Title: Using the Tolerance Principle to predict phonological change

Authors:

Betsy Sneller[1], Josef Fruehwald[2], and Charles Yang[3]

[1]Georgetown University [2]University of Edinburgh [3]University of Pennsylvania

Corresponding author:

Betsy Sneller

Department of Neurology

Building D, Suite 145

4000 Reservoir Road, N.W.

Washington D.C. 20057s

ABSTRACT

Language acquisition is a well-established avenue for language change (Labov, 2007). Given the theoretical importance of language acquisition to language change, it is all the more important to formulate clear theories of transmission-based change. In this paper, we provide a simulation method designed to test the plausibility of different possible transmission-based changes, using the Tolerance Principle (Yang, 2016) to determine precise points at which different possible changes may become plausible for children acquiring language. We apply this method to a case study of a complex change currently in progress: the allophonic restructuring of /æ/ in Philadelphia English. Using this model, we are able to evaluate several competing explanations of the ongoing change and determine that the allophonic restructuring of /æ/ in Philadelphia English is mostly likely the result of children acquiring language from mixed dialect input, consisting of approximately 40% input from speakers with a nasal /æ/ split. We show that applying our simulation to a phonological change allows us to make precise quantitative predications about the progress of this change. Moreover, it forces us to reassess intuitively plausible hypotheses about language change, such as grammatical simplification, in a quantitative and independently motivated framework of acquisition.

ACKNOWLEDGEMENTS

We would like to thank the audiences at MFM 2016 in Manchester, FWAV 2016 in New York City, and NWAV 2016 in Vancouver for their feedback on earlier versions of this project. The project has benefitted greatly from comments by the editor and three anonymous reviewers. This research was funded in part by National Science Foundation grants BCS-1251437 and BCS-162.

INTRODUCTION

One of the major loci for language change is in the transmission of language from caregiver to child. While most transmission results in a faithful replication of a child's input, language change may occur when a child imperfectly replicates the structure or form of their caregiver's input (Labov, 2007: 346). The original formulation of change via transmission in Labov (2007) proposes two causes of unfaithful replication. The first occurs when a child posits an innovative analysis of their input, thereby reanalyzing the input from their community into a new structural representation. Endogenous innovation may be affected by factors such as articulatory pressure, perception biases, social factors, grammatical simplification, or simply stochastic variability. The second type of change in transmission occurs when the input from a child's community is itself variable, as in cases of dialect contact and subsequent dialect leveling (e.g., Kerswill, 2000; Trudgill, 2004;). In this case, a child's input consists of several competing structural features, and that child must posit a grammar that can be maximally consistent with the heterogeneous data in their input, typically resulting in a dialect with new and levelled structural features.

Given the theoretical importance of language acquisition to language change, it is all the more important to formulate clear theories of transmission-based change. What grammatical changes are possible through parent-to-child transmission? How does variation in the input affect the formulation of a child's grammar? We argue that the application of rigorously formulated principles of language acquisition to the process of language change can begin to address these questions in a precise way, allowing specific quantitative predictions about language change to be made (e.g., Niyogi & Berwick, 1996; Yang, 2002).

In this paper we provide a computational framework designed to test the plausibility of different possible transmission-based changes, using the Tolerance Principle (Yang, 2016) to determine precise points at which different possible changes may become plausible for children acquiring language. We apply this method to a case study of a complex change currently in progress: the allophonic restructuring of /æ/ in Philadelphia English. Using this model, we are able to evaluate several competing explanations of the ongoing change and determine that the allophonic restructuring of /æ/ in Philadelphia English is mostly likely the result of children acquiring language from mixed dialect input, consisting of approximately 40% input from speakers with a nasal /æ/ split. We show that applying the Tolerance Principle to a phonological change allows us to make precise quantitative predications about the progress of this change. Moreover, it forces us to reassess intuitively plausible hypotheses about motivations for language change, such as grammatical simplification, in a quantitative and independently motivated framework of acquisition.

We begin by providing some background information on the allophonic change in /æ/ currently in progress, where a new nasal split for /æ/ is replacing the traditional allophonic rule. Section three describes the Tolerance Principle, which proposes a quantitative heuristic for the conditions under which a productive rule is acquirable. Section four introduces a simple and general model of rule learning under the Tolerance Principle given heterogenous input. We provide computational simulations which suggest that it is highly implausible for the allophonic change to endogenously emerge in Philadelphia as the result of child learners incrementally simplifying the traditional /æ/ system during language acquisition. In section five, we apply the model to contact-induced change. The model predicts that contact-induced change is the most plausible scenario for the course of this specific change, which is supported by preliminary evidence from the phonetic and demographic studies of the Philadelphia community. Throughout the paper, we demonstrate that the application of rigorously formulated theory makes clear, quantitative predictions about language change which can in turn be tested empirically.

RESTRUCTURING OF SHORT-A IN PHILADELPHIA

In this section, we briefly discuss the allophonic restructuring currently underway in Philadelphia. For a more in-depth analysis of the change with special reference to its current status, we refer the reader to Labov, Fisher, Gylfadottir, Henderson, and Sneller (2016) and Sneller (2018).

Traditional Short-a in Philadelphia

Like many older dialects of the Northeast and even the Midwest (Durian 2012, 256-277), Philadelphia English traditionally exhibits a division in its /æ/ phoneme into two forms which we will refer to as "tense" and "lax". The lax target is a nonperipheral low front vowel [æ] which is relatively short in duration (avg: 119 ms in the Philadelphia Neighborhood Corpus). The tense target is both fronted and raised, exhibits a longer average duration (130 ms.), and is often produced with an inglide ([æᵊ], [ɛᵊ], [eᵊ], or [iᵊ]). The traditional short-a split in Philadelphia is governed by a regular phonological rule, shown in (1) and henceforth referred to as PHL.

$$\text{(1) PHL: } æ \rightarrow \text{æ̞} / \_\_[+\text{ant}] \wedge \left([+\text{nasal}] \vee \begin{bmatrix} -\text{voice} \\ +\text{fricative} \end{bmatrix}\right)]_\sigma$$

Here, PHL is stated as a complex rule triggered by a disjunctive set of phonological conditions: /æ/ is tensed when it precedes nasals or voiceless fricatives which are also anterior and tautosyllabic. In other words, /æ/ is tensed when it precedes a tautosyllabic {m, n, f, θ, s}. This rule is best described as operating only at the stem level, following Bermúdez-Otero (2007), producing tense /æ/ in *ham* ([hæm]) and *plan+ing* ([[plæn]ɪŋ]) but lax /æ/ in *hammer* ([hæ.mɚ]). In addition to the regular PHL rule described in (1), short-a in Philadelphia also has developed some lexical specificity, with three affective adjectives exceptionally produced as tense (*mad, bad, glad;* c.f. lax *sad*), and a number of additional words exceptionally produced as lax[1]. The lexical specificity of this rule has also been subject to diachronic change, with *planet* becoming a member of the exceptionally tense class for some but not all speakers born around 2000 (Brody, 2011; Sneller, 2018). The complicated nature of PHL and the existence of lexical exceptions has prompted some analyses to describe this split as two distinct phonemes (e.g., Ferguson, 1972; Labov, 1989; Labov, Ash, & Boberg, 2006). However, recent work (e.g., Labov et al., 2016; Sneller, 2018) agrees with the position of Kiparsky (1995), finding evidence that speakers of Philadelphia English born after 1985 treat this split as an allophonic distinction rather than a phonemic one.

Incoming Nasal Split

The traditional PHL short-a split in Philadelphia is currently being overtaken by the geographically widespread and phonologically simpler nasal short-a rule, shown in (2) and henceforth referred to as NAS.

$$\text{(2) NAS: æ} \rightarrow \text{æh} / \_\_[+\text{nasal}]$$

As shown in (2), NAS is a simple allophonic rule, producing tense /æ/ before any nasal segment. NAS can be found in many areas of North America, including New Haven (Johnson, 1998), the

Midland region (Boberg & Strassel, 2000), Ohio (Durian, 2012), Indiana (Fogle, 2008), the St. Louis Corridor (Friedman, 2014), New York City (Becker & Wong, 2009), the West Coast (Hall-Lew, Cardoso, Kemenchedjieva, Wilson, Purse, & Saigusa, 2010), Michigan (Wagner, Nesbitt, & Savage, 2016), and New Jersey (Ash, 2002). In Philadelphia, tense /æ/ produced by NAS is phonetically very similar to the tense /æ/ produced by PHL. In both cases, the tense allophone has a longer duration than the lax allophone (187 ms. versus 139 ms. for NAS speakers in the Philadelphia Neighborhood Corpus), is raised along the front periphery, and is typically produced with an inglide. Figure 1 displays the *z*-score normalized productions of a PHL speaker (left) and a NAS speaker (right), demonstrating that the phonetic characteristics of both the tense and the lax allophones are similar for both allophonic rules.

Figure 1 about here

Labov et al. (2016) demonstrated that the prevalence of NAS has risen quickly across the Philadelphia community, beginning with speakers born after 1985, and that NAS is particularly prevalent amongst graduates of elite public high schools. In an analysis of three Philadelphia families, Fisher, Prichard, and Sneller (2015) furthermore demonstrated that this change occurs over the course of two generations, with the older generation producing PHL, and the youngest generation producing NAS. The intermediate generation produces what Fisher et al. (2015) termed "weak PHL" and Sneller (2018) analyzed as variable production of both systems. Given the wealth of sociolinguistic data demonstrating the existence of NAS in Philadelphia, we focus here on the question of *how* NAS came to exist in the speech of younger Philadelphians. Given that NAS is a formally simpler allophonic rule that can be described as a featural subset of PHL, we first investigate whether children were likely, according to the Tolerance Principle, to spontaneously reanalyze PHL as NAS given entirely PHL input. We begin by reviewing the features of the Tolerance Principle which are most relevant to this application of it.

THE TOLERANCE PRINCIPLE

As a model of language acquisition, Yang (2016) outlined a principle that determines the potential productivity of a rule given a set of input. This Tolerance Principle is shown in (3).

(3) **Tolerance Principle:**

Let *R* be a rule that is applicable to *N* items, of which *e* are exceptions. *R* is productive if and only if:

$$e \leq \theta_N \text{where} \theta_N := \frac{N}{\ln N}$$

The Tolerance Principle states that a rule is productive if the number of exceptions to that rule is less than the number of items the rule could apply to divided by the natural log of the number of items. For example, let's assume that a child has 10 verbs in their vocabulary. Some of these verbs take the regular –(*e)d* suffix to form a past tense (*walk, smile*), while other verbs in their vocabulary are exceptions to this regular rule (*run, fall*). The Tolerance Principle states that the regular past tense *–(e)d* rule can be productive for this child if their vocabulary has fewer than 10/ln(10), or 4.3, exceptions to the rule. In other words, if the child's vocabulary contains 4 or fewer irregular past tense verbs, then the regular past tense *–(e)d* rule can be a productive rule in their language.

It is important to stress that the Tolerance Principle applies over word types rather than tokens. This follows from its basis in the Paninian Elsewhere Condition (Anderson & Kiparsky, 1973): Functionally, it describes a speaker's linguistic processing as first traversing the exceptions to a rule before applying the Elsewhere Condition–the productive rule. These exceptions are listed in rank order of lexical frequency, and the time cost of traversing this list is +1 time unit per item listed. The Tolerance threshold identifies the point at which there are so many exceptions to the rule that it becomes more time efficient for a speaker to simply memorize all words. The reader is referred to Yang (2016: Chapter 3) for a review of reaction time studies that support the Elsewhere Condition as a model of morphological processing. This formulation means that token frequency only affects a lexical item's location in the list of exceptions, but does not affect the calculation of whether a rule is productive or not, since productivity is calculated over the number of items in said list. This means that despite robust evidence that word frequency is an important factor in language processing (Goldinger, 1998; Grainger, 1990; Segui, Mehler, Frauenfelder, & Morton, 1982), it does not play a role in the calculation of the productivity of a rule. While this follows from the architectural motivation of the Tolerance Principle, the conclusion that productivity is a function of type frequencies is broadly shared across different theoretical frameworks such as generative morphology (Aronoff, 1976), connectionist modelling (Plunkett & Marchman, 1991), corpus linguistics (Baayen & Renouf, 2018), and usage-based approaches (Bybee, 1995; Pierrehumbert, 2003). It predicts that a child would be able to learn a productive rule as long as the word types in their vocabulary fit the Tolerance Principle, regardless of the token frequencies of these words. For a full description of the Tolerance Principle and its derivation, we refer the reader to Yang (2016).

Here we highlight a key feature of the Tolerance Principle that is especially relevant for the present study. The threshold for exceptions is, perhaps for most readers, surprisingly high. Table 1 provides a range of values of N and the maximum number of exceptions that a rule defined over N items can tolerate.

Table 1 about here

It is clear that, as a proportion of $\underline{N}$, the tolerable number of exceptions decreases as $\underline{N}$ increases. This suggests that productive rules are relatively *easier* to learn when the learner has a *smaller* vocabulary, a conclusion that may have significant implications for the difference between child and adult language acquisition. In what follows, we simply assume the Tolerance Principle and use it to evaluate the viability of phonological rules in the face of exceptions. As we make clear in our methods, the general approach undertaken here of simulating grammar acquisition in a context of mixed input can be adapted to incorporate any theory of learning that, like the Tolerance Principle, provides a formal accounting of exceptions to a productive rule.

USING THE TP TO EVALUATE THE PLAUSIBILITY OF ENDOGENOUS CHANGE

A child acquiring their native language faces a monumental task. Given a wide range of phonetic input, they need to distinguish phonemic categories from their phonetic distribution as well as determine the allophonic rules that apply to each phoneme. This process of acquisition may result in children innovating new rules and generalizations to account for their input, resulting in transmission-based language change. Because PHL is not a surface-true generalization in the production of a traditional Philadelphia English speaker, we may expect that it will be particularly susceptible to a child reanalyzing or simplifying it. In this section, we use the Tolerance Principle to test the plausibility that a child receiving traditional Philadelphia English input will posit a NAS rule.

Can NAS tolerate PHL input in one fell swoop?

To apply the Tolerance Principle to the changing short-a systems in Philadelphia, we begin by asking whether a child receiving /æ/ input generated from the traditional PHL system could plausibly posit a NAS rule for that input. Here, we assume a child is receiving input generated only by the traditional PHL system, with its disjunctive featural specification, syllabic sensitivity, and lexical exceptions. This learner could possibly hypothesize that their target grammar is simply (2), tense before nasals, producing tense /æ/ in *ham, man*, etc. (e.g., via the inductive learning models referred to earlier). If they do so, they must somehow account for words they acquire that violate this generalization, like lax /æ/ in *bang* or tense /æ/ in *last*. If they maintain the generalization in (2), the child must treat these and all other words that violate the "tense before nasals" generalization as stored lexical exceptions. If the number of such exceptions (*e*) is lower than the Tolerance threshold for their vocabulary size, then it is plausible that learners in Philadelphia could endogenously hypothesize a NAS grammar given only PHL input. However, if the number of exceptions exceeds the Tolerance threshold, then some other source of the NAS grammar in Philadelphia must be sought. In this case, *N* will be the entire set of /æ/ words in a child's vocabulary, and *e* will be the list of words that violate *R*, where *R* = NAS.

To obtain these numbers, we use the CHILDES database (MacWhinney, 2000) to obtain a measure of the total *N* for a child's vocabulary. This database includes both child production data as well as caregiver production data, which provides an approximate view into the linguistic input given to a child. It doesn't matter for our purposes that the parents in the CHILDES database are not from the dialect area under consideration, as our goal is simply to obtain the statistical distribution of /æ/ words in a typical child vocabulary, rather than to measure any phonetic realization of these words. The CHILDES database contains $N = 1412$ word types containing at least one /æ/. Each word type was coded for its realization under the traditional Philadelphian input, under R = PHL, and under R = NAS. An example is shown in Table 2. Note that the mismatch between traditional input and PHL for *bad* reflects the fact that *bad* must be treated as a lexical exception, while PHL only captures the regular phonological generalization.

Table 2 about here

Using this coding system, we can then measure the total number of exceptions produced by positing either PHL or NAS as a rule. Using Table 2 as a dummy lexicon with $N = 5$ words, we can see that a child positing R = PHL would have to list $e = 1$ exception to that rule, because the expected realization of *bad* under R = PHL does not match the child's input. Because $1 \leq 3.11 = \frac{5}{\ln(5)}$, PHL emerges as a plausible rule for this dummy language. By contrast, a child positing R = NAS would have to list $e = 4$ exceptions, which does not pass the tolerance threshold of 3.11, rendering NAS an unproductive rule for the dummy language in Table 2.

Using the full list of /æ/ word types in CHILDES, we calculated whether the number of exceptions a child would need to list under R = PHL and R = NAS would pass the tolerance threshold of $e \leq 194.7 = \frac{1412}{\ln(1412)}$. We find that given the traditional Philadelphian input, a child positing R = PHL would have to store $e = 39$ lexical exceptions (mostly *mad, bad, glad*, strong verbs, and function words), well under the Tolerance threshold of 194.7. This, of course, is expected: children have been successfully acquiring PHL and its listed exceptions for well over 100 years (Labov et al., 2016; Labov, Rosenfelder, & Fruehwald, 2013). Turning to the question of whether NAS can be a productive rule given traditional input, we find that positing R = NAS requires listing a total of 324 exceptions (e.g., all tense /æ/ before anterior voiceless fricatives, all lax /æ/ before posterior nasals), well over the tolerance threshold.

Thus, despite being a formally simpler rule, and in fact a featural subset of PHL, NAS does not emerge as a plausible endogenous innovation of traditional Philadelphian input, using the Tolerance Principle as a metric of productivity. Positing NAS simply requires storing too many lexical exceptions for it to be productive.

Can NAS replace PHL incrementally over time?

It remains, however, that NAS is rapidly replacing PHL as the dominant allophonic rule for /æ/ in Philadelphia. Given the finding that *R* = NAS is not a plausible re-analysis of the traditional input, we can now turn to the question of incremental re-analysis. In other words, we ask whether it is possible that a child might posit an intermediate rule given traditional input, which might then be re-analyzed as a productive NAS rule by the subsequent generation of language learners. We take PHL, reproduced in (4), and break it down into its four constituent phonological components. *R* = PHL can be spelled out as /æ/ becomes tense when it precedes a stem-level (a) tautosyllabic (b) anterior (c) nasal or (d) voiceless fricative.

$$\text{(4) PHL: æ} \rightarrow \text{æh} / \_\_[+\text{ant}] \wedge ([+\text{nasal}] \vee [-\text{voice} + \text{fricative}])]_\sigma$$

At this point it is necessary to make a brief comment about the classification of lemmas, since the Tolerance Principle is calculated over word types rather than tokens. Because there is robust evidence that children acquire productive suffixes for plural, comparative, present tense, adjectival -*y* and diminutive fairly early (Brown, 1973), we take the perspective of a young language leaner and assume that words with these suffixes are classified under their stem-level lemma. The productive use of suffixes such as *-ify*, and those that involved learned vocabulary items generally, are not acquired until school age (Jarmulowicz, 2002; Tyler & Nagy, 1989). Therefore, we consider *class* and *classes* to belong to a single lemma, but *classify* to be distinct, as befitting the grammar of a young child.

Using the four main components of PHL, we construct six intermediate grammars between full PHL and NAS, beginning with excluding one component of PHL at a time and ending with excluding two components of PHL. We do not analyze intermediate forms of PHL that exclude the nasal trigger, since this would not produce an intermediate rule between PHL and NAS; NAS being the result of excluding every aspect of PHL except the nasal constraint. In Table 3, these intermediate rules are described as PHL minus the components that have been excluded. We note that some intermediate rules result in an expansion of the set of triggering segments (as in PHL-ant), while others result in a reduction (as in PHL-fric). The set of triggering phonological contexts resulting from each intermediate rule is shown in the third column of Table 3. We note finally that NAS is the same as PHL minus the tautosyllabic, anterior, and voiceless fricative components.

Table 3 about here

In addition to testing the intermediate rules shown in Table 3, we also consider the possible effects of a smaller vocabulary. As mentioned in Section 3, smaller vocabularies can tolerate a higher proportion of exceptions. This is particularly relevant to the question at hand: perhaps younger children with small vocabularies will be able to plausibly posit NAS as a productive rule for their traditional input. To test this, we also calculate the plausibility of NAS and intermediate PHL forms on several subsets of the most frequent words in CHILDES, with at least 20, 50, and 100 mentions in the corpus, so as to provide a rough approximation of learners' vocabulary composition at progressive stages of language development. The results are shown in Table 4.

Table 4 about here

As shown in Table 4, NAS does not emerge as a plausible analysis of traditional input, even with a limited vocabulary. However, we do find that the traditional input can plausibly be reanalyzed as any of the three intermediate rules that result from deleting one of the components of PHL. For example, a child could plausibly posit a phonological rule tensing /æ/ before all nasals and voiceless fricatives, including /ŋ/ and /ʃ/ (PHL-ant) without the number of lexical exceptions exceeding the tolerance threshold. Given the plausibility of *at least some* children positing these intermediate grammars, we must now turn to the question of whether *these* children could plausibly contribute enough to the linguistic environment that would cause subsequent language learners to plausibly posit NAS. To do so, we introduce the model of rule learning under heterogenous input.

Rule learning under a heterogenous input

Once an intermediate rule is plausible for a language learner, we must now turn to the question of whether it's *likely*. We expect, for instance, that while some children may posit intermediate rules, others will posit the traditional PHL rule. Therefore, the question of whether NAS is a plausible reanalysis of a child's input must then be reframed as "what proportion of intermediate input would a child need in order to plausibly posit NAS?"

To answer this question, we simulate a child's acquisition of /æ/ given heterogenous input: some PHL and some intermediate rule. We do this in the following way. First, we set *m* to represent the proportion of input from an intermediate grammar that a child receives during acquisition, and 1-*m* to represent the proportion of traditional input. We then construct a simulation of the plausibility of positing NAS, for values of *m* between 0 and 1 in steps of .01 for each of the three intermediate rules. We begin with the assumption that a child will store one form for each word type. For each run of the simulation, we generate a full mixed lexicon according to *m*. Each word is assigned lax or tense /æ/ on the basis of an intermediate rule or traditional input, according to *m*. For example, if *m* = .24, each word in the lexicon will have a 24% chance of its /æ/ allophone being determined by an intermediate rule. This assumption is motivated by empirical studies of how children deal with mixed input when each lexical item is subject to probabilistic variation at the level of token frequency (Hudson Kam & Newport, 2005, 2009). In the present case of mixed input, we assume that each word type has an *m* probability of being internalized in the child learner's vocabulary as a type produced by the intermediate grammar, and a probability of 1-*m* as a type produced by PHL. That is, the child regularizes a *probabilistic* mixture of tokens in the input as a *discrete* mixture of word types representing the two variant grammars. This is implemented by stochastically assigning each word type into one of two grammars with the associated probabilities.[2] We then evaluate the viability of the two grammars on the basis of the resulting lexicon.

It is worth stressing several important features of the learning model. First, it is crucial to note that this is an acquisition model of how a single learner evaluates possible rules when they are exposed to variable input. This is clear in the description of the model, where the sample lexicon for the learner is stochastically drawn from the distribution in their environment. By running the model many times, we can understand the outcome of learning for the speech community at large. Second, the model is agnostic as to the real-world source of the variable input: whether the variation is caused by dialect contact or endogenous innovation, it is treated identically by the model. An individual learner evaluates rules on the basis of the lexicon they acquire from the mixed environment, and it is immaterial how such a mixture is introduced in the first place; see Yang (2000) for additional discussion and applications to syntactic change. Third, the model also does not imply any particular time course for change. For a given mixture of input data, it estimates the probability that PHL or NAS will be a plausible grammar for a learner, but it does not predict what the rate of use would be for a speaker who has successfully acquired both systems. In other words, the outcome of this model tells us what proportion of speakers could posit an intermediate grammar, but this value is not necessarily the value of m for the next generation of learners. Fourth, we stress that this model does not address how a child may *generate* a possible rule; it is simply a model of how a child evaluates possible rules that have already been generated.

For each trial, which here represents an individual learner, we calculate whether an input lexicon comprised of a mixture of PHL and intermediate grammars would allow NAS to be a productive rule. 1000 trials were run for each value of *m* between 0 and 1 in steps of .01, for each intermediate grammar.

Figure 2 about here

Figure 2 presents the results of this simulation, with rates of *m* plotted along the x-axis and the proportion of trials that pass the tolerance threshold along the y-axis. Here, the y-axis represents the predicted proportion of children whose input would allow them to evaluate PHL (in stars) or NAS (in circles) as a plausible grammar for each value of *m*. Here, 1-*m* is the proportion of traditional input, while *m* is the proportion of intermediate rule input.

The outcome of this simulation provides two striking results. First, we see that PHL remains a plausible reanalysis of every intermediate rule's input, for all proportions of that intermediate input, up to 100% input from the intermediate rule. This result speaks to the stability of PHL in Philadelphia: even if speakers have been spontaneously positing intermediate rules throughout the history of the /æ/ split in Philadelphia, the output from these intermediate rules can still be reanalyzed as PHL by the next generation of speakers. Secondly, of the three intermediate rules that would be plausible re-analyses of traditional input, it is only the PHL-fricative rule that would allow NAS to be a plausible reanalysis of that intermediate rule. Furthermore, a NAS reanalysis of PHL-fricative only becomes possible when children are receiving approximately 73% of PHL-fricative input, which is the point at which the probability of accepting NAS becomes non-zero. That is, if at least 73% of Philadelphian children all posited and produced PHL-fricative output, then NAS could endogenously emerge as a consequence for the next generation of speakers. We note that this possibility mirrors the argument in Ash (2002), who models the change from PHL to NAS as occurring via an intermediate step of PHL-fricative.

However, we find this route of change to be highly implausible for Philadelphia, given the results of an empirical search for speakers exhibiting a PHL-fricative type grammar. Only 1 speaker out of 184 who had enough data to allow such an investigation was found.[3] The production of Jake S. is displayed in Figure 3, which shows his data is consistent with a PHL-fricative grammar: he produces lax /æ/ in PHL's pre-nasal conditions (in open syllables and before /ŋ/) as well as before all voiceless fricatives. While Jake's production suggests the possibility of some Philadelphian children positing PHL-fricative, finding only 1 in 184 speakers producing PHL-fricative does not come close to the 73% PHL-fricative input required for the next generation to posit NAS. Moreover, Jake's social profile suggests that he developed a PHL-fricative production as a result of NAS contact, rather than a modification of the traditional PHL input: Jake was born in 1992 and attended the elite Masterman public middle and high school before going on to attend the University of Pennsylvania. Labov et al. (2016) found that most of his peers–speakers born around 1992 who attended Masterman–already produced NAS, suggesting that Jake developed a PHL-fricative /æ/ system as a *response* to his predominately NAS surroundings rather than as a catalyst for NAS emerging. The lack of speakers producing PHL-fricative in Philadelphia suggests that language learners positing this intermediate rule was not the route by which NAS came into Philadelphia.

Figure 3 about here

To summarize the theoretical results so far, applying the Tolerance Principle to the change from PHL to NAS in Philadelphia has found that it is impossible for NAS to directly arise from a PHL input. We have further found that while it is conceivable that an intermediate grammar, specifically PHL-fricative, may eventually lead to NAS, the empirical data from the Philadelphia Neighborhood Corpus and the IHELP corpus finds this outcome to be highly unlikely.

ACQUIRING NAS THROUGH DIALECT CONTACT

Given the unlikelihood of and lack of empirical support for NAS emerging endogenously in Philadelphia, either through direct reanalysis of the original system or via a sequence of reanalyses, we now turn to the possibility of NAS emerging as a result of dialect contact between NAS and PHL.

Sociolinguistic background

The idea that Philadelphian children may be exposed to NAS speaking non-Philadelphians is not altogether unlikely. NAS has been found in the geographic area surrounding Philadelphia (Ash, 2002; Labov et al., 2006;); it is likely that some of these speakers may have access to and influence within Philadelphia. Furthermore, the conclusion in Labov et al. (2016) that NAS is an incoming change from above suggests that it is a change brought about through dialect contact with in-moving NAS speakers.

Theoretical analysis and predictions

Given that dialect contact with NAS speakers fits with the geographic and social patterns found in Philadelphia, we now turn to the question of how much contact with NAS speakers would be necessary for a Philadelphian child to posit NAS. Using the same simulation procedure described above, with NAS as the non-PHL input at proportion *m,* we tested for what proportion of NAS input would be necessary for a child to plausibly posit NAS. Figure 4 presents the results of this simulation, plotting the proportion of trials in which NAS emerged as a plausible rule (circles) and in which PHL emerged as a plausible rule (stars). Simulations were run for different size lexicons, from words with one mention to words with 100 mentions in CHILDES, in order to capture the effect of differently sized lexicons. The full results are displayed in Table 4, which presents the proportion NAS input necessary for NAS and PHL to be viable at all as well as viable 100% of the time.

Figure 4 about here

As expected, higher word frequency cutoffs produce shallower slopes; this is a reflection of the fact that these lexicons are smaller and therefore more proportionally tolerant of exceptions (see Table 1), resulting in a slightly higher proportion of trials that pass the tolerance threshold for each value of *m*. In contrast to the endogenously posited intermediate rules simulated in Section 4, we find that the heterogenous input of PHL and NAS would make positing NAS a highly plausible solution for a child receiving both inputs. Here we find that NAS becomes a plausible analysis of a child’s input if that child is receiving at least 32% NAS input.

Table 5 about here

The specific values of *m* for which NAS and PHL become viable are particularly applicable when the community network structure is taken into account. For instance, while *m* may be quite low over the entire speech community of Philadelphia, there may be local networks, which may be geographically or socially defined, in which the concentration of NAS speakers is quite high, which may lead to the rise of NAS in specific groups before it diffuses to the wider community. This is precisely the situation found in Labov et al. (2016), which finds the highest concentration of NAS speakers amongst the graduates of elite public high schools, with other school networks lagging behind in the change to NAS.

CONCLUSION

In this paper, we've demonstrated that applying the quantitative precision of the Tolerance Principle to the question of phonological change has allowed us to articulate a clearer model of the allophonic restructuring of /æ/ in Philadelphia in a way which would not be possible otherwise. Given a number of *prima facie* plausible hypotheses for the source of the NAS innovation (grammar simplification, endogenous reanalysis, and dialect contact), we have been able to determine that only dialect contact emerges as a likely source of this particular change. This specific finding bolsters the claim in Labov et al. (2016) that, based on community-wide social characteristics, the shift from PHL to NAS is a change from above through dialect contact with NAS speakers who are unevenly distributed across social networks.

More importantly, we've outlined a method for predicting phonological change that can itself be adapted to linguistic changes in other speech communities. The rise of NAS across the country provides a fertile testing ground for this method. Because NAS is emerging in different dialect areas, the calculations for the point at which NAS emerges as plausible for a language learner will be different for each dialect. Applying our model may allow researchers to determine whether NAS is likely spreading due to population movement and dialect contact or whether it may be spontaneously emerging simultaneously for multiple speech communities.

The specific findings reported here are quickly turned into their own empirical predictions. Our results for Philadelphia predict that a child who is receiving less than 32% NAS input will posit PHL, and a child who is receiving more than 70% NAS input to posit NAS. A child receiving roughly 50% NAS input is expected to learn both systems and produce variation between the two. This predicts that a child with one NAS-speaking caregiver and one PHL-speaking caregiver who receives roughly equivalent input from both will emerge as a variable speaker, at least before receiving input from their peer group when they begin to attend school. We note that this prediction aligns well with the empirical results of Payne (1980), who found children with only one PHL-speaking parent producing some /æ/ tokens that were inconsistent with PHL.

Variation, Stability, and Competition

In this paper, we have focused exclusively on what kind of input would be necessary in what mixture for a Philadelphian child to acquire a NAS grammar. The conclusion of the acquisition modelling is that across a broad range of mixtures, both PHL and NAS grammars are plausible. This raises two clear questions. First, is it possible that some learners acquire *both* PHL and NAS grammars as a result of dialect contact? Second, once both grammars are in use within the speech community, is it *inevitable* that one should replace the other, as is being observed?

We turn first to the question of co-existing variation as the outcome of learning. There is considerable acquisition evidence that even for fully native bilingual speakers, one of the phonemic systems appears dominant (Bosch & Sebastián-Gallés, 2003; Cutler, Mehler, Norris, & Sequi, 1989). The acquisition of the low-back merger system at the dialect boundary appears to be a case in point. When the unmerged system is dominant in the community, children acquire an unmerged system. This changes at a precise tipping point where the merged system becomes dominant in the community, resulting in the dramatic contrast in low-back system that can be found in siblings born just a few years apart (Johnson, 2010).

The present case, however, is somewhat different. At issue are two allophonic rules, rather than two phonemic systems, and previous research has found evidence of co-existence and subsequent competition of alternative phonological rules (Fruehwald, Gress-Wright, & Wallenberg, 2013). In the current case study, this would correspond to a young child learning both systems if exposed to them at a sufficiently young age. Once having acquired both systems, a child may then be able to variably access both systems.

Sneller (2018) provides evidence of individual speakers producing variation between PHL and NAS. Using a glm classifier to classify individual /æ/ tokens as having been produced by either PHL or NAS she analyzed 42 white Philadelphian speakers born to Philadelphian parents between the years 1983 and 2000. Of these 42 speakers, 10 produce PHL and 22 produce NAS. The remaining 10 speakers produce robust variation between PHL and NAS in all phonological contexts that differentiate NAS from PHL, ranging in proportion from Beth G. (who produces 58% PHL and 42% NAS tokens) to Maeve F. (who produces 33% PHL and 67% NAS). These 42 speakers represent the results of distinct local school networks discussed in (Labov et al. 2016), where elite public schools provide enough NAS input to typically override a child's PHL system, local diocesan schools provide enough PHL input to reinforce a child's PHL system, and the elite Catholic schools provide enough input from both PHL and NAS to allow both systems to remain viable.

Conclusions

The specific empirical predictions that result from our approach here provides a guideline for a more detailed look into the Philadelphia speech community. More generally, it is our hope that our computational framework based on an established theory of language acquisition may be used for future investigations of phonological changes in many speech communities. Because young children may postulate different grammars from the one provided in the environment and the success of potential postulations can be formally characterized by the Tolerance Principle, our model can be applied to evaluate the viability of innovations based on endogenous change and dialect contact in any dialect.

NOTES

1. This set of lax exceptions has been classically described as consisting of weak words (*and, am, an, than, can*), class 3 strong verbs (*ran, swam, began*), truncated words (*exam, math*), and learned words (*ascot, carafe*) (Labov, 1989). Sneller (2018, ch. 3) notes that there are exceptions to these generalizations as well as interspeaker variation in which words are exceptionally lax, and identifies 39 exceptionally lax forms.
2. Lignos (2013) provides an earlier attempt at using the Tolerance Principle to study language variation and change.
3. Using data from the Philadelphia Neighborhood Corpus (Labov & Rosenfelder, 2013) and the IHELP corpus (Labov, 2015)**,** we analyzed every white speaker who produced at least 5 /æ/ tokens in both the fricative environment and the lax nasal environment. The search was restricted to white speakers, as African American and Latinx speakers in Philadelphia traditionally produce a neutral /æ/ system, produced as a raised lax form **[ɛ:]** for all phonological categories (Fisher et al., 2015; Labov & Fisher, 2015)**.**

REFERENCES

Anderson, Stephen, and Paul Kiparsky. 1973. "``Elsewhere'' in Phonology." In *A Festschrift for Morris Halle*, 93--106.

Aronoff, Mark. 1976. *Word Formation in Generative Grammar*. MIT Press.

Ash, Sharon. 2002. "The Distribution of a Phonemic Split in the Mid-Atlantic Region: Yet More on Short A." *University of Pennsylvania Working Papers in Linguistics* 8 (3): Article 2. http://repository.upenn.edu/pwpl/VOL8/iss3/2.

Baayen, Harald, and Antoinette Renouf. 2018. "Linguistic Society of America Chronicling the Times : Productive Lexical Innovations in an English Newspaper" Author (s): R . Harald Baayen and Antoinette Renouf Published by : Linguistic Society of America Stable URL : Https://Www.Jstor.Org/Stable/416794" 72 (1): 69–96.

Becker, Kara, and Amy Wing-mei Wong. 2009. "The Short-a System of New York City English: An Update." *University of Pennsylvania Working Papers in Linguistics* 15 (2): 1–12.

Bermúdez-Otero, Ricardo. 2007. "Diachronic Phonology." In *The Cambridge Handbook of Phonology*, edited by Paul de Lacy, 497–517. Cambridge University Press.

Boberg, Charles, and Stephanie Strassel. 2000. "Short-a in Cincinnati: A Change in Progress." *Journal of English Linguistics* 28: 108--126.

Bosch, Laura, and Núria Sebastián-Gallés. 2003. "Simultaneous Bilingualism and the Perception of a Language-Specific Vowel Contrast in the First Year of Life." *Language and Speech* 46: 217--243.

Brody, Stefanie. 2011. "Transmissions of Philadelphia Short-a: The Changing Gravity of 'Planets'" *MA Thesis,* University of Pennsylvania.

Brown, Roger. 1973. *A First Language: The Early Stages*. London: George Allen & Unwin.

Bybee, Joan. 1995. “Regular Morphology and the Lexicon.” *Language and Cognitive Processes* 10 (5): 425--455.

Cutler, Anne, Jacques Mehler, Dennis Norris, and Juan Sequi. 1989. “Limits on Bilingualism.” *Nature* 340: 229--230.

Durian, David. 2012. “A New Perspective on Vowel Variation across the 19th and 20th Centuries in Columbus, OH.” The Ohio State University.

Ferguson, Charles. 1972. “``Short A’’ in Philadelphia English.” In *Studies in Linguistics in Honor of George L. Trager*, edited by M Estellie Smith, 259–74. Mouton.

Fisher, Sabriya, Hilary Prichard, and Betsy Sneller. 2015. “The Apple Doesn’t Fall Far from the Tree: Incremental Change in Philadelphia Families.” In *University of Pennsylvania Working Papers in Linguistics* 21 (2): Article 7. https://repository.upenn.edu/pwpl/vol21/iss2/7/

Fogle, Deena. 2008. “Expansion or Approximation: The Low-Back Merger in Indianapolis.” In *Presented at the 37th Meeting of the New Ways of Analyzing Variation Conference*. Houston.

Friedman, Lauren. 2014. “The St. Louis Corridor: Mixing, Competing, and Retreating Dialects.” *PhD Thesis,* University of Pennsylvania.

Fruehwald, Josef, Jonathan Gress-Wright, and Joel Wallenberg. 2013. “Phonological Rule Change: The Constant Rate Effect.” In *Proceedings of the 40th Annual Meeting of the North East Linguistic Society*, edited by Seda Kan, Claire Moore-Cantwell, and Robert Staubs.

Goldinger, Stephen. 1998. “Echoes of Echoes? An Episodic Theory of Lexical Access.” *Psychological Review* 105 (2): 251–79.

Grainger, Jonathan. 1990. "Word Frequency and Neighborhood Frequency Effects in Lexical Decision and Naming." *Journal of Memory and Language* 29 (228--244). https://doi.org/10.1016/0749-596X(90)90074-A.

Hall-Lew, Lauren, Amanda Cardoso, Yova Kemenchedjieva, Kieran Wilson, Ruaridh Purse, and Julie Saigusa. 2010. "San Francisco English and the California Vowel Shift." In *Proceedings of Meetings on Acoustics, Volume 9*, 1--10.

Hudson Kam, Carla, and Elissa Newport. 2005. "Regularizing Unpredictable Variation: The Roles of Adult and Child Learners in Language Formation and Change." *Language Learning and Development* 1 (2): 151–95. https://doi.org/10.1080/15475441.2005.9684215.

Hudson Kam, Carla, and Elissa Newport. 2009. "Getting It Right by Getting It Wrong: When Learners Change Languages." *Cognitive Psychology* 59 (1): 30–66. https://doi.org/10.1016/j.cogpsych.2009.01.001.

Jarmulowicz, Linda. 2002. "English Derivational Suffix Frequency and Children's Stress Judgments." *Brain and Language* 81 (1–3): 192–204. https://doi.org/http://dx.doi.org/10.1006/brln.2001.2517.

Johnson, Daniel Ezra. 1998. "Tensing and Laxing of Short 'a' in New Haven, Connecticut." 1998.

Johnson, Daniel Ezra. 2010. *Stability and Change along a Dialect Boundary: The Low Vowels of Southeastern New England*. Durham, NC: Duke University Press.

Kerswill, Paul. 2000. "Creating a New Town Koine : Children and Language Change in Milton Keynes." *Language in Society* 29 (1): 65–115.

Kiparsky, Paul. 1995. "The Phonological Basis of Sound Change." In *Handbook of Phonological Theory*, edited by John Goldsmith. Blackwell.

Labov, William. 1989. "The Exact Description of a Speech Community." In *Language Change and Variation*, edited by Ralph Fasold and Deborah Schiffrin, 1–57. John Benjamins.

Labov, William. 2007. "Transmission and Diffusion." *Language* 83: 344–87.

Labov, William. 2015. "Influence of Higher Education on Local Phonology Corpus."

Labov, William, Sharon Ash, and Charles Boberg. 2006. *The Atlas of North American English: Phonetics, Phonology and Sound Change*. Mouton de Gruyter.

Labov, William, and Sabriya Fisher. 2015. "African American Phonology in a Philadelphia Community." In *The Oxford Handbook of African American Language*, edited by Jennifer Bloomquist, Lisa J Green, and Sonja L. Lanehart, 1–27. Oxford. https://doi.org/10.1093/oxfordhb/9780199795390.013.34.

Labov, William, Sabriya Fisher, Duna Gylfadottir, Anita Henderson, and Betsy Sneller. 2016. "Competing Systems in Philadelphia Phonology." *Language Variation and Change* 28 (03): 273–305. https://doi.org/10.1017/S0954394516000132.

Labov, William, and Ingrid Rosenfelder. 2013. "Philadelphia Neighborhood Corpus." Philadelphia.

Labov, William, Ingrid Rosenfelder, and Josef Fruehwald. 2013. "One Hundred Years of Sound Change in Philadelphia: Linear Incrementation, Reversal and Reanalysis." *Language* 89: 30–65.

Lignos, Constantine. 2013. "Modeling Words in the Mind". *PhD Thesis*, University of Pennsylvania.

MacWhinney, Brian. 2000. *The CHILDES Project: Tools for Analyzing Talk*. 3rd ed. Mahwah, NJ: Lawrence Erlbaum.

Niyogi, Partha, and Robert Berwick. 1996. "A Language Learning Model for Finite Parameter

Spaces." *Cognition* 61 (1--2): 161--93.

Payne, Arvilla C. 1980. "Factors Controlling the Acquisition of the Philadelphia Dialect by Out-of-State Children." In *Locating Language in Time and Space*, edited by William Labov. New York: Academic Press.

Pierrehumbert, Janet. 2003. "Probabilistic Phonology: Discrimination and Robustness." In *Probabilistic Linguistics*, edited by Rens Bod, Jennifer Hay, and Stefanie Jannedy, 177--228. Cambridge: MIT Press.

Plunkett, Kim, and Virginia Marchman. 1991. "U-Shaped Learning and Frequency Effects in a Multi-Layered Perception: Implications for Child Language Acquisition." *Cognition* 38 (1): 43–102. https://doi.org/10.1016/0010-0277(91)90022-V.

Segui, Juan, Jacques Mehler, Uli Frauenfelder, and John Morton. 1982. "The word frequency effect and lexical access." *Neuropsychologia* 20 (6): 615--627.

Sneller, Betsy. 2018. "Mechanisms of Phonological Change." *PhD Thesis,* University of Pennsylvania.

Trudgill, Peter. 2004. *New-Dialect Formation: The Inevitability of Colonial Englishes*. New York: Oxford University Press.

Tyler, Andrea, and William Nagy. 1989. "The Acquisition of English Derivational Morphology." *Journal of Memory and Language*, 649--667.

Wagner, Suzanne E, Monica Nesbitt, and Matt Savage. 2016. "Reversal and Re-Organization of the Northern Cities Shift in Michigan." *Penn Working Papers in Linguistics* 22 (2).

Yang, Charles. 2000. "Knowledge and Learning in Natural Language." Cambridge: Massachusetts Institute of Technology. http://link.springer.com/content/pdf/10.1007/978-1-4419-1428-6%7B_%7D837.pdf.

Yang, Charles. 2002. *Knowledge and Learning in Natural Language*. New York: Oxford University Press.

Yang, Charles. 2016. *The Price of Linguistic Productivity: How Children Learn to Break the Rules of Language*.

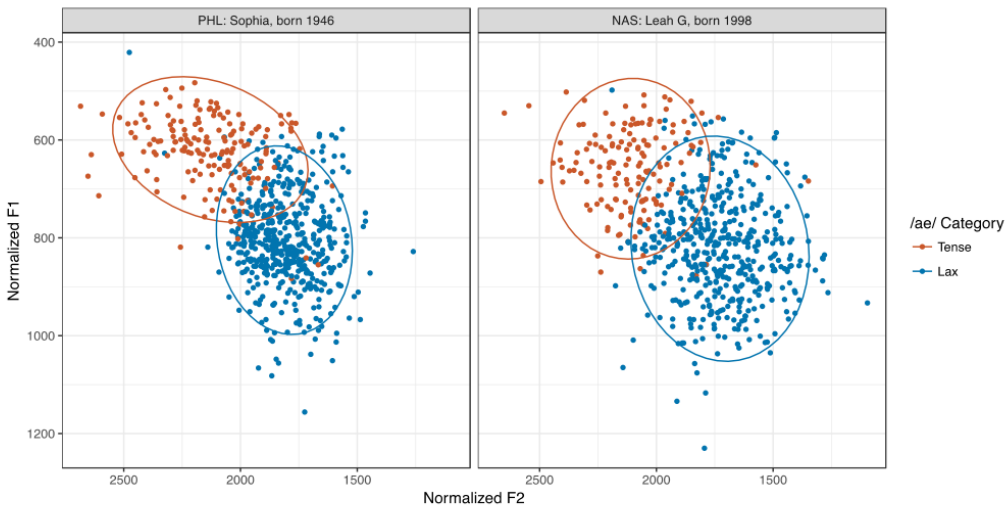

FIGURE 1. Similar phonetic characteristics for tense and lax /æ/ allophones for PHL speaker (left) and NAS speaker (right).

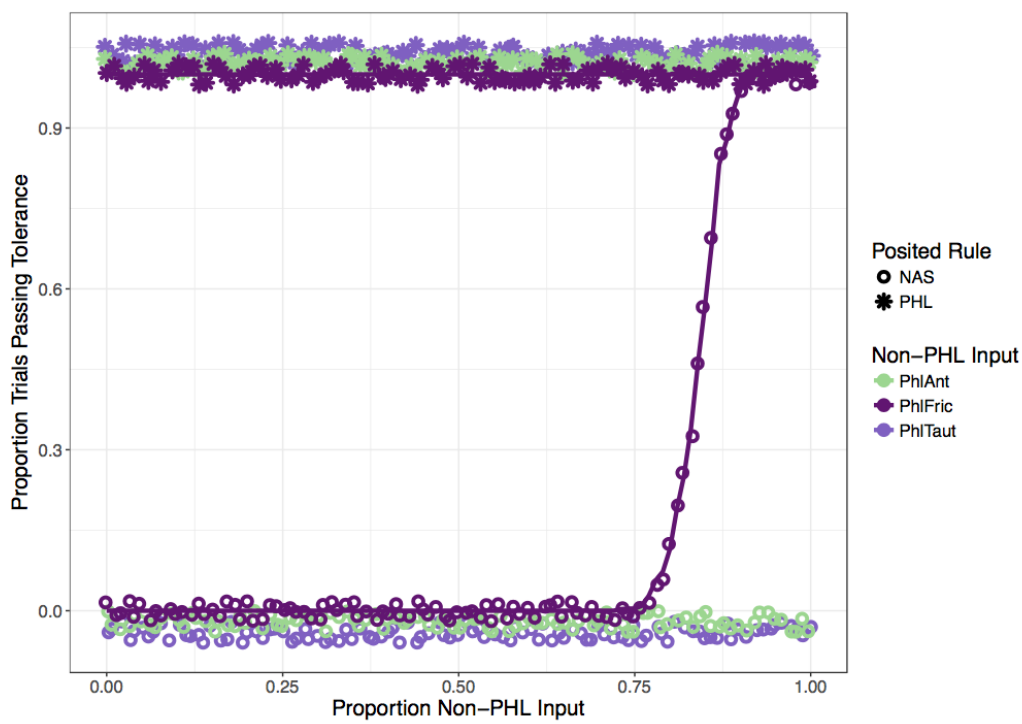

FIGURE 2. Proportion trials which pass the tolerance threshold for each proportion of intermediate rule input for positing NAS or PHL.

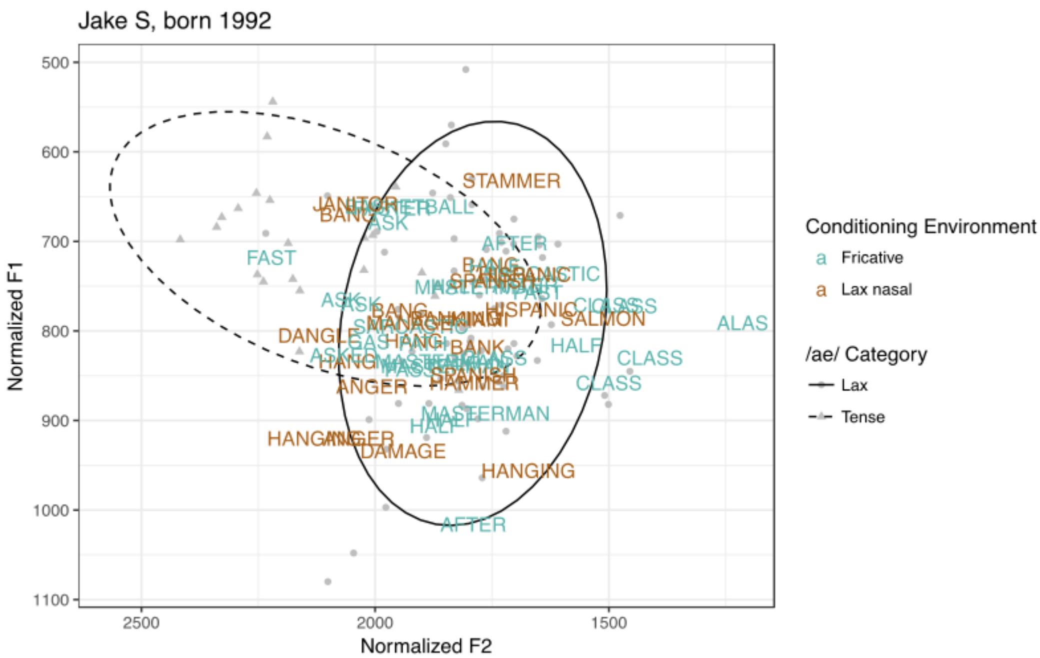

FIGURE 3. PHL-fricative production in the speech of Jake S.

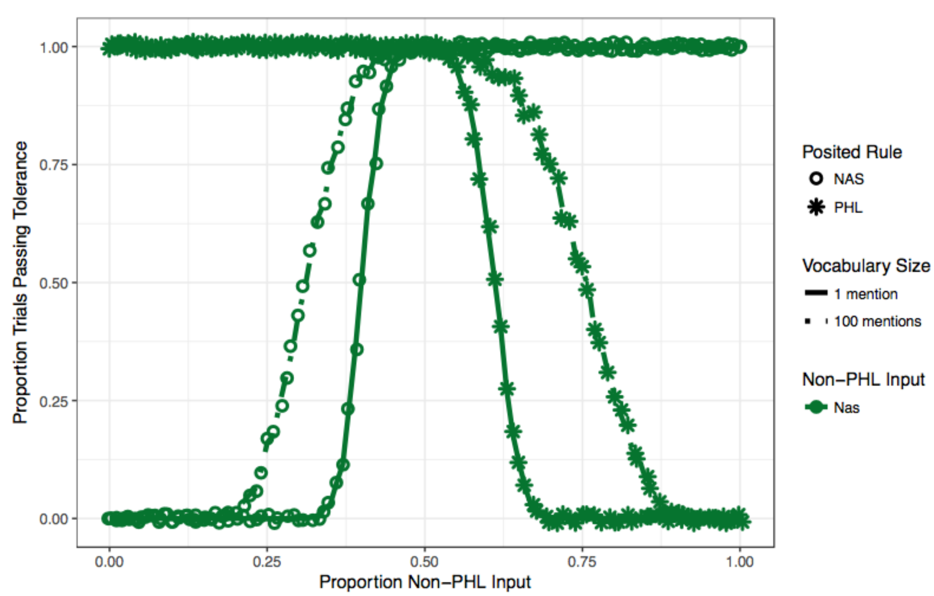

FIGURE 4. Proportion trials that pass the tolerance threshold for NAS (circles) and PHL (stars) for different proportions of NAS input.

TABLE 1. *Number and percent of total lexicon tolerated as exceptions (*e*) by lexicons of* N *size*

n	*e*	%
10	4	40
20	7	35
50	13	26
100	23	23
200	38	19
500	80	16
1,000	145	14.5

TABLE 2. *Input realizations of /æ/ compared to expected /æ/ realizations for PHL and NAS. Expectations incompatible with traditional input highlighted*

Word	**Traditional input**	**PHL expectation**	**NAS expectation**
bad	Tense	Lax	Lax
hammer	Lax	Lax	Tense
cat	Lax	Lax	Lax
fast	Tense	Tense	Lax
bang	Lax	Lax	Tense

TABLE 3. *Intermediate grammars between PHL and NAS*

Name	Rule	Triggering Segments
PHL-ant	/æ/ → tense/__([+nasal] ∨ $\begin{bmatrix}-\text{voice}\\+\text{fricative}\end{bmatrix}$)σ	{m, n, ŋ, f, θ, s, ʃ}σ
PHL-taut	/æ/ → tense/__[+ant] ∧ ([+nasal] ∨ $\begin{bmatrix}-\text{voice}\\+\text{fricative}\end{bmatrix}$)	{m, n, f, θ, s}
PHL-fric	/æ/ → tense/__[+ant] ∧ [+nasal]σ	{m, n}σ
PHL-ant-taut	/æ/ → tense/__([+nasal] ∨ $\begin{bmatrix}-\text{voice}\\+\text{fricative}\end{bmatrix}$)	{m, n, ŋ, f, θ, s, ʃ}
PHL-ant-fric	/æ/ → tense/__[+nasal]σ	{m, n, ŋ}σ
PHL-taut-fric	/æ/ → tense/__[+ant] ∧ [+nasal]	{m, n}

TABLE 4. *Exceptions required for each intermediate rule for vocabularies consisting of words with 1, 20, 50, and 100 mentions in CHILDES. Implausible grammars shaded*

Rule	Ex-1m N = 1412 T=194.7	Ex-20m N = 498 T=80.2	Ex-50m N = 334 T=57.5	Ex-100m N = 239 T=43.6
PHL	**39**	**19**	**15**	**11**
PHL-ant	244	**60**	**42**	**31**
PHL-taut	**155**	**55**	**36**	**25**
PHL-fric	**155**	**64**	**48**	**38**

PHL-ant-taut	273	94	63	45
PHL-ant-fric	237	93	67	51
PHL-taut-fric	240	92	65	50
NAS	324	121	84	63

TABLE 5. *Proportion NAS input at which NAS and PHL become variably viable and categorically viable*

Vocabulary size	NAS leaves 0% viable	NAS reaches 100% viable	PHL leaves 100% viable	PHL reaches 0% viable
1 mention	.32	.48	.53	.7
20 mentions	.25	.46	.52	.82
50 mentions	.2	.47	.54	.86
100 mentions	.17	.48	.54	.9