BUCLD 35 Proceedings
To be published in 2011 by Cascadilla Press
Rights forms signed by all authors

The Roles of Phonotactics and Frequency in the Learning of Alternations

Gaja Jarosz

1. Introduction

Several recent studies have examined the acquisition of morpho-phonological alternations, with apparently conflicting findings. Production and comprehension studies with both real and nonce words suggest that 3½ year old Dutch-learning children do not yet have productive knowledge of the morpho-phonological voicing alternation (Zamuner, Kerkhoff, Fikkert, and Westrek 2005; Zamuner, Kerkhoff, and Fikkert 2006; Zamuner, Kerkhoff, and Fikkert 2007). Although five year olds perform well on real alternating words (Kerkhoff 2007), even seven year olds are reluctant to extend alternations to novel forms (Kerkhoff 2004). At the same time, in an artificial language learning task, White, Peperkamp, Kirk, and Morgan (2008) find evidence that one year olds are already learning morpho-phonological alternations.

This paper presents the results of computational simulations that may help explain these divergent findings. Given data representative of the Dutch voicing distribution, the computational model predicts learning curves consistent with the Dutch acquisition findings, with a long delay for alternating forms. The focus of this paper is on analyzing the behavior of the model to determine what properties of the input or assumptions about the learning process underlie the observed effects. The analyses reveal that two independent properties of the input distribution in Dutch conspire against the alternating segments in Dutch. A major focus of the analyses is on the interaction of prior phonotactic learning with the learning of alternations. A number of researchers have observed that phonotactic learning may be helpful in the learning of alternations (Hayes 2004; Prince and Tesar 2004; Jarosz 2006; Tesar and Prince 2003 / 2007). Despite the delay in the learning of voicing alternations in Dutch, the analyses suggest the observed effects are consistent with early phonotactic learning that aids in subsequent learning of alternations.

The rest of the paper is organized as follows. Section 2 presents the learning model. Section 3 reviews experimental findings on the acquisition of the Dutch

---

* Gaja Jarosz, Yale University, Department of Linguistics, gaja.jarosz@yale.edu. I would to thank Anne-Michelle Tessier and Tania Zamuner for particularly helpful discussion. For feedback on portions of this work, I am also grateful to the members of the CLAY (Computational Linguistics at Yale) research group and the participants at the 2009 Learning Meets Acquisition Workshop in Osnabruck, Germany.

voicing alternation and presents the results of the Dutch simulation. In Section 4, the predicted effects of the model are examined in three analyses of the model. Finally, Section 5 discusses the results of the analyses with respect to the acquisition findings.

2. The Learning Model

This section describes the learning model (Jarosz 2006), focusing on the assumptions and properties of the model most relevant to its predictions for the learning of phonological alternations and the interaction with phonotactic learning. See Jarosz (2006) for an in-depth description and technical details.

Within a generative framework, learning a morpho-phonological alternation involves the learning of at least four major interacting components of morpho-phonological knowledge. These include the learning of phonotactics, the learning of phonological mappings, the learning of underlying representations (URs), and the morphological analysis itself. For example, in order to learn the Dutch voicing alternation, the learner must determine that Dutch phonotactic knowledge includes a prohibition against word-final [d] but not [t] while allowing both [t]s and [d]s in other positions. The learner must determine the morphological relation between surface forms such as [bɛt] – 'bed' and [bɛdən] - 'beds' and decompose complex forms like [bɛdən] into their component morphs. The learner must also determine that the unfaithful mapping responsible for the alternations is devoicing word-finally. Finally, the learner must identify a single underlying representation for morphologically related forms, setting URs of alternating forms to /d/ and the URs of non-alternating forms to /t/.

All simulations presented here rely on a learning model (Jarosz 2006) cast within a probabilistic variant of Optimality Theory (OT; Prince and Smolensky 1993/2004). The model deals with the learning of three of these major subproblems: the phonotactics, the phonological mappings, and the URs. Like other work on learning in OT, the model makes the simplifying assumption that the morphological analysis is available to the learner. It is a probabilistic generative model with two components, a grammar and a lexicon of URs, both of which must be learned. Generation works just like in standard OT except that the production process is probabilistic. In order to generate a word, the learner first probabilistically selects a UR from the lexicon, then, given the UR, the learner uses the grammar to probabilistically map the UR to the surface form. The grammar is a probabilistic ranking of constraints (a probability distribution over rankings), and the lexicon encodes for each morpheme a probability distribution over possible underlying representations. The probability allows the model to represent uncertainty (such as at the beginning of learning or in cases of free variation), but in cases where surface forms are predictable, the lexicon and grammar can settle on unique URs and rankings, respectively.

The goal of learning is to reproduce the learning data and its distribution. During learning both the grammar and lexicon are updated gradually in response to the learning data, making detailed predictions about phonological learning over time. The learning of the grammar and lexicon is divided into two stages: phonotactic learning and the learning of phonological alternations. During the **phonotactic learning stage**, the task of the model is to learn a probabilistic ranking of constraints that reproduces the observable distribution of surface forms, to learn the (probabilistic) phonotactics of the language. This is done by gradually updating the (initially uninformed) grammar but holding constant the lexicon, which at this point is a flat distribution over potential URs for each unanalyzed word. Then, prior to the second stage, the (idealized) morphological analysis takes place, associating each morpheme with its own (initially uninformed) distribution of potential URs. By hypothesis, this point during learning corresponds to the onset of production. During the **phonological learning stage**, the learner gradually settles on the URs for each morpheme and also refines the grammar to capture the mappings from URs to surface forms. The examination of the phonological learning stage is the primary focus of this paper since this is the stage that makes predictions for production.[1]

3. The Dutch Simulation

This section describes the Dutch simulation, which is modeled after a picture-naming task with Dutch children (Zamuner *et al.* 2005) and (Zamuner *et al.* 2007). This task - involving production of real words whose frequencies are known - provides the opportunity to examine the effects of frequency on learning of alternations in the model and to compare these predictions to the experimental results.

3.1. Zamuner et al (2005, 2007)

Zamuner *et al.* (2005) and Zamuner *et al.* (2007) examined Dutch children's production of intervocalic [t]s and [d]s in a picture-naming task at two age groups, 2;7 and 3;7. The findings of the two studies were similar; the results of Zamuner *et al.* (2007: "ZKF") are summarized here. In order to examine the acquisition of alternations, ZKF compared production of stops in two morphological conditions. In the mono-morphemic condition, stops occurred intervocalically in morphologically simple nouns, e.g. [wɑtər] - 'water' and [rɪdər] – 'knight'. In the bi-morphemic condition, stops also occurred in intervocalic position but were the stem-final stops of plural nouns, e.g. [bɛdən] – 'beds' and [pɛtən] - 'caps'. Thus, the voiced stops in the bi-morphemic condition include the alternating /d/s. They also tested production of voicing word-finally in the corresponding singulars, e.g. [bɛt] – 'bed' and [pɛt] - 'cap'.

[1] The implementation of the model used here makes the simplifying assumption that the grammar and lexicon are represented as lists of URs and rankings, respectively, with associated probabilities. The algorithm used to gradually update these representations during both stages of learning is Expectation-Maximization (Dempster, Laird, and Ruben, 1977), adapted appropriately; see Jarosz (2006) for further details and discussion.

ZKF found that voicing was always produced correctly for the singulars, with both age groups only producing voiceless word-final stops. The accuracy of voicing in intervocalic position was more varied. Their results showing the average proportion of correct medial voicing in each condition are summarized in Table 1. In addition to a significant effect of Age, ZKF found significant effects of both morphology and voicing. Overall production of [t]s was significantly more accurate than production of [d]s (voicing effect), a result that is consistent with several other studies on the acquisition of voicing in Dutch (Kager, van der Feest, Fikkert, Kerkhoff, and Zamuner 2007; Kerkhoff 2007; van der Feest 2007). Also, production was overall significantly more accurate in the mono-morphemic condition than in the bi-morphemic condition (morphology effect). Thus, accuracy was lower for alternating /d/s than for non-alternating /d/s, and accuracy on bi-morphemic /t/s was lower than for mono-morphemic /t/s.

Table 1 - Proportion of Correct Voicing Intervocalically (from ZKF)

	Bi-morphemic		Mono-morphemic	
	[t]	[d]	[t]	[d]
Age 2;7	83%	5%	100%	28%
Age 3;7	97%	13%	99%	75%

Sections 3.2 - 3.3 show that these effects of morphology, voicing, and age are predicted by the probabilistic, generative model. Section 4 then examines the underlying causes of these effects in the model, revealing that the observed effects are a consequence of the Dutch distribution and the interaction of phonotactic learning with the learning of the voicing alternations.

3.2. Simulation Method

Table 2 summarizes the six types of items and their frequencies as provided to the learning model with examples (from ZKF) of each of the six types from Dutch. The simulation is based on ZKF's picture-naming task and uses the same six types of items: singular nouns with /t/s and /d/s word-finally, corresponding plurals with intervocalic /t/s and /d/s, and mono-morphemic nouns with intervocalic /t/s and /d/s. The frequencies of the six item types used for the simulation are the total token frequencies of the lexical items from the picture-naming task in child-directed speech as reported by ZKF. Importantly, however, the frequencies used in the simulation are consistent with the distribution of voicing in child-directed speech in Dutch overall. Of particular relevance to the simulation are that facts that plural /d/s are very infrequent, that there are more medial /t/s than medial /d/s overall, and that there are more singulars than plurals (Kerkhoff 2004; Kerkhoff 2007).

The actual data provided to the learning model are shown in the last two (shaded) lines of Table 2. In order to ensure that only the relevant factors of voicing, morphology, frequency, and phonological context (word-final vs.

medial) affected learning of stop voicing in the different types, all data items were of the form rɛ{t,d}(ə). The subscripts 1-5 are arbitrary labels for the 5 morphemes in the input data: the four stems and one suffix. The subscripts encode the morphological analysis of the surface forms, telling the model which words are morphologically related to one another and which segments correspond to which abstract morphemes.

Table 2 – The Learning Data and Frequency of Stops by Type (from ZKF)

Singulars		Bi-morphemic Plurals		Mono-morphemic	
/t/→[t]	/d/→[t]	/t/→[t]	/d/→[d]	/t/	/d/
[pɛt] -'cap'	[bɛt] -'bed'	[pɛtən] -'caps'	[bɛdən] -'beds'	[wɑtər] -'water'	[ridər] -'knight'
11.8%	55.5%	8.6%	0.5%	19.3%	4.4%
$rɛt_1$	$rɛt_2$	$rɛt_1ə_5$	$rɛd_2ə_5$	$rɛtə_3$	$rɛdə_4$

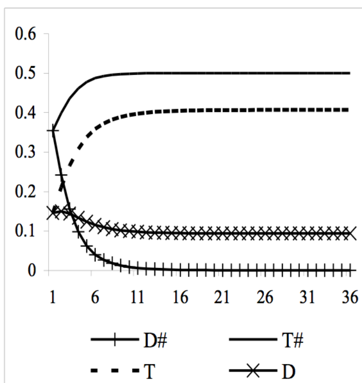
The standard OT constraints used in the simulation are defined below in (1). This set of constraints makes available a range of analyses of the input data that the learning model must successfully navigate. For example, given these constraints, both intervocalic voicing and final devoicing are possible explanations of voicing alternations. Also, the presence of MAX in the constraint set makes it possible to avoid voicing word-finally by deleting rather than devoicing. Crucially, an appropriate ranking of these constraints will capture the Dutch pattern of voicing contrast intervocalically and devoicing word-finally.

(1) Constraints
  a. *VOICE – No voiced obstruents
  b. *VOICECODA – No voiced obstruents in syllable coda position
  c. *VTV – No voiceless obstruents intervocalically
  d. IDENT[VOICE] – No change in voicing from input to output
  e. MAX – No deletion

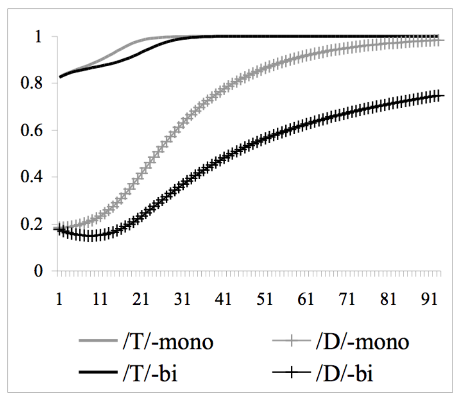
Given the learning data and frequency distribution shown in the last two lines of Table 2 and the constraints in (1), the learning model for this simulation must ultimately find the URs for each of the 5 morphemes and a ranking of constraints that accounts for phonotactic restrictions and the phonological mappings. During phonotactic learning the model ignores the morphological analysis (the subscripts in the last line of Table 2). It starts with an uninformed probabilistic ranking of constraints, with all constraints ranked equally such that any total ranking is equally likely. Its task during phonotactic learning is to find a probabilistic ranking that reproduces the distribution of surface forms, including word-final voicing neutralization with contrast intervocalically. Then, prior to the second stage of learning, morphological analysis takes place, and each of the 5 morphemes is associated with its own probabilistic UR. The UR distributions are initially uninformed with /t/ and /d/ being equally likely for each morpheme. During phonological learning, the model adjusts the UR distributions gradually as it settles on a UR for each of the morphemes: its task is to determine for example that the URs of morphemes 1 and 2 must have a /t/ and /d/, respectively. The model also gradually adjusts the probabilistic ranking learned during the first stage to ensure URs map correctly to their surface forms.

3.3. Simulation Results

The model is successful, and the learning curves for both stages are shown in Figure 1. Figure 1(A) shows how the relative probabilities of the surface forms change gradually during the phonotactic learning stage. The graph shows that the model learns that forms with word-final voicing (D#) are illegal in Dutch as the probability of such forms drops quickly to 0. The model also learns that the other surface forms are possible in Dutch, and their relative probabilities at the end of phonotactic learning reflect the Dutch distribution, with word-final [t]s (T#) being most probable, followed by medial [t]s (T), and finally, medial [d]s (D).



(A) Phonotactic Learning:
Probability of Surface Forms



(B) Phonological Learning:
Probability of Correct Medial Voicing

Figure 1 – Predicted Learning Curves for Dutch

It is with this phonotactic knowledge that the model enters the phonological stage of learning. The graph in Figure 1(B) shows how the probability of correctly producing voicing changes over the course of learning for the four data items with medial stops. Since production relies on first probabilistically selecting a UR from the lexicon and then probabilistically mapping that UR to the surface form using the grammar, these production probabilities reflect the interaction of grammatical knowledge and knowledge of the URs as both are gradually updated over time. The curves for the singulars (rɛt_1 and rɛt_2) are not shown in the graph: the probability of correct voicing for these is already at 100% at the beginning of phonological learning due to the phonotactic restriction against final-voicing. Examination of the curves for the medial stops reveals that learning is fastest for mono-morphemic /t/ (/T/-mono), followed closely by bi-morphemic /t/ (/T/-bi). Correct voicing for /d/ takes much longer, with faster learning for mono-morphemic /d/ (/D/-mono) than for bi-morphemic /d/ (/D/-bi).

The learning curves in Figure 1(B), predicted by a probabilistic, generative model with early phonotactic learning, closely mirror the acquisition findings. Given data representative of the Dutch voicing distribution, the model predicted the effects of age, morphology, and voicing, as well as the lack of voicing word-finally for singulars, matching the experimental results discussed above (Zamuner *et al.* 2007). The age effect follows trivially from the gradual (and improving) nature of learning. More interesting are the effects of morphology (an overall delay for bi-morphemic forms) and voicing (an overall delay for voiced [d]). The correspondence between the predicted learning curves and the experimental results is striking: the predicted curves at (approximately) iterations 15 and 30 closely match the younger and older age groups, respectively. A particularly noteworthy finding is the difference between mono-morphemic /t/s and bi-morphemic /t/s in both the model and the experimental results. Since both /t/s are always realized as [t]s on the surface, what explains the difference? The following section explores the causes underlying these and other effects in the model.

4. Model Analysis

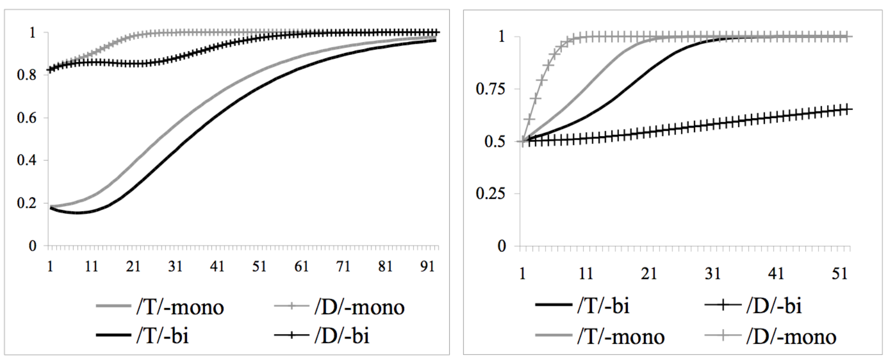
By manipulating the computational simulation in various ways, it is possible to determine what properties of the model or aspects of the Dutch distribution underlie these various effects. Understanding what underlies the observed effects in the model can provide a possible explanation for the delayed learning of voicing alternations in Dutch. This section presents analyses of the above simulation results, exploring in turn the voicing effect, the morphological effect, and the relationship between learning of phonotactics and alternations.

4.1. The Voicing Effect

The model predicts slower learning for voiced /d/ than for voiceless /t/ overall: indeed, both mono-morphemic and bi-morphemic medial /t/s are learned more quickly by the model than the mono-morphemic and bi-morphemic medial /d/s. As work on acquisition of voicing in Dutch has noted, one possible explanation for the slower acquisition of [d] may be the low frequency of voiced consonants in Dutch (Kager *et al.* 2007; Kerkhoff 2007; van der Feest 2007). The input distribution (Table 2) provided to the model reflects this statistical bias: medial /t/s have an overall frequency of 27.9% (19.3% + 8.6%) while medial /d/s have an overall frequency of 4.9% (4.4% + 0.5%). In the computational model, it is possible to examine the effect of this statistical bias directly by altering the input distribution. Figure 2(A) shows the learning curves during phonological learning for ‘Anti-Dutch’, which is identical to the one for

Dutch in Figure 1(B) above except that the frequencies of medial /t/s and /d/s are reversed. Specifically, in Anti-Dutch the frequencies of [rɛt$_1$ə$_5$], [rɛtə$_3$], [rɛd$_2$ə$_5$], and [rɛdə$_4$], are 0.5%, 4.4%, 8.6%, and 19.3%, respectively. This means the Anti-Dutch input distribution has the same overall proportions of mono-morphemic, bi-morphemic, and singular forms as the Dutch distribution, but medial /d/s (27.9%) are more frequent than medial /t/s (4.9%) overall.

In contrast to the simulation for Dutch, the curves for Anti-Dutch in Figure 2(A) show slower learning for voiceless (/T/s) than for voiced (/D/s). This effect obtains even though the overall frequency of [t]s is higher than the frequency of [d]s if all positions are considered. This is because the effects of frequency are context-specific and grammatical in nature. Because both Dutch and Anti-Dutch prohibit voicing word-finally, in both simulations the model learns during phonotactic learning that *VOICECODA must be highly ranked. High ranking of *VOICECODA has no impact on medial voicing, however. So also during phonotactic learning, the model ranks *VOICE highly in Dutch and *VTV highly in Anti-Dutch to account for opposite biases favoring medial [t]s in Dutch and medial [d]s in Anti-Dutch. These statistical phonotactic biases then shape the learning curves during phonological learning.



(A) Probability of Correct Medial Voicing for Anti-Dutch

(B) UR Learning in Dutch

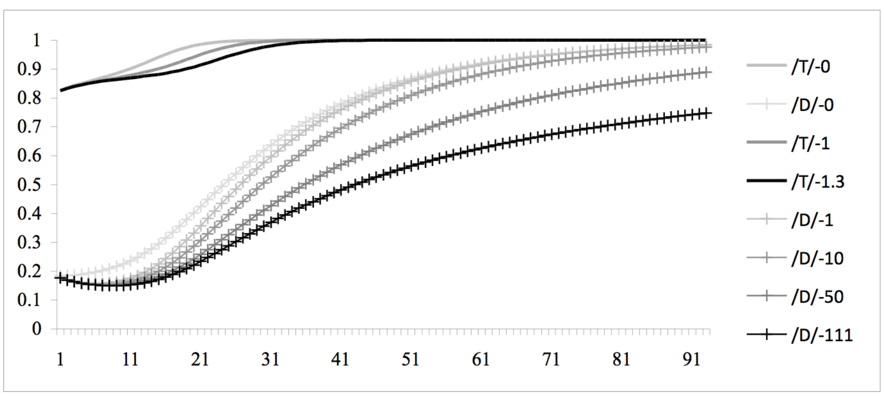
Figure 2 – Learning Curves for Anti-Dutch (A) and Dutch URs (B)

In sum, this manipulation reveals that the model's delayed learning of [d]s can be attributed to the high frequency of medial [t]s in the Dutch distribution, resulting in a probabilistic grammatical preference for medial [t]s. The manipulation reveals this is a big part of the explanation for the slow learning of alternating /d/s. In Anti-Dutch, which has the opposite frequency bias, alternating /d/s do not show a particularly slow learning curve. Importantly, however, the morphological effect in Anti-Dutch remains the same as in the original Dutch simulation, with slower learning for bi-morphemic stops than for corresponding mono-morphemic stops. Thus, the morphological effect must be due to some other property of the model or input; this is addressed next.

4.2. The Morphological Effect

While the voicing effect can be attributed to a global grammatical preference, the explanation of the morphological effect can only be attributed to a difference in the learning of URs for bi-morphemic forms as opposed to mono-morphemic forms. This is because the constraints used in the simulation make no reference to morphological status in any way. The grammar, therefore, cannot distinguish morphologically complex forms from morphologically simple forms, and any morphological difference can only be attributed to the URs. This fact can be confirmed by examining the learning curves for the URs themselves for the original Dutch simulation, shown in Figure 2(B). The figure reveals that the mono-morphemic forms are quicker to converge on their target URs than the bi-morphemic forms. Specifically, it takes the model longer to settle on the underlying /t/ for [rɛt_1ə$_5$] than for [rɛtə$_3$], and longer to settle on underlying /d/ for [rɛd_2ə$_5$] than for [rɛdə$_4$].

Why does the model predict slower UR learning for bi-morphemic stops? The slower learning of underlying /t/ for [rɛt_1ə$_5$] than for [rɛtə$_3$] is particularly illuminating. Both /t/s always surface as [t], and the overall frequency of morphemes 1 and 3 are comparable, with both occurring about 20% of the time in the data. The only difference is that some of the occurrences of morpheme 1 are in the singular, placing the stem-final /t/ in word-final position. In contrast, all occurrences of the /t/ in morpheme 3 occur medially. This suggests learning of underlying voicing in word-final position is slower than in medial position.



Phonological Learning: Probability of Correct Voicing
Figure 3 – Dutch Learning Curves by Proportion of Word-Final Context

The simulation results pictured in Figure 3, showing learning curves for different rates of occurrence in word-final context, confirm this. This simulation is a variant of the original Dutch simulation with several distinct lexical items making up the bi-morphemic /t/ and /d/ types. The frequencies of the lexical items were chosen in such a way as to maintain the same overall frequencies for the six types as in the original simulation, but the various lexical items within each type varied in the relative frequency of occurrence in word-final context. Specifically, the overall frequency of plural /d/ and singular /d/ were still 0.5% and 55.5%, respectively. However, one of the lexical items (stems) with target /d/ (labeled /D/-10 in the figure) occurred 10 times more frequently in the singular, word-final context than in the plural, medial context. Items /T/-0 and /D/-0 never occur in word-final context and thus correspond to the mono-morphemic /t/ and /d/, respectively. /T/-1.3 and /D/-111 are the bi-morphemic curves from the original Dutch simulation since 1.3 and 111, respectively, are the relative proportions in the original data. The results show that varying the rate of word-final context derives a continuum of learning curves, with slower learning for higher rates of word-final context.

This manipulation shows that the different rates of occurrence in word-final context are responsible for different rates of UR learning in mono-morphemic versus bi-morphemic forms. It is not an accident that the phonological context in which UR learning is slower is one in which voicing contrast is neutralized in Dutch, while the context in which UR learning is faster is one that permits a voicing contrast. It is precisely because of the model's existing phonotactic knowledge that learning of URs is slower in a neutralizing context. This effect of phonotactic knowledge is examined directly next.

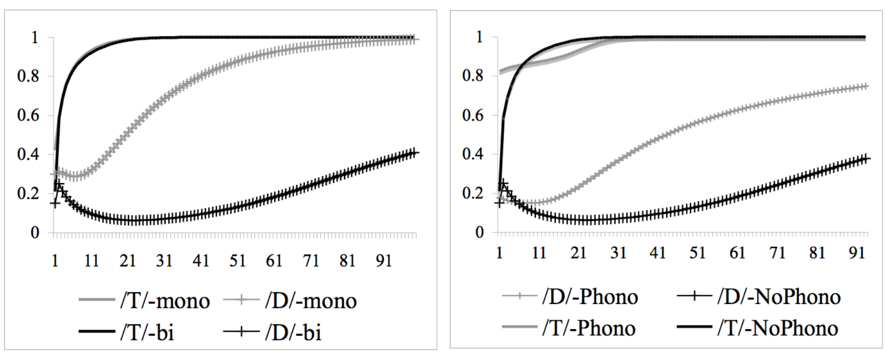
4.3. The Effect of Early Phonotactic Learning

The effect of existing phonotactic knowledge can be examined directly in the model by examining the impact of the early stage of phonotactic learning. An early stage of phonotactic learning is built-in to this implementation of the model, but its impact can be minimized by allowing phonotactic learning to continue for only one iteration[2]. The results of just such a simulation, which is otherwise identical to the original Dutch simulation, are shown in Figure 4(A). Reducing the impact of early phonotactic learning has a number of noteworthy effects. Although all the curves for medial stops are plotted in this graph, it is impossible to distinguish the curves for mono-morphemic and bi-morphemic /t/s because they are right on top of one another. The curves for the /d/s are lower, with bi-morphemic /d/s learned most slowly. Thus, although bi-morphemic /d/s are still learned most slowly, there is no longer a difference predicted between mono-morphemic /t/s and bi-morphemic /t/s. The morphological effect, which applied to both /t/s and /d/s in the original simulation, is no longer observed. In general, as compared to the original simulation, the learning curves in Figure 4(A) are consistent with a phonotactically 'naïve' learner that learns phonological alternations primarily on the basis of the surface realizations, penalizing /d/s for their low frequency and additionally penalizing bi-morphemic /d/s for their unfaithful surface realizations in the singular.

---

[2] One iteration corresponds to one pass through the data and a small grammatical update as shown in Figure 1(A).

This overall pattern of results can be understood in terms of the consequences of phonotactic knowledge of neutralization for the learning of underlying representations. Specifically, the word-final context is less informative about underlying voicing to a learner with knowledge of word-final voicing neutralization than to a phonotactically naïve learner. Knowledge of voicing neutralization tells the learner that both /d/ and /t/ are possible underlying representations for a word-final [t]. Therefore, occurrences of stops word-finally are not informative about underlying voicing to a phonotactically informed learner. In contrast, a phonotactically naïve learner will be more likely to take the surface realization of voicing word-finally as evidence for an underlying voiceless specification. As a result, the phonotactically naïve learner, as compared to the phonotactically informed learner, should learn bi-morphemic /t/s *more* quickly and bi-morphemic /d/s *less* quickly. This is exactly what happens. Figure 4(B) compares the learning curves for bi-morphemic /t/s and /d/s with early phonotactic learning (from the original simulation in Figure 1(B)) to the learning curves for bi-morphemic /t/s and /d/s with minimal early phonotactic learning from Figure 4(A). As this figure shows, early phonotactic learning dramatically *improves* learning of alternating /d/ and results in somewhat *slower* learning of bi-morphemic /t/.



(A) Phonological Learning without Prior Phonotactic Learning

(B) Phonological Learning of Bi-Morphemic Stops with and without Prior Phonotactic Learning

Figure 4 – The Effect of Early Phonotactic Learning

In sum, the model's early phonotactic learning enables the learner to differentiate phonological contexts according to how informative they are about underlying voicing. Surface realizations in neutralizing contexts are not informative about underlying voicing whereas surface realizations in contrastive contexts are. Accordingly, a phonotactically informed learner is less likely to be mislead by the surface voicelessness of alternating /d/s word-finally. In this way, phonotactic knowledge *helps* with the learning of alternating /d/s. In contrast, since surface realizations of /t/s word-finally match the target underlying specification, ignoring these realizations *hurts* the learning of bi-morphemic /t/s.

5. Discussion

This paper has shown that a probabilistic, generative model given input data representative of the Dutch voicing distribution predicts a delay in the learning of voicing alternations. The analyses of the model show that the delay for alternating /d/s results from two cumulative effects, both of which derive from the frequency distribution of Dutch. The voicing effect is a grammatical effect due to the statistical bias favoring medial [t]s to medial [d]s in the Dutch data. The morphological effect results from the interaction of input frequency, specifically the frequency of stem-final stops word-finally, with early phonotactic learning. Early phonotactic learning causes the model to treat word-final context as uninformative about underlying voicing, and the high relative frequency of alternating /d/s in word-final context delays their subsequent learning. In spite of the delays in learning of alternating /d/s, computational simulations show that early phonotactic learning helps the learning of alternations, with even slower learning expected in the absence of early phonotactic learning.

Understanding the behavior of the model can help explain the delay in the learning of voicing alternations in Dutch. The analyses show that even with phonotactic knowledge, the learning of alternating /d/s takes a long time from the onset of phonological learning. This delay occurs even though the model assumes morphological analysis is complete and treats the alternation as a regular, fully productive process. The implication of the model is that Dutch children's delay in the learning of voicing alternations may likewise be due, at least in part, to the frequency distribution in Dutch. If, as recent experimental results from artificial grammar learning suggest (White *et al.* 2008), learning of alternations is already underway by the time children are one year old, the simulations presented here help explain why this process takes so long in Dutch. The analyses highlight in particular the potential role of the low frequency of alternating /d/s in medial context in Dutch. If the model is on the right track, voicing alternations in nouns that occur more frequently in the plural should be acquired earlier than alternations in nouns with low frequency of occurrence in the plural. This is a concrete prediction derived from the model that can be tested in future acquisition studies. More generally, the model predicts that acquisition of phonotactically driven alternations should be slower in neutralizing contexts than in contrastive contexts. Another prediction of the model is that statistical biases in the lexicon favoring one of the alternating segments can interact with the learning of alternations. A statistical bias against voiced medial segments in Dutch leads to an overall preference for voiceless segments in the model. However, Section 4.1 reveals that learning of alternations should be faster in the absence of such a bias. These predictions can be checked against findings from languages with different voicing distributions.

As discussed in Section 2, this model does not learn to morphologically analyze the phonological surface forms, a process that is likely to interact with the subproblems of phonological learning examined here. Further modeling work is needed to determine how morphological analysis interacts with the learning of morpho-phonological alternations and in what ways it depends on the input distribution. Nonetheless, the simulations presented here illustrate that the interactions of phonotactic learning, learning of phonological mappings, and the learning of URs already create a complex system with predictions that can be hard to foresee. Computational modeling generates these predictions, helping to explain developmental findings and develop hypotheses for further testing.

References

Hayes, Bruce. 2004. Phonological Acquisition in Optimality Theory: The Early Stages. In *Fixing Priorities: Constraints in Phonological Acquisition*, ed. René Kager, Joe Pater, and Wim Zonneveld, 245-91. Cambridge University Press.

Jarosz, Gaja. 2006. Rich Lexicons and Restrictive Grammars - Maximum Likelihood Learning in Optimality Theory. Ph.D. Dissertation, The Johns Hopkins University, Baltimore, MD.

Kager, René, Suzanne V.H. van der Feest, Paula Fikkert, Annemarie Kerkhoff, and Tania S. Zamuner. 2007. Representations of [Voice]: Evidence From Acquisition. *Amsterdam Studies in the Theory and History of Linguistics Science Series 4* 286:41.

Kerkhoff, Annemarie. 2007. Acquisition of Morpho-Phonology: The Dutch Voicing Alternation. Ph.D. Dissertation, Utrecht: LOT. University of Utrecht.

Kerkhoff, Annemarie. 2004. Acquisition of Voicing Alternations. In *Proceedings of GALA*, ed. S. Baauw and J. van Kampen, 269-80. Utrecht: LOT.

Prince, Alan and Paul Smolensky. 1993/2004. *Optimality Theory : Constraint Interaction in Generative Grammar*. Malden, MA : Blackwell Pub. (Original work published 1993)

Prince, Alan and Bruce Tesar. 2004. Learning Phonotactic Distributions. In *Fixing Priorities: Constraints in Phonological Acquisition*, ed. Rene Kager, Joe Pater, and Wim Zonneveld, 245-91. Cambridge University Press.

Tesar, Bruce and Alan Prince. 2003/2007. Using Phonotactics to Learn Phonological Alternations. In *Proceedings of the Thirty-Ninth Conference of the Chicago Linguistics Society*, 209-37. Chicago: Chicago Linguistics Society. (Original work published 2003)

van der Feest, Suzanne V.H. 2007. Building a Phonological Lexicon: The Acquisition of the Dutch Voicing Contrast in Perception and Production. Ph.D. Dissertation, Radboud University Nijmegen.

White, Katherine S., Sharon Peperkamp, Cecilia Kirk, and James L Morgan. 2008. Rapid Acquisition of Phonological Alternations by Infants. *Cognition* 107:238 - 265.

Zamuner, Tania, Annemarie Kerkhoff, and Paula Fikkert. 2006. Acquisition of Voicing Neutralization and Alternations in Dutch. In *Proceedings of the 30th Annual Boston University Conference on Language Development. Somerville, MA: Cascadilla Press*.

Zamuner, Tania, Annemarie Kerkhoff, and Paula Fikkert. 2007. Children's Knowledge of How Phonotactics and Morphology Interact*, Ms.* Radboud University Nijmegen and Utrecht University, The Netherlands.

Zamuner, Tania, Annemarie Kerkhoff, Paula Fikkert, and Ellen Westrek. 2005. Dutch Children's Acquisition of Morpho-Phonological Alternations in Plural Formation. *IASCL* [Poster]. Berlin, Germany.