Learning Allophones: What Input Is Necessary?

Caitlin Richter

1. Introduction

The discovery of phonemes and allophonic relations, as surface alternations accumulate in children’s growing vocabulary, is a vital element of language acquisition. This paper quantitatively models allophone learning, when the child’s input triggers a merger of surface-distinct segments into a single underlying phoneme, illustrated here for the case of grouping US English alveolar flaps [r] together with voiceless alveolar stops [t] as allophones of an underlying /T/.

The Tolerance Principle (Yang, 2016), grounded in considerations of processing efficiency, defines for the general case the tipping point at which any grammar becomes unsustainable for a child, given the linguistic items that they know a grammatical rule should apply to and how many of them evidently violate it. Integrating this cognitively motivated model of grammar evaluation with a principled model for the nature of the phonological system the child will be evaluating at some non-final acquisitional stage (Ringe & Eska, 2013), it is possible to trace a portion of the development of the phonological inventory, as it is continually re-evaluated with growing vocabulary and potentially ultimately rejected in favour of another grammar. Therefore, the present model of allophone acquisition proposes firstly that the learner’s default initial grammar is roughly ‘what you hear is what you get’ in terms of the language’s surface contrasts, as specified by Invariant Transparency (an acquisitional hypothesis soundly implied by patterns of historical language change, §1.1; Ringe & Eska, 2013); and secondly, that this initial grammar is subject to the Tolerance Principle, such that if the child has too much violating input in which it is obvious that what is heard on the surface is not what is underlying then the initial grammar will become unsustainable. Input violates the initial grammar when e.g. a child hears eat [it] ~ eating [iRIN], knows both words, and understands the morphology, so they are aware how the flap in eating and stop in eat relate (compared to [v] in drive ~ driving).

Abandoning the default grammar is identical with learning an allophone; although the child of course may not immediately move to the adult grammar, if a
minimally-abstract grammar in which one surface segment corresponds exactly to one underlying segment is an impossible grammar, then any other grammar the child moves on to will necessarily be more abstract for this special case of the first move away from the Invariant Transparency default. The model therefore tracks the fundamental events in the basic shape of the child’s phonemic inventory, that is, changes in the number of different underlying forms that surface-contrastive segments are distributed into, or the initial discovery of allophones. These changes to underlying representation are, as postulated by Stoel-Gammon (2011), a function of expanding vocabulary, and the cognitive load of maintaining an increasingly inefficient phonology (Yang, 2016). The Tolerance Principle alone, applied to a full native-speaker vocabulary, could model whether a certain phonological system (one with or without a particular allophonic relationship) is possible for speakers of a language; however, finding an initial state for the learner based on Invariant Transparency increases the granularity at which the Tolerance Principle can be applied, for the first time modelling development iteratively throughout acquisition as the child’s knowledge of language changes.

1.1. Background

Alveolar flapping in North American English is an allophonic rule (De Jong, 2011) according to which intervocalic alveolar stops surface as flaps before unstressed syllables: *water* [wɔrə], *putting* [pərɪŋ], *alligator* [ælɪɡərə]. Children’s productions in flapping contexts follow a U-shaped developmental curve characteristic of rule learning: When children first show the articulatory control to differentiate [r] and [t], they produce flap-type realisations (ex. 1a-d) fairly regularly where required.

\begin{enumerate}
\item a. *Peter* that’s *Mickey Mouse* [ˈpiərə] (CHILDES: Davis/Cameron, 1;8.25)
\item b. *At the cottage?* [ˈkʌrɪʤ] (CHILDES: Weist/Emily, 2;6.06)
\item c. *Spongebob’s a kitty cat* [ˈkɪri] (CHILDES: Weist/Emily, 2;6.20)
\item d. *he’s eating strawberries* [ˈɪrɪŋ] (CHILDES: Weist/Emily, 2;9.02)
\end{enumerate}

Later in development, erroneous alveolar stops can become more prevalent in flapping contexts (ex. 2a-b), with production of [t] in obligatory flap contexts increasing in frequency from <36 months to 36–60 months (Klein & Altman, 2002).
Finally, children seem to approach competence in an adult-like system and avoid outright ungrammatical productions rather late, only around the age when they learn to read (Eddington, 2007). From there, gradual progress may continue towards full adult control over variation and optionality, in speech varieties with a complex system of flap usage (Labov, 2011).

In the classical U-shaped (or ‘non-linear’) trajectory for phonological learning, the child’s initial correct productions will be of the most common lexemes; then, pattern generalisation or rule learning occurs later on, and initiates the signature period of decreased accuracy in the child’s production (Vihman, 2013). For the case of learning the [r]∼[t] allophone, children early on produce obligatory [r] with good accuracy in a relatively small vocabulary, but after learning of an underlying relationship between [r] and [t] they gain the ability to produce errors that critically depend upon having that allophonic relationship (in addition to some further source of error, either in the grammar or in performance), such as frequently producing [t] where [r] has been the obligatory surface form in their input. The present model is concerned with this first downwards inflection in the U-shaped trajectory, its critical non-linearity. Although flap realisation for adults is variable in some words or contexts (Byrd, 1994), children’s errors of excessive stop use include definitive errors at obligatory flap contexts; furthermore, productions like (2b) and like Florida with ‘d’ [r] spoken as voiceless stop [t] (De Jong, 2011, p.2725), where [t] is produced in place of an obligatory [r] that for adults is considered underlyingly /D/, indicate clear instances of learners treating the [r] they hear as an allophone of alveolar stop /T/. Therefore, the pattern of increased stops in obligatory flap contexts after an earlier period of more accurate obligatory flap production is an instance of typical U-shaped or non-linear phonological development, with these production errors providing evidence of when an allophonic relationship must have been acquired.

Invariant transparency is a descriptive proposal from historical linguistics: learners posit abstract phonemes only when motivated to account for surface alternations in the data, and otherwise rely on projecting surface contrasts of the language’s distinctive features directly into underlying contrasts (Ringe & Eska, 2013). Although the proposal does not specify a mechanism or explanation for learners’ behaving this way, repeated and thoroughly-attested patterns of historical change indicate that something along these lines must be the case, that is, an initial surface-contrast-based ‘what you hear is what you get’ phoneme inventory becomes a default when not motivated to change by subsequent input. In particu-
lar, occurrences of secondary split\(^1\) require that prior to the visible secondary split, there must be a covert phonemic split in which ‘at least some native learners learn non-alternating outputs of phonological rules as underlying segments’ rather than merging them into one underlying form (Ringe & Eska, 2013, p.98). Invariant transparency does not identify the exact conditions of the tipping point at which linguistic input would (not) motivate a learner to posit merged abstract forms, but it is an empirically testable prediction of the proposal that there is such a point.

The independently motivated acquisitional model of Yang (2016) quantitatively defines a tipping point at which a grammar or rule is acceptable or impossible to acquire from particular input, according to the Tolerance Principle. A productive rule is cognitively tolerable if the number of exceptions to it is lower than a certain threshold, which is determined by the total number of linguistic items potentially subject to that rule. The threshold is a result of processing efficiency for the language user, given general assumptions of frequency-correlated lexical access and an Elsewhere Condition (exceptional forms processed first before broader generalisations are otherwise applied; Yang, 2016). If a potential rule or generalisation speeds processing despite some number of memorised exceptions, then it is tolerable, but if processing the available linguistic input is overall less efficient using the rule plus listed exceptions, then it will not be tolerable to learners as a productive part of their mental grammar even if it is an accurate descriptive generalisation over some of the input (Yang, 2016). While phonological learning can also be sensitive to distributional information (Peperkamp et al., 2006; White et al., 2008), this model particularly addresses the role of cognitive cost to the learner, amid their developing semantic and morphological knowledge, in shaping the emerging phonology.

The learner’s grouping of surface-contrasting segments into underlyingly connected allophones of the same phoneme, starting from the Invariant Transparency grammar of default projection of distinct surface segments into underlying contrasting phonemes, is modelled by measuring the tolerability of the default grammar given the child’s state of linguistic knowledge. With enough exceptions to the default, it can eventually be abandoned, leading to acquisition of an allophonic relationship between segments and the possibility of making allophone-aware production errors. Non-creation of allophonic relationships is modelled correspondingly; the initial phonological structure of separate phonemes like those projected from \([k]\) and \([f]\) will persist in the absence of input evidence to reject the initial grammar and merge them. The aims of the present work are to give a general model for the relationship between a child’s input and the fundamental shape of their phonemic inventory, illustrated with \([r]\sim[\text{t}]\) allophony, and to ex-

---

\(^1\)A type of phonological split so named when the original conditioning environment of some surface-contrastive rule-conditioned allophones becomes lost or obscured, yet the surface contrast persists as an underlying split in the absence of its conditioning environment - the phonemic split is observed secondary to the loss of the original allophonic conditioning (Hoenigswald, 1960).
tend the psychological granularity at which Tolerance Principle is applied. Being essentially a model of downwards inflection points for U-shaped allophone learning trajectories, this is not a complete model of development towards a complex adult phonological grammar; the model quantitatively predicts whether surface-distinct segments are underlyingly contrastive or allophonic given current input, and thereby captures the critical points at which new allophones are acquired and the phonemic inventory shape changes. The ‘initial’ grammar modelled here is of course not the child’s very earliest state of (proto)-phonological knowledge, as it requires already sorting out some linguistic system of surface contrasts from highly variable acoustic input, but Invariant Transparency posits that this type of surface-based phonological grammar forms part of a basic developmental pathway (Ringe & Eska, 2013), modelled here as a relatively early stage. The Tolerance Principle’s predictions for productivity of generalisations given input data have been rigorously validated in empirical case studies spanning generalisations of phonology, morphology, and syntax (Yang, 2016); for example, individual children’s (non)acquisition of productive morphology in artificial language learning is as quantitatively predicted based on the composition of their personally learned vocabulary items (Schuler et al., 2016). While previously the Tolerance Principle has been applied only as a one-time (or time-abstracted) model, to evaluate learnability of grammars once the period of acquisition is complete, the introduction of Invariant Transparency means that the Tolerance Principle can now be applied iteratively to evaluate a hypothesised grammar through earlier acquisition stages and predict grammatical changes parallel to those indicated by children’s productions during development. This initial implementation of the finer-grained progressive model for allophone learning is validated mainly by qualitative comparison to the non-linear path of acquisition.

2. English alveolar flapping

Intervocalic flapping is used here to name a phonological phenomenon found in North American English, whereby /T/ and /D/ in certain environments are realised as an alveolar flap [r] (which is itself slightly perceptually confusable with [d] for English speakers; Eddington, 2007; Herd et al., 2013). For this paper, an intervocalic flapping rule will be defined as follows:

(3) Intervocalic /T/ and /D/ are realised as [r] when preceding an unstressed vowel.

Alveolar flapping has received decades of attention from several subfields of linguistics due to its prevalence across dialects, frequency in speech, and variation along many dimensions (phonetic, social, etc.) making it seem an attractive test case for general linguistic hypotheses. However, a substantial portion of the research to date has actually sought to determine the basic nature of alveolar flapping itself: despite the initial promise of this ubiquitous and seemingly categorical
phenomenon, it is characterised by variation at all levels and as a study object it proves tricky to define and work with. Here I briefly sketch the research conclusions necessary to give this application of the acquisitional model, which is fully general to phonemic inventory learning and instantiated here with flapping for a concrete illustration, sufficient grounding in the empirical facts of the phenomenon.

2.1. Phonology

A central concern for this modelling endeavour is the status of flaps as genuinely allophonic (a categorical phenomenon), rather than a non-phonological epiphenomenon such as merely one end of a continuum of articulatory reduction. The literature emerging from a variety of methodologies and theoretical interests converges on the assessment that English flapping does involve true flap allophones of alveolar stops. De Jong (2011) concludes that despite recent attempts (including his own; 1998) to explain flapping as a non-categorical byproduct of gradient phenomena, in the end there is no coherent relationship between articulatory, acoustic, and perceptual facts that manages to avoid treating ‘full’ /T/ and flapped /T/ as some kind of categorically distinct linguistic objects (allophones). Additionally, style-based patterns of flap variation contrast in several ways with the style-based patterns of non-phonemic reduction phenomena that are observed in the production of other English consonants (Warner et al., 2009; Warner, 2005), further differentiating the flap from non-phonemic reduction. A relative consensus has also emerged that flapping does not neutralise the distinction between /T/ and /D/, although the differentiation of these underlying consonants is chiefly expressed in the sonorants that precede them rather than being expressed in the flap itself. It has been more difficult to find objective measures to tell whether any given alveolar realisation (/T/ or /D/) is a flapped realisation or not; however, this may confound linguists (viewing acoustic and articulatory data) more than language-learners. Distinctions among these sounds are most often investigated by recording word pairs like *kitty*/kiddie* or *putting*/pudding*, and measuring a variety of acoustic and articulatory parameters (Fox and Terbeek, 1977; Herd et al., 2010; De Jong, 1998; Malécot and Lloyd, 1968; Zue and Laferriere, 1979). Many of these studies include perceptual experiments on whether people reliably perceive underlying differences, and results are mixed; however, relatively few experimental paradigms have so far been used to address this question, and all of them use isolated word presentation in artificial situations like a forced-choice task (Braver, 2014; Riehl, 2003; Warner and Tucker, 2011). Referring to complex articulatory and acoustic data, Fukaya and Byrd (2005), echoed by Son (2008), suggest that a constellation of individually gradient articulatory phenomena gives rise to perceptually quantised acoustic signals that listeners categorically interpret as either full /T,D/ stop allophones or flap allophones.
Ultimately, multiple lines of evidence seem to support a categorical mental construct of some nature, in which flapped /T/ and /D/ are representationally distinct from non-flapped /T/ and /D/. This justifies the appropriateness of the present model treating [t] and [r] as mentally distinct segments (whether allophones or separate phonemes, according to developmental progress), not merely varied phonetic expressions of a single mental object.

2.2. Acquisition

The acquisitional problem of North American English word-medial alveolar stop allophones poses substantial challenges to learners (Klein et al., 1998), as there are a variety of conditioned allophones, two phonemes to learn, and a rapid articulatory gesture needed to produce the flap allophones. A progressive model of flap acquisition requires an accurate enough picture of the course of children’s flap development to be faithful to.

Klein et al. (1998) and Song et al. (2015) present valuable time-course data showing the acquisitional progress of children learning to use alveolar stop allophones. Song et al. (2015) look at earliest stages of alveolar stop pronunciation from first words through 2-year-olds; children did not produce any flaps until near the upper age bounds of the study. Rimac & Smith (1984) compare the acoustic detail of older children’s flaps to adults’ flaps in putting/pudding pairs, measuring durations of the consonants and of the preceding vowels. Children had consistently longer flap closure durations than adults in absolute terms, but there were relative differences in preceding vowel length from /T/ to /D/ flaps in both age groups. Therefore, even at a stage when children differ from adults in some details of pronunciation (especially the quick articulation of flapping), they are already systematically sensitive to the differential effects of underlying voicing from /T/ and /D/ when producing respective flaps.

Klein et al. (1998) explore how data for children’s articulations should best be coded, and directly compare the value of perceptual and acoustic coding for English intervocalic flap. They argue that while perceptual judgements can vary by person and be subject to perceptual bias, perceptual transcription for alveolar variants is still superior to categorical coding by supposedly ‘objective’ acoustic/spectral measures: the acoustic measures are selected circularly for their tendency to correlate with perceptual judgement, while any basic algorithmic combination of them fails to categorise flap allophones as well as natural human perception does (Klein et al., 1998). In their own acquisitional study of child pronunciations, Klein & Altman (2002) track 15 total variants for alveolar stops (some of which are not adult target allophones for any alveolar stop, but occur in child data), with perceptual coding. This production study follows 4 children ages 2-5, eliciting word-medial alveolar stop allophones in 2-syllable words; even by age 5 none had achieved an overall adult-like pattern of alveolar stop allophone competence.
Figure 1: Developmental course of productions in obligatory alveolar flap contexts by one child (CHILDES: Providence/Lily)

A sample development course of flap production from one child (CHILDES: Providence/Lily) ages 1;4 to 4;4 is presented in Figure 1, based on perceptual coding of flap tokens for 25 high-frequency intervocalic obligatory flap words, as correct (flap) or incorrect (something else). Data was coded by a single judge; while the absolute percentages of flapping perceived in each age interval might vary slightly with different judges, the relative shape of the child’s flap accuracy changing across time should be faithfully represented by perceptions of a single judge, and for present purposes this contour is the most important feature of the data in Figure 1. The flap is produced with decent accuracy in the earliest recordings while articulation is still an active challenge, then produced with high accuracy, before abruptly dropping to an accuracy of only 51% of flaps produced in obligatory flap contexts at age 3, and trending gradually somewhat upwards from there. The productions summarised in Figure 1 strikingly illustrate the non-linearity described by the U-shaped trajectory of phonological rule learning, indicating that by the marked drop around age 3 (that has not been fully recovered from by 4;4) this child had learned an allophonic relationship involving the flap.

3. Methods
3.1. Formal model

The formal model identifies whether a child’s current vocabulary motivates acquisition of an allophonic relationship between distinct surface segments [r] and
The tolerance principle (Yang, 2016) describes the number of exceptions that a grammatical rule can tolerate, relative to the number of linguistic items subject to the rule, in order to exist at all as a productive rule in a grammar. A rule \( R \) is a statement of the form

\[
\text{(4) } R: \text{If } A \text{ then } B
\]

where \( A \) is a ‘structural description’ of the trigger for applying the linguistic transformation \( B \), such as the environment \textit{intervocalic preceding an unstressed vowel} in (3) triggering transformation of an underlying alveolar phoneme into surface expression \([r]\). Then, ‘If \( R \) is a productive rule applicable to \( N \) candidates, then the following relation holds between \( N \) and \( e \), the number of exceptions that could but do not follow \( R \):’

\[
\text{(5) } e < \theta_N \text{ where } \theta_N := \frac{N}{\ln(N)}
\]

If the given relation does not hold between \( N \) and \( e \) - there are too many exceptions whose form matches the structural description \( A \) to trigger \( R \) but clearly violates the expected outcome \( B \) of its application - then \( R \) would not be cognitively efficient and cannot be a productive rule.

The principle of invariant transparency, as formulated by Ringe and Eska (2013), proposes that:

\[
\text{(6) } \text{‘Native learners in the developmental window for NLA do not posit abstract forms if there is no alternation to account for; instead they project surface segments, defined by the language’s system of distinctive features, into underlying forms.’}
\]

For the case of alveolar flapping, I interpret Invariant Transparency to mean that the learner’s default grammatical hypothesis, projecting surface segments to underlying forms, is the following rule:

\[
\text{(7) } H_0: \text{If } \text{surface flap then underlying flap}
\]

This rule applies from the perspective of a listener/comprehender; that is, the learner whose grammar contains \( H_0 \) considers every \([r]\) they hear to be a candidate subject to \( H_0 \). The formula of the Tolerance Principle calculates when the child will abandon this \( H_0 \), after exposure to a sufficiently pervasive pattern of surface alternations, as the number of exceptions to \( H_0 \) becomes too great relative to the number of potential applications \( N \). An exception \( e \) to \( H_0 \) is counted when the learner has morphological and semantic evidence to motivate understanding a candidate with surface \([r]\) as underlingly not \([t]\). A child whose vocabulary includes both \textit{eat} \([it]\) and \textit{eating} \([i\eta]\), and knows how they are related by the suffix \textit{-ing}, must memorise the surface \([t]\) and \([r]\) alternation as an exception to their rule.
H0, for example by storing duplicate forms /it/ and /it/ in their lexicon in violation of mutual exclusivity. (A grammar like adults’ incorporating one phoneme with appropriately conditioned allophones, not needing duplication in the lexicon, naturally avoids this cost of violating mutual exclusivity.) More exceptions than are tolerable will motivate the learner to abandon H0 and proceed with some other, necessarily more abstract, grammar.

The present model is concerned only with identifying the tipping point at which the learner finds sufficient evidence to discard Invariant Transparency’s H0, not what happens subsequently. The (‘highly individual and notoriously variable’; Vihman, 2013, p.269) paths by which children, over several years, reach full adult competence of alveolar allophones and phonologically conditioned flapping will inevitably involve somehow proposing and evaluating quite a number of hypotheses, and the hypothesised grammar generation/selection mechanism(s) are not treated in this model other than the initial special case specified by Invariant Transparency. The tipping point for Invariant Transparency-based hypotheses with the form of (7) H0 is of central interest because whether or not this point it is reached - whether there are too many exceptions for H0 to be tolerable, or not - pinpoints the critical factor in the shape of the phonemic inventory, whether surface-distinct segments are distinct phonemes or are merged into one underlying form.

### 3.2. Data

The input available to the learner during acquisition is approximated according to word frequency in about 3 million English words of the CHILDES corpus (MacWhinney, 2000). Language knowledge at eight aquisitional ‘stages’ was excerpted from this corpus by taking the 50 most frequent words to represent the child’s knowledge at an early stage, the 100 most frequent words as a somewhat later stage, and so on in varying intervals up to 1500 words as the latest stage modelled. Naturally this is a rough approximation, and each arbitrary ‘stage’ does not represent any particular aquisitional benchmark, but relative comparisons are justifiable; certainly the 500-word stage models a child further along in their language acquisition than the 100-word stage.

The CMU Pronouncing Dictionary (Weide, 1998) served as the source of phonetic input for this model, with pronunciations represented by the 39-phoneme ARPAbet (which has one T covering all /T/ allophones including the flap) and three levels of lexical stress marked on vowels (primary, secondary, unstressed). The North American English flapping rule is generally described with reference only to a binary division of ‘stressed’ and ‘unstressed’ syllables. Treating CMU secondary stress as unstressed for purposes of computing the flapping rule in obligatory flapping words yielded accurate pronunciations for the most frequent 1500 words.
3.3. Procedure

The acquisitional model operates by assuming an early stage where surface-contrastive segments are known to the learner and not underlyingly merged, evaluating the pervasiveness of exceptions for this phonology relative to compatible flaps given the child’s input, and requiring instead a merged grammar (with allophone) if the unmerged phonology is evaluated as cognitively unacceptable at some stage of input.

To model the evaluation of whether Invariant Transparency’s H0 for [r] is tolerable to a learner, based on the total number of candidate lexemes subject to H0 and the number of exceptions, each word in the child’s linguistic knowledge at the various acquisitional ‘stages’ is processed and added as appropriate to counts for:

\[ N: \text{ words that H0 is applicable to, that is, words containing surface flap} \]
\[ e: \text{ exceptions; H0 is applicable but [r] seems underlyingly something else} \]

and used in the Tolerance Principle formula (5). Coding exceptions \( e \) is based on the child’s entire vocabulary at a given stage. Counting an exception for a word with a surface flap requires that the child also know a corresponding non-flapped word form, and can be reasonably assumed to understand the morphological relationship between the two forms. For example, while the word daddy \([dæri]\) is present from early vocabulary stages, it is not counted as an exception until much later when dad \([dæd]\) enters the vocabulary. Without a semantically and morphologically connected word giving specific evidence for an exception to H0, daddy in isolation is compatible with H0 and incurs no cognitive cost, as the child’s lexicon can store a flap in the sole entry for the word without violating mutual exclusivity.

4. Results

Table 1 models tolerability of the default rule H0 as vocabulary grows and exceptions accumulate.

The rule is tolerable early on, but by 750 words the learner has crossed a tipping point where H0 is rejected, due to accumulation of sufficient \([t]\sim[r]\) alternations relative to the total instances of \([r]\). The relevant words \( N \) for this stage are given in (8) with exceptions \( e \) starred.

\[
\]
Table 1: Tolerability of default rule $<[r] \text{ IS UNDERLYING FLAP}>$ through vocabulary stages.

<table>
<thead>
<tr>
<th>Vocabulary</th>
<th>Flap words</th>
<th>Exceptions to H0</th>
<th>Tolerable exceptions</th>
<th>H0 tolerable?</th>
</tr>
</thead>
<tbody>
<tr>
<td>50</td>
<td>0</td>
<td>0</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td>100</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>yes</td>
</tr>
<tr>
<td>200</td>
<td>7</td>
<td>1</td>
<td>3.60</td>
<td>yes</td>
</tr>
<tr>
<td>300</td>
<td>11</td>
<td>3</td>
<td>4.59</td>
<td>yes</td>
</tr>
<tr>
<td>500</td>
<td>22</td>
<td>6</td>
<td>7.11</td>
<td>yes</td>
</tr>
<tr>
<td>750</td>
<td>33</td>
<td>11</td>
<td>9.44</td>
<td>NO</td>
</tr>
<tr>
<td>1000</td>
<td>47</td>
<td>16</td>
<td>12.21</td>
<td>NO</td>
</tr>
<tr>
<td>1500</td>
<td>74</td>
<td>24</td>
<td>17.19</td>
<td>NO</td>
</tr>
</tbody>
</table>

Earlier than the tipping point, children posit an isolated underlying flap phoneme in accordance with H0, while once H0 is not tolerable in the grammar, children must recognise an allophonic relationship where surface $[r]$ can systematically be underlingly something else (perhaps $/T/$). This accords with the familiar U-shaped curve of phonological acquisition, in which an an early period of relatively successful flap production is followed by an apparent decline in linguistic performance as children’s new allophonic knowledge enables them to produce errors they were not previously capable of (Vihman, 2013).

5. Discussion

The model for allophone learning developed in this paper has qualitatively traced the acquisition course of $[r]$ allophone development, in accordance with flap production patterns and with interpretable cognitive motivation, while extending the psychological granularity at which the Tolerance Principle can be applied.

This model describes phoneme learning as guided by simultaneous development of vocabulary and other linguistic knowledge, as appropriate for children with the vocabulary sizes modelled (Stoel-Gammon, 2011). In particular, the alternations-based trigger for ‘exceptions’ views phonological acquisition in the context of semantic, lexical, and morphological development (Brown, 1973; Bergelson & Swingley, 2015); here, principal contributors to triggers for flap allophone discovery seem to be the verbal lexicon and morphology, e.g. -$\text{ing}$, -$\text{n’t}$.

Sensitivity of developing phonology to surface alternations is also directly evident in child production data: Examining error production patterns in children up to 5 years, Klein & Altman (2002) find that obligatory flap is absent most often in the morphological alternation context -$\text{ding}$ (reading, riding, feeding), while it is most consistently present in the rarely morphologically alternating context -$\text{ty}$ (Betty, dirty, kitty). These production tendencies complement this model’s interpretation by showing cognitive relevance of surface alternations for learners during this stage of development.
What has been proposed is a general, framework-independent, phonemic inventory learning model, with no strong commitment to any particular ‘rule’ formalism. While the present illustration of flap learning includes some formal commitments for the sake of a concrete implementation, in principle any inductive learning scenarios may be evaluable under the Tolerance Principle (Yang, 2016), and Invariant Transparency’s default projection of surface segments into underlying forms (6) may likewise be represented in a wide range of phonological formalisms (Ringe & Eska, 2013). The model requires only initially projecting surface contrasts as underlying contrasts, retaining this grammar (however it is represented) as long as tolerable, or rejecting it if motivated by reaching an input threshold for obviously incompatible surface alternations. It yields a quantitative prediction of a critical factor in the shape of a phonemic inventory: whether surface-distinct segments are underlyingly contrastive or allophonic.

References


