

## Unnatural Classes and Phonological Generalization in Dialect Formation

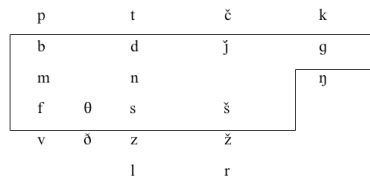
Aaron J. Dinkin, University of Pennsylvania  
*dinkin@babel.ling.upenn.edu*

**The Northern Cities Shift** begins with **general tensing**: /æ/ raises, fronts, and diphthongizes in all environments.

Labov et al. (2006): “This general raising pattern appears to be the type of simplification that often occurs in situations of radical **dialect mixture** with rapid population growth: a koineization.... In such a situation, it is not unusual for different conditioning factors, sub-categories and sub-rules to disappear in favor of the **simplest possible** treatment.”

**Dialect mixture?**: Individuals from New York City and New England, with different short-*a* patterns, settling in the same area (western New York State).

**New York’s “split” short-*a* system** (diagram from Labov 2005):



Tensing occurs: before voiceless fricatives, voiced stops, and nonvelar nasals, **but**:

- lax in morphologically simple function words (*can, and*)
- lax in open syllables without level-2 morphological boundary (*animal, hammer*)
- various lexical exceptions (tense *avenue*, lax *alas*)

**New England’s nasal short-*a* system**: tensing occurs before all nasals, regardless of syllabic or morphological status.

**Nasal system is default American short-*a* pattern**:

- It occurs in disjoint and unrelated regions (Labov et al. 2006):
  - dominant system in NE, northern N.J., W. Pa., Atlanta, Charleston SC, Fla.
  - also very frequent in Midland and West, along with “continuous” short *a*
- Communities that retreat from Phila. split system end up with nasal system (Ash 2002, Friesner & Dinkin 2006)

**So if koineization is a retreat to the least marked system, why would the result be other than the nasal pattern?**

Three categories of short-*a* words:

- **Class 1**: Tense in both nasal system and NY split system (*plan, hamper*)
- **Class 2**: Tense in split system and lax in nasal system (*class*) or vice versa (*animal*)
- **Class 3**: Lax in both systems (*trap, cat*)

**Exemplar theory** is disfavored by Labov et al.’s account:

Bybee (1999): “Segment inventory can be derived” from “repeated sets of coordinated gestures”—i.e., phonemes aren’t stored as part of underlying phonological representation; they arise from distribution of lexical items.

Exemplar theory predicts Class 3 words would remain lax in a dialect-contact situation:

Each word’s evolution is determined by its own exemplars, and Class 3 has no tense exemplars to begin with.

**Therefore**: the koineization account of general tensing needs to assume abstract underlying phonemic representations.

**Phonemic status of the split system and the nasal system**

**Split system**:

- Conditioning rules are complicated—several interacting morph/phonological criteria
- **But** tense and lax short *a* are separate phonemes:
  - Late-learned words often disobey distributional constraints
  - Payne (1976): NY natives learning Phila. system ignore phonological regularities
- Suggests: NY speakers **don’t regard tense and lax *a* as synchronically related** (contra Kiparsky 1995)
- Complicated conditioning rules **needn’t be part of grammatical knowledge**.

**Nasal system**:

- Conditioning rule is very simple; only one phonetic criterion
- Single phoneme with **easy-to-acquire allophonic rule** (cf. Labov 2006)

**What happens when they’re combined?**

A child growing up with input from both split and nasal systems would hear **tense and lax tokens** of Class 2 words: **variable tensing** for these words.

If there’s only one word in Class 2, then the child could decide that Class 1 and Class 3 are separate phonemes, and that one word has variable representation. **But**:

**How many Class 2 words are there?**

513 most frequent short-*a* words in the Brown Corpus of Standard American English (frequency data source: <http://www.edict.com.hk/textanalyser/wordlists.htm>):

<b>Class 1</b>	117 words	<b>23%</b>
<b>Class 2</b>	139 words	<b>27%</b>
<b>Class 3</b>	257 words	<b>50%</b>

With a large number of words in Class 2 with apparent variable tensing, the learner can conclude that there is **one phoneme** with tense and lax short *a* as possible realizations, related by some **variable rule** (cf. Kroch et al. 2000 for a similar phenomenon in syntax).

But **what is the rule relating the allophones?**

- Tensing is **obligatory** in Class 1: before **tautosyllabic nasals**.
  - Tensing is **variable** in Class 2:
    - before **heterosyllabic** nasals;
    - before **tautosyllabic** voiced stops;
    - before tautosyllabic **voiceless** fricatives.
- No natural class of environments!

Since the environments in which variable tensing occurs are heterogeneous and have few features in common, learners may **overgeneralize** and attribute variable tensing to a **wider class of environments** in order to simplify the rule (again cf. Kroch et al. 2000). The smallest phonological natural class that encompasses all the Class 2 environments is the class of **all environments!**

**Summary:** Learners in the mixed community are trying to:

- acquire a **complicated and unnatural** distribution of tensing (like NY speakers)
  - interpret the distribution as a **single phoneme** (which NY speakers don't have to do)
- ...and this leads them to simplify the rule and overgeneralize.

**Caveats:**

- This account of how koineization could have occurred is necessarily speculative.
- Labov et al. (2006)'s hypothesis of koineization is itself speculative:
  - More information is needed about 19<sup>th</sup>-C. New England and New York dialects.
  - Hanley records (ADS 1931–7)?

**So what is the point?**

This phonological account provides a model under which koineization is plausible as an explanation for general tensing, in the absence of direct evidence.

Potential contributions of dialectology research to phonological theory:

- Exemplar theory vs. traditional models of phonology
- Markedness of phonological rules: how unnatural can rules be and still be learnable?

**References:**

- American Dialect Society (1931–1937). *American Dialect Society Collection*. Collection number AFC 1984/011, Archive of Folk Culture, American Folklife Center, Library of Congress, Washington, D.C.
- Ash, Sharon (2002). "The Distribution of a Phonemic Split in the Mid-Atlantic Region: Yet More on Short A". Penn Working Papers in Linguistics 8.3.
- Bybee, Joan (1999). "Usage-Based Phonology". In Michael Darnell, Edith Moravcsik, Frederick Newmeyer, Michael Noonan, & Kathleen Wheatley (eds.), *Functionalism and Formalism in Linguistics, Volume I: General Papers*, 211–242. John Benjamins, Amsterdam.
- Friesner, Michael & Aaron Dinkin (2006). "The Acquisition of Native and Local Phonology by Russian Immigrants in Philadelphia". Penn Working Papers in Linguistics 12.2.
- Kiparsky, Paul (1995). "The Phonological Basis of Sound Change". In John A. Goldsmith (ed.), *The Handbook of Phonological Theory*. Blackwell, Cambridge, Mass.
- Kroch, Anthony, Ann Taylor, & Donald Ringe (2000). "The Middle English Verb-Second Constraint: A Case Study in Language Contact and Language Change". In Susan Herring, Lene Schoesler, & Pieter van Reenen (eds.), *Textual Parameters in Older Language, Current Issues in Linguistic Theory 195:353–391*. John Benjamins, Philadelphia.
- Labov, William (2005). "Transmission and Diffusion". Available as [www.ling.upenn.edu/~wlabov/Papers/TD.pdf](http://www.ling.upenn.edu/~wlabov/Papers/TD.pdf)
- Labov, William (2006). "The Binding Force in Segmental Phonology". Talk presented at the 2006 LSA annual meeting, Albuquerque, N.M.
- Labov, William, Sharon Ash, & Charles Boberg (2006). *Atlas of North American English*. Mouton/de Gruyter, Berlin.
- Payne, Arvilla Chapin (1976). *The Acquisition of the Phonological Systems of a Second Dialect*. Ph.D. dissertation, University of Pennsylvania.
- Pierrehumbert, Janet (2002). "Word-Specific Phonetics". In Carlos Gussenhoven & Natasha Warner (eds.), *Laboratory Phonology VII*: 101–140. Mouton/de Gruyter, Berlin.

**Acknowledgements:**

Thanks to Michael Friesner for pointing out to me the question addressed in this paper, and to William Labov and Tony Kroch for their advice while I was revising and preparing it.