Rapid learning of morphologically conditioned phonetics: vowel nasalization across a boundary[*]

Rebecca Morley
The Ohio State University

That widely attested phonological processes and measurable articulatory and perceptual processes are highly related seems clear. But how that correlation is realized is an issue of much debate. One type of account situates phonetic effects, rooted in the universal properties of peripheral auditory and articulatory systems, as the common source of possible phonological grammars, thus explaining part of the typological commonality of those grammars. Leaving aside the question of how the transition from phonetics to phonology might then be made, a necessary (but not sufficient) condition for this class of theories is that learners both be able to encode this ever-present gradient information, as well as integrate it into an existing categorical grammar. An artificial grammar learning experiment tested the ability of participants to learn an association that was conditioned on a morphological boundary (and was thus grammatical), but that consisted of acoustic information that was sub-phonemic in nature (degree of nasalization on a pre-nasal vowel, never contrastive in English). Success in learning indicates that even over short time periods (roughly 15 minutes of training) listeners can assimilate novel phonetic cues. Furthermore, the results show that grammatical and sub-grammatical components of the linguistic system have the ability to interact. That a distinction between inter-morpheme and intra-morpheme environments can be made at the phonetic level is further consistent with a common distinction in phonological systems between processes that only apply (or only fail to apply) at morphological boundaries (a particular kind of derived environment).

1 Introduction

The origin of phonological grammars remains a wide open question in the field, one that is often left unexplored, but that has been gaining increasing attention. The speed and ease with which young children converge on seemingly complicated and abstract linguistic knowledge has long been taken as support for the hypothesis that many aspects of grammar must be innate (Chomsky 1986). However, there has been a growing body of work showing success in pattern extraction by associationist models (cf. Rumelhart & McClelland 1986, Elman 2003), as well as statistical

---

[*] I would like to thank Paul Smolensky, Colin Wilson and the members of the Johns Hopkins IGERT lab.

learning by infants and adults (cf. Jusczyk *et al.* 1999, Maye *et al.* 2002, Saffran & Thiessen 2003, Newport & Aslin 2004, Wilson 2006). This work has re-opened the question of how much information is in fact contained within the auditory input, and how much of that can be attended to and extracted by listeners. Related to the learning question is how languages might change over time, that is, with certain phonological rules being adopted, or lost, as speech is transmitted from speakers to listeners.

I will examine one theory of how such changes come about, Evolutionary Phonology, which situates language universals in the universal properties of auditory and articulatory mechanisms, and ties language change directly to language acquisition. Central to this theory is the role of phonetic cues in the development of phonological rules. This article will focus on one such cue (degree of nasalization), and show that it can be encoded and generalized by adult learners in an artificial grammar learning task, adding to our understanding of what listeners can both extract from the speech signal, as well as use to condition associations. In Sections 2 and 3, I will describe the experiment in detail, and I will argue that success in learning phonetically conditioned alternations across a morpheme boundary provides necessary support for a phonetic origin of this type of domain restricted process. In Section 4 I will consider more detailed analyses of the experimental results, looking carefully at the degree of phonetic nasality in individual stimulus items. Finally, in Section 5 I will summarize my conclusions, and situate them within an Evolutionary Phonology account, suggesting a special role that morphological decomposition might play in the process of phonologization.

1.1 Theoretical Background

Although what exactly is extractable from the speech signal (that is, how impoverished the stimulus is) is a question of considerable debate, it is clear that a large amount of language-specific information is learned and therefore learnable. For example, speakers must acquire not only lexical items, but information regarding their appropriate context of use, and the conditioning environment for alternations which might depend on idiosyncratic morphological, syntactic and semantic structure. Furthermore, the non-universal, and yet consistent, conditioning of acoustic cues at the phonetic level requires listeners to be sensitive to the degree to which their language nasalizes vowels before nasal consonants, for example, or lengthens low vowels relative to high ones (Keating 1985, Sole 1992, Beddor & Krakow 1999). There is, additionally, considerable evidence that speakers can, at least for certain tasks, access highly detailed representations of particular words and sounds (Summerfield 1981, Goldinger 1996, Remez *et al.* 1997, Clopper & Pisoni 2004, Allen & Miller 2004).

That this fine-grained acoustic and articulatory level information also plays a central role in how languages change over time is a tenet of the theory of Evolutionary Phonology, which situates language change in the process of transmission of the physical speech signal (Ohala 1981, 1990, 1993, Blevins 2004). Actual speech cannot readily be broken up into its constituent elements; adjacent sounds overlap and sometimes merge or disappear completely when produced at fast, or even normal speaking rates. The listener, in successfully reconstructing the speaker's intended utterance, must somehow be able to subtract out this 'noisiness' in the channel. This task, in turn, must rely on a familiarity with the way acoustic cues are likely to appear in particular environments: a degree of phonetic expertise. For example, some features possess a certain long-range character, such that, in production, they can be distributed over multiple segments in the acoustic signal. This is the case for the nasalization that occurs on vowels adjacent to nasal consonants (phonetically [ṽn]). The listener who is, at some level, aware of this property will be able to hypothesize that the only underlying nasal feature belongs to the final nasal consonant, phonemically /vn/. This ability on the part of the listener might be broadly classed as compensation for coarticulation, and various experimental results support our ability to perform it (Mann & Repp 1980, Alfonso & Baer 1982, Whalen 1984).

In general we seem to be very good at this, except when we're not. If, for whatever reason, we fail to correctly compensate, either by under-correcting or over-correcting, that is where change can occur (Blevins 2004). As I understand it, the stage at which acoustics, or phonetics, becomes part of the phonology is the point at which what the listener perceives runs counter to their phonetic expectation, either because that expectation fails in some way, or because perception does. To make this process more concrete, let us consider the case of nasal vowels. In some languages, such as Portuguese, nasality is contrastive on vowels, and minimally distinct word pairs can be found: /vi/, meaning 'saw', and /vĩ/, meaning 'came'; /mudu/, meaning 'mute', and /mũdu/, meaning 'world'. In other languages, such as English, nasal vowels have a completely predictable distribution; they occur only in environments before nasal consonants, and have an allophonic, or phonetic status. The standard account for the historical emergence of phonemic nasal vowels attributes their origins to these former allophones. Furthermore, it is usually assumed that the conditioning nasal (via some independent sound change) must be lost in order for the nasality on the vowel to become phonemicized (see Ohala 1993).

In terms of the Evolutionary Phonology account outlined above, listeners who are aware of the expected degree of phonetic nasalization may use it to predict the presence of a following nasal consonant. Conversely, the presence of the nasal consonant also allows them to assess the quality of the preceding. The case in which the phonetic nasality is present on the vowel but the conditioning nasal is missing represents a mismatch between perception and expectation. There is no source in the signal for the nasality perceived on the vowel, other than the vowel itself. Thus, a phonemically nasalized vowel is born. This argument gains support from experimental work by Kawasaki (1986) who found that, for a series of one syllable stimuli, linguistically naïve English speaking participants rated the vowel as more nasal, the lower the amplitude (and thus lower the perceptibility) of the final nasal consonant.

If an Evolutionary Phonology account is the right story about the origins of phonological grammars then it needs to be able to explain the types of systems that are attested in the world's languages. The basic premise is that every perceivable acoustic difference has the potential to be realized, at some point in time, and in some language, as a phonological distinction. Reversing this relationship, the enterprise is to show that common phonological distinctions possess consistent phonetic correlates. This paper will investigate that hypothesis as it relates to a particular class of phonological phenomena: derived environment effects.

I will take derived environment effects here to refer to any process that is conditioned by morphological domain: either only applying, or only failing to apply, across a morpheme boundary. For example, all syllables in Desano are required to be completely oral or completely nasal. The rule extends farther than that, however, such that combinations of oral and nasal syllables are disallowed: *bonã, *ãgomĩ (Kaye 1974). It is not the case, however, that all syllables within a word must agree on nasality, since some words, in particular, multimorphemic words, are allowed to contain both nasal and oral syllables – as long as they are within different morphemes. In this case the rule, or constraint, only applies within morphemes, and not across a morphological boundary. This is illustrated by the data in (1).

(1) /pea/ + /mĩ/ → [peamĩ] 'he breaks'
/kõmẽ/ + /da/ → [kõmẽda] 'wire'

Another type of morphologically-limited process is exemplified by Korean palatalization. Stop consonants preceding high front vowels surface as affricates as shown in (2a). This rule, however, fails to apply when the conditioning environment and the undergoer occur within the same morpheme, as shown in (2b) (from McCarthy 2002).

(2a) /patʰ/+/i/ → [patʃʰi] 'field-COP'
/mat/+/i/ → [matʃi] 'eldest-NOM'
(2b) /mati/ → [mati] 'knot'
/katʃʰi/ → [katʃʰi] 'value'

2 Experiment

The present experiment is designed to investigate the relation between phonetic and phonological patterns, in particular, the hypothesis that phonologization occurs when perception breaks down, or listener expectation is not met. An example of this, as discussed previously, is the genesis of phonemically nasal vowels (/ṽ/) from phonetically nasal vowels ([ṽn]) that have lost their final nasals. The experiment described below will also focus on nasalized vowels; these vowels will retain their conditioning environments, but listener expectation will be established by training on two different degrees of nasalization depending on morphological position. This experiment will address the question of whether listeners are able to extract these kinds of sub-phonemic cues, as well as associate their presence with a particular kind of grammatical structure.

Previous experimental work has found both production and perception differences related to the presence or absence of morphological boundaries. Measurements of Korean speakers show that there is a difference in amount of variability in gestural timing with regards to palatalization across versus within a morpheme boundary (Cho 2001). Work on English has demonstrated a correlation between morphological boundary strength and degree of phonetic reduction (Hay 2003). And work by Frazier (2005), also in English, shows reliable differences in vowel length for monosyllabic words depending on whether they are monomorphemic or bimorphemic (e.g., passed/past). These effects have also been shown, in some cases, to be accessible to hearers. Frazier reports a perceptual difference correlated with vowel length in terms of participants' likelihood of selecting the mono- or bi-morphemic variants in a forced choice task. Another set of experiments have examined the effect of different boundary types (phonological phrase, prosodic word) in word disambiguation (Salverda *et al.* 2003, Christophe *et al.* 2004), providing evidence that listeners can make use of phonetic cues associated with domain structure, cues which can include differences in segment length, pitch accent and degree of coarticulation.

The perception experiments described above employed experimental tasks that were centered around the disambiguation of semantically distinct minimal pairs using contrasts within the native language of the participants. In contrast, the current experiment involves explicit training on a novel morphological alternation (the presence of absence of the suffix –m), and only implicit training on the associated (redundant) phonetic difference of interest (the degree of nasalization on pre-nasal vowels). Furthermore, the task is discrimination between two words which are phonologically identical, but one of which is the correct word *phonetically* (given the participants' training), and one of which is not.

Similarly, work described in the preceding section has shown that listeners are sensitive to differences in the realizations of the phonetic cues associated with the productions of different speakers, and in different environments. But that work dealt with cues which were otherwise contrastive in the given language (such as VOT differences), or were robust phonetic indicators of phonemic contrast (such as vowel length differences, which signal voicing distinctions on stops in English). In the current case, however, the feature under investigation is nasality – never contrastive on English vowels, and, as far as I am aware, almost always redundant in signaling a subsequent nasal consonant.

The present paradigm is an approach that combines the power of the statistical, or artificial grammar, learning paradigm (cf. Newport & Aslin 2004) – the ability to carefully control the listener's input, and test associations learned implicitly – with the type of experimental phonology advocated by Ohala (1974, 1981) and exemplified in Kawasaki's (1986) work investigating the acoustic-level correlates of language change. This combined approach is a very promising avenue to testing a number of hypotheses about language learning and change.

2.1 Procedure

The experimental design was as follows. Participants were told that they would be hearing words in a new language, words spoken by somebody named 'Frank'[1], and that they would later be asked questions about those words. Both plural and singular versions of these words would be heard. In the singular the words would appear with a picture of a single object; in the plural they would appear with a picture of two of the same object.

What followed was a passive training stage in which participants listened to words over the headphones and looked at pictures on the computer monitor. There was a 1200 ms pause between each picture-word pair. All training items were presented in singular-plural pairs, with the singular presented first. The singular and the plural differed in that the plural ended with the suffix –m. For example, participants heard the word [skimtu] over the headphones at the same time they saw a picture on the screen of a single key; this was followed by the word [skimtum] heard over the headphones, accompanied by a picture of two keys.

Participants were trained on 12 distinct singular-plural word pairs, repeated in 6 randomized blocks. Once mid-way through training, and again after training had completed, a practice block

---

[1] The speaker was identified in the hopes of priming speaker identification – a task known to support the encoding of sub-phonemic cues.

occurred. Each practice block consisted of presentation of 12 pictures (a random selection from the set of all singular and plural items seen during training). Each of these pictures was presented once, and accompanied by two auditorily presented words. 600 ms after the picture appeared, the first word was played; this was followed 800 ms later by the second word. Participants were instructed to select (via key press) the spoken word that matched the picture. This was a test of the singular/plural distinction, such that of the two word choices per picture, one was a singular inflection, and the other a plural inflection. As soon as the participants pressed a key, the picture disappeared. Participants received feedback during these practice trials, seeing either ‘correct’ or ‘incorrect’ appear on the screen, and hearing a buzzer noise in the latter case. 200 ms later the next picture appeared. Participants’ performance in the second practice block was used as a criterion test for inclusion of their results.

The alternation of interest related to the behavior of pre-nasal vowels. All plural words contained such a vowel before the plural suffix (-m), that is, at the morpheme boundary. This boundary was demarcated by greater (or less) regressive nasalization than observed in a context that was identical except for the absence of the boundary (stem-internal pre-nasal vowels). In training for the ‘LO-HI’ condition, there was 0% (LOw) regressive nasalization within morphemes and 100% (HIgh) across; in the ‘HI-LO’ condition, those values were reversed. See Table I for example stimuli. It should be noted that the ‘LO-HI’ condition presents an alternation to the learner on the word final vowel, whereas the ‘HI-LO’ condition shows no alternation.

Half of the stems ended in /i/ and half in /u/[2]. Half of the words also contained /m/ in the stem, which was preceded in 3 stems by /i/ and in the other 3 by /u/. Thus, in both conditions, subjects heard 6 instances of /im/ and 6 of /um/ word-internally, and 6 instances of each word-finally; what differed was whether it was in the word-internal (tautomorphemic) or word-final (heteromorphemic) /Vm/ sequences that the vowel was highly nasalized.
At test, subjects were asked to identify which of two words was the one spoken by ‘Frank’. A two-alternative forced-choice task consisted of two auditorily presented words, one with the high degree of nasal coarticulation, and one with the low, but otherwise identical. Test items included

---

[2] The full list of stems used in training and test:

Old Stems				Novel Stems			
hǽzi	tʃú	skímtu	agímdu	dí	fí	ʒadímfu	tʃúmgu
óski	spí	jatúmbi	ədʒúmpu	plóksu	ðípi	húmʃi	glaudúmki
hú	έgθu	θúmzi	twímtʃi	ríldu	stú	frabímsi	ímdʒi

both old words, and new words, both singular and plural. These words were accompanied by pictures, as in the training phase. The six stems that lacked an internal nasal were only tested in the plural; the six nasal stems were tested both in the singular and plural. Each test item was presented twice in either order, for a total of 12 singular test items and 24 plural test items each for old and new words. The order of items was randomized across participants. Table 1 gives example test stimuli for each condition. Participant responses are shown in Figure 1.

LO-HI	Sing.	Plural	Test	HI-LO	Sing.	Plural	Test
non-nasal	oski	oskĩm	oskĩm/oskim	non-nasal	oski	oskim	oskĩm/oskim
nasal	skimtu	skimtũm	skimtũm/skimtum[†] skimtu/skĩmtu	nasal	skĩmtu	skĩmtum	skĩmtũm/skĩmtum[†] skimtu/skĩmtu

Table 1
Example training and test items for the two experimental conditions. ĩ and ũ indicate vowels with a high degree of nasalization; i and u indicate vowels with a low degree of nasalization.

The phonetic cues present in the stimuli are natural (regressive nasalization associated with nasal consonants) and redundant (always occurring with the cue of the accompanying nasal consonant). Furthermore, vowel nasalization of any degree is non-contrastive in English, the native language of the experimental participants. For these reasons, we might not even expect listeners to reliably hear differences along this dimension. To check for accuracy of perception, the final task of the experiment was an AXB test to assess participants' auditory discrimination of the phonetic nasalization cue. Test items consisted of a subset of the words participants had been tested on earlier in the experiment. For each triple, participants had to choose which two words were identical; either the first word was the same as the second, or the third word was the same as the second. The non-identical token differed only in degree of nasalization, e.g. [skimtu] [skĩmtu] [skĩmtu]. Singular and plural words were analyzed separately and were used as a criterion test for inclusion in the analysis. 70% accuracy or better was required for each type.

2.2 Stimuli

The stimuli consisted of words of varying length and syllable structure. The nasal vowel at the morpheme edge was always post-tonic (except for the one syllable roots); the nasal vowel within

---

[†] The plural nasal-stem test items were different across the 'LO-HI' and 'HI-LO' conditions. This was so that only one position was being tested at a time. In the plural nasal-stem items the trained the correct degree of nasalization always appeared on the stem-internal vowel, and only the degree of nasalization on the stem-final vowel alternated across the test items.

the morpheme was always tonic, but sometimes occurred in the first and sometimes in the second syllable. Natural speech tokens were used; each stimulus was recorded separately. Also recorded were monomorphemes with final oral consonants, and monomorphemes with final nasal consonants, e.g., /hæzim/, /hæzi/, /zib/, /zim/. High nasalization vowel tokens were created by splicing from a stressed nasal coda environment ([z**i**m], generating [hæz**ĩm**]) such that the entire vowel was nasalized. This was determined by visual inspection, verifying that the second nasal formant (around 1000Hz) was visible throughout the duration of the vowel. Low nasalization tokens were created by splicing from a non-nasalized environment, e.g., [z**i**b], generating [hæz**i**m]. For these items, no nasal formant was visible in any part of the vowel. Vowels were normalized for length and intensity. All words were produced by a phonetically trained male native American English speaker who was instructed to pronounce unstressed vowels as full vowels (rather than reducing them). All stimuli were recorded in a sound-attenuated booth at a 22kH sampling rate.

2.3 Participants

58 undergraduates at Johns Hopkins University were given course credit to complete the 30 minute experiment. All participants reached criterion in the second practice block (> 95% correct). 32 participants failed the AXB threshold and were left out of the analysis[3].

2.4 Results

The results discussed here are for responses in the 2AFC task. Only participants who performed above threshold on the subsequent AXB discrimination task are included. Test items consisted of two types: singular or plural. The choice at test was always between a high nasalization variant and a low nasalization variant (that is, the two possible responses differed only in nasality on the critical vowel). Two conditions were contrasted: 'LO-HI': high nasalization across boundary, low within; and 'HI-LO': low nasalization across boundary, high within. Each participant was run in only one condition.

The dependent variable was the percent of the responses for which participants selected the high nasalization variant (as opposed to the low). Under successful learning, the value of this variable should be large for plural items in the 'LO-HI' condition, and small for singular items (small for

---

[3] The large number of participants who failed threshold on the AXB task is not necessarily surprising, given the low-level nature of the phonetic distinction at test. To forestall concerns, however, about the possibility that removing such a large number of participants from consideration could skew the results, I examined the learning data from the below-threshold participants as well. Pooling the data from all 58 participants (29 in each condition) produced exactly the same results as the above-threshold participants considered alone; all significant effects remained significant. No new effects emerged.

plural items in the ‘HI-LO’ condition, large for singular items). As just described, the prediction is that there should be a significant interaction between condition and test type. See Figure 1 for the results of the current experiment.

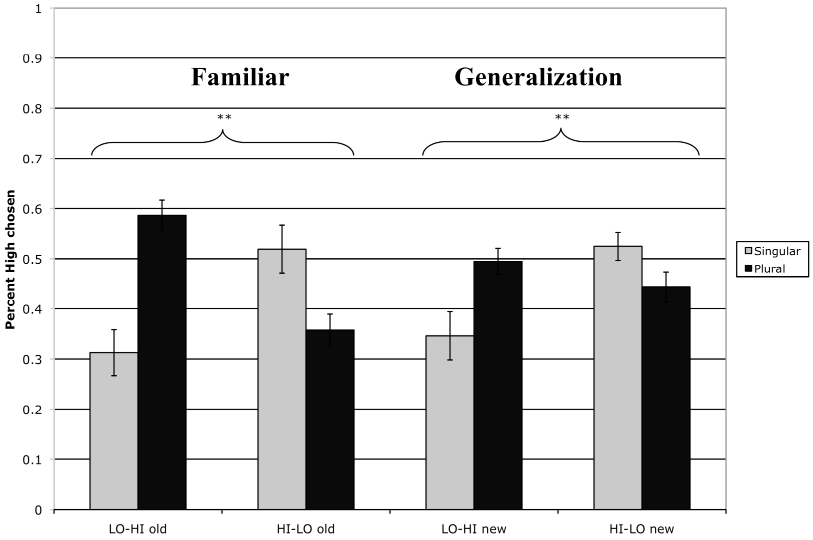

Figure 1
Experiment1: Percent HIgh (=100% nasalization) variant chosen by Condition (=LO-HI or HI-LO) and Type (=Singular or Plural), plotted separately for old and new words.

Since the dependent variable was a proportion response, varying between 0 and 1, a logistic regression analysis was performed (Jaeger, in press, Agresti 1996). This was done separately for old and new items, with model terms for Condition (LO-HI or HI-LO) and Type (plural or singular), as well as a term for the interaction between Condition and Type (the critical test of learning). Each model term was assessed for its reduction in the residual deviance of the logistic fit as compared to the model without that term. The significance of the reduction was then evaluated using a chi-square test of significance, producing the p values shown in the final column for each analysis below.

Old Items: Considering just the additive model for Condition + Type, there was no main effect of Type, but Condition was significant. Adding the critical interaction term improved the model fit by a significant factor, as revealed by a sequential analysis of variance over each model term.

New Items: There was no main effect for Condition or Type. Adding the critical interaction term improved the model fit by a significant factor. See Table 2.

OLD	Df	Dev	Resid.	Resid.	P(>|Chi|)	**NEW**	Df	Dev	Resid.	Resid.	P(>|Chi|)

			Df	Dev					Df	Dev	
NULL			51	120.23		NULL			51	61.91	
Cond	1	6.38	50	113.85	0.01*	Cond	1	0.61	50	61.30	0.43
Type	1	2.73	49	111.11	0.09	Type	1	0.98	49	60.32	0.32
Cond:Type	1	39.92	48	71.18	2.64e-10*	Cond:Type	1	11.46	48	48.85	0.001*

Table 2: Experiment 1 HI response fit to Condition, Test Type and Condition x Type

Response levels hovered near chance for trained High (high nasalization) items. A regression model fitting degree of nasalization trained (High or Low) against accuracy found a significant difference. Responses were more likely to be accurate for trained Low (‘LO-HI’ singular and ‘HI-LO’ plural) than for trained High (‘LO-HI’ plural and ‘HI-LO’ singular), for both old and new items. See Table 3.

OLD	Df	Dev	Resid. Df	Resid. Dev	P(>|Chi|)	**NEW**	Df	Dev	Resid. Df	Resid. Dev	P(>|Chi|)
NULL			51	82.56		NULL			51	60.11	
TrainLow	1	8.45	50	74.10	.004*	TrainLow	1	6.67	50	53.44	.010*

Table 3 Experiment 1 response accuracy, fitted to type of training item (High or Low)

2.5 Discussion

A robust interaction effect (Condition x Type) indicates an effect of training on participant response. For old items, learners were able to encode detailed phonetic representations for words they had heard before, representations that included information about sub-phonemic vowel quality features for at least two different positions in the word. This same interaction effect for new items indicates that learners were able to make an association between those phonetic features and some other property of the training words such that they were able to correctly generalize to novel test items[4].

That this association is purely morphological cannot be determined from these results alone. Since the morpheme in question is a single-segment suffix, the boundary condition also coincides with the final position in the word. The stem-internal pre-nasal vowel, however, occurs sometimes in the initial, sometimes, the medial syllable of the word. This locus, however, is always the bearer of stress. Morpheme boundary vowels are post-tonic for the most part, but this is not completely confounded, as 1/6th of the words contain only a single syllable, putting

---

[4] One concern is that this result could be explained by participants learning to make only one association for a given word: over-all high nasalization on pre-nasal vowels, or over-all low nasalization. If that is the case, their responses should be accurate for both non-nasal plurals and nasal singulars. For nasal plurals, however, there are two conflicting degrees of nasalization; participants are predicted to be at chance on those items. Analysis of the nasal-stem items alone, however, reveal that the critical interaction holds. This indicates that the single-association hypothesis is incorrect, and that there are indeed two different positions within a given word that are encoded.

this vowel in stressed position[5].

A main effect of Condition for old items indicates that high nasalization was less likely to be chosen overall for the ‘HI-LO’ condition than the ‘LO-HI’ condition. This main effect goes away, however, for new items. As a result, it is not entirely clear how to interpret this finding. Furthermore, accuracy modeled against condition for both old and new items indicates that there was no difference in performance between the ‘LO-HI’ and the ‘HI-LO’ condition (p>.5).

There was, however, a consistent difference in accuracy between low and high items, an asymmetry such that learners were more permissive of low forms when they were trained on high; but were less likely to accept high forms when trained on low. This might reflect a bias such that, prior to any training, there is an expectation for a low (or close to zero) degree of nasal coarticulation in all contexts. To test this hypothesis, a control experiment was run. The results are described in the next section.

3 Experiment 2: Control

Experiment 2 was run as a control condition in which participants were exposed to variable input (equal numbers of low and high nasalization tokens in both singular and plural contexts). The results will show whether a bias exists for English speakers to pick low nasalization forms, thus accounting for the asymmetry found in Experiment 1. If the bias hypothesis is correct, then low items should be chosen significantly more often than high in Experiment 2. Otherwise, in the absence of such a bias, we should expect participants to be at chance.

The stimuli and procedure were the same as for Experiment 1, except now participants heard both variants (HIgh and LOw) of each word type. For nasal-stem plurals, there were four possible variants: e.g., {skimtum, skĩmtum, skimtũm, skĩmtũm}; for singulars, there were two: {skimtu, skĩmtu}. Over 4 blocks of training data, participants heard each possible nasal-

---

[5] In the present experiment, stress is not particularly salient for our English-speaking participants since all vowels are unreduced and morpheme final pre-nasal vowels are somewhat lengthened. However, for completeness, I tested the hypothesis that learners may be using stress rather than morphological position to condition the alternation. In order to do so I divided the test stimuli into two sets: one in which the prediction for the degree of nasalization of the stress-alignment hypothesis were identical to the predictions of the morphology-alignment hypothesis, and one set in which the predictions differed. The second set consisted of the one-syllable plurals: {hum,dim,stum,fim,tʃum,spim}, and the first set, of all other plurals. If participants were using the stress-alignment hypothesis then we should expect them to treat the pre-nasal vowels in the ‘divergent’ set of words as they treat stem-internal vowels, and thus differently from how they treat other stem-final, and thus unstressed vowels. The regression analysis in which the model factor considered was membership in the ‘divergent’ set showed no difference in participant response. This indicates that the hypothesis about a response strategy based on word stress must be discarded.

stem plural once, and each possible singular twice. Twenty one Johns Hopkins undergraduates were given course credit to participate in the 30 minute experiment. The thirteen with the highest scores on the second practice trial (> 90% correct) and the AXB task (> 75 % correct overall; average = 88%) were included in analysis. The results are shown in Figure 2.

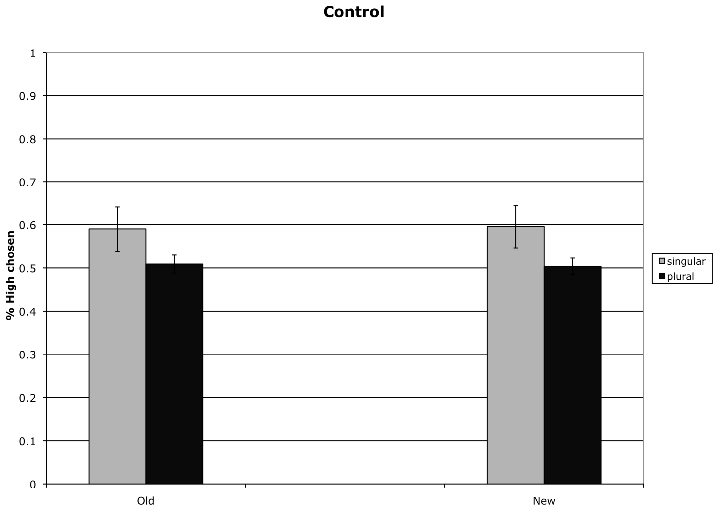

Figure 2
Experiment 2: Percent HIgh (=100% nasalization) variant chosen by Condition (=LO-HI or HI-LO) and Type (=Singular or Plural), plotted separately for old and new words.

The results show that there is no low-nasalization bias for this group of English speaking participants. If anything, there's a slight tendency to pick the high-nasality variant – at least for singular items, that is, stressed vowels. This is consistent with other work characterizing English as a language with a relatively high degree of phonetic nasalization (Sole 1992, Tanowitz & Beddor 1997). To determine if participants were better than Control in Experiment 1 three separate logistic regression analyses were performed; in each the model factor considered was experimental condition. These three regressions were: Old Control plural vs. Old 'LO-HI' plural; Old Control Singular vs. Old 'HI-LO' singular. See Table 4. The same three regressions were performed for the novel test items: New Control plural vs. New 'LO-HI' plural; New Control singular vs. New 'HI-LO' singular. In each case, it is participants' response in Experiment 1 to the items presented with high nasalization during training that are being compared to the corresponding control group. I find that there is no significant difference due to experimental condition in any of these analyses (although 'LO-HI' plural Old items are marginally different from Control). Thus, participants seem to act as though they had received no consistent training for those items, and the learning effect found in Experiment 1 is carried by the Low nasalization items. In order to determine the reason for this asymmetry, I took a closer look at the phonetic characteristics of the experimental stimuli in question.

‘LO-HI’ Plural	Df	Dev	Resid. Df	Resid. Dev	P(>|Chi|)	‘HI-LO’ singular	Df	Dev	Resid. Df	Resid. Dev	P(>|Chi|)
NULL			25	23.85		NULL			25	40.09	
Control Plural	1	3.78	24	20.06	.0516.	Control Singular	1	1.88	24	38.21	.170

Table 4: Experiment 1 High items compared to Control; Each condition analyzed separately. OLD items only.

4 Degree of Nasalization

As described in previous sections, high nasalization items were created by splicing nasalized vowels from a stressed pre-nasal environment. The only criterion used was that a nasal formant (around 1000 Hz) be visible in the spectrogram throughout the length of the vowel. This measure was based on previous work referring only to the proportion of a vowel that was nasalized (rather than to degree) (Tanowitz & Beddor 1997, Beddor & Krakow 1999). This method, however, relying as it does on visual inspection, is clearly quite crude. Furthermore, there has been work suggesting a standardized approach to measuring degree of vowel nasalization. In the following sections I will be following the methodology of Chen (1997).

4.1 Measurements

Coarticulation with a neighboring nasal segment can be observed in nasal formants which are often visible in the spectrogram of a nasalized vowel. The amplitude of these formants is correlated with degree of nasalization. Also correlated with nasalization is lowering of the amplitude of the vowel’s first oral formant. Chen’s metric combines these two cues by considering the amplitude of the first nasal peak in the vicinity of the first oral formant (~950 Hz): P1, and the amplitude of the first oral formant: A1 (Chen 1997). A higher value of P1 (as compared to the oral vowel) indicates a higher level of nasalization, whereas a higher value of A1 (as compared to the oral vowel) indicates a lower level of nasalization. The quantity A1-P1 is computed for both members (oral and nasal) of each pair of tokens, with the difference, Δ, giving the relative amount of nasalization.

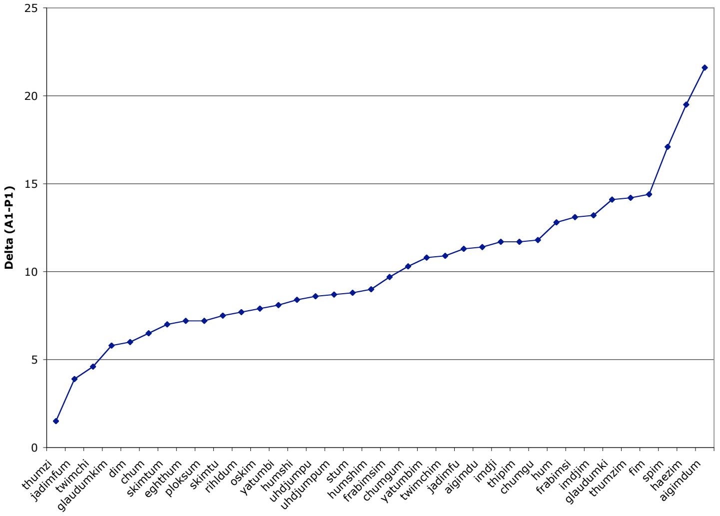

Figure 3
Degree of nasalization: Δ(A1-P1) by stimulus item (in ascending order)

This measure was taken for the stimuli of Experiments 1 and 2, calculating the nasalization difference between the HI and LO variants for each word. The results are plotted in Figure 3 (where nasalization for nasal-stem plurals was measured on the final vowel). The plot is arranged so that nasalization increases monotonically to the right. As can be seen, degree of nasalization varied considerably, with a minimum value of 1.5 dB (θ**u**mzi) and a maximum of 21.6 dB (aɪgimd**u**m) (the measured values fall within the range observed by Chen in her study of the natural productions of 8 English speakers (monosyllables, either completely oral or nasal in context, e.g., ‘bed’ v. ‘men’)). Although the average nasalization for [i] vowels was somewhat higher than that for [u] vowels (11 dB vs. 9 dB), a t-test reveals that this difference was not significant (p>.1). The measurement results provide support for the hypothesis that participants are responding to High tokens as high variability tokens because they do, in fact, exhibit variation in degree of nasalization. Even though this set of training items (‘LO-HI’: plural; ‘HI-LO’: singular) never exhibit ∅ nasalization (Low), the Low choice at test time may be within bounds as a possible variant of the “High” items, leading to a relatively large number of Low responses.

5 Conclusions

This paper has been a description of initial experimental work within a paradigm that combines the strengths of the artificial grammar learning apparatus with the insights of work investigating the phonetic bases of phonological sound changes. The results of Experiments 1 and 2 suggest that listeners can perceive and encode different phonetic associations for boundary versus non-boundary environments. This might be characterized as degree of coarticulation, or perhaps degree of variability of coarticulation. What is more, adult speakers can readily learn new associations of this type with very little training (less than 30 minutes). This result may seem surprising for a couple of reasons. One of these is the body of work that suggests that discriminating and producing phonemic distinctions in a second language that are allophonic in the native language is quite difficult (Goto 1971, Dupoux *et al.* 1998). In the second case, making an association based on morphology requires a level of abstraction on the part of the learner, one that is not always observed in artificial grammar learning experiments in which participants fail to generalize to novel segments or novel words, or unseen members of a natural class (see Peperkamp (2003) for a review of some of the literature).

For these reasons alone, this result is a significant one. However, it also has a broader relevance. Listeners may use the phonetic cue of nasality on a preceding vowel to predict the nasality of a following consonant, or the nasality of a following consonant to assess the quality of the preceding vowel. And on an Evolutionary Phonology account it is a mismatch between an expected and an observed degree of this coarticulation which could lead to the genesis of a phonemic nasal vowel over time. The experimental results presented here satisfy a seemingly necessary condition for this basic misparsing story – an associable phonetic cue for a phonological process – at least in regards to the broad class of derived environment effects. Furthermore, the fact that, typologically, both systems of over-application and under-application in derived environments are attested, is supported experimentally by the learnability of associations in both directions: in the 'LO-HI' condition where nasal harmony applies across morphemes but not within, and the 'HI-LO' condition: where nasal harmony applies within morphemes, but not between. However, there is much further to go in developing a complete theory of the route from phonetics to phonology.

One question that immediately comes to mind is under what circumstances we might expect compensation for coarticulation to fail. In other words, what factors might make it more likely for a mismatch to arise between the listener's expectation and their perception of the speech signal? Intuitively, these circumstances seem to be provided in cases in which the conditioning environment is somehow lost. But what about the situations in which the conditioning environment remains? The derived environment effects in Korean, discussed at the beginning of this paper in (3), would be an example of this type.

We could start by assuming, for the moment, that with perfect knowledge of the correct class of element (the correct boundary) and the expected range of feature spread (or coarticulation) due to that boundary, we are practically perfect in our ability to correctly reconstruct the constituent phonemes. If, on the other hand, our ability to recognize the correct boundary is diminished in some way, then our ability to apply the appropriate degree of compensation for coarticulation is automatically compromised. In this way might a listener attribute a phoneme to a different category than that intended by the speaker, either by interpreting its degree of coarticulation with the neighboring segments as insufficient (and thus subtracting the relevant feature), or analyzing the degree as exceeding expectation (and thus adding the relevant feature).

The way in which morphological junctures might become important to this story beyond demarcating a specific class of phonological processes is by providing a mechanism for initiating the process of phonologization. That is, we could advance the preliminary hypothesis that all internal phonological change originates at the morpheme boundary. This move allows us to make use of independent work in psycholinguistics related to the question of word level processes. In this framework, the ability to reconstruct a morpheme boundary can be related to the representational status of the morphologically complex word. If a particular complex form, due to its high frequency of use, achieved a lexicalized, undecomposed status, then the original morpheme boundary could be thought of as weakening or disappearing.

Lexical access models of morphologically complex words often describe a competition between two routes to word meaning. One route achieves access via composition of the constituent morphemes, and the other through the word as a whole. Sensitivity of response times to word frequency in lexical decision tasks is taken as evidence for access via the whole-word route. This is expected to occur above a certain frequency threshold, such that the compositional route only wins when the whole word frequency is relatively low (see Gordon & Alegre (1999) for discussion of these models).

We don't yet know how sub-phonemic information might enter into the picture. But we could imagine a situation in which an expected high degree of nasalization across a boundary, for example, might fail to be subtracted from the signal by a listener for whom the whole word, rather than the compositional route had become the predominant one. This story may not be the right one, but it provides us with some account of how this original failure to correctly reconstruct the source of the speech signal might systematically occur, and it provides part of the motivation for examining morphological environments under sound change.

Agresti, A. (1996). *An Introduction to Categorical Data Analysis*. New York, Wiley.

Alfonso, P. and T. Baer (1982). Dynamics of vowel articulation. *Language and Speech* **25**: 151-173.

Allen, J. S. and J. L. Miller (2004). Listener sensitivity to individual talker differences in voice-onset-time. *Journal of the Acoustical Society of America* **115**(6): 3171-3183.

Beddor, P. S. and R. A. Krakow (1999). Perception of coarticulatory nasalization by speakers of English and Thai: Evidence for partial compensation. *Journal of the Acoustical Society of America* **106**(5): 2868-2887.

Beddor, P. S., R. A. Krakow, et al. (1986). Perceptual constraints and phonological change: a study of nasal vowel height. In C. Ewen and J. Anderson (eds.) *Phonology Yearbook 3*. Cambridge: Cambridge University Press. 197-217.

Blevins, J. (2004). *Evolutionary Phonology: the emergence of sound patterns*. New York, Cambridge University Press.

Chen, M. (1997). Acoustic correlates of English and French nasalized vowels. *Journal of the Acoustical Society of America* **102**(4): 2360-2370.

Cho, T. (2001). Effects of morpheme boundaries on intergestural timing: evidence from Korean. *Phonetica* **58**: 129-162.

Chomsky, N. (1986). *Knowledge of Language: its nature, origin and use*. New York, Praeger.

Christophe, A., S. Peperkamp, et al. (2004). Phonological phrase boundaries constrain lexical access: I. Adult data. *Journal of Memory and Language* **51**: 523-547.

Clopper, C. G. and D. B. Pisoni (2004). Some acoustic cues for the perceptual categorization of American English regional dialects. *Journal of Phonetics* **32**(1): 111-140.

Dupoux, E., K. Kakehi, et al. (1999). Epenthetic vowels in Japanese: a perceptual illusion. *Journal of Experimental Psychology: Human Perception and Performance* **25**(6): 1568-1578.

Elman, J. L. (2003). Generalization from sparse input. Proceedings of the 38th Annual Meeting of the Chicago Linguistics Society.

Frazier, M. (2005). Output-Output faithfulness to moraic structure: evidence from American English. North East Linguistics Conference, U Mass, Amherst.

Goldinger, S. D. (1996). Words and voices: episodic traces in spoken word identification and recognition memory. *Journal of Experimental Psychology: Learning, Memory and Cognition* **22**: 1166-1183.

Gordon, P. and M. Alegre (1999). Is there a dual system for regular inflections? *Brain and Language* **68**: 212-217.

Goto, H. (1971). Auditory perception of normal Japanese adults of the sounds 'l' and 'r'. *Neuropsychologia* **9**: 317-323.

Hardcastle, W. J. (1985). Some phonetic and syntactic constraints on lingual coarticulation in stop consonant sequences. *Speech Communication* **4**: 247-263.

Hay, J. (2003). *Causes and Consequences of Word Structure*. Routledge.

Holst, T. and F. Nolan (1995). The influence of syntactic structure on [s] to [S] assimilation. In B. Connell and A. Arvaniti (eds.) *Phonology and Phonetic Evidence: Papers in Laboratory Phonology IV*. Cambridge University Press. 315-333.

Jaeger, T. F. (in press). Categorical data analysis: Away from ANOVAs (transformation or not) and towards logit mixed models. *Journal of Memory and Language*.

Jusczyk, P. W., M. B. Goodman, et al. (1999). Nine-month-old's attention to sound similarities in syllables. *Journal of Memory and Language* **40**: 62-82.

Kawasaki, H. (1986). Phonological universals of vowel nasalization. In J. J. Ohala and J. J. Jaeger (eds.) *Experimental Phonology*. Orlando, FL.: Academic Press, Inc. 81-98.

Kaye, J. (1974). Morpheme structure constraints live! *Montreal Working Papers in Linguistics* **3**: 55-62.

Keating, P. A. (1985). Universal phonetics and the organization of grammars. In V. A. Fromkin (ed.) *Phonetic Linguistics: Essays in honor of Peter Ladefoged*. Orlando: Academic Press. 115-132.

Mann, V. and B. H. Repp (1980). Influence of vocalic context on perception of the [S]-[s] distinction. *Perception and Psychophysics* **28**: 213-228.

Maye, J., J. E. Werker, et al. (2002). Infant sensitivity to distributional information can affect phonetic discrimination. *Cognition* **82**: B101-B111.

McCarthy, J. J. (2002). Comparative Markedness. In A. C. Carpenter, A. W. Coetzee and P. de Lacy (eds.) *Papers in Optimality Theory II*. Amherst, MA. 171-246.

Newport, E. L. and R. N. Aslin (2004). Learning at a Distance I. Statistical Learning of Non-adjacent Dependencies. *Cognitive Psychology* **48**: 127-162.

Ohala, J. (1974). Experimental historical phonology. In J. M. Anderson and C. Jones (eds.) Historical Linguistics II. *Theory and description in phonology*. Amsterdam: North Holland. 353-389.

Ohala, J. (1981). The Listener as a source of sound change. Parasession on Language and Behavior: Chicago Linguistics Society, Chicago.

Ohala, J. (1990). There is no interface between phonology and phonetics: a personal view. *Journal of Phonetics* **18**: 153-171.

Ohala, J. (1993). The phonetics of sound change. In C. Jones (ed.) *Historical Linguistics*. 237-278.

Peperkamp, S. (2003). Phonological acquisition: Recent attainments and new challenges. *Language and Speech* **46**(2-3): 97-113.

Pisoni, D. B., R. N. Aslin, et al. (1982). Some effects of laboratory training on identification and discrimination of voicing contrasts in stop consonants. *Journal of Experimental Psychology: Human Perception and Performance* **8**(2): 297-314.

Remez, R. E., J. M. Fellowes, et al. (1997). Talker identification based on phonetic information. *Journal of Experimental Psychology: Human Perception and Performance* **23**: 651-666.

Rumelhart, D. E. and J. M. McClelland (1986). On learning the past tenses of English verbs. In *Parallel distributed processing: explorations of the microstructure of cognition, vol 2: psychological and biological models*. Cambridge, MA: MIT press. 216-271.

Saffran, J. R. and E. D. Thiessen (2003). Pattern induction by infant language learners. *Developmental Psychology* **39**(3): 484-494.

Salverda, A. P., D. Dahan, et al. (2003). The role of prosodic boundaries in the resolution of lexical embedding in speech comprehension. *Cognition* **90**: 51-89.

Sole, M.-J. (1992). Phonetic and phonological processes: the case of nasalization. *Language and Speech* **35**(1,2): 29-43.

Summerfield, Q. (1981). Articulatory rate and perceptual constancy in phonetic perception. *Journal of Experimental Psychology: Human Perception and Performance* **7**: 1074-1095.

Tanowitz, J. and P. S. Beddor (1997). Temporal characteristics of coarticulatory vowel nasalization in English. *The Journal of the Acoustical Society of America* **101**: 3194A.

Whalen, D. H. (1984). Subcategorical phonetic mismatches slow phonetic judgments. *Perception and Psychophysics* **35**: 49-64.

Wilson, C. (2006). Learning Phonology with Substantive Bias: An Experimental and Computational Study of Velar Palatalization. *Cognitive Science* **30** (5): 945-982.