is governed by lexical specification (see §§6.2.1 and 6.2.4). First, the phonological activity of voiceless obstruents, to the exclusion of voiced obstruents and sonorants, demonstrated in section 6.2, points to a relatively abstract level of representation, which is underspecified for laryngeal features. Second, we discovered two further properties of harmony, its directionality (anticipatory nature) and its non-local nature (passing across an intervening vowel), which both point to lexical representation as the relevant level. Both properties are reminiscent of processes which are assumed to be sensitive to lexical representation, such as speech errors involving place of articulation and voice (see Stemberger 1991a,b).

To further test our hypothesis that lexical specification, not surface realization, is the relevant level conditioning harmony, we checked whether /B...B/ target monosyllables whose final obstruent was produced voiceless due to final devoicing had an increased chance of undergoing onset devoicing as compared to /B...B/ targets whose final consonant remained voiced. As expected, we found no such effect. Thus, for predicting the likelihood for initial devoicing in a /B...B/ item, surface specification of the final obstruent was about equally as informative as its lexical representation. Tentatively, lexical specification alone (not surface realization) may account for laryngeal harmony.

Interestingly, we also found that likelihood of onset devoicing increased as a function of unfaithfulness in the coda, regardless of whether this involved deletion, devoicing, or other segmental changes. This ‘unfaithfulness effect’ was found for /B...B/ targets ( $\chi^2 = 13.61$ ,  $p \leq 0.001$ ), as well as /B...P/ targets ( $\chi^2 = 6.11$ ,  $p \leq 0.025$ ). It need not have a phonological interpretation; instead we suggest a role for general factors affecting accuracy of realization of the word or utterance as a whole.<sup>17</sup>

The tentative conclusion that lexical representations are involved in initial devoicing is supported by evidence from Seth’s monosyllables that shows that onset devoicing is, to some extent, sensitive to lexical frequency of individual items.<sup>18</sup> To test for a correlation between item frequency and onset devoicing, we placed all forty /B...P/ and /B...B/ targets in a rank-order by frequency in Seth’s corpus, split the list into two halves (of most frequent and least frequent items), and calculated the error rates for each list. We found that initial devoicing occurs less often in the most frequent items: only 5.1% of the most frequent items underwent initial devoicing, versus 10.7% of the least frequent items. The frequency effect ( $\chi^2 = 4.01$ ,  $p \leq 0.05$ ) is, of course, compatible with developing lexical representations.

## 7. *Conclusions*

This study of the acquisition of laryngeal contrast in three Germanic languages with binary laryngeal contrasts (Dutch, German, and English) offers evidence supporting the language-specific selection of laryngeal features [voice] and [spread glottis]. The Multiple Feature Hypothesis correctly predicts differences between Dutch and German in children's error patterns in initial obstruents, as the result of neutralization to the unmarked value depends on the language-specific laryngeal feature: loss of [voice] for Dutch, and loss of [spread glottis] for German. However, we observed that the Articulatory Effort Hypothesis could also explain the asymmetry between Dutch and German, assuming that prevoicing and aspiration both pose articulatory challenges to the young child, which are avoided by devoicing (in Dutch) and de-aspiration (in German), respectively. For Dutch and German, we argued that the error patterns in acquisition cannot be fully explained on the basis of input frequency.

Focusing on evidence from the acquisition of English, we argued that an account based on the phonological feature [spread glottis] explains the observed asymmetry between voiceless obstruents and other consonants. In Seth's early productions, we found anticipatory devoicing, a case of laryngeal harmony, triggered by following voiceless obstruents but not by following voiced obstruents, nor by sonorants. This finding was interpreted as to support the Multiple Feature Hypothesis, i.e. languages with binary laryngeal contrasts differ in their 'active' laryngeal features, either [voice] or [spread glottis]. For English, a language which selects [spread glottis] as its active laryngeal feature, this correctly predicts that only voiceless obstruents trigger harmony.

Finally, we argued that articulatory effort alone cannot account for observed effects of anticipatory devoicing because of its non-local nature and the abstractness of the specification involved. Anticipatory devoicing arguably involves an abstract level of featural organization, that of contrastive specification. We presented additional evidence to support the hypothesis that the level of representation that is relevant for laryngeal harmony is lexical representation: its non-locality, its sensitivity to lexical frequency, and its insensitivity to the presence of the triggering segment in the output.

Perhaps the main interest of harmony patterns in children's productions resides in the possibility of testing the nature of lexical representations in early childhood. Even within a single language, laryngeal contrasts may be realized by rather different articulatory gestures (which correspond to different acoustic parameters) in syllable onset and coda. Evidence from children's early productions for laryngeal harmony between coda and onset, two positions which differ in the articulatory implementations of the laryngeal specification, suggests that young children (starting round the age 1;6) already construct phonological representations that abstract away from the phonetic detail which differentiates specific positions.

### ***Acknowledgements***

We gratefully acknowledge comments on earlier versions by Joe Stemberger and three anonymous reviewers, as well as the suggestions and remarks from audiences at the Old World Conference in Phonology I, Leiden University, January 2003, the DGfS Meeting at Universität München, February 2003, the workshop Voicing in Dutch, Leiden University, June 2003, and the Child Phonology Conference, University of British Columbia, July 2003. Any errors of fact and interpretation are, of course, solely our responsibility. We wish to thank Ann Peters and Brian MacWhinney for making the CHILDES database of Seth's utterances available to us in a digital format. Special thanks to Fieke Noordzij for assistance during the phonetic transcription of Seth's data.

### ***Notes***

1. An alternative to Iverson & Salmons' multiple laryngeal features is that of Avery & Idsardi (2001) who express the features in more phonetic terms in three dimensions: Glottal Width (for aspiration languages like English), Glottal Tension (for voicing languages, such as Dutch) and Larynx Height for languages which have ejectives or implosives in their segment inventories. In this paper, we use the terminology [voice] and [spread glottis] to express the phonological contrasts.
2. Here we focus on word-initial stops.
3. VOT interacts with place of articulation features (Lisker & Abramson 1964): dorsal stops have a longer VOT than both coronal and labial stops; in turn, coronal stops have a longer VOT than labial stops.
4. Other monovalent features may also be used to capture laryngeal contrasts cross-linguistically, such as [constricted glottis] or [stiff vocal cords]. Languages may have more than one feature; for example, Thai, which has a three-way laryngeal contrast, employs both features [voice] and [spread glottis].
5. One could also look at the frequency of voiceless vs. voiced stops collapsed across different prosodic positions, i.e. collapsed across word-initial, word-medial and word-final position (see Zamuner 2004). Based on these types of calculations, one finds that voiceless stops are overall more frequent than voiced stops. Input frequencies would then match the patterns of production seen in Dutch acquisition data.
6. Data in the Nijmegen database were collected and transcribed by Susan Powers, Jürgen Weisenborn, Wolfgang Klein, Heike Behrens, and Max Miller
7. The stronger trend toward voicing in tokens can be attributed to the fact that many prefixed words start with /bə/ or /gə/. A related issue, which we leave for future work, is how prosodic factors (mainly, stress) affect the salience of laryngeal contrasts in the input. For example, contrasts in onsets of stressed syllables may be more salient than those in unstressed syllables, such as prefixes.
8. As far as we know, no gestural accounts have been proposed which assign a single laryngeal gesture to onset and coda, thus spanning an entire syllable. Such accounts would necessarily be more abstract than the standard accounts, moving closer to a phonological representation.
9. Dutch and German have word-final neutralization, hence monosyllables cannot be used to test predictions on laryngeal harmony. Logically, laryngeal harmony might affect initial and medial consonants in polysyllables in these languages, but unfortunately, relevant cases in the Dutch and German databases were too rare to base any conclusions on.
10. Analysis is necessarily based on types, since Smith (1973) does not present token counts.
11. Amahl's mother spoke English as her fourth language, after Hindi, Bengali and Marathi. According to Smith (1973:7-8), her speech was characterized by 'fuller voicing of voiced obstruents'.

12. Stop-initial function words rarely occurred during the early stages of Seth's development. Moreover, targets for function words were difficult to establish.
13. If a single type occurred as both voiced and voiceless, it was coded as unfaithful. For example, bark was coded as [P...] because it occurred with both voiced and voiceless initial stops.
14. A third question, which we will briefly address at the end of this section, concerns item-specificity: to what extent is the devoicing effect restricted to particular target words?
15. Of course, we predict that if right-to-left (perseverative) laryngeal harmony were to be found, then it should be of the devoicing type, rather than of the voicing type, since English employs [spread glottis] as its active feature. However, coda voicing effects were generally too rare in Seth's productions for this prediction to be testable.
16. Other realizations, such as those involving onset deletion, were left out of consideration.
17. Some /B...P/ items were found in which initial devoicing occurred even though the final obstruent was left deleted, for example, [pli] 'blink' and [kiy] 'geese' [1;7]. Although such cases may seem to constitute strong evidence for the relevance of lexical representation, their relevance is somewhat undermined by the observation that deletion in /B...B/ items also increased chances of devoicing in onset. Both cases can be explained by the general unfaithfulness factor discussed above.
18. Thanks to Joe Stemberger for suggesting this to us.

### References

- Alphen, Petra M. van. This volume. "Prevoicing in Dutch Initial Plosives: Production, Perception, and Word Recognition".
- Avery, Peter. 1996. *The Representation of Voicing Contrasts*. Ph.D. dissertation, University of Toronto.
- Avery, Peter & William J. Idsardi. 2001. "Laryngeal Dimensions, Completion and Enhancement". *Distinctive Feature Theory* ed. by Tracy A. Hall, 41-69. Berlin: Mouton de Gruyter.
- Beers, Mieke. 1995. *The Phonology of Normally Developing and Language-Impaired Children*. Ph.D. dissertation, University of Amsterdam.
- Braunschweiler, Norbert. 1997. "Integrated Cues of Voicing and Vowel Length in German: A Production Study". *Language and Speech* 40.353-376.
- Cho, Taehong & Peter Ladefoged. 1999. "Variations and Universals in VOT: Evidence from 18 Languages". *Journal of Phonetics* 27.207-229.
- Cho, Young-Mee Yu. 1990. *Parameters of Consonantal Assimilation*. Ph.D. dissertation, Stanford University (revised version published in 1999 as *Lincom Studies in Theoretical Linguistics* 15, Munich).
- Clark, Eve V. & Melissa Bowerman. 1986. "On the Acquisition of Final Voiced stops". *The Fergusonian Impact. Volume 2* ed. by Joshua A. Fishman, Andrée Tabouret-Keller, Michael Clyne, Bhadriraju Krishnamurti & Mohamed H. Abdulaziz, 51-68. Berlin: Mouton de Gruyter.
- Cruttenden, Alan. 2001. *Gimson's Pronunciation of English*. London: Arnold.
- Davis, Katharine. 1995. "Phonetic and Phonological Contrasts in the Acquisition of Voicing: Voice Onset Time Production in Hindi and English". *Journal of Child Language* 22.275-305.

- Ernestus, Mirjam. 2000. *Voice Assimilation and Segment Reduction in Casual Dutch. A Corpus-Based Study of the Phonology-Phonetics Interface*. Ph.D. dissertation, Vrije Universiteit Amsterdam.
- Feest, Suzanne V.H. van der. 2004. "Acquiring Voicing in Dutch: The Role of Function Words". *Linguistics in the Netherlands 2004* ed. by Leonie Cornips & Jenny Doetjes, 47-57. Amsterdam & Philadelphia: John Benjamins.
- Feest, Suzanne V.H. van der. 2007. *Building a Phonological Lexicon. The Acquisition of the Dutch Voicing Contrast in Perception and Production*. Ph.D. dissertation, Radboud University Nijmegen.
- Ferguson, Charles A. 1978. "Fricatives in Child Language Acquisition". *Papers on Linguistics and Child Language* ed. by Vladimir Honsa & Martha J. Hardman-Bautista, 93-115. The Hague: Mouton.
- Fey, Marc E. & Jack Gandour. 1982. "Rule Discovery in Phonological Acquisition". *Journal of Child Language* 9.71-81.
- Fikkert, Paula. 1994. *On the Acquisition of Prosodic Structure*. Ph.D. dissertation, Leiden University.
- Fikkert, Paula & Clara C. Levelt. 2002. "Putting Place into Place". Paper presented at the GLOW Workshop on Phonological Acquisition. Utrecht, April 12-13, 2002.
- Fromkin, Victoria A., ed. 1973. *Speech Errors as Linguistic Evidence*. The Hague: Mouton.
- Grijzenhout, Janet & Sandra Joppen-Hellwig. 2002. "The Lack of Onsets in German Child Phonology". *The Process of Language Acquisition* ed. by Ingeborg Lasser, 319-339. Berlin/Frankfurt am Main: Peter Lang.
- Hansson, Gunnar Ó. 2001. *Theoretical and Typological Issues in Consonant Harmony*. Ph.D. dissertation, UC Berkeley.
- Iverson, Gregory K. & Joe C. Salmons. 1995. "Aspiration and Laryngeal Representation in Germanic". *Phonology* 12.369-96.
- Iverson, Gregory K. & Joe C. Salmons. 2003. "Legacy Specification in the Laryngeal Phonology of Dutch". *Journal of Germanic Linguistics* 15.1-26.
- Jakobson, Roman. 1941. *Kindersprache, Aphasie und Allgemeine Lautgesetze*. The Hague: Mouton. Translated by Allan R. Keiler (1968) as *Child Language, Aphasia and Phonological Universals*.
- Jessen, Michael. 1989. "Laryngeale Mechanismen in der Phonetik und Phonologie des Deutschen. Auslautverhärtung, Simmassimilation und Vokaleinsätze aus universal-grammatischer Sicht". Ms., University of Bielefeld [revised version of 1989 M.A. Thesis, University of Bielefeld].
- Jessen, Michael. 1996. *Phonetics and Phonology of the Tense and Lax Obstruents in German*. Ph.D. Dissertation, Cornell University.
- Jessen, Michael & Catherina O. Ringen. 2002. "Laryngeal Features in German". *Phonology* 19.189-218.
- Keating, Patricia A. 1984. "Phonetic and Phonological Representation of Stop Consonant Voicing". *Language* 60.286-319.
- Kewley-Port, Diane K. & Malcolm S. Preston. 1974. "Early Apical Stop Production". *Journal of Phonetics* 2.195-210.

- Kuijpers, Cecile T.L. 1993a. *Temporal Coordination in Speech Development. A Study on Voicing Contrast and Assimilation of Voice*. Ph.D. dissertation, University of Amsterdam.
- Kuijpers, Cecile T.L. 1993b. "The Voiced-Voiceless Distinction in Speech Development of Young Dutch Children". *Journal of Phonetics* 21.313-327.
- Ladefoged, Peter. 1973. "The Features of the Larynx". *Journal of Phonetics* 1.73-83.
- Levelt, Clara C. 1994. *On the Acquisition of Place*. Ph.D. dissertation, Leiden University.
- Levelt, Clara C., Niels O. Schiller & Willem J.M. Levelt. 2000. "The Acquisition of Syllable Types". *Language Acquisition* 8.237-264.
- Lisker, Leigh & Arthur S. Abramson. 1964. "A Cross-Language Study of Voicing in Initial Stops: Acoustical Measurements". *Word* 20.384-422.
- Lombardi, Linda. 1991. *Laryngeal Features and Laryngeal Neutralization*. Ph.D. dissertation, University of Massachusetts, Amherst.
- Lombardi, Linda. 1995. "Laryngeal Neutralization and Syllable Wellformedness". *Natural Language and Linguistic Theory* 13.39-74.
- Lombardi, Linda. 1996. "Restrictions on the Direction of Voicing Assimilation: An OT account". *University of Maryland Working Papers in Linguistics* 4.88-102.
- Lombardi, Linda. 1999. "Positional Faithfulness and Voicing Assimilation in Optimality Theory". *Natural Language and Linguistic Theory* 17.267-302.
- Lombardi, Linda. 2001. "Why Place and Voice are Different: Constraint-Specific Alternations in Optimality Theory". *Segmental Phonology in Optimality Theory – Constraints and Representations* ed. by Linda Lombardi, 13-45. Cambridge: Cambridge University Press.
- MacEachern, Margaret R. 1997. *Laryngeal Cooccurrence Restrictions*. Ph.D. dissertation, University of California, Los Angeles.
- Macken, Marlys A. & David Barton. 1980a. "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.
- Macken, Marlys A. & David Barton. 1980b. "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, Brian J. 1999. The CHILDES System. *Handbook of Child Language Acquisition* ed. by William C. Ritchie & Tej K. Bhatia, 457-494. New York: Academic Press.
- Menn, Lisa. 1971. "Phonotactic Rules in Beginning Speech". *Lingua* 26.225-241.
- Mester, R. Armin & Junko Ito. 1989. "Feature Predictability and Underspecification: Palatal Prosody in Japanese Mimetics". *Language* 65.258-294.
- Pater, Joe & Adam Werle. 2001. "Typology and Variation in Child Consonant harmony". *Proceedings of HILP5* ed. by Caroline Féry, Anthony Dubach Green & Ruben van de Vijver, 119-139. Potsdam: University of Potsdam.
- Pater, Joe & Adam Werle. 2003. "Direction of Assimilation in Child Consonant Harmony". *Canadian Journal of Linguistics* 48.385-408.

- Rooy, Bertus van & Daan Wissing. 2001. "Distinctive [Voice] Implies Regressive Voicing Assimilation". *Distinctive Feature Theory* ed. by Tracy A. Hall, 294-334. Berlin: Mouton de Gruyter.
- Rose, Sharon & Rachel Walker. 2001. "A Typology of Consonant Agreement as Correspondence". Ms., University of California, San Diego & University of Southern California.
- Salmons, Joe C. & Gregory K. Iverson. To appear. "Laryngeal Enhancement in Germanic". *Phonology*.
- Shattuck-Hufnagel, Stefanie. 1979. "Speech Errors as Evidence for a Serial-Ordering Mechanism in Sentence Production". *Sentence Processing: Psycholinguistic Studies Presented to Merrill Garrett* ed. by William E. Cooper & Edward C.T. Walker, 295-342. Hillsdale, New Jersey: Lawrence Erlbaum.
- Slis, Iman H. & Bert van Heugten. 1989. "Voiced-Voiceless Distinction in Dutch Fricatives". *Linguistics in the Netherlands 1989* ed. by Hans Bennis & Ans van Kemenade, 123-132. Foris: Dordrecht.
- Smith, Neil V. 1973. *The Acquisition of Phonology: A Case Study*. Cambridge: Cambridge University Press.
- Stemberger, Joseph P. 1991a. "Apparent Anti-Frequency Effects in Language Production: The Addition Bias and Phonological Underspecification". *Journal of Memory and Language* 30.161-185.
- Stemberger, Joseph P. 1991b. "Radical Underspecification in Language Production". *Phonology* 8.73-112.
- Steriade, Donca. 1995. "Underspecification and Markedness". *Handbook of Phonological Theory* ed. by John Goldsmith, 114-174. Cambridge, Mass.: Blackwell.
- Steriade, Donca. 1997. "Phonetics in Phonology: The Case of Laryngeal Neutralization". Ms., University of California, Los Angeles.
- Stoel-Gammon, Carol. 1985. "Phonetic Inventories, 15-24 Months: A Longitudinal Study". *Journal of Speech and Hearing Research* 28.505-12.
- Stoel-Gammon, Carol & Eugene H. Buder. 1999. "Vowel Length, Post-Vocalic Voicing and VOT in the Speech of Two-Year Olds". *Proceedings of the XIIIth International Conference of Phonetic Sciences* 3, 2485-2488.
- Van de Velde, Hans, Marinel Gerritsen & Roeland van Hout. 1996. "The Devoicing of Fricatives in Standard Dutch. A Real-Time Study Based on Radio Recordings". *Language Variation and Change* 8.149-175.
- Van de Velde, Hans & Roeland van Hout. 2001. "The Devoicing of Fricatives in a Reading Task". *Linguistics in the Netherlands 2001* ed. by Ton van der Wouden & Hans Broekhuis, 225-236. Amsterdam & Philadelphia: John Benjamins.
- Vaux, Bert. 1998. "The Laryngeal Specification of Fricatives". *Linguistic Inquiry* 29.497-512.
- Vihman, Marilyn M. & Charles A. Ferguson. 1987. "The Acquisition of Final Consonants". *Proceedings of the Eleventh International Congress of Phonetic Sciences* ed. by Ülle Vks, 381-384. Tallinn, Estonia.

- Weijer, Joost C. van de. 1998. *Language Input for Word Discovery*. Ph.D. dissertation, Max Planck Institute for Psycholinguistics, Nijmegen. Max Planck Series in Psycholinguistics 9.
- Wetzels, W. Leo & Joan Mascaró. 2001. "The Typology of Voicing and Devoicing". *Language* 77.207-244.
- Wilson, Bob & Ann M. Peters. 1988. "What are you Cookin' on a Hot?" *Language* 64.249-273.
- Zamuner, Tania S. 2003. *Input-Based Phonological Acquisition*. New York: Routledge.
- Zamuner, Tania S. 2004. "Acquiring Phonotactic Knowledge of Neutralization". Ms., Radboud University Nijmegen.
- Zamuner, Tania S., Louann A. Gerken & Michael Hammond. 2005. "The Acquisition of Phonology Based on Input: A Closer Look at the Relation of Cross-Linguistic and Child Language Data". *Lingua* 10.1403-1426.