To Appear in A. Yu (ed.), *Origins of Sound Change: Approaches to Phonologization*. Oxford: Oxford University Press.

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

Rebecca Morley

Word count: 7,295

* Affiliation: The Ohio State University

Mailing Address:

Linguistics Department
The Ohio State University
200 Oxley Hall
1712 Neil Ave.
Columbus, OH 43210
morley@ling.osu.edu

8.1 Introduction

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 learning by infants and adults (cf. Jusczyk *et al.* 1999, Maye *et al.* 2002, 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.

The fact that many aspects of phonetics cannot be attributed to universal processes of articulation and motor planning, but that individual languages adopt individual phonetic implementations, is evidence that these phonetic facts must be learned, and therefore, that learners must be able to induce them from the speech stream. Research suggests that among the relations that listeners must encode are the degree to which their language nasalizes vowels before nasal consonants, and 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).

8.1.1 Evolutionary Phonology

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 reconstruction 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).

If, however, listeners fail to correctly compensate, either by under-correcting or over-correcting, that is where change can occur. The stage at which acoustics, or phonetics, becomes part of the phonology is the point at which a persistent difference arises between the original source of the signal (the speaker's intention), and what the listener encodes. This may happen either because what the listener perceives runs counter to their phonetic expectation, or because the signal is inherently ambiguous in some way. Consider again the case of nasalized 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 are usually attributed to the phonetics rather than the phonology (see Cohn (1993) for articulatory evidence of gradient nasalization in English). The standard account for the historical emergence of phonemic nasal vowels attributes their origins to these phonetic allophones. Furthermore, it is usually assumed that the conditioning nasal 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 can be used to assess the quality of the preceding vowel. The case in which 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, in the absence of the originally conditioning nasal consonant, the nasality cues on the vowel may first become more salient to the listener, leading them to be encoded as part of the underlying representation of the vowel, and ultimately to an oral/nasal vowel contrast in the language as a whole. This argument gains support from experimental work by Kawasaki (1986) who found that, for a series of one syllable stimuli, English speaking participants rated the vowel as more nasal, the lower the amplitude (and thus the lower the perceptibility) of the final nasal consonant[1].

The ultimate story, of course, must be significantly more complicated than this, given the multifarious character of natural language. As one example of this complexity, consider the prosodic hierarchy, nested levels of constituent structure each of which may have its own associated phonetic and phonological rules (see Selkirk (1980), Byrd and Saltzman (1998)). Furthermore, a large body of psycholinguistic work on processing of sentence, word and morpheme size units exists to motivate a cognitively plausible model of phonological competence. As far as I am aware, there exists no study of how these known factors should or could be incorporated into a diachronically-based theory. The current work represents a first step in integrating these linguistic areas, laying the groundwork for consideration of phonetic cues within structured domains and how they might become phonological cues over time.

The particular cue investigated in this article is vowel nasalization, and the methodology is experimental – an artificial grammar learning task in which vowel nasality is linked with morphological inflection. 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.

8.1.2 Morphemes & Derived Environment Effects

Evidence for the activity of phonological rules comes largely from alternations; many of these alternations occur, in turn, at morphological boundaries. In some cases the morphophonology reflects rules or constraints that are also observable in the static phonotactics of the language. In other cases, the alternation reflects a restriction that does not apply to mono-morphemes (over-application). Finally, the absence of an expected alternation indicates a restriction that holds only for simplex and not complex words (under-application).

Consider the following example of an over-application derived environment effect. Stop consonants preceding high front vowels surface as affricates in Korean as shown in (1a). This rule, however, fails to apply when the conditioning environment and the undergoer occur within the same morpheme, as shown in (1b) (from McCarthy 2002).

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

Translating this phenomenon into Evolutionary Phonology terms, we could hypothesis a point in the history of Korean at which a discrepancy arose between the amount of palatalization perceived across the morpheme boundary, and that anticipated by the listener. An additional factor that makes this hypothesis an interesting one, is the general assumption of productivity in morphology. A difference between off-line processing (monomorphemes), and on-line processing (polymorphemes) could provide the basis for derived environment effects. These, in turn, positing a gradual spread of the rule or constraint from one environment to another, could potentially provide the basis for the emergence of any new phonological pattern (see the Conclusion for further discussion of this idea).

8.2 Experiment

The present experiment is designed to investigate the relation between phonetic and phonological patterns, in particular, the hypothesis that phonologization is initiated due to a discrepancy in perception or production, either of which can lead to a difference between the listener's analysis of their input and the speaker's intended output. 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. A clearly necessary pre-condition for a failure of expectation is the establishment of such an expectation in the first place. The experiment described here will be concerned with testing the hypothesis that such an expectation can be induced – that listeners are able to attend to sub-phonemic vowel nasalization, as well as learn novel rules that link those cues with the grammatical structure of morpheme boundaries.

Previous experimental work has found both production and perception differences related to the presence or absence of morphological boundaries. Measurements of Korean speakers' productions 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 differences have also been shown, in some cases, to be accessible to hearers. Frazier reports a perceptual effect 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) on word selection (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.

These perception experiments employed experimental tasks that were centered around the explicit disambiguation of semantically distinct minimal pairs. The current experiment, on the other hand, involves explicit training on a novel morphological alternation (the presence or 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, the work described above 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), 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 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.

8.2.1 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[2], and that they would later be asked questions about those words. Each word would appear in the singular, accompanied by a picture of a single object, and then in the plural, accompanied by 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 such pairs, with the singular appearing 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 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 ('1' for the first word; '2' for the second). 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. Since all stems were vowel-final, all plural words contained such a vowel before the plural suffix (-/m/), that is, at the morpheme boundary. Half of the words also contained stem-internal nasals. In both conditions the degree of regressive nasalization on the vowel contrasted in these two environments. In training for the ORAL-NASAL condition, there was 0% regressive nasalization within morphemes and 100% across; in the NASAL-ORAL condition, those values were reversed. See Table 1 for example stimuli. It should be noted that the ORAL-NASAL condition presents an alternation to the learner on the word final vowel, whereas the NASAL-ORAL condition shows no alternation.

Half of the stems ended in /i/ and half in /u/[3]. 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 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 without any nasalization on the pre-nasal vowel, but otherwise identical. Test items included both old words (heard during training), 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 8.1 gives example stimuli for each condition.

The phonetic cues present in the stimuli are natural (regressive nasalization associated with nasal consonants) and redundant (only 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 triplet, 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].

[Table 1 here]

8.2.2 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 morpheme was always tonic, but sometimes occurred in the first and sometimes in the second syllable. Natural speech tokens were used; each token was recorded separately. 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. Also recorded were monomorphemes with final oral consonants, and monomorphemes with final nasal consonants, e.g., /hæzim/, /hæzi/, /zib/, /zim/. Nasalized vowel tokens were created by splicing a portion of the vowel from a stressed nasal coda environment ([zĩm], generating [hæzĩm]). Since these stressed vowels were longer in duration it was possible to select only the part of the vowel that was nasalized. This was determined by visual inspection, verifying that a nasal formant (around 1000Hz) was visible throughout the duration of the vowel. See Fig. 8.1 for an example spectrogram. Non-nasalized tokens were created by splicing from a non-nasalized environment, e.g., [zib], generating [hæzim]. For these items, no nasal formant was visible in any part of the vowel. Vowels were normalized for length, such that root-internal front vowels did not significantly differ across oral and nasal tokens (similarly for root-final front vowels, root-internal back vowels, and root-final back vowels). To reduce auditory artifacts, splicing was always done at zero crossings, and effort was taken to produce a smooth intensity and frequency contour by splicing from multiple parts of the replacement vowel (beginning, middle, and end). Intensity was adjusted where necessary to avoid the percept of stress on the spliced vowel.

[Fig. 8.1 here]

8.2.3 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). Only the results from the 13 participants in each condition who reached threshold on the AXB task (> 70% singular and plural items, separately) are plotted. See Fig. 8.2. Analyses were subsequently carried out for the entire set of participants (adding back in the 32 who fell below threshold).

8.2.4 Results

The results discussed here are for responses in the 2AFC task. Only participants who performed above threshold on the AXB discrimination task are included for the first analyses. Test items consisted of two types: singular and plural. The choice at test was always between a nasalized variant and a non-nasalized variant (that is, the two possible responses differed only in nasality on the critical vowel). Two conditions were contrasted: ORAL-NASAL: nasalization across boundary, none within; and NASAL-ORAL: no nasalization across boundary, nasalization within. Each participant was run in only one condition.

[Fig. 8.2 here]

The dependent variable in the first analysis was the percentage of responses for which participants selected the nasalized variant (as opposed to the non-nasalized). Under successful learning, the value of this variable should be large for plural items in the ORAL-NASAL condition, and small for singular items (small for plural items in the NASAL-ORAL condition, large for singular items). As just described, the prediction is that there should be a significant interaction between condition and test type. See Fig. 8.2a.

Since the dependent variable was a proportion response, varying between 0 and 1, a logistic regression analysis was performed (Jaeger in press, Agresti 1996). 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 here and in the final column of the tables in the Appendix.

For the first analysis two separate regressions were performed, one for old items, to establish firstly discriminability and attention, and one for new items, to show true generalization. Model terms were Condition (ORAL-NASAL or NASAL-ORAL) and Type (plural or singular), as well as a term for the interaction between Condition and Type (the critical test of difference). Old Items: There was no main effect of Type, but Condition was significant ($\chi^2(1)=6.38$, $p<.05$). Adding the critical interaction term improved the model fit by a significant factor ($\chi^2(1)=39.92$, $p<.0005$). New Items: There was no main effect for Condition or Type. Adding the critical interaction term improved the model fit by a significant factor ($\chi^2(1)=11.46$, $p<.005$).

Alternatives to the morphological hypothesis were also considered. One possibility is that the participants were learning a single degree of nasalization (rather than 2 differing by morphological location); another possibility is that nasalization was encoded by stress rather than morphological condition (since stress location is largely confounded). To test these hypotheses, a second regression model was run by item for plurals alone, with participant accuracy as the dependent variable, and model terms for stem type (nasal stem/oral stem, which provides a test of the single-association hypothesis by comparing performance on nasal stem plurals with other plurals), and stress location (pre-final/final). No effect was found for the interaction of stem type with stress location.

As can be seen from Fig. 8.2b, response levels hovered near chance for items that were nasalized during training. A regression model of participant accuracy with a term for Training type showed that responses were significantly more likely to be accurate for trained Oral (ORAL-NASAL singular and NASAL-ORAL plural) than for trained Nasal (ORAL-NASAL plural and NASAL-ORAL singular) ($\chi^2(1)=12.33$, $p<.0005$). No effect, however, was found for condition, and the results of the two conditions are combined in Fig. 8.2b.

A final analysis involved pooling the data from all 58 participants (including those previously excluded due to their below-threshold performance on the AXB task). The results of analysis 1 hold when performed over all participants, implying that learning has taken place. This in turn suggests that the 70% discrimination threshold is an unnecessarily stringent criterion. Discrimination level (averaged over all test types in the AXB task) was added as a continuous term to the regression model of accuracy. This term was a significant predictor of accuracy over all test items ($\chi^2(1)=7.48$, $p<.05$). This shows, perhaps unsurprisingly, that the better participants were at the AXB task (at reliably discriminating the difference to be learned), the more accurately they performed at test.

8.2.5 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.

A main effect of Condition for old items indicates that the nasalized variant was less likely to be chosen overall for the NASAL-ORAL condition than the ORAL-NASAL 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 ORAL-NASAL and the NASAL-ORAL condition ($p>.5$).

There was, however, a consistent difference in accuracy between Oral and Nasal items, an asymmetry such that learners were more permissive of oral forms when they were trained on nasal; but were less likely to accept nasal forms when trained on oral. 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.

8.3 Experiment 2: Control

Experiment 2 was run as a control condition in which participants were exposed to variable input (equal numbers of oral and nasal tokens in both singular and plural contexts). The results will show whether a bias exists for English speakers to pick oral forms, thus accounting for the asymmetry found in Experiment 1. If the bias hypothesis is correct, then oral items should be chosen significantly more often than nasal 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 (Nasal and Oral) 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-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 Fig. 8.3.

[Fig. 8.3 here]

The results show that there is no oral bias for this group of English speaking participants. If anything, there's a slight tendency to pick the nasal 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). Regression analyses comparing the results of Experiment 1 to Experiment 2 show that there is no significant difference between the trained Nasal items and participants' responses after hearing unpredictable nasalization (although ORAL-NASAL plural Old items are marginally different from Control). In other words, 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 trained Oral 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.

8.4 Degree of Nasalization

As described in previous sections, nasalized 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 spliced vowel portion. 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).

8.4.1 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 a 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. A1 and P1 were measured in each instance using a spectral window over the entire length of the vowel. A pair of example spectral windows are shown in Fig. 8.4b for the token ‘oskim’ (cf. Fig. 8.1), with approximate P1 location indicated.

This measure was taken for the stimuli of Experiments 1 and 2, calculating the nasalization difference between the Nasal and Oral variants for each word. Box plots of the results are given in Fig. 8.4a (where nasalization for nasal-stem plurals was measured on the final vowel). As can be seen, degree of nasalization varied considerably, with a minimum value of 1.5 dB (θu̲mzi vs. θũ̲mzi) and a maximum of 21.6 dB (aɪgimdu̲m vs. aɪgumdũ̲m)[4]. Average nasalization for [i] vowels was somewhat higher than that for [u] vowels (11 dB vs. 9 dB).

If Chen’s measurement can be taken as an objective degree of nasalization then the spread in Fig. 8.4a indicates a non-uniform distribution over the Nasal tokens participants heard in each of the experimental conditions. The oral bias reported in the results section becomes less surprising when viewed in this light. Training conditions are more accurately characterized as Oral-Variable Nasal, and Variable Nasal-Oral. The lack of measurable variation in the nasality of the Oral items correlates with participants’ more consistent rejection of nasal test variants. Oral variants, however, are more acceptable responses to a category of training items that exhibited a range of nasalization degrees. Another way to couch these results is as participants’ sensitivity to a departure from English-like coarticulation – found only in the oral tokens.

[Fig. 8.4 here]

8.5 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 the development of an expectation for a degree of coarticulation, or perhaps degree of variability of coarticulation. What is more, adult speakers can 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 which 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 can lead to the genesis of a phonemically nasal vowel over time. The experimental results presented here satisfy a necessary condition for this basic misparsing story – the ability to develop such an expectation in the first place, in particular, an association tied to the linguistically active domain of the morpheme boundary. 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 situations in which the conditioning environment remains? The derived environment effects in Korean, discussed at the beginning of this paper in (1), 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 can imagine a situation in which a speaker produces the appropriate high degree of nasalization across a boundary, but a listener for whom the whole word, rather than the compositional route has become the predominant one fails to compensate for the boundary effect. 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, an element often missing from historically-based accounts. Furthermore, it raises an intriguing hypothesis about the relation between domain-limited effects in linguistics and general processes. If the former is a necessary stage on the way to the latter, then a more careful study of a seemingly marginal phenomenon like derived environment effects might prove fruitful for insights into linguistic phenomena of all kinds.

Acknowledgments

This work was supported by a Department of Education Javits fellowship as well as a National Science Foundation IGERT grant to the Johns Hopkins Cognitive Science department. I would like to thank Paul Smolensky, Colin Wilson and the members of the Johns Hopkins IGERT lab.

References

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. (2009). A Coarticulatory Path to Sound Change. Language 85(4): 785-821.

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.

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

Byrd, D. and E. Saltzman (1998). "Intragestural dynamics of multiple prosodic boundaries." Journal of Phonetics 26: 173-199.

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.

Cohn, A. C. (1993). Nasalisation in English: phonology or phonetics. Phonology 10: 43-81.

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.

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

Jaeger, T. F. (2008). Categorical data analysis: Away from ANOVAs (transformation or not) and towards logit mixed models. *Journal of Memory and Language* **59**(4): 434-446.

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.

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.

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.

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.

Selkirk, E. O. (1980). Prosodic domains in phonology: Sanskrit revisited. Juncture. M. Aronoff and M.-L. Kean, Anma Libri: 107-129.

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.

Appendix

Table A1: Experiment 1: % Nasal response fit to Condition, Test Type and Condition x Type; above threshold participants only

OLD	Df	Dev	Resid. Df	Resid. Dev	P(>|Chi|)	**NEW**	Df	Dev	Resid. Df	Resid. Dev	P(>|Chi|)
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 A2: Experiment 1: % Correct response by item for plurals only; model fit to Stress location (pre-final/final), Nasality of Stem (2 nasals/1 nasal); above-threshold participants only

OLD & NEW combined	Df	Dev	Resid. Df	Resid. Dev	P(>|Chi|)
NULL			23	32.282	
NasalStem	1	0.58031	22	31.702	0.4462
Stress	1	0.04912	21	31.653	0.8246

Table A3: Experiment 1: % Correct response fit to Condition, Test Type, Training type (Oral/Nasal), and AXB results; all participants

OLD & NEW combined	Df	Dev	Resid. Df	Resid. Dev	P(>|Chi|)
NULL			231	278.42	
Cond	1	0.3761	230	278.04	0.54
Type	1	1.0147	229	277.02	0.31
Training	1	12.3306	228	264.69	0.0004 ***
AXB results	1	7.4753	227	257.22	0.006**

---

[1] But see Beddor (2009) for an alternative account that situates this type of phonologization in the inherent ambiguity of the signal, where trading relations in the localization of phonetic cues are the source of the discrepancy between speaker and listener (this closely resembles Blevin's (2004) CHANCE route to change).

[2] The speaker was identified in the hopes of priming speaker identification – a task known to support the encoding of sub-phonemic cues (Remez *et al*. 1997, Allen & Miller 2004).

[3] The full list of stems:

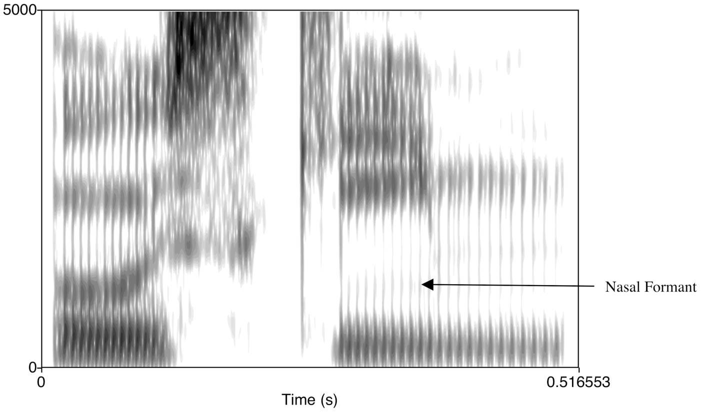
Old Stems (heard both in training and test)				New Stems (heard only at test)			
ˈhæzi	ˈtʃu	ˈskimtu	aiˈgimdu	ˈdi	ˈfi	ʒaˈdimfu	ˈtʃumgu
ˈoski	ˈspi	jaˈtumbi	əˈdʒumpu	ˈploksu	ˈðɪpi	ˈhumʃi	glauˈdumki
ˈhu	ˈɛgθu	ˈθumzi	ˈtwimtʃi	ˈrɪldu	ˈstu	fraˈbimsi	ˈimdʒi

[4] 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').

Table 8.1: Example training and test items for the two experimental conditions.

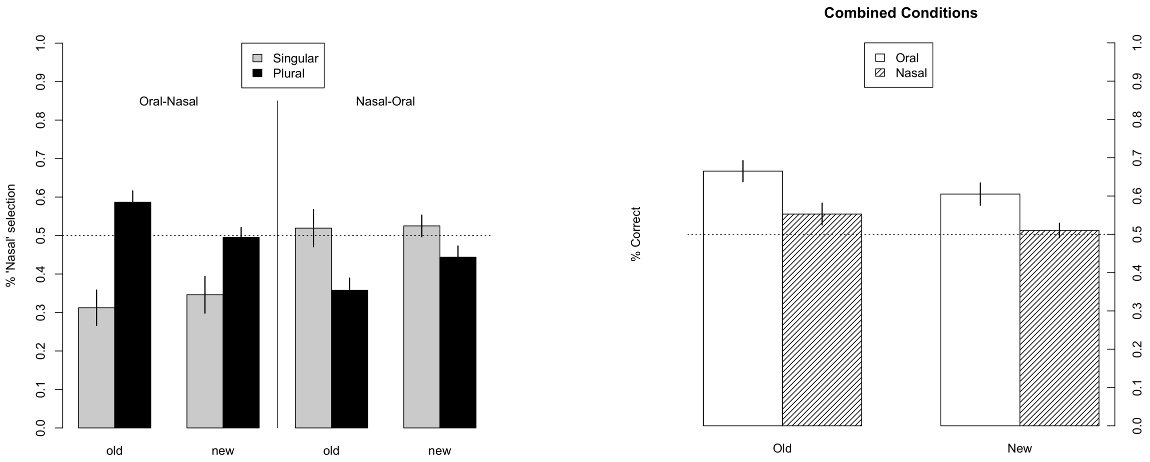
ORAL-NASAL			**NASAL-ORAL**		
Training		Test	Training		Test
Singular	Plural		Singular	Plural	
oski	oskĩm	oskĩm/oskim	oski	oskim	oskĩm/oskim
skimtu	skimtũm	skimtũm/skimtum[a]	skĩmtu	skĩmtum	skĩmtũm/skĩmtum[a]
		skimtu/skĩmtu			skimtu/skĩmtu

[a] The plural nasal-stem test items were different across the ORAL-NASAL and NASAL-ORAL conditions. This was so that only one position was being tested at a time. In the plural nasal-stem items the trained 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.



[oskĩm]

Fig. 8.1
Example spectrogram showing nasal formant



(a)

(b)

Fig. 8.2
Experiment 1:
(a) Percent Nasal variant chosen by Condition (=ORAL-NASAL or NASAL-ORAL) and Type (=Singular or Plural), plotted separately for old and new words.
(b) Percent Correct chosen by Type, plotted separately for old and new words; Conditions combined.

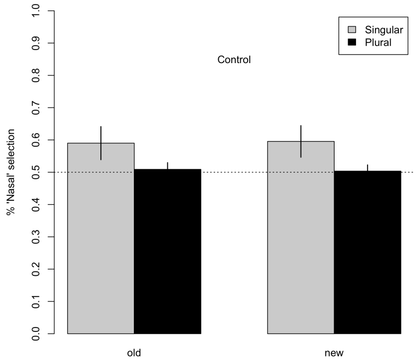
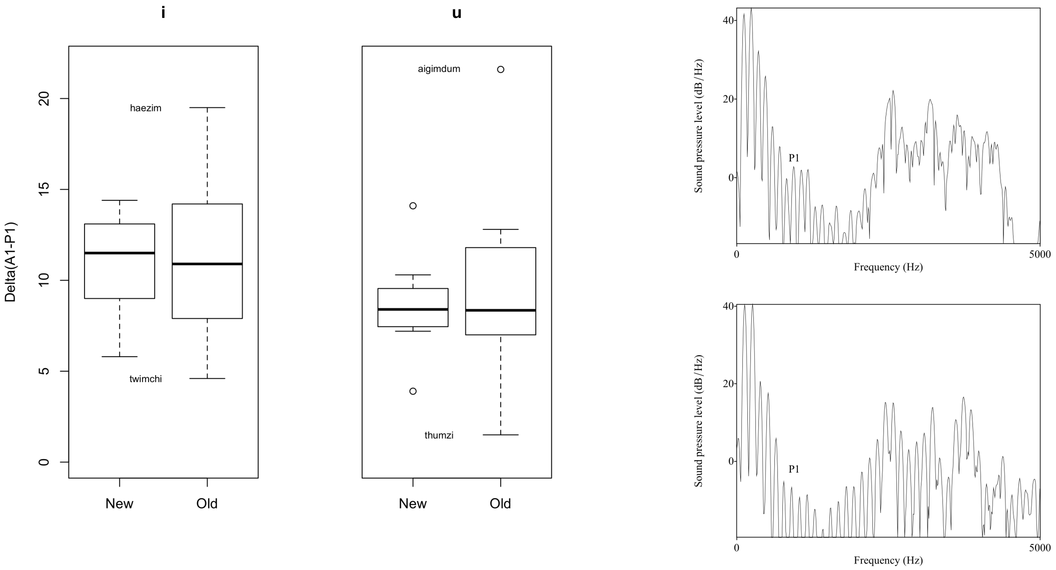


Fig. 8.3
Experiment 2: Percent Nasal variant chosen by Type (=Singular or Plural), plotted separately for old and new words.



(a)

(b)

Figure 8.4

(a) Box plot of Degree of nasalization: Δ(A1-P1). Separately by vowel (i or u), and separately by Old and New words. The extreme outlier tokens are indicated.

(b) Example Spectral Slice from token ‘oskim’ (see Fig. 1). Top: Nasal; Bottom: Oral. The location of the nasal formant is indicated by P1.