Chapter 3

Allophonic Analysis of Traditional Philadelphia /æ/

While Chapter 2 provided empirical support for the traditional Philadelphia /æ/ system as an allophonic rather than phonemic split, here I provide an in-depth theoretical account of the allophonic status of the traditional PHL /æ/ split. I argue that PHL is a productive allophonic rule with a limited set of lexical exceptions. I appeal specifically to the Tolerance Principle (Yang, 2016) to define the upper limit of lexical exceptions; I note, however, that my analysis of PHL as allophonic is compatible with any treatment of productive rules that allow for a precise and limited number of lexical exceptions to that rule.

3.1 Lexical Exceptions in Productive Phonological Processes

Determining whether two sounds in a language hold an allophonic or a phonemic relationship is not always a straightforward task. In generative frameworks (e.g., Chomsky and Halle, 1968; Stampe, 1979; Kiparsky, 1982), defining a phonemic relationship is typically an all-or-nothing undertaking, with segments either considered to be perfectly contrastive or not contrastive at all. Phonologists have traditionally relied on a number of criteria to determine which of these two relationships hold (see, for example Steriade, 2007; Hall, 2013, for an extensive list). The two criteria most commonly appealed to and held up as the most important are *Predictability*, defined as it traditionally has been in 3.1, and *Contrastiveness*, defined in 3.2 (both adapted from Hall 2013).

(3.1) **Predictability:**

Two sounds A and B are considered to be contrastive if, in at least one phonological environment in the language, it is impossible to predict which segment will occur. If in every phonological environment where at least one of these segments can occur, it is possible to predict which of the two segments will occur, then A and B are allophonic.

(3.2) **Contrastiveness:**

Two sounds A and B are contrastive when the substitution of A or B in a given phonological environment causes a change in the lexical identity of the words they appear in. If the use of A as opposed to B causes no change in the identity of the lexical item, A and B are allophonic.

The underlyingly binary approach to phonological classification suggested by the criteria above, in which a phonological relationship is either productive or contrastive and not something in between, holds a great deal of theoretical interest for phonologists who subscribe to a view of phonology in which phonological forms and processes are categorical. There are, however, a number of phonological relationships which are not clearly defined using these criteria or which would even be given contrasting definitions based on these two criteria. The problem of intermediate phonological relationships has been taken up by phonologists for quite some time (e.g., Gleason, 1961; Goldsmith, 1995; Harris, 1990, 1994), with varying degrees of importance given to this problem.

In this chapter, I propose that the primary problem in so-called "intermediate relationships" is not in the resulting classification but rather in the definitions of the criteria used to define phonological relationships. In what follows, I begin by highlighting a synchronic and a diachronic case of lexical specificity in otherwise regular phonological processes. In §3.1.2, I discuss previous solutions to the problem of lexical specificity in regular phonology, and in §3.3 I present my definition of *Predictability* using Yang (2016)'s Tolerance Principle to determine an upper limit to lexical exceptions in productive phonology. In §3.4 I apply this metric to the traditional PHL rule, demonstrating that under all configurations, PHL emerges as sufficiently *Predictable* and therefore as a productive allophonic rule.

3.1.1 Lexical Specificity in Productive Phonological Processes

Here, I outline just a few examples of lexical specificity in otherwise productive phonological processes.

Synchronic Lexical Specificity in the Scottish Vowel Length Rule

The Scottish Vowel Length Rule (SVLR) provides a classic case of lexical specificity (Aitken, 1981). The SVLR is a generally productive phonological process found in Scottish English, whereby vowels are produced as short allophones when they precede voiceless stops, voiceless fricatives, voiced stops, nasals, or /l/. Long allophones of these vowels occur preceding voiced fricatives, /r/, and when in an open syllable. This results in short duration *bead* and *beet* ([bid], [bit]) but long duration *bee* and *beer* ([biː], [biːr]). In addition to this set of conditioning factors triggering the SVLR, the phonological targets of this rule are also somewhat complicated and may vary: the SVLR applies to /i, ʉ/ and /ai/, does not apply to /ɛ, ʌ/ or /ɨ/, and other vowels remain disputed (Scobbie et al., 1999; Ladd, 2005).

In an analysis of the large-scale Glasgow Speech Project, Scobbie and Stuart-Smith (2008) report an additional complication on the SVLR which is most applicable here: lexical specificity in its application, which for some lexical items varies by speaker. Table 3.1 reproduces their findings of young female speakers' production of /ai/ in a word list for words that typically would be produced short under the SVLR. In Table 3.1, each row represents a single speaker, with the top four speakers from a middle-class suburb of Glasgow (Bearsden) and the bottom four from a largely working class area of the city (Maryhill). Cells with a 's' follow the expected pattern, while empty cells represent lexical exceptions to the SVLR. Cells with n/a represent a lack of data due to subject error in reading the word.

Here we see lexical specification within individual speakers, so that Bearsden Speaker 3 has the

	bible	sidle	libel	micro	nitro	hydro	title	tidal	pylon	crisis	miser
Bearsden											
1	s	s	n/a		s		s	s		s	
3	s	s		s	s		s	s			
4	s	s		s			s	s		s	
5	s			s	n/a		s			s	
Maryhill											
1		s	n/a				s	s		s	
2		s	n/a				s	n/a		s	n/a
3	s	n/a					s	s			
4	s	s		s	s		s	s		s	

Table 3.1: Lexical specificity in SVLR for young female subjects. Adapted from Scobbie and Stuart-Smith (2008). Cells with ‘s’ were produced as short (expected pattern), cells with ‘n/a’ were not produced or were errors, and blank cells were produced as long.

following lexical exceptions to the SVLR: *libel, hydro, tidal, miser.* For this individual, who in large part follows the SVLR, there remain some lexical exceptions. Under the strict binary approach to classification presented at the beginning of the chapter, this data raises a problem. Does Speaker 3 now have a phonemic contrast in what is otherwise a productive allophonic process simply because four words are lexically specific? Complicating the picture are speakers like Bearsden Speaker 5, who in addition to six lexical exceptions also produces a marginal minimal pair between *title* [taitl̩] and *tidal* [taiːdl̩]. Under the classic definitions of phonemic classification, Speaker 5’s SVLR is a phonemic relationship in length while Speaker 3’s SVLR is unclear.

Additionally, while there is interspeaker variation in the lexical specificity of the SVLR presented in Table 3.1, a more general community trend also emerges. Across the community as a whole, *bible, sidle, title, tidal* and *crisis* are generally produced as expected, while *libel, nitro, hydro, pylon* and *miser* are exceptionally long. Here we see interspeaker variation aligning overall to produce a larger community-level pattern that may in turn be learned, perhaps with varying degrees of faithfulness in which lexical items are exceptional, by the next cohort of speakers.

The problem of classification for all speakers lies in the overwhelmingly productive nature of the SVLR: while there are a few lexical exceptions for speakers, the pattern is overwhelmingly followed. In following a tradition of analyzing morphological conditioning as a marginal contrast, Scobbie and Stuart-Smith (2008) analyze the SVLR as a *Quasi-Phoneme* with what they term *Fuzzy Contrast* which is morphologically predictable save for a few lexical exceptions. Analyzing the SVLR as a stem-level application, which I do here, accounts straightforwardly for the apparent morphological conditioning; what we are left with is a productive stem-level rule with some lexical specificity.

Diachronic Lexical Specificity in Philadelphia /ay/-Raising

The problem of lexical specificity in phonological processes has also been taken up by scholars of language change, most notably debated by historical linguists in the late 1970s and early 1980s. This debate, dubbed the "Neogrammarian Controversy", debated the relative roles of lexical diffusion and regular sound change in language change. The traditional Neogrammarian position holds that sound change is phonetically gradual and lexically abrupt, affecting all segments in the language that share the same phonological target equally. Lexical diffusionists (e.g., Wang, 1969; Chen and Wang, 1975) hold that sound change is phonetically abrupt but lexically gradual, with segments in only particular words at a time abruptly changing in phonetic output until all words in the language with that segment have changed. Labov (1981) attempts to resolve the Neogrammarian Controversy by proposing two distinct types of changes: Neogrammarian changes, which are lexically abrupt and phonetically gradual, and Lexical Diffusion changes which are lexically gradual but phonetically abrupt. Labov (1981) further proposes that these changes have typical target profiles: that Neogrammarian changes will affect phonological features like raising and fronting (features associated with what I consider to be surface phonological representations), while Lexical Diffusion changes affect the underlying phonological representation, causing a "redistribution of some abstract class into other classes." This predicts that certain changes, like phonemic mergers or secondary allophonic splits, may proceed with lexical diffusion, while other regular sound changes, like /u/-fronting, do not.

While Labov (1981)'s solution carries theoretical appeal, providing both an explanation of seemingly disparate facts and predictions for future sound changes, such a discrete separation between two types of sound changes is not borne out in empirical data on sound change. Take, for instance, Fruehwald (2013)'s analysis of /ay/-raising in Philadelphia, where the nucleus of the PRICE diphthong undergoes raising when it precedes a voiceless segment but remains low elsewhere. On the surface, this appears to be a classic example of Neogrammarian change, with a regular phonological conditioning rule of /ay/ raising before all phonologically voiceless segments and remaining low before all phonologically voiced segments, as shown in Figure 3.1. In the middle of this quite regular sound change, Fruehwald (2013) outlines several lexical items which abruptly change categories from low [aɪ] to raised [ʌɪ]: *Snyder* (a street name in Philadelphia), *cider*, and *spider*. In Figure 3.2, the height of these tokens are plotted against the background of /ay/ raising overall. Each point represents the mean of a single speaker's production of these words, with the size of the point representing the number of tokens per speaker. The baseline community production for /ay/-raising before voiceless segments is plotted in blue, and non-raised tokens before voiced segments is plotted in red.

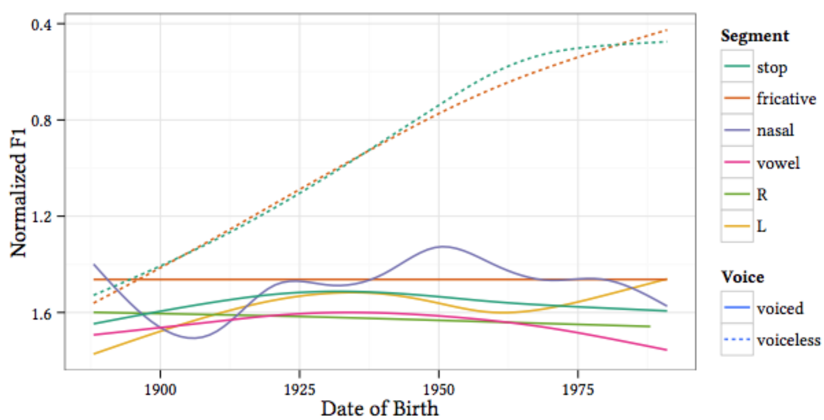

Figure 3.1: PRICE raising by phonological context. From Fruehwald (2013).

Figure 3.2 displays a clear jump in the production of these three words, with most tokens produced low (as predicted by phonological context) near the beginning of the corpus but produced with a raised nucleus near the end of the corpus. The emergence of lexical specificity in the middle of an otherwise regular sound change raises a challenge to the hypotheses laid out in Labov (1981).

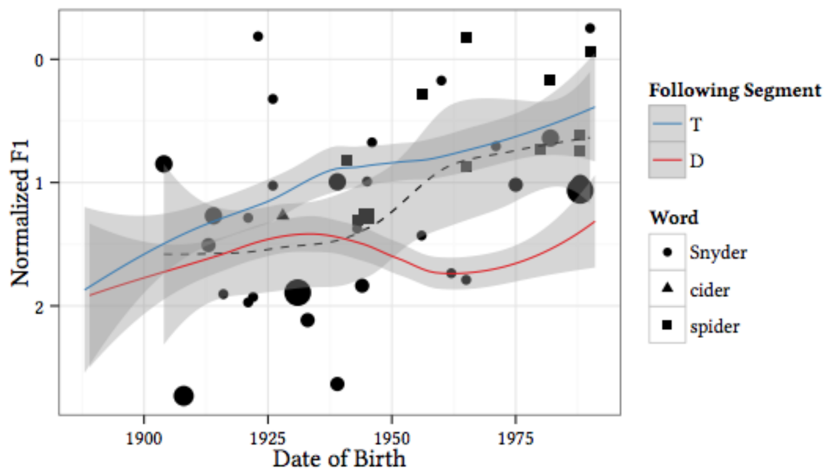

Figure 3.2: Lexical Exceptions in PRICE raising. From Fruehwald (2013).

Here, we see an instance of lexical specificity in the *allophonic* representation of words rather than in the underlying representation. Under Labov (1981), lexical specificity in sound change is hypothesized to occur at the level of underlying specification. This can be potentially resolved by positing that *Snyder, spider* and *cider* did in fact undergo lexical diffusion in their underlying representation, with speakers born after 1940 having re-analyzed the neutralized /ɾ/ as an underlying voiceless /t/. Under this analysis, the lexical specification in the diachronic raising of PRICE is simply an instance of lexical diffusion occurring concurrently with a regular sound change, not an instance of lexical specificity in the allophonic raising rule. However, this solution does not hold for all speakers. Fruehwald also found examples of speakers raising in voiced contexts that do *not* exhibit neutralization between a voiced and a voiceless underlying segment in the output: *tiger* and *cyber*. For these speakers, this lexical specificity cannot be driven by a re-analysis of the underlying form and must instead be accounted for as lexical exceptions to the otherwise regular raising rule.

3.1.2 Solving the Problem of Lexical Specificity

Given that lexical specificity is a well-documented problem in phonology, it is perhaps unsurprising that a number of solutions to these intermediate-type relationships exist. Broadly speaking, these solutions have fallen into one of two main camps. The first camp posits an additional intermediate layer to the phonological architecture to handle these ill-behaved phonological relationships, the idea being that an intermediate relationships is a phonological reality existing between allophonic and phonemic which may be diachronically a step along the way to phonemicization (Kiparsky, 2015). A number of solutions have been brought forward in this vein, with nearly an equal number of distinct labels given to intermediate relationships (e.g. *semi-phonemic, hemi-phoneme, quasi-phoneme, weak contrast, mushy phonemes, marginal contrast* – see Hall 2013, for a robust review). This approach allows for the existence of relationships which would be classified as intermediate under the criteria listed above. There are however, two main critiques to be given to this approach, which fall under an empirical and a theoretical frame. From an empirical perspective, the predictions made by an intermediate phonological category differ from proposal to proposal and often do not make any distinct predictions at all between how an allophonic relationship compared to an intermediate relationship should behave in production (though, see Kiparsky, 2015, for a discussion of quasi-phonemes as a distinct stage in diachronic phonologization).

The second camp takes a gradient view of phonology, arguing that amongst these intermediate relationships, there are some that are more allophonic and some that are more phonemic. This is the view offered in Boulenger et al. (2011), which proposes a Gradient Phonemicity Hypothesis on the basis of gradient responses in an ERP experiment, and in Hall (2013), which redefines the Predictability criterion as a gradient measure of predictability based on the entropy score of a phonological rule. In other words, under both Boulenger et al. (2011) and Hall (2013), the more predictable a pair of sounds is, the more allophonic that pair is and the less predictable a pair of sounds is, the more phonemic that pair is. While this approach provides an overall solution to the problem of intermediate phonological relationships, it introduces fundamental problems to a categorical view of phonology. It predicts, for example, that given two intermediate relationships with different entropy scores, the higher more predictable one will behave more like a productive rule. It is not immediately clear how we may expect this to be borne out in empirical data: perhaps a more predictable intermediate relationship will exhibit more regularity in e.g. nonce word production than a less predictable relationship will. From a theoretical perspective, it is also difficult to incorporate a gradient distinction between allophones and phonemes into a view of phonology that relies on categorical segments and categorical processes, as does any view of phonology consistent with the modular feed-forward approach adopted in this dissertation.

Here I submit a third solution to phonological classification, which is to redefine the definition of *Predictability*. This solution will allow phonological relationships to remain categorical, by enabling alternations previously classified as intermediate to be strictly defined as either allophonic or phonemic. In general terms, my point is simple: that productive allophonic rules may allow a limited number of lexical exceptions. In this dissertation, I specifically invoke the Tolerance Principle (Yang, 2016) to define an upper limit to the number of lexical exceptions a productive phonological rule may allow. This principle was derived independently from phonology, as a model of language acquisition. For a detailed description of the derivation of the Tolerance Principle and numerous examples of it working particularly well to explain lexical exceptions in morphology and phonology cross-linguistically, I refer the reader to Yang (2016). In §3.4 I provide a full account of PHL and its lexical exceptions, demonstrating that it falls well below the threshold for excessive exceptions and therefore is a plausible productive rule.

3.2 Philadelphia /æ/

The phonological conditioning of the traditional PHL split is repeated in (9) below. In (9), the Philadelphia /æ/ split is represented as a rule triggered by a disjunctive set of phonological conditions: nasals or voiceless fricatives which are also interior and syllable final. This produces tense *man*, where /æ/ is followed by a syllable final anterior nasal /n/, but lax *manner*, where the following /n/ is syllabified as the onset of the following syllable.

(9) **PHL:** $\text{æ} \rightarrow \text{æh} \;/\; \_\left[\ \text{+anterior}\ \right] \cap (\left[\ \text{+nasal}\ \right] \cup \begin{bmatrix} \text{-voice} \\ \text{+fricative} \end{bmatrix})]\ \sigma$

I note briefly that disjunction in the featural representation of segments that trigger a productive phonological rule is necessary for a number of cross-linguistic phonological processes (see Mielke, 2008, for an extensive review); as such, the disjunction in PHL is does not in itself present a challenge to PHL as an allophonic rule. We can conceive of PHL as an example of emergent features, where the segments triggering tensing in PHL become classified as a set of similar features by the speakers of the language, which can be represented as in (10). Here, I employ the stratal aspect of the modular feed-forward approach (Bermúdez-Otero, 2007), in which phonological rules may apply at the stem level, word level, or phrase level. I analyze PHL as a productive rule that applies at the stem level of a word, so that an /æ/ followed by an open syllable in the stem (e.g., *manage*) is produced as lax but an open syllable created by an inflectional morpheme (e.g., *man+ning the ship*) is not relevant to the rule, as it has already been applied at the stem level and is also applied at the word or phrase level.

(10) **PHL:** æ → æh / _{m, n, f, θ, s}] σ

The general PHL rule shown in 9 accounts for much of the Philadelphia /æ/ data. However, there are a number of lexical exceptions to this rule which results in a lack of perfect predictability based on phonological context. For example, while most words with /æ/ followed by a tautosyllabic /d/ (such as *dad* and *fad*) follow the rule and are produced as lax, there are three lexical exceptions which are produced as tense: *mad, bad* and *glad.* There are far more lexical exceptions produced as lax, in which words with an /æ/ followed by a tautosyllabic anterior nasal or voiceless fricative are produced as lax (such as *asterisk, ran, than, carafe*). The total number of lexical exceptions to the general rule is extensive, and includes some words whose status as a lexical exception is dependent on the individual speaking. For example, *planet* follows the traditional rule and is produced as lax by many speakers in Philadelphia, but produced as a lexical exception to tense by a number of speakers born in the 1990s (Brody, 2011). The exact number of lexical exceptions required by a Philadelphia English speaker is the focus of §3.4.

This lack of predictability has made the classification of /æ/ in Philadelphia English historically controversial. Since its first treatment in descriptive dialectology literature by Ferguson (1972), PHL has been sometimes analyzed as phonemic (Ferguson, 1972; Labov, 1989; Dinkin, 2013) and sometimes analyzed as allophonic (Kiparsky, 1995; Labov et al., 2016; Sneller, 2018), with each of these works also acknowledging the controversial nature of the classification of PHL. Proponents of a phonemic analysis have almost categorically appealed to the lack of perfect predictability in the distribution of the two sounds, and to the possible existence of one minimal pair (auxiliary *can* produced as lax and noun *can* produced as tense) as evidence for the phonemic analysis of PHL. Proponents of an allophonic analysis have pointed to the *mostly* predictable distribution of the tense and lax versions and more recently to the community-level competition between PHL and NAS (Labov et al., 2016) as evidence for an allophonic analysis.

In what follows, I demonstrate that applying the Tolerance Principle as a diagnostic of productive phonological processes results in an analysis of PHL as a plausible productive allophonic rule with a number of lexical exceptions.

3.3 Tolerance Principle approach to productive rules

As a model of language acquisition, Yang (2016) outlines a principle that determines the productivity of a rule given a set of input. This principle is shown in (11).

(11) **Tolerance Principle:**

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

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

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 potentially apply to divided by the natural log of that number of items. For example, let's assume that a child has 10 verbs in her vocabulary. Some of these verbs take the regular -(e)d suffix to form a past tense (e.g., *walk, smile*), while some of these verbs are exceptions to this regular rule (e.g., *run, fall*). The Tolerance Principle states that the regular past tense -(e)d rule can be productive for this child if her vocabulary has fewer than $10/ln(10)$, or 4.3, exceptions to this 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 her language.

It is important to stress that the Tolerance Principle applies over word *types* rather than tokens. This means that despite evidence that word frequency is an important factor in language processing (Goldinger, 1998; Grainger, 1990; Seguie et al., 1982), it does not play a role in the calculation of the productivity of a rule (modulo the fact that high-frequency words are more likely to be acquired by children and thus more likely to be involved in the calculation of N and e). This predicts that a child would be able to learn a productive rule as long as the word types in her vocabulary fit the Tolerance Principle, regardless of the token frequencies of these words.

Here I highlight a few key features of the Tolerance Principle that are especially relevant for the present dissertation. First, the threshold for exceptions is surprisingly high. Table 3.2 gives a range of values of N and the maximum number of exceptions that a rule defined over N items can tolerate, along with the percentage of total N tolerated as lexical exceptions.

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

Table 3.2: Number and percent of total lexicon tolerated as exceptions (e) by lexicons of N size.

As shown in Table 3.2, as N increases, the tolerable proportion of exceptions (e) decreases. 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. Second, the Tolerance Principle has proved effective in accounting for a wide range of problems in language acquisition ranging from phonology and morphology to syntax (see Yang, 2016, , which provides a discussion of over 100 successful applications of the Tolerance Principle). An artificial language learning study (Schuler et al., 2016) found near-categorical support for the Tolerance Principle. In this study, children between the ages of 5 and 6 learned an artificial language comprised of 9 total nouns. According to the Tolerance Principle, such a language can support up to 4 exceptions ($\theta_9 = 4.1$); Schuler et al. (2016) found that children learned a generalized suffix rule when there were only 4 exceptions but failed to learn the rule when the number of exceptions exceeded the tolerance threshold.

In what follows, I will simply assume the correctness of the Tolerance Principle as a diagnostic of productivity and use it to evaluate the viability of PHL as a productive allophonic rule in the face of exceptions.

3.4 Calculating the tolerance threshold for /ae/ in Philadelphia

To analyze the Philadelphia /æ/ split using the Tolerance Principle, we must first determine the value of N. That is, we must determine the total number of lexical items containing /æ/, which will be the total number of lemmas a tensing rule could apply to. In what follows, I outline a number of choices that must be made with regards to calculating N, and provide an analysis of PHL based on both a conservative approach to each of these choices (i.e., bringing PHL closer to not passing the tolerance threshold) as well as what I believe to be a more accurate description of PHL.

Procedurally, once N has been determined, lexical exceptions are then calculated as those words that violate the productive rule. An example is provided Table 3.3, which presents the expected realization (PHL Expectation) and the actual ralization (Traditional Input) for seven lexical items containing /æ/. In the final column, each lexical item is evaluated for whether the actual realization is an exception to the PHL rule or not. Here, we can see that *mad* must be treated as a lexical exception to the regular PHL rule, as its traditional realization does not match the expected output of the regular rule. Once the total number of lexical exceptions (e) has been determined, we can then calculate whether $e \leq$ the tolerance threshold of θ_N. If the lexical items in Table 3.3 were the entirety of a child's /æ/ words, N would be 7, θ_N would be $\frac{7}{ln(7)}$, or 3.59, and e would be 3. Since 3 < 3.59, PHL would emerge as a productive rule in this dummy language.

Here, I calculate the values of N, θ_N, and e under different assumptions about PHL. In all cases, I obtain the total number of lexical items containing /æ/ from the *CHILDES* database (MacWhinney, 2000) to obtain a measure of the total N for a child's vocabulary. This database includes both child and caretaker production data, which gives an approximation of the linguistic input given to a child.

Word	Traditional Input	PHL Expectation	Exception?
hand	tense	tense	no
mad	tense	lax	yes
cat	lax	lax	no
ran	lax	tense	yes
hammer	lax	lax	no
laugh	tense	tense	no
swam	lax	tense	yes

Table 3.3: Input realizations of /æ/ compared to expected realization under PHL.

3.4.1 Productive Morphology

The first major decision that must be made is the role of productive suffixes. Because the Tolerance Principle applies to word types and not tokens, the crucial calculation is over lemmas. This is particularly relevant to calculating lexical exceptions to PHL: while *laugh* [leːᵊf] straightforwardly follows the productive rule, some suffixes (such as -ing) result in resyllabification of the following /f/, producing a surface-level exception to the productive rule: *laughing* [leːᵊ.fɪŋ] is produced with a tense /æ/ despite the /f/ being intervocalic.

Counting pairs like *laugh* and *laughing* as two distinct lemmas has a large impact on the calculations of both the total *N* as well as the total number of exceptions. Because there is robust evidence that children acquire productive suffices for plural, comparative, present and past tense, adjectival -*y* and diminutive fairly early (Brown, 1973), I posit that words with these suffixes are classified as their stem-level lemma. The productive use of suffixes such as *-ify* and those that involve learned vocabulary items generally are not acquired until school age (Jarmulowicz, 2002; Tyler and Nagy, 1989). In other words, I consider *class* and *classes* to belong to a single lemma *class* which is produced with a tense /æ/ following the tensing rule, but *classify* to be a distinct lemma produced with a lax /æ/ following the tensing rule. I note that this formulation of phonology as allowing children to categorize inflected forms under a single lemma fits well with the stratal view of phonology that I adopt throughout this dissertation, in which phonological processes may apply at the stem, word, or phrase level. Under a stratal view of phonology, the discussion above can simply be read as a statement that the PHL rule applies only at stem level.

3.4.2 Status of /ae/ before /l/

A second decision must also be made regarding the status of /æ/ preceding /l/. In the oldest speakers recorded in the Philadelphia Neighborhood Corpus (PNC), we see a noncontroversial production of lax /æl/. However, as noted by Dinkin (2013) and Labov et al. (2013), the production of /æl/ has been increasingly phonetically tensed beginning with speakers born around 1945, in what appears to be a gradual phonetic process rather than a phonological one. In other words, some speakers produce /æl/ in an intermediate phonetic production between their tense and their lax targets, rather than the expected result of lexical diffusion in which some /æl/ tokens would be produced in line with a speaker's tense target. This suggests that /l/ has not simply been added to the PHL rule as an additional tensing environment, since speakers are not producing /æl/ in line with their own tense target. Additional evidence that /l/ has not been added to the set of triggering environments lies in the fact that *all* /æl/ tokens display phonetic raising, not just the tautosyllabic ones. In other words, both *pal* and *palace* display this gradual raising, where only *pal* would be expected to raise if /l/ were part of the PHL rule.

Dinkin (2013) notes further that this raising of /æl/ coincides with the phonetically gradual fronting and raising of /aw/ (as in *owl*) in Philadelphia, and results in a number of misunderstandings between the /awl/ and /æl/ classes: *owl* with *Al*, *vowel* with *Val*, *Powell* with *pal*. Dinkin (2013) argues that the phonetically gradual behavior of raising /æl/, its phonetic realization tracking the realization of awl as it first peripheralizes then retreats in phonetic space, and the large number of misunderstandings between /æl/ and /awl/ suggests that /æl/ has undergone a phonological reanalysis in these younger speakers, in which words traditionally transcribed with /æl/ have been merged phonologically with awl.

The phonological status of /æl/ is important for calculating both N and e; if /æl/ is phonologically /awl/, then all /æl/ forms should be excluded from both calculations. If /æl/ is still underlyingly part of the /æ/ class, then all /æl/ tokens should be counted as part of N as well as part of e.

3.4.3 Patterns in the lexical exceptions

Finally, it should be noted that in most treatments of /æ/ in Philadelphia, the lexical exceptions have been noted to follow certain patterns. Setting aside "patterns" that follow straightforwardly from the discussion about productive morphology in §3.4.1 (which serve as the primary evidence in Ferguson 1972 for PHL as a phonemic distinction), the remaining patterns have been described, following Labov (1989), as:

1. Truncations of /æ/ words in originally open-syllable position retain lax /æ/ regardless of surface syllable structure: *math* [mæθ] from *mathematics*, *exam* [ɛgzæm] from *examination* [3].
2. Function words that can be reduced to schwa are lax: *and, am, than,* auxiliary *can.*
3. Ablaut past tense forms are lax: *ran, swam, began*, the archaic but marginally productive *wan* (past tense of *win*).
4. Rare and late-learned words are lax: *asp, daft, gaffe, carafe* [4].
5. Polysyllabic words with zero onset before voiceless fricatives are lax: *aspirin, Africa* [5].
6. Affective adjectives *mad, bad, glad* are tense [6].

While these patterns can be identified by linguists (though not without their own exceptions, as highlighted in the footnotes), my account here takes on the perspective of the language learner by simply listing all exceptions in a nonhierarchical list. I do this for several reasons. First, this is the more conservative approach: The Tolerance Principle clearly allows for recursive rules, and analyzing these lexical exceptions as the product of additional rules decreases the number of actual lexical exceptions that must be listed. Analyzing them instead as a flat list as I do here makes an allophonic result less likely. Second, this approach takes child language into account: while there is robust evidence that children learn productive derivational suffixes (-ed, -er, -ly, -ing) early on (Brown, 1973), inflectional suffixes (-ify, -ic) and classifications like "Class 3 Strong Verb" are learned quite late, if at all. So for a young child acquiring Philadelphia English, learning

[3] Though note *gas* from *gasoline* does not follow this pattern: [geːᵊs]. Additionally, individual speakers vary with regards to this pattern, with some speakers producing tense *math* [meːᵊθ] and *exam* [ɛgzeːᵊm]

[4] Though note the "late-learned" effect varies by speaker, with some speakers realizing some of these words as tense

[5] Though note *athlete, afternoon* are tense

[6] Though note the affective adjective *sad* is lax

a phonological pattern based on a classification that that child has not yet acquired is somewhat nonsensical. Finally, it is both unnecessary and inaccurate to consider these patterns as rule: an exception can be found for nearly every pattern described above, and listing recursive exceptional rules as part of the phonological rule unnecessarily complicates the productive phonology.

Instead, I posit that all the lexical exceptions to PHL are simply listed in two lists: L_{tense}\{*mad, bad, glad*\} and L_{lax}\{*math, exam, ran, and*, …\}. This way of listing lexical exceptions means there is no problem in listing some truncations as exceptionally lax (*math, exam*) while leaving other lemmas that were historical truncations to follow the rule (tense *gas*). Additionally, using two lists of lexical exceptions (one for exceptionally tense lemmas and one for exceptionally lax lemmas) easily allows for diachronic additions and subtractions from these lists without expecting changes to affect other words. For example, *planet* is free to join the lexically tense list for the children reported in Brody (2011), then leave it again for speakers reported in Sneller (2018) without any complication to the phonological architecture.

3.4.4 PHL is Productive under All Calculations of N and e

Table 3.4 presents the calculations of N, θ_N and e for all configurations of PHL. As was discussed in §3.3, the tolerance threshold is proportionally higher for smaller vocabularies. This raises the possibility that PHL would be emerge as a productive rule for very young children whose vocabularies are small and therefore more proportionally tolerant of lexical exceptions, but not be productive for older speakers with larger vocabularies. To test this, I calculated N, θ_N and e for different vocabulary sizes. Here, I use the frequency of words defined by the number of instances that word appeared in CHILDES (1, 20, 50, or 100 times) as an approximation of learners' vocabulary at progressive stages of language development. As the frequency value goes up, the total vocabulary goes down; this can be seen in the N values for each row. In each cell, N, θ_N and e are reported, and any cell in which $e \leq \theta_N$ successfully passes the tolerance threshold and is a plausible productive rule.

Table 3.4 reports the results for whether N is calculated with PHL as a stem-level rule (allowing *laugh* and *laughing* to be considered a single lemma) or a surface-level rule under three evaluations

Freq	Surface Rule	Stem Rule	Tense /æl/ Surface Rule	Tense /æl/ Stem Rule	/æl/ as /awl/ Surface Rule	/æl/ as /awl/ Stem Rule
1	$N = 2161$ $\theta_N = 281.4$ $e = 68$	$N = 1412$ $\theta_N = 194.7$ $e = 39$	$N = 2161$ $\theta_N = 281.4$ $e = 165$	$N = 1412$ $\theta_N = 194.7$ $e = 116$	$N = 2064$ $\theta_N = 270.4$ $e = 68$	$N = 1335$ $\theta_N = 185.5$ $e = 39$
20	$N = 660$ $\theta_N = 101.7$ $e = 23$	$N = 487$ $\theta_N = 78.7$ $e = 19$	$N = 660$ $\theta_N = 101.7$ $e = 49$	$N = 487$ $\theta_N = 78.7$ $e = 42$	$N = 634$ $\theta_N = 98.3$ $e = 23$	$N = 464$ $\theta_N = 75.6$ $e = 19$
50	$N = 413$ $\theta_N = 68.6$ $e = 17$	$N = 330$ $\theta_N = 56.9$ $e = 15$	$N = 413$ $\theta_N = 68.6$ $e = 31$	$N = 330$ $\theta_N = 56.9$ $e = 28$	$N = 399$ $\theta_N = 66.6$ $e = 17$	$N = 317$ $\theta_N = 55$ $e = 15$
100	$N = 282$ $\theta_N = 49.9$ $e = 12$	$N = 239$ $\theta_N = 43.6$ $e = 11$	$N = 282$ $\theta_N = 49.9$ $e = 21$	$N = 239$ $\theta_N = 43.6$ $e = 20$	$N = 273$ $\theta_N = 48.7$ $e = 12$	$N = 232$ $\theta_N = 42.6$ $e = 11$

Table 3.4: PHL is productive under all configurations of productive morphology and /æl/ analysis.

of /æl/. The first two columns calculates values based on /æl/ as a lax production of /æ/. The second two columns calculates /æl/ as a tense production of /æ/, and the final two columns calculate values based on /æl/ as no longer belonging to the /æ/ class but rather merged with /awl/. As shown in Table 3.4, there is no configuration of exceptions under which e exceeds θ_N for PHL. In other words, regardless of whether PHL is a stem-level rule or a surface-level rule, and regardless of whether tense forms of /æl/ constitute lexical exceptions for /æ/ or have undergone a secondary split and merged with /awl/, PHL emerges as a plausible productive allophonic rule. A full list of lexical exceptions is provided in Appendix A.

3.4.5 Marginal Contrast in *I can* and *tin can*

While the Tolerance Principle clearly identifies PHL as a plausible productive rule, there is one final sticking point regularly held up as evidence of a phonemic contrast: the marginal contrast between lax auxiliary *can* and tense content *can*.

I have little to say about this contrast, other than to say that whatever mechanism accounts for homophony can be used to account for this contrast. Because the formulation of lexical exceptions relies on a speaker knowing the lexical identity of a word, there is nothing surprising about a speaker being able to distinguish between auxiliary *can* and content *can* underlyingly. In this case, auxiliary *can* is added to the list of exceptions produced as lax, while content *can* remains a regular, unspecified lexical item that is fed straightforwardly through the tensing rule.

3.5 Formulation of PHL as an allophonic rule

As shown in §3.4.4, PHL emerges as a plausible productive rule with some lexical specificity for any configuration of productive morphology and /æl/, using the Tolerance Principle as a measure of productivity. This raises the inevitable question of how to formulate an allophonic rule that has lexical specificity, as well as specifically how I analyze PHL according to the options discussed above.

First, to the problem of representing lexical specificity. Adopting the Tolerance Principle to phonology provides a framework for representing lexical specificity in a productive rule. This principle is formulated as an evaluation metric that "quantifies real time language processing" (Yang, 2016, pg. 40), specifically drawing on the Pāṅinian Elsewhere Condition. To optimize the time-efficiency of representation, the Tolerance Principle argues that speakers list lexical exceptions (w) in order of lexical frequency (w_1 through w_e). When going to process or produce a word containing /æ/, speakers will run through their rules, which are organized first as a rule for each lexical exception ranked by frequency followed by the productive rule and finally the Elsewhere Condition. This is demonstrated in (12), which is adapted from Yang (2016).

(12) IF $w = w_1$ THEN ...
IF $w = w_2$ THEN ...
...
IF $w = w_e$ THEN ...
Apply R

Here, the word w the speaker is processing is evaluated against listed exceptions (w_1 through w_e). If w finds a match, then the relevant exceptional clause is triggered. If the list of exceptions is exhausted without finding a match for w, then rule R applies. The key claim behind this formulation is that the computation of productive rules and their exceptions is a serial rather than an associative process, and that it is the computational search for exceptions that contributes to the cost of real-time processing. While this operation may appear on the surface to be an unwieldy account of processing, Yang (2016) argues that the time cost of adding a lexical exception is minimal and can only be identified in languages where additional processing effects such as neighborhood density and priming do not play a large role. I refer the reader to Yang (2016) for a full derivation and defense of the Tolerance Principle. As for applying the Tolerance Principle to PHL, we can formulate the productive rule as a series of frequency-ranked lexical exceptions followed by the productive rule. This is shown in (13), which applies the computation of frequency ranked lexical exceptions followed by the productive rule R.

(13) **PHL**:

1. IF w = *and* THEN /æ/ → lax

2. IF w = *can* THEN /æ/ → lax

...

39. IF w = *gaffe* THEN /æ/ → lax

40. $\text{æ} \rightarrow \text{æh} \ / \ \_[\ \text{+anterior}\ ] \cap ([\ \text{+nasal}\ ] \cup \begin{bmatrix} \text{-voice} \\ \text{+fricative} \end{bmatrix})]\ \sigma$

Following evidence in Chapter 4 that speakers who vary between the productive rules of PHL and NAS also exhibit similar rates of variation in their lexical exceptions, I consider the entire series of computations listed in (13) to be the allophonic rule PHL. This formulation, notably, can accommodate speakers across the speech community having slightly different lexical exceptions and numbers of lexical exceptions, which may be based on differences in exposure to lexical exceptions during acquisition. This would account straightforwardly for the variation that we find between speakers in lexical exceptions. This also allows for diachronic changes to the list of lexical exceptions: when speakers add *planet* to their list of lexical exceptions as a tense production, these speakers simply add a line for *planet* to their lexical exceptions processes. Speakers are only limited by the number of lexical exceptions they may represent, which is capped at θ_N. For my analysis of PHL, $\theta_N = 194.7$.

As to which words qualify as lexical exceptions to PHL, here I take into account the fact that children acquire productive derivational morphology at a relatively young age. This is equivalent to postulating that PHL is a rule that applies at stem-level only, which is the analysis I consider to be accurate. Secondly, while Dinkin (2013) found evidence that /æl/ has merged with /owl/ in the Philadelphia Neighborhood Corpus, the data from speakers in the IHELP corpus (Chapter 2) and the IMPC corpus (Chapter 4) find speakers producing lax tokens of /æl/, in line with their CAT class tokens. For this reason, my analysis of PHL is that it applies at stem level, and includes /æl/ as part of the CAT class of tokens. In other words, I adopt the second column of Table 3.4 as my analysis of PHL. For a full description of my analysis of lexical exceptions, see Appendix A.

3.6 Discussion

Here, I've presented an in-depth analysis of one of the most contested intermediate phonological relationships using the Tolerance Principle to define the upper limit to lexical exceptions. In all formulations, we find that the traditional PHL rule emerges as a productive analysis for language learners. The specific repercussion of this analysis on the dissertation is a confirmation of the position taken by Labov et al. (2016) and Sneller (2018) that PHL is an allophonic rule. The extensions of this approach, however, are far more wide-reaching. This approach provides a solution to phonological relationships previously analyzed as intermediate or problematic, and also brings with it additional empirical predictions.

The first main prediction is that any intermediate relationship classified under the Tolerance Principle as productive should behave like an allophonic relationship rather than a phonemic one. In other words, allophonic rules with lexical exceptions are still expected to be productive: nonce words are expected to follow the regular rule. In any other task that differentiates allophones from phonemes, we would expect allophonic rules with lexical exceptions to also behave like allophones rather than phonemes. One potential additional piece of evidence may come from a phoneme alteration task. It seems to be more difficult for naïve speakers to produce a nonconforming allophonic production of a sound than to produce a different phoneme. Asking a NAS system speaker to produce a lax form of *man* often results in a production more aligned with their /ɑ/ target than their /æ/ target ([mɑːn] rather than [mæn]), but asking a speaker to swap out phonemic productions appears to be easier. We may expect that intermediate relationships classified as phonemic will be easily produced in a production alteration task while those classified as allophonic will be more difficult for speakers to target.

Secondly, this analysis predicts a precise tipping point between an allophonic and phonemic analysis, at the tolerance threshold of θ_N. If a productive rule held enough lexical exceptions to be near this threshold, it is easy to see how phonemicization may differentially affect speakers whose input is comprised of a different set of lexical items. For example, if a speaker of Philadelphia English acquired all the lexical exceptions in their input but through an accident of exposure was not exposed to enough lexical items that conformed to PHL, this speaker would posit a phonemic analysis of PHL while their peers, having been given a more representative vocabulary, would posit an allophonic analysis of PHL. This possibility both reinforces the importance of the individual in phonological change and provides a clear pathway for a productive rule to become phonologized into a phonemic distinction.