Stages of Acquisition without Ranking Biases: the Roles of Frequency and Markedness in Phonological Learning*

Gaja Jarosz

Yale University

1. Introduction

A growing body of literature within Optimality Theory (OT; Prince and Smolensky 1993/2004) focuses on the potential for formal computational learning algorithms to model aspects of the human acquisition process, including order of acquisition effects and intermediate stages (Boersma and Levelt 2000; Tessier 2006; Jesney and Tessier, this volume). This work characterizes learning as a gradual transition between an initial constraint ranking and the target ranking. In this and other work on acquisition and learnability within OT, initial (or persistent) explicit ranking biases are employed to structure the learner's path (Tesar 1995; Smolensky 1996; McCarthy 1998; Smith 2000; Gnanadesikan 1995/2004; Boersma and Hayes 2001; Hayes 2004; Prince and Tesar 2004; Tesar and Prince to appear). These ranking biases are used to establish an initial state from which subsequent learning of the ranking proceeds. This serves two principal functions: it aids in the identification of restrictive end-state grammars, and it enables modeling of the acquisition process as a transition from this initial state to the target grammar. In this paper, I apply MLG (Maximum Likelihood Learning of Lexicons and Grammars), the theory of phonological learning proposed in Jarosz (2006a, 2006b), to the problem of acquisition modeling. MLG models the learning of phonological grammars and lexicons of underlying forms relying on two general principals: richness of the base and likelihood maximization, but no explicit ranking biases are assumed. In related work (Jarosz to appear), I show that the first role of ranking biases is subsumed by MLG, which identifies restrictive grammar and lexicon combinations even in cases where ranking biases do not suffice. In this paper, I show that MLG also predicts order of acquisition effects and intermediate stages without relying on ranking biases.

This paper also investigates the relative roles of markedness and frequency in phonological acquisition from both theoretical and empirical perspectives. In particular,

* I would like to thank Joe Pater, Karen Jesney, and Kathryn Flack for helpful comments on this paper.

the paper discusses the roles of markedness and frequency in MLG and presents empirical evidence of the causal connection between frequency and order of acquisition. While it is generally accepted that both markedness and frequency play some role in child acquisition, the proper characterization of the relative roles of markedness and frequency is a topic of ongoing debate (Levelt 1994; Levelt and Van de Vijver 1998; Levelt et al. 2000; Demuth 2004; Stites et al. 2004; Kirk and Demuth 2005; Zamuner et al. 2005). In MLG (see also Boersma and Levelt 2000), universal markedness is embodied in the constraint set, which defines the space of possible grammars cross-linguistically[1]. Frequency plays a secondary role, selecting the language-particular learning path based on the relative frequencies of various configurations in the ambient language. This prediction of MLG is examined in a series of case studies focusing on the acquisition of onset and coda clusters in Dutch, English, and Polish. Corpus analysis reveals that these three languages have distinct relative proportions of onset and coda clusters in adult, child-directed speech. The predictions of MLG are discussed with respect to developmental orders observed in children acquiring these languages. The observed developmental orders, evidenced by previous work on Dutch and English and novel data from Polish presented here, are consistent with the predictions of MLG.

2. The Learning Theory

Maximum Likelihood Learning of Lexicons and Grammars (MLG) is a theory of phonological learning that accounts for the learning of phonological grammars and lexicons from unstructured overt phonological forms. This section presents an overview of the learning theory and discusses its treatment of markedness and frequency[2].

2.1. General Structure of the Model

MLG is a generative, probabilistic model of the acquisition of a phonological grammar and lexicon of underlying forms. As such, it relies on a formal, probabilistic characterization of both the grammar and lexicon: the grammar and lexicon are both probabilistic entities. The grammar is a probability distribution over rankings of OT constraints and assigns a conditional probability to possible surface realizations of a given underlying form[3]. The lexicon is probabilistic as well and associates each morpheme with a set of possible underlying forms, each with its own likelihood. These probabilistic components can express uncertainty (as in the initial stages of learning) or variation by spreading probability over multiple rankings or underlying forms, and they can express certainty (as in the final stages of learning) by assigning to a single ranking

[1] Although I assume the constraint set is universal throughout this paper, I leave open the question of whether it is entirely innate. It is possible that the constraint set itself, or portions of it, can be learned from universal, shared experience; see Hayes (1999) and Flack (2007; this volume) for proposals along these lines.

[2] For a more in-depth presentation of the structure and properties of MLG see Jarosz (2006b).

[3] Various probabilistic variants of OT proposed in previous work, such as Stochastic OT, Partial Order Grammars, and Floating Constraints, are all examples of such a probabilistic grammar (Boersma 1998; Anttila 1997; Reynolds 1994).

or underlying form a probability of one. Together, the grammar and lexicon assign a likelihood, or probability, to the overt forms of the language.

Learning in MLG relies on two general principles: *richness of the base* (ROTB) and *likelihood maximization*. According to ROTB, the set of possible underlying forms is universal: there are no systematic, language-particular restrictions on underlying representations, and therefore all language-specific restrictions must be handled by the grammar (Prince and Smolensk 1993/2004). MLG incorporates a probabilistic formulation of ROTB into the learning model. The second learning principle, likelihood maximization, defines the correct grammar and lexicon combination as the one that maximizes the likelihood, or probability, of the overt forms. In other words, likelihood maximization requires that the grammar and lexicon combination generate all and only the observed forms of the target language with high probability, a standard generative perspective cast in a probabilistic setting.

These general principles form the foundation of MLG and are incorporated into a learning model with two stages of learning, phonotactic and morphophonemic learning (see also Prince and Tesar 2004; Hayes 2004):

(1) Two Stage Learning in MLG
- a Phonotactic Learning
  - i A fixed, universal rich base is assumed
  - ii No morphological awareness
  - iii Grammar learning but no lexicon learning
- b Morphophonemic Learning
  - i Words are analyzed into component morphemes
  - ii Learning of morpheme specific underlying forms occurs
  - iii Further learning of the grammar to account for alternations

The phonotactic stage of learning occurs before morphological awareness and prior to the development of a phonological lexicon; it is during this stage that learning of a language-specific phonotactic grammar takes place. Formally, this stage consists of gradual learning of a grammar that maximizes the likelihood of the overt forms, given a (fixed) rich base. The rich base is a representation of all possible underlying forms, each with roughly equal likelihood. This base may be characterized as the expected, unbiased distribution over phonological forms given by the free combination of phonological elements. Under this characterization, phonotactic learning involves maximizing the likelihood of the observed distribution of overt forms, given the expected distribution.[4]

During morphophonemic learning, words are analyzed into component morphemes, and each morpheme is associated with its own probabilistic lexical entry. During this stage, the grammar gradually transitions between the phonotactic grammar

[4] A formally equivalent characterization of phonotactic learning in MLG involves an identical rich base associated with each morphologically unanalyzed overt form. Because the base is identical and unchanging, the overt forms are effectively generated from the same base.

learned in the first stage and the target grammar, while the lexicon gradually converges on the target lexicon. Formally, during morphophonemic learning the *grammar and lexicon combination* that maximizes the likelihood of the overt forms is gradually learned. Thus, the crucial difference between the two stages resides in the role of the rich base and lexical learning.

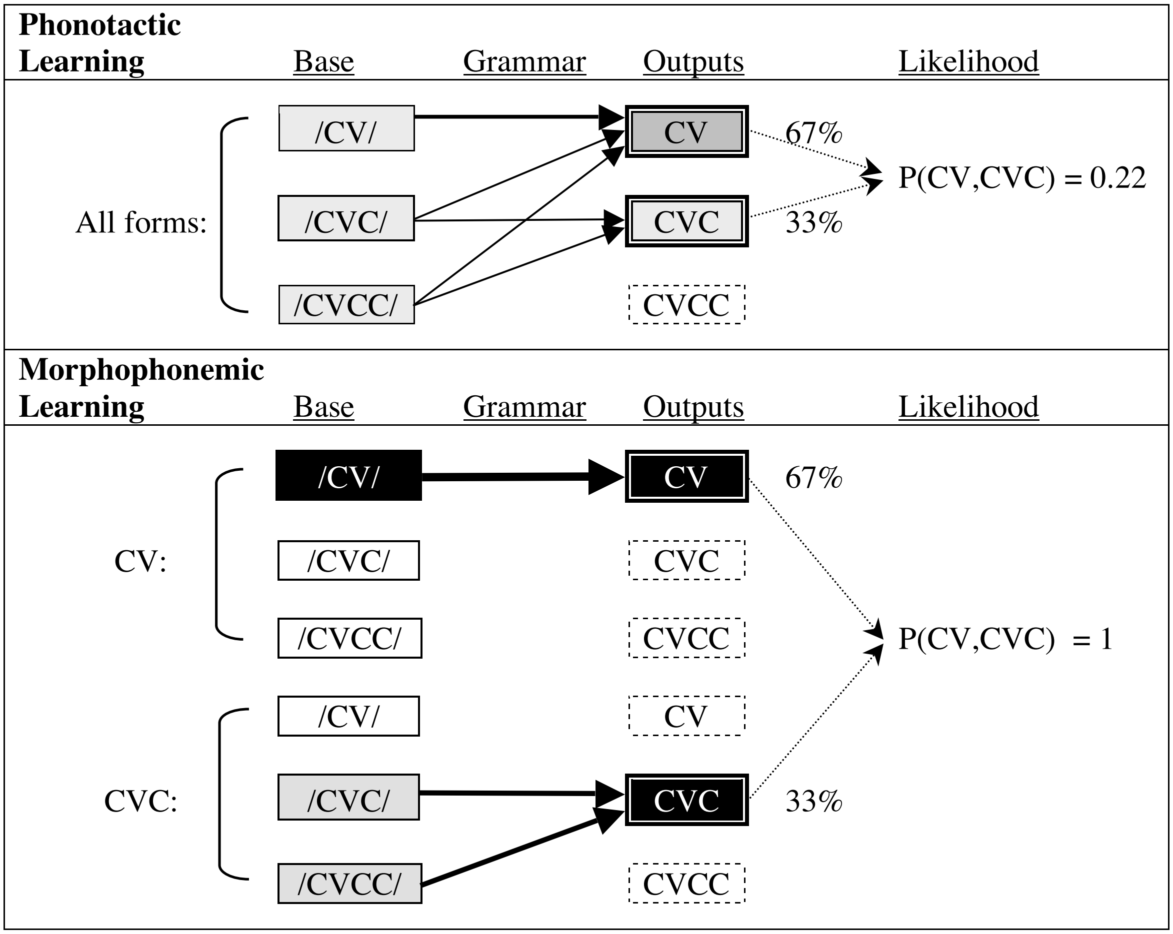


Figure 1: Outcomes of Phonotactic and Morphophonemic Learning in MLG

Figure 1 illustrates the outcomes of the two stages of learning with a simple example of a language with two overt forms, CV and CVC, the former occurring twice as often as the latter. During phonotactic learning, a single, equally-distributed rich base is held constant while the grammar that maximizes the likelihood of the overt forms is gradually learned. Phonotactic learning results in a restrictive grammar that matches the frequencies of the overt forms, given the rich base. This grammar is shown in the figure as a mapping from the rich base to the overt forms, with outputs generated in proportion to their frequency of occurrence (darker shading corresponds to higher likelihood). During morphophonemic learning, each morpheme (corresponding to each overt form in this case) is associated with its own probabilistic lexical entry. The lexical entries and grammar are gradually updated until they converge on the target lexicon and grammar shown in the lower portion of the figure. The target lexicon and grammar generate all and only the correct forms for each overt form, as indicated by the black shading of the correct output forms.

Although the division into phonotactic and morphophonemic learning in MLG follows previous work (Prince and Tesar 2004; Hayes 2004), the outcome and properties of phonotactic learning in particular are quite distinct from the earlier work. As Figure 1 illustrates, in MLG, unlike in previous work, the final phonotactic grammar is not an identity map that faithfully maps all legal configurations in the target language to themselves. The outcome of phonotactic learning in MLG is a restrictive grammar with a statistical bias against infrequent, marked configurations. This property of MLG is discussed in detail in Section 2.4.

2.2. Linking Hypotheses to Development of Production

The division into the phonotactic and morphophonemic stages corresponds to children's phonological development. A large body of literature shows that children acquire at least some phonotactic knowledge by approximately 9 months of age (Jusczyk et al 1993; Friederici and Wessels 1993). On the other hand, learning of alternations and the lexicon occurs much later, roughly between the ages of 2 and 4.5 years, and in some cases even later (Berko 1958; Stager and Werker 1997; Pater 1997; Pater, Stager and Werker 2004; MacWhinney 1978). The two stages of MLG are based on this overall developmental progression.

Since phonotactic learning in MLG occurs prior to the development of a phonological lexicon and corresponds to development that occurs by 9 months of age, it is hypothesized that production does not occur until the morphophonemic stage. In particular, the onset of morphophonemic learning, when each form (or morpheme) is associated with an individual lexical entry, is hypothesized to correspond to the onset of production in children. In other words, the phonotactic grammar learned during the first stage serves as the initial production grammar for the model, and it is from this phonotactic grammar that further learning of the grammar proceeds. In sum, the path taken by the model between the initial production grammar and the target grammar corresponds to the predicted acquisition path. It is possible to model grammatical development in this way because phonotactic learning in MLG does not result in an identity map grammar.

This paper examines the effects of frequency and markedness on grammatical development. Since production is hypothesized to coincide with the onset of morphophonemic learning, the focus of this paper is on the grammatical progression predicted by the model during the morphophonemic stage and its correspondence to observed order of acquisition, as evidenced by production data.

2.3. The Implementation

Given the overall structure of the MLG model, there are a number of possible implementations of the actual learning, or likelihood maximization, procedure. In the simulations described here, I employ the standard Expectation Maximization (EM) algorithm (Dempster et al. 1977). EM is a general-purpose algorithm for likelihood maximization with hidden variables, and it has some properties that make it a suitable candidate for the present task. First, EM is guaranteed to converge on a (local) maximum, and second, it adjusts the grammar and lexicon gradually. In other words, EM transitions gradually from the initial state of the grammar and lexicon to the target states, enabling an examination of the gradual learning path it predicts.

In these simulations, I make the simplifying assumption that the grammar and lexicon are lists of rankings and underlying forms, respectively, with associated probabilities. This simplifies the maximization step of the EM algorithm used here but is not an intrinsic aspect of MLG and crucially does not determine the overall, qualitative predictions of the theory. For a discussion of how more sophisticated representations of the grammar and lexicon may be implemented in MLG, see Jarosz (2006b).

2.4. Deriving Stages in MLG: the Role of Frequency

To understand MLG's predictions for production, it is necessary to understand the final phonotactic grammar, which, as discussed above, is hypothesized to be the initial production grammar. The intermediate stages the grammar passes through on its way to the target grammar depend on the grammar's starting point and its update procedure. The central idea is that integrating a rich base into the phonotactic learning stage results in a final phonotactic grammar that is probabilistically biased against marked and infrequent forms. Essentially, the rich base results in a phonotactic grammar with a probabilistic (not absolute) Markedness » Faithfulness bias. This phonotactic grammar serves as the initial production grammar for the learner, and during the course of learning, the grammar gradually transitions from the initial Markedness » Faithfulness state to the target grammar.

To illustrate the properties of the phonotactic grammar and the predicted acquisition path, consider again the simple syllable structure example, reproduced in Figure 2 below. In this example, the target language consists of two overt forms, CV and CVC, with CV being twice as frequent as CVC. The universal rich base includes CV and CVC as well as a third form CVCC, with all three forms being equally likely.

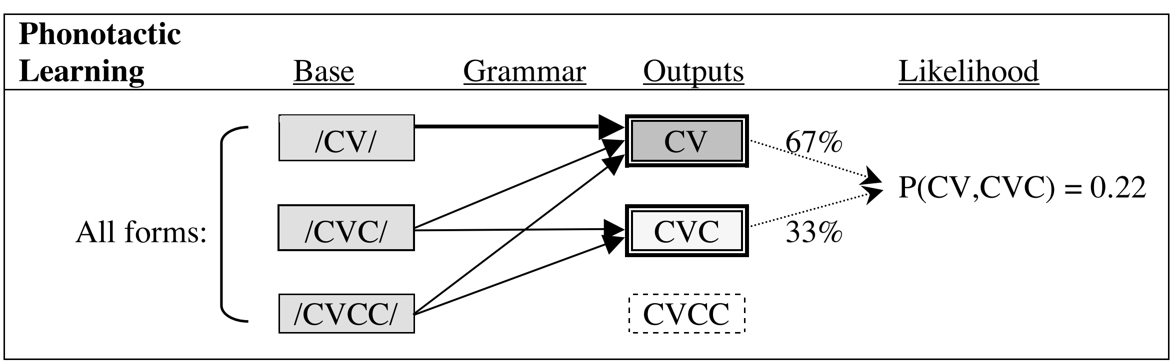


Figure 2: Derived Biases in Phonotactic Learning

Given the three constraints shown in (2)a, it is clear that the target morphophonemic grammar for this language is *COMPLEXCODA » MAX » NOCODA since complex codas are not permitted, while singleton codas are. This target grammar serves as the end-point of morphophonemic learning, while the starting point for morphophonemic learning is the final phonotactic grammar.

(2) Phonotactic Grammar Illustration

a Constraint Set:

- MAX: Segments may not be deleted
- NOCODA: Syllable codas are not permitted
- *COMPLEXCODA: Complex syllable codas are not permitted

b Target (Morphophonemic) Grammar:

- *COMPLEXCODA » MAX » NOCODA

c Phonotactic Grammar:

- P(*COMPLEXCODA » NOCODA » MAX) = 50%
- P(*COMPLEXCODA » MAX » NOCODA) = 50%

As mentioned above, the phonotactic grammar exhibits a statistical bias against marked, infrequent forms. This statistical bias results from the pressure to maximize the likelihood of the overt forms *given the rich base*. In order to maximize the likelihood given the rich base in this example, the final phonotactic grammar must assign less than full probability to the mapping /CVC/ → CVC. Specifically, the maximum likelihood phonotactic grammar ranks *COMPLEXCODA highest and ranks NOCODA and MAX variably, such that the likelihood of either relative ranking is 50%, as shown in (2)c. This means that the mapping of /CVC/ → CV is just as likely as the mapping /CVC/ → CVC in the final phonotactic grammar.

To see why this is the maximum likelihood phonotactic grammar, it helps to think of the phonotactic grammar as a mapping between the distribution over underlying forms in the rich base and the distribution over overt forms in the ambient language. The maximum likelihood grammar is the mapping that does the best job of matching the frequencies of overt forms. In order to do so, the phonotactic grammar redistributes the probabilities associated with underlying forms in the base among the various overt forms. In this example, the probabilities distributed equally among /CV/, /CVC/ and /CVCC/ must be redistributed by the phonotactic grammar such that the likelihood of CV is two-thirds and the likelihood of CVC is one-third. Ranking *COMPLEXCODA highest means that the 33.3% of the probability mass associated with /CVCC/ in the rich base will be mapped to and distributed among CV and CVC. Ranking NOCODA and MAX equally means that /CVCC/ mapping to CV is equally likely as mapping to CVC, as depicted by the arrows pointing away from /CVCC/ in the figure. Under this variable ranking, /CVC/ itself will map to CV and CVC with equal likelihood, shown as arrows pointing away from /CVC/. In other words, the overt form CV gets one-third of the total, redistributed probability mass from underlying /CV/, and one-sixth of the probability mass from each of /CVC/ and /CVCC/. This redistribution is illustrated in the figure as arrows pointing to

CV. As a result, two-thirds of the probability mass ends up on CV and the remainder on CVC, matching the target distribution exactly.

The final phonotactic grammar encodes the legal and illegal phonotactics of the target language, mapping illegal configurations such as CVCC to legal ones. However, among the legal forms admitted by the grammar, there is a statistical bias, a soft preference, for unmarked and frequent configurations. In sum, in order to match the target distribution, the phonotactic grammar exhibits a bias against the marked infrequent forms by mapping some (or all, in case of illegal forms) of their input probability variably to the more frequent, less marked forms. This variable grammar serves as the starting grammar for morphophonemic learning.

At the onset of morphophonemic learning, learning of underlying representations begins: each morpheme (or overt form) is assigned its own lexical entry, which is gradually adjusted during this stage. As a result, overt forms no longer share the rich base as they do during phonotactic learning, and the maximum likelihood morphophonemic grammar generates for each overt form the correct output form(s) from their correct underlying representations[5].

In the present example, the morphophonemic stage consists of a gradual transition between the phonotactic grammar in (2)c and the adult, target grammar in (2)b. Because the phonotactic grammar maps /CVC/ to CV half the time, it displays a probabilistic Markedness » Faithfulness bias relative to the target grammar. That is, NOCODA is probabilistically higher ranked than MAX in the phonotactic grammar relative to the target grammar. In this simple example, the learning path involves a gradual transition from equally ranked NOCODA and MAX to a ranking with MAX dominating NOCODA completely. This means the learner begins with a grammar in which codas are not reliably produced and gradually converges on the target grammar that generates codas faithfully.

This probabilistic bias against marked forms is a consequence of the target language distribution and rich base and is itself not specified anywhere in the model. The strength of the initial bias depends on a number of factors, including the target language distribution and the particular choice of rich base. In general, the lower the frequency of a marked configuration, the stronger the initial bias against it will be. For example, if the target distribution in the above example exhibited 80% CV forms, then a stronger NOCODA » MAX bias would be required to match the target distribution. Thus, the strength of the bias rests in part on the presence of a statistical bias against the marked configuration in the target language relative to the rich base. However, no amount of statistical evidence in favor of a marked configuration over an unmarked configuration

---

[5] The maximum likelihood morphophonemic grammar (target grammar) does not match frequencies because the overt forms during morphophonemic learning have their own lexical entries, and probability mass is no longer shared between distinct lexical entries. In the maximum likelihood morphophonemic grammar, frequency matching only occurs among forms in free variation, forms generated from the same lexical entry.

can ever result in a bias in favor of marked forms. For instance, given the implicational universals embodied in the constraint set, no ranking characterizes a language that requires codas, and therefore no phonotactic grammar (or intermediate grammar) where codas are favored or required is possible.

The contents and distribution of the rich base also affect the strength of the bias. If a larger base consisting additionally of CVCCC were used, then a lower probability would be associated with CV in the base, and the NoCoda » Max bias would need to be stronger in order to match the high frequency associated with CV in the target distribution. In the simple example discussed above, the bias is not particularly strong, but as the simulations in the next section will show, when even marginally more complex examples are designed and applied to target distributions representative of actual languages, the biases are strong enough to predict a complex and realistic series of acquisition stages.

In sum, the morphophonemic stage in MLG models the gradual mastery in production of the various configurations in the target language. The starting point for production is a grammar that encodes the phonotactics of the target language probabilistically, assigning probability to legal configurations only and exhibiting a statistical bias against marked, infrequent configurations. The probabilistic nature of MLG enables phonotactic learning to result in a grammar that corresponds to an early stage of development in children when language-particular phonotactics have been acquired but production is unmarked relative to the target language. The predicted acquisition path corresponds to the transition between this language-particular unmarked state and the final, adult grammar.

2.5. The Relative Roles of Markedness and Frequency in MLG

While the role of frequency in MLG is significant, its effects are secondary to that of markedness. In essence, universal markedness is embodied in the constraint set, which defines the space of possible learning paths, and language-specific frequency selects among those paths. This view of the relative roles of markedness and frequency is not new to MLG (Levelt 1994; Levelt and Van de Vijver 1998; Levelt et al. 2000; Boersma and Levelt 2000). What is new to MLG is that the initial, unmarked state of the grammar is derived from general learning principles rather than being stipulated. In MLG, as in the previous work cited above, order of acquisition is influenced, but not determined by frequency. Frequency can only influence the classes of phonological units embodied in the constraint set: the frequency of individual forms plays no direct role in grammatical learning. Each overt form contributes to the frequency of the phonological classes of which it is part, and it is the frequency of these general phonological classes that shapes learning. Phonological class membership is hierarchical and complex, however, and any attempt to calculate the effects of frequency on a single configuration must take into account the net effects of many interacting factors.

In sum, in MLG markedness and frequency both have important roles, markedness on a universal level, and frequency on a language-particular level. In particular, if any implicational relationships are entailed by universal markedness, these implications cannot be reversed in the learning path. For example, since no OT grammar can describe a language that admits complex codas but not singleton codas (complex codas entail singleton codas in universal markedness), a learning path where complex codas are learned before singleton codas is not possible. On the other hand, when markedness considerations are ambivalent or contradictory, the effects of frequency can emerge. Taking the example of coda and onset clusters that will be taken up later in the paper, since markedness does not determine a relative order for these two configurations, relative frequency is able to do so, favoring the more frequent of the two. This does not mean, however, that all types of the more frequent configuration will necessarily be acquired before all types of the less frequent configuration. It is entirely possible for other frequency and markedness factors to penalize a particularly marked subtype of the more frequent type. For example, even though complex codas may in general be acquired earlier than complex onsets, some complex codas, such as clusters with particularly difficult segments like affricates or retroflexes, may be acquired later than some complex onsets as long as there are markedness constraints against these segments.

2.6. Comparison with Explicit Ranking Biases

In general, the order of acquisition predictions of MLG and learning models with an explicit Markedness » Faithfulness ranking bias are not that different[6] (Boersma and Levelt 2000; Tessier 2006; Pater, Jesney, and Tessier to appear). One important difference between the intermediate stages predicted by MLG and theories with ranking biases is that the phonotactic grammar in MLG that serves as the initial production grammar already embodies a sort of language-specific markedness. The phonotactic grammar encodes the legal phonotactics of the target language in a way such that phonotactically illegal configurations are no longer allowed by the grammar. Therefore, the learning path will involve a gradual transition from the language-particular unmarked initial state to the language-particular target state, with only phonotactically legal productions predicted[7]. Gradual learning from a general Markedness » Faithfulness initial bias need not be so constrained.

A simple example can illustrate this point. Consider a language with intervocalic voicing, with two competing markedness constraints on voicing, *VOICE, prohibiting voicing in general, and *VTV, prohibiting voicelessness intervocalically. A general Markedness » Faithfulness bias may be implemented by ranking *VOICE and *VTV equally high, well above IDENTVOICE. Since *VOICE and *VTV are equally ranked, they will interact in the initial stages with some productions of phonotactically illegal voiceless intervocalic segments predicted[8]. In MLG, on the other hand, *VOICE is already

[6] Although I have not comprehensively explored this, the effects of other biases, such as Specific Faithfulness » General Faithfulness, should in principle be derivable as well if the ambient language supports a statistical bias in their favor relative to the rich base.

[7] It would be possible for phonotactic restrictions to be overpowered by factors that become active only in the morphophonemic stage, such as constraints that rely on morphological or lexical information.

[8] In weighted constraint systems, this prediction holds only if intervocalic voicing is an active phonological process conditioning alternations (and not a static regularity) that requires underlying voiceless intervocalic

lower ranked in the phonotactic grammar, and phonotactically illegal productions are not predicted. This seems to be a desirable prediction overall since children's non-adult-like productions in general abide by the phonotactic restrictions of the target language (Zamuner et al., in prep; though see Jesney and Tessier, this volume, for some exceptions).

3. Case Studies: Syllable Structure Acquisition

This section describes a series of case studies that explore the role of markedness and frequency in the acquisition of syllable structure in three languages with distinct syllable type frequency profiles. The studies focus on the relationship between relative frequency and the relative order of acquisition of complex onsets and complex codas.

From a markedness perspective, there is no implicational relationship between complex onsets and complex codas. Some languages allow complex onsets but not complex codas (Spanish), while others allow complex codas but not complex onsets (Finnish). Examining the relative order of acquisition and its relationship to relative frequency in languages that permit both types of clusters can shed light on the role of frequency in phonological acquisition.

The purpose of this section is twofold. The first goal is to illustrate the capacity of MLG to predict stages of acquisition when applied to input distributions extracted from adult, child-directed speech. The input distributions in the three languages differ significantly, which provides an opportunity to examine how cross-linguistic differences in acquisition order arise in MLG.

The second goal is to examine the hypothesis that frequency plays a causal role in shaping the acquisition process by examining the relationship between relative frequency and relative acquisition orders in three languages. Previous findings are indicative of a connection between frequency and acquisition order (Ingram 1988; Levelt 1994; Levelt and Van de Vijver 1998; Levelt et al. 2000; Roark and Demuth 2000; Kirk and Demuth 2005); nonetheless; further work examining this relationship cross-linguistically for various structures is necessary before a causal connection can be established.

Previous research on acquisition of consonant clusters has examined the relative order of acquisition and relative frequencies of complex codas and onsets in English and Dutch. The predominant order attested in English-speaking children is acquisition of complex codas before complex onsets (Kirk and Demuth 2005; Templin 1957). In English, complex codas are also significantly more frequent than complex onsets in adult, child-directed speech (Kirk and Demuth 2005). Dutch-speaking children, on the other hand, show variation in the order of acquisition, with a preference for acquiring complex codas first (Levelt 1994; Levelt and Van de Vijver 1998; Levelt et al. 2000). As shown

---

segments to be mapped to voiced correspondents. Otherwise, if the voicing is underlying, lower-weighted IDENTVOICE could settle the tie between the higher-weighted constraints in favor of the phonotactically legal variant.

by Levelt et al. (2000), the relative frequencies of various syllable types in Dutch corresponds to the order of acquisition of these syllable types. The frequencies of complex onsets and codas are roughly equivalent in adult, child-directed speech (with complex codas being slightly, but not significantly, more frequent), to which the authors attribute the variation in attested acquisition orders. More precisely, the version of the frequency hypothesis under investigation in the previous work and the present study is summarized in (3).

(3) The Frequency Hypothesis (based on Levelt and Van de Vijver (1998)):
Universal markedness constrains possible learning paths cross-linguistically while language-particular learning paths are driven by the relative frequencies of output configurations in the ambient language.

These findings are consistent with a frequency-based explanation, but they leave open the possibility of alternative explanations as well. Overall, these two languages still show a preference for complex codas to be acquired first; therefore, an explanation in terms of universal factors, such as the structural complexity of onset clusters, or factors common to the two languages, such as articulatory difficulty, is viable as well (Kirk and Demuth 2005). In addition, in both languages word-final clusters contain more morphological information than word-initial clusters, leaving open the possibility of a morphological explanation.

To provide a more thorough test of the frequency hypothesis, an examination of the acquisition of syllable structure in a language for which the frequency hypothesis makes the opposite prediction is required. An example of such a language is Polish. As the corpus analysis presented in Section 3.1.3 reveals, complex onsets are significantly more frequent than complex codas in adult, child-directed speech in Polish. Consequently, the frequency hypothesis and the MLG simulation predict an order of acquisition favoring earlier acquisition of complex onsets. The final portion of the section presents novel data from Polish suggesting that complex onsets are indeed acquired before complex codas in child Polish.

3.1. Syllable Structure Simulations

The three MLG simulations modeling the acquisition of syllable structure in Dutch, English and Polish are identical except for the frequencies of various syllable types that are provided to the learner. The frequencies of various syllable types are estimated from adult, child-directed speech in each language as described in the corresponding sections.

The design of the simulation is based on Boersma and Levelt (2000). All simulations model the acquisition of nine syllable types: CV, CVC, CVCC, V, VC, VCC, CCV, CCVC, and CCVCC. All simulations employ the set of five standard syllable structure constraints shown below.

(4) Constraint Set for Syllable Structure Simulations:
ONSET: No vowel initial syllables
NOCODA: No consonant final syllables
*COMPLEXONSET: No syllable-initial consonant clusters
*COMPLEXCODA: No syllable-final consonant clusters
MAX: No deletion

The rich base used in the simulations consists of all syllable types, each with equal likelihood, shown in (5). Since the base is assumed to be universal, the exact same base is used during phonotactic learning for all languages. As discussed in the Section 2.4, many of the precise numerical predictions of MLG depend on the exact nature of the rich base. Since this base is clearly a major simplification of the universal rich base, the predictions of the simulations should be interpreted broadly and qualitatively[9]. Nonetheless, predictions about relative order of acquisition within each language can be examined, and comparisons between languages can be made.

(5) Rich Base Employed in All Simulations:

CV	CVC	CVCC	V	VC	VCC	CCV	CCVC	CCVCC
0.11%	0.11%	0.11%	0.11%	0.11%	0.11%	0.11%	0.11%	0.11%

As discussed in Section 2, predictions of MLG for production derive from the path taken by the grammar during the morphophonemic learning. Accordingly, the presentation of the simulations focuses on the learning path of the grammar during morphophonemic learning and the associated stages of acquisition predicted by the learning theory.

3.1.1. Dutch

As discussed above, complex onsets and complex codas in Dutch occur in roughly equal proportions. As can be seen in the distribution of syllable types in child-direct speech in Table 1 (Boersma and Levelt 2000), the combined frequency of syllables with complex onsets is 3.7%, while the combined frequency of syllables with complex codas is 4%. This difference is not statistically significant (Levelt and van de Vijver 1998). Under the frequency-based hypothesis, this results in an order of acquisition which is variable for these two syllable types. Indeed, Levelt (1994) observed two acquisition orders in a study of twelve Dutch-speaking children:

(6) Attested Acquisition Paths in Dutch (Levelt 1994):

CV → CVC → V → VC → { CVCC → VCC → CCV → CCVC / CCV → CCVC → CVCC → VCC } → CCVCC

---

9 It is worth noting that if the base were made more detailed (with distinct types of consonants for example), the Markedness » Faithfulness bias in the phonotactic grammar would become stronger overall. This is because expanding the number of segment types places more probability mass in the rich base on longer, more marked structures overall, and the phonotactic grammar must overcome this effect to match the distribution of the language.

Comparing the order of acquisition in (6) to the relative frequencies in Table 1, it is clear that there is a strong correspondence between order of acquisition and relative frequency; however, where frequency is roughly equal, variation in acquisition orders is found.

Table 1: Relative Frequencies of Syllable Types in Dutch

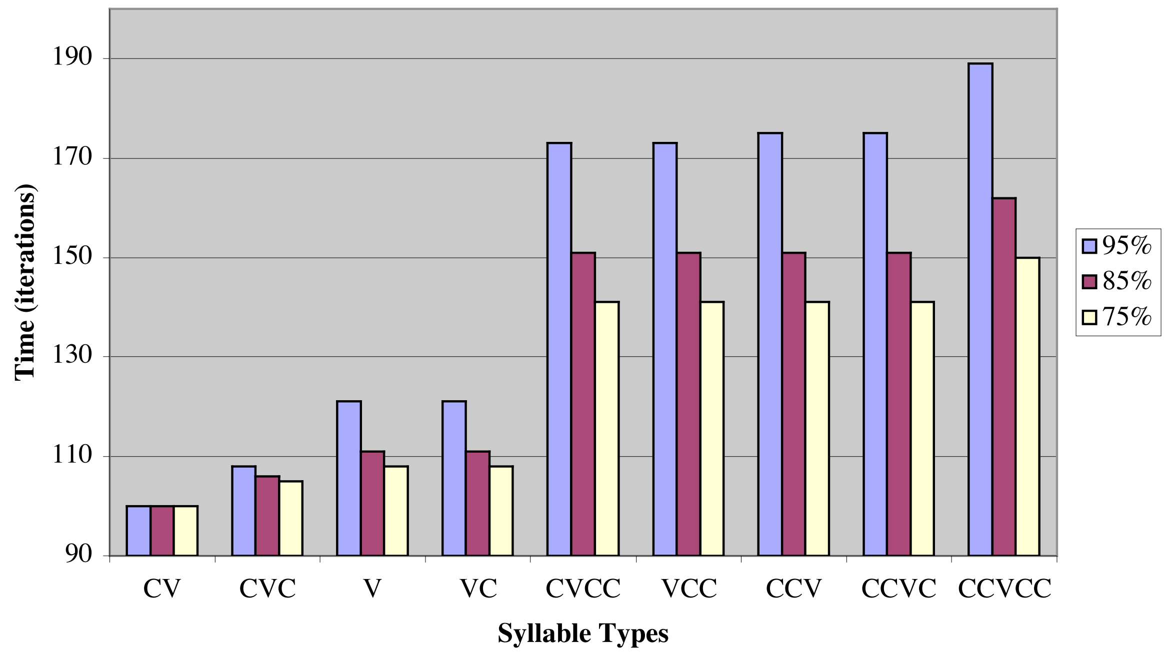
CV	44.8%
CVC	32.1%
CVCC	3.3%
V	3.9%
VC	12.0%
VCC	0.4%
CCV	1.4%
CCVC	2.0%
CCVCC	0.3%

This target distribution is used in the MLG simulation, and the resulting acquisition order during morphophonemic learning is depicted in the chart below. The chart shows the relative time at which three thresholds of production accuracy (95%, 85%, 75%) were reached by the learner for the various syllable types. The syllable types are ordered according to their relative learning order. As shown in the chart, the acquisition order of complex onsets and codas is predicted to be extremely close: both are produced with approximately 85% accuracy by the 150th iteration of learning.

Since the difference between the frequencies of onset and coda clusters is not significant in the sample overall, it is likely that different children are exposed to slightly different distributions, with some children hearing complex codas more frequently and others hearing complex onsets more frequently. If this simulation were repeated with slightly different relative frequencies of clusters, variation in acquisition order would be predicted over multiple trials, with earlier acquisition order corresponding to higher relative frequency in the sample. It is also possible that variation results from randomness in the learning procedure itself, as suggested by Levelt et al. (2000) and Boersma and Levelt (2000)[10]. Further work examining the effect of frequency on order of acquisition within a single language is needed to answer this question.

[10] The implementation adopted here employs the EM algorithm, which, in its standard form, is a deterministic algorithm. Noise could be introduced into the algorithm in any number of ways, however, and this would result in variable order of acquisition of syllable types whose acquisition orders are very close in the standard version of the algorithm.

Chart 1: Predicted Order of Acquisition for Dutch



The order of acquisition of the other syllable types is consistent with the attested acquisition orders. As discussed in the previous section, frequency guides the order of acquisition in the learning model, but it does not directly translate into acquisition order. For example, the relative frequencies of V and VC, 3.9% and 12%, respectively, do not translate into a relative acquisition order with VC being learned earlier. In this constraint system, V and VC are treated as a class due to their violation of ONSET, while NOCODA can no longer distinguish between the two types because it has been ranked well below MAX and plays no active role in the grammar. This appears to coincide with children's development: children do not acquire VC earlier, despite its dramatically higher frequency[11]. This illustrates the primary role of markedness in MLG: frequency cannot drive the earlier acquisition of a more marked configuration.

3.1.2. English

Previous research has established that complex codas are significantly more frequent than complex codas in child-directed English and shown that complex codas tend to be acquired earlier (Kirk and Demuth 2005; Templin 1957). In order to set up an MLG simulation for English comparable to the above Dutch simulation, however, the relative

[11] The observation that VC is actually acquired later than V by children who have already acquired CVC, as reported in Levelt (1994), appears to be a kind of cumulativity effect that is problematic for all OT models. However, interestingly, the very same simulation shows that cumulativity of a sort is possible in OT. The acquisition order for the syllable type with violations of both *COMPLEXONSET and *COMPLEXCODA, CCVCC, is later than the acquisition of syllable types with only one of these violations. This is because acquisition order is defined in terms of reaching a threshold of accuracy, which may be reached later for forms that are subject to more unfaithful mappings in the developing grammar.

frequencies of all syllable types in English are required. To get an estimate of the relative token frequencies of the various syllable types in English child-directed speech, the primary stressed monosyllabic words were extracted from the CHILDES Parental Corpus (MacWhinney 2000; Li and Shirai 2000). The CHILDES Parental Corpus consists of 2.6 million word tokens and 24,156 word types of which 67% were monosyllabic, primary stressed words used in this analysis. The syllables were phonemicized using the CMU Pronouncing Dictionary (Weide 1994), and the relative frequencies of the various syllable types were tallied. This estimate of the distribution of the various syllable types in English is shown in Table 2[12].

Table 2: Relative Frequencies of Syllable Types in English

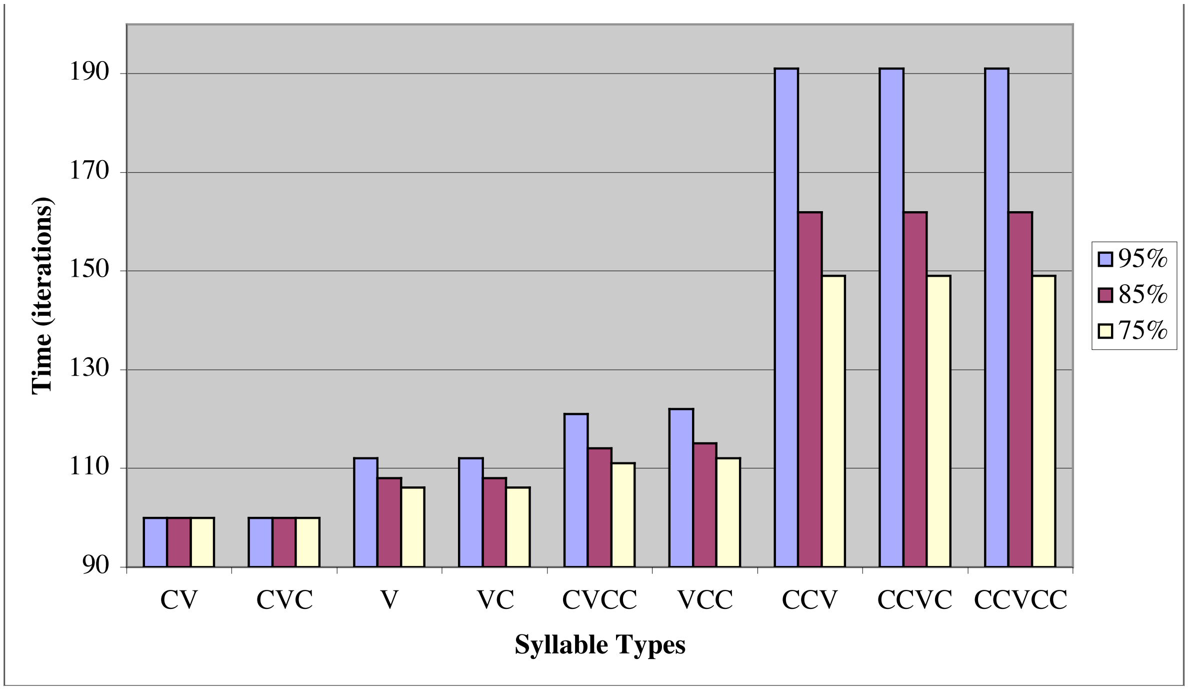
CV	24.4%
CVC	40.5%
CVCC	10.1%
V	4.7%
VC	13.0%
VCC	3.5%
CCV	0.9%
CCVC	2.2%
CCVCC	0.6%

As can be seen from these data, the overall frequency of complex codas is 14.2% while the frequency of complex onsets is 3.7%. In the MLG simulation, this difference in relative frequencies results in a predicted acquisition order with complex codas acquired much earlier than complex onsets. Given the input distribution in Table 2, no reasonable amount of noise is likely to allow variation in the order of acquisition. In other words, MLG predicts that complex codas should be acquired before complex onsets by English-speaking children.

Overall, the predicted order for English is very similar to the order in Dutch, except for the added restriction that complex codas be acquired before complex onsets. In both languages, onsetless syllables are acquired before complex margins. This is not terribly surprising as the English and Dutch distributions are very similar, and the two languages are closely related. The next simulation looks at a more distantly related language: Polish.

[12] This estimate could certainly be improved by including all stressed syllables, but what is crucial for present purposes is that the relative proportion of complex codas is notably higher than the proportion of complex onsets, consistent with previous findings. Kirk and Demuth (2005) find the overall proportion of complex codas to complex onsets in child-directed speech to be 2:1, suggesting that the difference in frequencies estimated here may be somewhat overestimated.

Chart 2: Predicted Order of Acquisition for English[13]



3.1.3. Polish

A corpus analysis of parental speech found in the Szuman corpus of Polish, available in CHILDES (Smoczynska 1985; MacWhinney 2000), reveals that complex onsets are more frequent than complex codas in Polish child-directed speech. Monosyllabic words were extracted from the corpus and automatically phonemicized based on the orthography (which reliably corresponds to pronunciation), resulting in a corpus of 118,701 syllables. The proportions of various syllable types are shown in Table 3. In this estimate of syllable proportions, the relative frequency of complex onsets is 10.1%, while the relative frequency of complex codas is 6%[14].

---

[13] Though this is not the focus of this study, careful readers may note that according to this simulation, CV and CVC syllable types should be acquired at roughly the same time. This prediction results from the simplifying assumptions about the rich base made in this simulation, and should not be taken as a prediction of MLG in general.

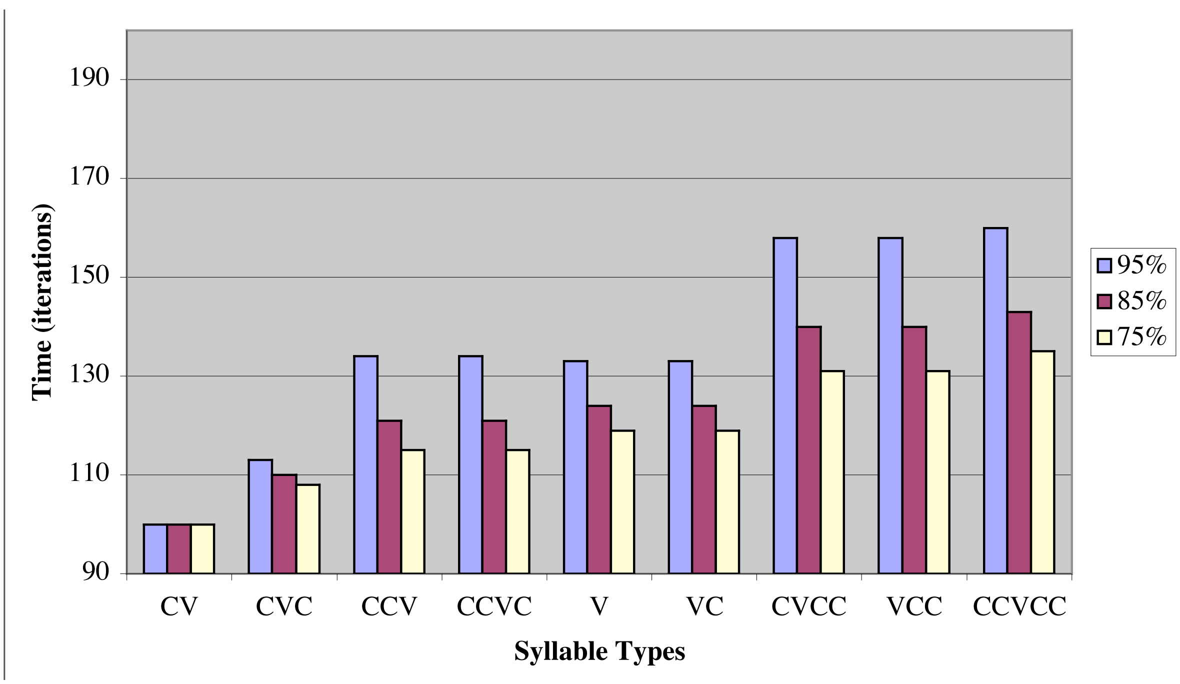
[14] To confirm the robustness of the difference between the relative frequencies of coda and onset clusters, two other corpora analyses were performed. One estimated the relative proportions of the various syllable types from the proportions of word-final and word-initial clusters in parental speech. Another analysis estimated the proportions of syllable types from monosyllabic words extracted from the IPI PAN corpus of written Polish. Both analyses revealed an even stronger difference between the two cluster types than found in the analysis used above.

Table 3: Relative Frequencies of Syllable Types in Polish

CV	54.7%
CVC	19.0%
CVCC	5.3%
V	9.0%
VC	1.6%
VCC	0.3%
CCV	4.4%
CCVC	5.3%
CCVCC	0.4%

These frequencies were submitted to the MLG learner, and the predicted order of acquisition is shown below. The difference in relative frequencies between the two cluster types results in a clear difference in the predicted order of acquisition. Unlike the Germanic languages in the previous two simulations, Polish exhibits a higher proportion of complex onsets, which in turn corresponds to the earlier acquisition of complex onsets. Another notable difference between the Germanic languages and Polish is the relative proportion of onsetless syllables to complex margins of either kind. As noted above, in English and Dutch, onsetless syllables are predicted to be acquired prior to either complex margin type. On the other hand, onsetless syllables are relatively infrequent in Polish and are therefore predicted to be acquired late relative to the other languages.

Chart 3: Predicted Order of Acquisition for Polish



The remaining question is, of course, whether the prediction of the frequency hypothesis concerning the relative order of acquisition of cluster types corresponds to actual acquisition orders exhibited by children acquiring Polish. Previous research addressing this question is limited although one brief description of the acquisition of clusters by three Polish children suggests these children acquired initial clusters before final clusters (Dziubalska-Kołaczyk 1998). Providing a comprehensive answer to this question is beyond the scope of this paper; however, Section 3.2 discusses the results of a preliminary analysis that suggests the predicted order may be preferred in Polish.

3.2. Developmental Order in Child Polish

This section describes a study investigating the relative order of acquisition of coda and onset clusters in child Polish.

3.2.1. Method

The data are from spontaneous productions of three Polish-speaking children from the Weist corpus of child Polish, available as part of the CHILDES corpus (Weist and Witkowska-Stadnik 1986; Weist et al. 1984; MacWhinney 2000). The digitized recordings were phonologically transcribed using the ChildPhon program (Rose 2003). The ages of the children range from 1;7 to 2;1.

For each child, target word-final and word-initial CC clusters were identified and coded as either correct or incorrect. Focusing on word edge clusters avoids the complication of the inherent ambiguity in word-medial syllabification in Polish, especially since syllabification may not be adult-like. Target clusters were coded as correct if they were produced as clusters, and coded as incorrect otherwise. Consonant-to-consonant substitutions were not counted as errors; however, less than 2% of the clusters coded as correct involved substitutions that involved a change in sonority level. In other words, when the clusters were produced, they were very close to the adult targets. The vast majority of errors (about 90%) involved deletion, although coalescence (for stop-fricative sequences to affricates) and epenthesis were also observed.

3.2.2. Results

The percent correct production for each child and each cluster type is shown in Table 4. As shown in the table, the percent correct production on initial clusters ranges from 47.7% to 71.4% while the percent correct production on final clusters ranges from 0% to 40.5%. Even with this small sample of productions, the results are significant: each of the children has a significantly higher proportion of correct complex onsets than correct complex codas (as measured by Fisher's Exact Test).

While complex onsets have a significantly higher proportion for all the children, the highest accuracy on complex codas relative to complex onsets is found in Bartosz's productions. In other words, the accuracy in production of complex onsets and codas is closest in this data. Intriguingly, the relative proportions of complex onsets to complex codas extracted just from Bartosz's parental input differ from the proportions found in parental speech sample from the three children grouped together. The relative frequencies of complex onsets and complex codas in Bartosz's parental input are 10.3% and 9.5%, respectively, which is a significant difference (Fisher's Exact Test, $p < 0.01$) from the more distinct relative frequencies found in the sample overall.

Table 4: Children's Production Accuracy on Onset and Coda Clusters in Polish

	Age	#CC Targets	#CC % Correct	CC# Targets	CC# % Correct	Fisher's Exact Test
Marta	1;7	86	47.7%	11	0%	$p < 0.01$
Bartosz	1;11	99	61.6%	42	40.5%	$p < 0.05$
Kubus	2;1	56	71.4%	14	21.4%	$p < 0.01$

3.2.3. Discussion

While these results are certainly not definitive, they suggest that children acquiring Polish do in fact show a tendency to acquire complex onsets before complex codas. The higher accuracy in production of complex onsets in the data above suggests that mastery (as defined by some threshold of accuracy) will be achieved earlier for complex onsets than for complex codas in subsequent stages. However, to fully understand the developmental progression, a longitudinal analysis examining the development of clusters in a larger sample of children is required. One potential issue with the present study is that the child productions in the corpus are spontaneous, not controlled in any way. As a result, many of the targets are represented by a small number of lexical items produced multiple times. For example, a majority of the complex coda targets are tokens of the word [jɛst] – "is" (3rd person, singular). As a result, it is not clear how representative the results are of coda clusters in general. Further work testing the accuracy of cluster production in an experimental setting will likely be required to overcome the limitations of this methodology.

Nonetheless, the results do support the frequency hypothesis and the predicted order of acquisition in the simulation, especially when considered together with the predictive success of the frequency hypothesis in English and Dutch. Since the results of this study suggest a different relative order of acquisition of complex onsets and codas in different languages, the explanations that appeal to universal factors such as structural complexity are not supported. The data presented here do not speak to an articulatory explanation. A finer-grained analysis, looking at order of acquisition of various clusters relative to their relative articulatory difficulty in the three languages, is needed to examine the articulatory explanation.

The results may also support the morphological explanation. This is because while Dutch and English final clusters contain more morphological information than initial clusters, the reverse is true in Polish. Indeed, a study examining the acquisition of clusters morpheme-internally and across morpheme boundaries in Polish found a tendency for clusters that cross morpheme boundaries to be produced more accurately earlier (Zydorowicz 2007). On the other hand, a comparable analysis for English found no effect of morphological complexity (Kirk and Demuth 2005). These mixed findings indicate that further work exploring the morphological explanation is needed as well.

An interesting possibility is raised by the apparent individual differences in the present study. The correspondence between Bartosz's input distribution and his relatively high coda cluster accuracy may be accidental. However, if there are in fact significant differences between the input distributions for different children within a single language, then the frequency hypothesis predicts that individual differences in acquisition order could arise as a result. Whether significant individual differences in input distribution exist between children acquiring the same language and, if so, whether these differences correspond to differences in acquisition order, are questions for further research. An advantage of examining the frequency hypothesis within a single language is that other factors corresponding to alternative explanations (such as the morphological, articulatory, and structural explanations) are constant.

4. Conclusion

In conclusion, this paper has examined the relative roles of frequency and markedness from the perspectives of computational modeling and child acquisition. MLG, a theory that relies on two general principles to solve a number of learnability problems, is shown to be capable of modeling acquisition order without relying on built-in ranking biases. A strength of MLG is that productions are predicted to abide by language-particular phonotactic restrictions. The predicted acquisition orders in MLG are sensitive to markedness and frequency, with markedness constraining possible acquisition paths on a universal level and frequency constraining acquisition paths on a language-particular level.

The predictions of MLG for the acquisition order of syllable structure are examined in three languages with distinct distributions of syllable types. The distributions of syllable types in the three languages, estimated from adult, child-directed speech, are found to be predictive of the relative order of acquisition in these languages, as evidenced by previous work on Dutch and English and a study on Polish cluster acquisition presented here. These findings provide support for the causal role of frequency in acquisition, and in particular, for a learning theory in which markedness and frequency interact as they do in MLG.

References

Anttila, Arto. 1997. *Variation in Finnish Phonology and Morphology*. Ph.D. thesis, Stanford Univ.

Berko, Jean. 1958. The child's learning of English morphology. *Word* 14, 150-177.

Boersma, Paul. 1998. *Functional Phonology*. Doctoral Dissertation, University of Amsterdam. The Hague: Holland Academic Graphics.

Boersma, Paul and Bruce Hayes. 2001. Empirical tests of the Gradual Learning Algorithm. *Linguistic Inquiry* 32(1):45-86.

Boersma, Paul and Clara C. Levelt. 2000. Gradual constraint-ranking learning algorithm predicts acquisition order. In *The proceedings of the thirtieth annual child language research forum*, ed. Eve V. Clark. Stanford: CSLI.

Dempster, Arthur P., Nan M. Laird, and Donald B. Rubin. 1977. Maximum Likelihood from incomplete data via the EM Algorithm. *Journal of Royal Statistics Society*. 39(B):1-38.

Dziubalska-Kołaczyk, Katarzyna (1999): "Early L1 clusters and how they relate to universal phonotactic constraints" in Proceedings of the 14th International Congress of Phonetic Sciences, San Francisco. San Fransisco: University of Berkeley, 317-320.

Flack, Kathryn. 2007. *Sources of Phonological Markedness*. Ph.D. dissertation, University of Massachusetts, Amherst.

Flack, Kathryn. to appear. Inducing functionally grounded constraints. In M. Becker (ed.), *UMass Occasional Papers in Linguistics* 36.

Friederici, Angela D. and Jeanine E. Wessels. 1993. Phonotactic knowledge of word boundaries and its use in infant speech perception. *Perception and Psychophysics* 54, 287-295.

Gnanadesikan, Amahlia. 1995/2004. Markedness and faithfulness constraints in child phonology. In R. Kager, W. Zonneveld, J. Pater, eds., *Fixing Priorities: Constraints in Phonological Acquisition*, Cambridge, Cambridge University Press.

Hayes, Bruce. 2004. Phonological acquisition in Optimality Theory: the early stages. Appeared in Kager, Rene, Pater, Joe, and Zonneveld, Wim, (eds.), *Fixing Priorities: Constraints in Phonological Acquisition*. Cambridge University Press.

Ingram, David. 1988. The Acquisition of Word-Initial [v]. *Language and Speech*. 31(1). 77-85.

Jarosz, Gaja. 2006a. Richness of the Base and Probabilistic Unsupervised Learning in Optimality Theory. *Association for Computational Linguistics: Proceedings of the Eighth Meeting of the ACL Special Interest Group in Computational Phonology*.

Jarosz, Gaja. 2006b. *Rich Lexicons and Restrictive Grammars - Maximum Likelihood Learning in Optimality Theory*. Ph.D. dissertation, Johns Hopkins University..

Jarosz, Gaja. to appear. Restrictiveness in Phonological Grammar and Lexicon Learning. *43rd Annual Meeting of the Chicago Linguistics Society*, Chicago, Illinois.

Jesney, Karen and Anne-Michelle Tessier. to appear. ???. In M. Becker (ed.), *UMass Occasional Papers in Linguistics* 36.

Jusczyk, Peter W., Angela D. Friederici, Jeanine M.I. Wessels, Vigdis Y. Svenkerud, and Ann Marie Jusczyk. 1993. "Infants' sensitivity to the sound patterns of native language words," *Journal of Memory and Language* 32, 402-420.

Kirk, Cecilia, and Katherine Demuth. 2005. Asymmetries in the Acquisition of Word-initial and Word-final Consonant Clusters. *Journal of Child Language* 32: 709-734.

Levelt, Clara. 1994. On the Acquisition of Place. Doctoral Dissertation Leiden University, HIL Dissertation Series 8.

Levelt, C., N. Schiller, and W. Levelt. 2000. The Acquisition of Syllable Types. *Language Acquisition* 8(3), 237-264.

Levelt, Clara and Ruben Van de Vijver. 1998. Syllable Types in Cross-Linguistic and Developmental Grammars. Rutgers Optimality Archive 265.

Li, P., and Shirai, Y. 2000. *The acquisition of lexical and grammatical aspect*. Berlin & New York: Mouton de Gruyter.

MacWhinney, B. 1978. The acquisition of morphophonology. *Monographs of the Society for Research in Child Development*, 43, Whole no. 1.

MacWhinney, B. 2000. *The CHILDES project (3rd Edition)*. Mahwah, NJ: Lawrence Erlbaum.

McCarthy, John J. 1998. Morpheme structure constraints and paradigm occultation. Appeared in M. Catherine Gruber, Derrick Higgins, Kenneth Olson, and Tamra Wysocki, eds., *Proceedings of the Chicago Linguistic Society 5*, Vol. II: The Panels, Chicago, CLS.

Pater, Joe, Karen Jesney and Anne-Michelle Tessier. 2007. Phonological acquisition as weighted constraint interaction. In Alyona Belikova, Luisa Meroni and Mari Umeda (eds.), *Proceedings of the Conference on Generative Approaches to Language Acquisition - North America (GALANA 2)*. Somerville, MA: Cascadilla Proceedings Project.

Pater, J., C. Stager, and J. Werker. 2004. The Perceptual Acquisition of Phonological Contrasts. *Language* 80.3.

Prince, Alan, and Paul Smolensky. 1993/2004. Optimality Theory: Constraint interaction in generative grammar. Technical Report, Rutgers University and University of Colorado at Boulder, 1993. Revised version published by Blackwell, 2004.

Prince, Alan and Bruce Tesar. 2004. Learning Phonotactic Distributions. Appeared in Kager, Rene, Pater, Joe, and Zonneveld, Wim, (eds.), *Fixing Priorities: Constraints in Phonological Acquisition. Cambridge University Press.*

Reynolds, Bill. 1994. *Variation and phonological theory*. Doctoral Dissertation, University of Pennsylvania.

Roark, B. and Katherine Demuth. 2000. Prosodic constraints and the learner's environment: A corpus study. *Proceedings of the Boston University Conference on Language Development*, 24, 597-608.

Rose, Yvan. 2003. ChildPhon: A Database Solution for the Study of Child Phonology. *Proceedings of the 26th Annual Boston University Conference on Language Development*. Somerville, MA: Cascadilla Press.

Smith, Jennifer L. 2000. “Positional faithfulness and learnability in Optimality Theory,” in Rebecca Daly and A. Rehl, eds., *Proceedings of ESCOL99*, Ithaca, CLC Publications.

Smoczynska, M. (1985). *The acquisition of Polish*. *The crosslinguistic study of language acquisition*. D. I. Slobin. Hillsdale, N. J., Erlbaum: 595-686.

Smolensky, Paul. 1996. The Initial State and ‘Richness of the Base’. Technical Report JHU-CogSci-96-4.

Stager, Christine, and Janet Werker. 1997. Infants listen for more phonetic detail in speech perception than in word-learning tasks. *Nature* 388. 381-382.

Stites, Jessica, Katherine Demuth, and Cecilia Kirk. 2004. Markedness vs. Frequency Effect in Coda Acquisition. In Alejna Brugos, Linnea Micciulla, & Christine E. Smith (eds.), *Proceedings of the 28th Annual Boston University Conference on Language Development*, pp. 565-576.

Templin, Mildred. 1957. *Certain Language Skills in Children: Their Development and Interrelationship* (Monograph Series No. 26). Minneapolis: University of Minnesota, The Institute of Child Welfare.

Tesar, Bruce. 1995. *Computational Optimality Theory*. Ph.D. dissertation, University of Colorado, Boulder.

Tesar, Bruce, and Alan Prince. to appear. Using phonotactics to learn phonological alternations. Revised version will appear in *The Proceedings of CLS 39*, Vol. II: The Panels. ROA-620.

Tessier, Anne-Michelle. 2006. *Biases and stages in phonological acquisition*. Ph.D. dissertation, University of Massachusetts, Amherst.

Weide, Robert L. 1994. CMU Pronouncing Dictionary. http://www.speech.cs.cmu.edu/cgi-bin/cmudict.

Weist, R., & Witkowska-Stadnik, K. (1986). Basic relations in child language and the word order myth. *International Journal of Psychology*, 21, 363–381.

Weist, R., Wysocka, H., Witkowska-Stadnik, K., Buczowska, E., & Konieczna, E. (1984). The defective tense hypothesis: On the emergence of tense and aspect in child Polish. *Journal of Child Language,* 11, 347–374.

Zamuner, T.S., Gerken, L.A., and M. 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.

Zamuner, T.S., Kerkhoff, A., & Fikkert, P. (in prep.). Children's knowledge of how phonotactics and morphology interact.

Zydorowicz, Paulina. 2007. Polish Morphonotactics in First Language Acquisition. In Menz, Florian and Marcus Rheindorf (eds.), *Weiner Linguistische Gazette* 74, 24-44.

Department of Linguistics
Yale University
370 Temple St.
New Haven, CT 06511

gaja.jarosz@yale.edu