Andries W. Coetzee, University of Michigan
May 2017

Grammar, Variation and Typology in Phonology

In this statement, I will focus on the relationship of between-language phonological variation (what I will call "typology" below) and within-language phonological variation (what I will call "variation" below). I will argue for a tight connection between typology and variation, such that it should be possible to map directly from patterns of variation to typological patterns, and vice versa. Focusing more on variation, I will then argue that it is influenced by both grammatical and non-grammatical factors. Although grammatical factors are typically dominant over non-grammatical factors, it may be possible for non-grammatical factors to sometimes trump grammar. In those instances, the tight link between within-language variation and typology may be broken.

1. A tight connection between typology and within-language variation

Early generative phonology allowed for phonological variation. Chomsky and Halle (1968), for instance, not only include several optional rules in *Sound Pattern of English* (*SPE*), but also acknowledge variation as a part of the design of their grammatical model: "It is not necessarily the case that each deep structure determines a single phonetic representation; if the grammar contains optional rules or analyses, a given deep structure can underlie two or more phonetic transcriptions." (1968:294; see also Postal, 1966:185; 1968:14-15.) Although early generative grammar did therefore allow for variation, variation was nonetheless usually relegated to the edges of grammar and not considered a core concern for linguistic theory—variation was often considered to be a result of "performance" factors rather than of core "competence", and were ascribed to things such as the articulatory externalization of linguistic categories, the impact of memory limitations on production, real world knowledge, etc. (Chomsky, 1965:3-4, 10-15; see Newmeyer, 2003:695-698 for a more recent but similar opinion).

Due to the marginalization of variation, there was little to no consideration of the relationship between variation and typology in the early generative linguistic tradition. This situation changed with the reconceptualization of variation not as rule optionality but as competition between multiple grammars (for phonological variation, see Anttila, 1997; 2002; Kiparsky, 1993; for morphosyntactic variation, see Kroch, 1989; 1994). If variation is the result of the co-existence of more than one grammar in an individual speaker, and if each of the competing grammars could also constitute a single language, then there is a much more direct link between within-language variation and typology.

In the phonological tradition, the competing grammars model of variation was developed in most detail in constraint-based frameworks, whether that be in the tradition of Optimality Theory (OT; Anttila, 1997, 2002, 2007; Boersma, 1997; Boersma & Hayes, 2001; Coetzee, 2004, 2006; Kiparsky, 1993; Reynolds, 1994) or the closely related (noisy) Harmonic Grammar (HG; Coetzee, 2009a; Coetzee & Kawahara, 2013; Jesney, 2007; Jesney & Tessier, 2011). In these developments, the link between within-language variation and typology is even stronger and more direct. Specifically, three core assumptions of (classic) OT, make this conceptualization of grammar supremely typologically constrained:

(i) The set of constraints (CON) is universal.

(ii) Every ranking between the constraints in CON constitute the grammar of a possible human language.

(iii) A language whose grammar cannot be expressed as a ranking between the constraints in CON is not a possible human language.

If phonological variation is the result of the co-existence of two grammars (that is, two possible rankings between the constraints in CON), it follows necessarily that variation can be viewed as

the co-existence of two typologically possible grammars in an individual speaker. However, the multiple grammars conceptualization of variation in OT implies an even stronger connection with typology. Because constraints stand in specific relations to each other, the presence of a certain phonological process in some language often implies the presence of a different phonological process in that language, thereby enabling this model of grammar to also capture Greenbergian style implicational universals. From this follows a set of strong generalizations about the possible patterns of variation that can co-exist in a single language.

Assume a Greenbergian style generalization in terms of which the presence of process p_1 in some language implies the presence of process p_2. An example could be consonant cluster simplification before a following consonant or a following vowel. In general, if a language has cluster simplification before a vowel, it will necessarily also simplify clusters before a consonant. The implication does not hold in the opposite direction, however. That is, the three kinds of languages given in (1a) are possible languages: no simplification, simplification everywhere, and simplification before consonants only. A language with simplification before vowels only, as in (1b), is not a possible language.

(1)

		'west bank'	**'west end'**
a.	**Nowhere**	west bank	west end
	Everywhere	wes_ bank	wes_ end
	Pre-C only	wes_ bank	west end
b.	**Pre-V only**	west bank	wes_ end

This generalization can be captured in OT by formulating constraints such that deletion in pre-vocalic context violates a superset of the faithfulness constraints violated by deletion in pre-consonantal context.[1] A typical representation of this kind of system is given in (2), based on an account developed in more detail by Coetzee and colleagues in various publications (Coetzee, 2004, 2012; Coetzee & Kawahara, 2013; Coetzee & Pater, 2011). This tableau does not represent the grammar of any given language, but rather shows the violation profiles of the deletion and non-deletion candidates (see the references just above for evidence that these constraints do result in the pattern of possible and impossible languages in (1)). *Ct# is a markedness constraint that penalizes a word ending on a consonant cluster with a final /t/ or /d/. MAX is a general faithfulness constraint that penalizes deletion in any context, and is hence violated by both pre-vocalic and pre-consonantal deletion. MAX-PRE-V, on the other hand, is a position specific constraint and is violated only by deletion in pre-vocalic context.,

(2)

		*Ct#	MAX	MAX-PRE-V
/west bank/	[west bank]	*		
	[wes_bank]		*	
/west end/	[west end]	*		
	[wes_ end]		*	*

[1] This follows De Lacy's proposal for using constraints that are in stringency relation to capture implicational universals (de Lacy, 2002, 2004) rather than earlier approaches under which such generalizations were captured through stipulating a fixed ranking between constraints (Prince & Smolensky, 1993). Although variation can be modeled under both approaches (see Reynolds, 1994 for modeling variation when some constraints are assumed to be in a fixed ranking relationship), De Lacy's stringency related constraints are more consistent with the core assumptions of OT by allowing all rankings between constraints as possible grammars.

Variable deletion can be expressed in this model by assuming crucial non-ranking between the markedness constraint (*Ct#) and one or more of the faithfulness constraints.[2] Due to the stringency relation between MAX and MAX-PRE-V, however, it follows that any language with variable deletion in pre-vocalic context will also have deletion in pre-consonantal context—that is, the implicational universal for differences between languages also hold for variation within a language. The ranking conditions that must hold for deletion in each of pre-vocalic and pre-consonantal contexts are stated in (3) below. As (3) shows, the conditions for variable deletion in pre-consonant context is a subset of those required for variable deletion in pre-vocalic context from which follows the implicational relation between variable deletion in these two contexts.

(3) Crucial non-ranking that must hold for variable deletion
   a. Pre-consonantal context: *Ct#, MAX
   b. Pre-vocalic context: *Ct#, MAX, MAX-PRE-V

Within this conceptualization of the grammar of variation, the relation between typology and patterns of within-language variation is, in fact, even more intricately organized. An assumption that is shared by most constraint-based models of phonological variation (whether those be OT or HG models) is that the frequency with which a specific variant is observed follows directly from the relationship between the constraints. In the classic OT competing grammars model of variation (Anttila, 1997; Kiparsky, 1993), the assumption is that each of the competing grammars (possible rankings between a set of crucially unranked constraints) have equal likelihood of being used.[3] In a language with variable deletion in both pre-vocalic and pre-consonantal contexts (i.e. a language with a grammar such as that in (3b)), there are six competing grammars. The table in (4) shows whether deletion will be observed in pre-vocalic and pre-consonantal context under each of these competing grammars. Since the set of rankings under which pre-consonantal deletion is possible is a superset of the set of rankings under which pre-vocalic deletion is observed (see (2) and (3) above), it follows that more deletion is predicted in pre-consonantal than pre-vocalic contexts.

(4) Deletion in each of the competing grammars

Competing Grammar	Pre-Consonantal	Pre-Vocalic
*Ct# ≫ MAX ≫ MAX-PRE-V	Yes	Yes
*Ct# ≫ MAX-PRE-V ≫ MAX	Yes	Yes
MAX ≫ *Ct# ≫MAX-PRE-V	No	No
MAX ≫ MAX-PRE-V ≫ *Ct#	No	No
MAX-PRE-V ≫*Ct# ≫ MAX	Yes	No
MAX-PRE-V ≫ MAX ≫ *Ct#	No	No
Total	3/6	2/6

We can therefore summarize the relation between implicational typological generalizations and within-language patterns of variation as in (5).

[2] Crucial non-ranking between two constraints is equivalent to having two competing grammars. That is, a grammar with constraints C_1 and C_2 unranked relative each other, is equivalent to a situation where the grammars $C_1 \gg C_2$ and $C_2 \gg C_1$ are in competition.

[3] Similar assumptions hold under other conceptualizations of constraint interactions, such as stochastic OT (Boersma, 1997; Boersma & Hayes, 2001) or noisy HG (Coetzee & Kawahara, 2013; Jesney, 2007), and these generalizations are therefore not specific the Anttila/Kiparsky conceptualization of variation.

(5) Assume an implicational typological generalization whereby the presence of process p_1 (pre-V deletion) in a language implies the presence of process p_2 (pre-C deletion). The possible patterns of variable realization of p_1 and p_2 are then predicted to be:

a. If process p_1 (pre-V deletion) applies variably in language *L*, then process p_2 (pre-C deletion) will also apply in language *L*.[4]

b. If process p_1 (pre-V deletion) applies variably in *L* with a frequency of *x*, then p_2 (C-deletion) will apply in *L* with a frequency greater than *x*.

2. Grammatical architecture and phonological variation

As mentioned in §1 above, early generative models of phonology relegated variation to the edges of grammar or even excluded it from grammar proper and ascribed it to "performance" or "phonetic implementation". The models of variation that were developed in the constraint-based traditions, on the other hand, all assume that grammar is centrally responsible for variation, and accounts for all aspects of variation through grammar alone (both for what the possible variants are and for the frequency with which different variants are observed). Phonological theory therefore went from one extreme (assuming that variation was not under the purview of grammar) to the other extreme (assuming that grammar alone accounts for all aspects of variation).[5]

This grammar-only approach to phonological variation runs counter to decades of evidence from the sociolinguistic and speech production literatures, both of which have documented repeatedly that variation is influenced by many factors of which grammar is but one. In fact, Bayley (2002:118) identifies "the principle of multiple causes" as one of the pillars of sociolinguistics that should direct the study of variation. Any formal model that purports to be a realistic model of the cognitive system underlying phonological variation therefore has to allow for both grammatical and non-grammatical factors to co-determine the patterns of variation.

Inspired by ideas found in earlier work such as Boersma and Hayes (2001: Appendix C) and van Oostendorp (1997), Coetzee (2009a, 2009b, 2016) and Coetzee and Kawahara (2013) develop a model of grammar that allows for both grammatical and non-grammatical factors to co-determine the nature of phonological variation. Their model is developed in a constraint-based architecture of grammar, and therefore shares the basic results about models with this architectures with regard to variation and typology, as laid out in §1 above. However, their model incorporates the contribution of non-grammatical factors by allowing such factors to move the faithfulness constraints up or down in the grammar of some language.[6] Moving faithfulness constraints down results in the application of variable processes becoming more likely, and moving them up results in such processes becoming less likely.

Because their model does not allow non-grammatical factors to add constraints to the grammar or remove constraints from the grammar, but only to move constraints up or down, all of the basic results about the relation between typology and within-language variation laid about in §1 still holds true of their model. Their model of variation, though it allows for both grammatical and non-grammatical factors to co-determine the realization of phonological variation in a specific

[4] Process p_2 could apply variable or invariably in *L*. Since per the generalization (5b), p_2 will always apply more frequently *L*, it follows that variable application of p_1 and invariant/exceptionless application of p_2 can co-exist in a single language.

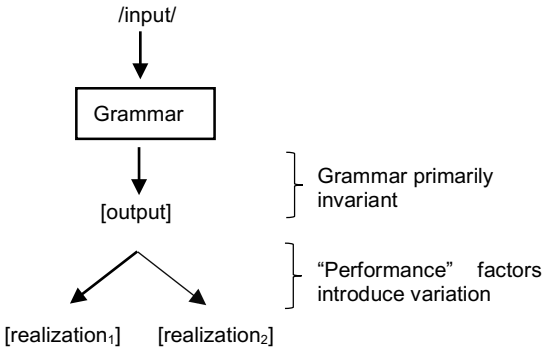
[5] See Coetzee & Kawahara (2013:§1) and especially Coetzee (2016:§1) for a more detailed review of the historical development of these different conceptualizations of the relationship between grammar and variation.

[6] Coetzee (2009b) develops the model in OT so that faithfulness constraints are moved up or down in their rankings in that model. Coetzee (2009a, 2016) and Coetzee and Kawahara (2013), however, develop the model in HG so that the weights of faithfulness constraints are moves up or down. The differences between these two conceptualizations are minimal and not relevant for the purposes of the current statement.

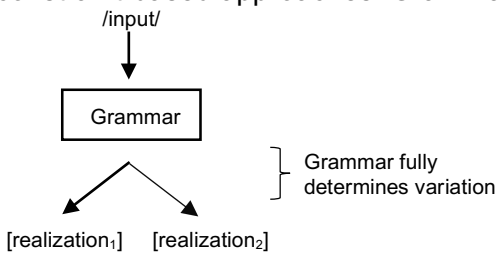
language, is still “grammar dominant”: “Grammar defines the space of possible or grammatical forms, and the non-grammatical factors can only affect how frequently the forms that grammar allows are observed.” (Coetzee, 2016:238).

That grammar dominance holds is easy to illustrate with the example developed in §1 above. The rankings that must hold for deletion to be observed in pre-vocalic position (*Ct# ≫ {MAX, MAX-PRE-V}) form a proper superset of the rankings that must hold for deletion to be observed in pre-consonantal context (*Ct# ≫MAX). This relation is determined by the definition of the faithfulness constraints, not by their ranking in the grammar. Allowing non-grammatical factors to move faithfulness constraints up or down in the grammar of a language will not change the stringency relation between the constraints, so that the results from §1 above that followed from this relationship between the constraints still hold. The figure in (6) below visualizes the three different conceptualizations for the architecture of phonological grammar. See Coetzee and Kawahara (2013) and Coetzee (2016) for a full exploration of the “grammar dominance” properties of the model in (6c).

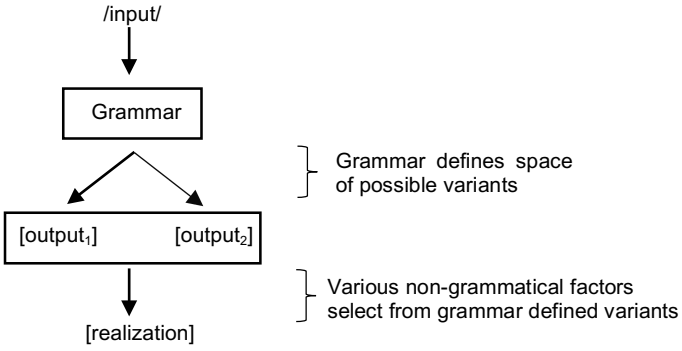
(6) a. Early generative approaches: Variation outside of grammar



b. Classic constraint-based approaches: Grammar alone responsible for variation



c. Grammar dominant models (Coetzee & Kawahara)

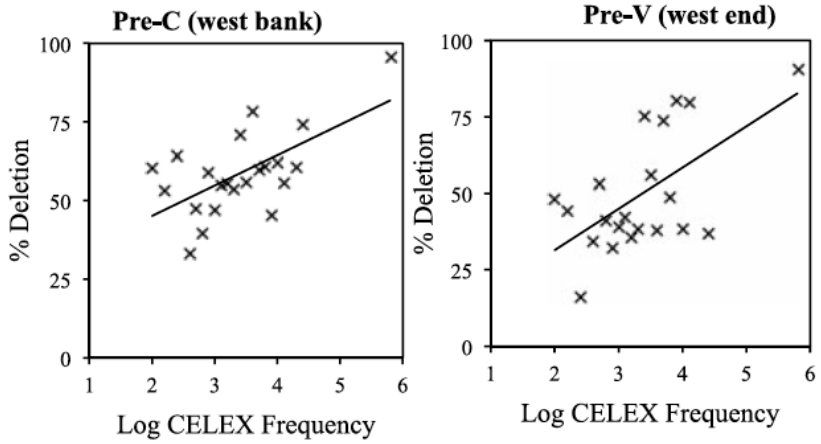


3. Near absolute grammar dominance

Coetzee and Kawahara base their assumption of grammar dominance on the observation form the Labovian variationist literature that so-called "internal" (grammatical) factors and so-called "external" (non-grammatical) factors are typically independent from each other (Bayley, 2002:120; Guy, 1991:5). More specifically, the assumption in this literature is that internal factors set the basic patterns that hold true of some speech community while external factors can impact how a variable pattern is realized within the limits determined by the internal factors. In his classic New York City *r* study, for instance, Labov (1966) found although speakers of different socio-economic classes differ in how likely they are overall to delete coda *r*, all speakers are more likely to delete pre-consonantal *r* (*fourth*) than pre-vocalic *r* (*foor*). The grammatical factor (pre-vocalic or pre-consonantal context) therefore sets the boundaries of possible variation (more *r*-deletion pre-consonantally than pre-vocalically). The non-grammatical factor (here socio-economic status) does not invert this grammar specific ordering, but can only change the overall rate of *r*-deletion.

Coetzee and Kawahara (2013) then demonstrate this dominance of grammar determined patterns of possible variation over non-grammatical factors in terms of usage frequency and *t/d*-deletion in English as spoken in Columbus, Ohio. The plots in (7), based on Figure 3 from Coetzee and Kawahara (2013), show the deletion rate in pre-consonantal and pre-vocalic contexts in the "Buckeye Corpus" of Columbus English (Pitt et al., 2007) on the *y*-axis, plotted against the log frequency of words in CELEX (Baayen, Piepenbrock, & Gulikers, 1995) on the *x*-axis. The solid line represents the result of a linear regression over these data, and the broken horizontal lines mark the overall deletion rate in each of the two contexts. In agreement with the observations about this process made in §1 above, the overall deletion rate is higher in pre-consonantal than in pre-vocalic contexts (80% vs. 76%). Additionally, the deletion rate is significantly impacted by usage frequency such that deletion is more likely for words of higher than lower usage frequency (i.e. more deletion from *best* than *jest*). Grammar therefore applies differently to words that differ in usage frequency. Crucially, however, the grammatically specified relation between pre-consonantal and pre-vocalic context holds for words of the same usage frequency (at any specific usage frequency, deletion is always more likely in pre-consonantal than pre-vocalic context). The non-grammatical factor of frequency does not invert the grammatically specified relation.

(7) *t*/*d*-deletion in Columbus, Ohio, English



Coetzee (2016) demonstrates similar dominance of grammatical factors over non-grammatical factors in terms of variable cross-word nasal place assimilation in English. He shows that although assimilation is overall more likely at faster than slower speech rates, at all speech rates assimilation is more likely in pre-velar than pre-labial contexts (i.e. always more assimilation in *green card* than in *green box*).

Given the assumption of grammar dominance that is at least tacitly assumed in the sociolinguistic variationist literature and results such as those of Coetzee and Kawahara mentioned above, the evidence for grammar dominance seems fairly strong. Even so, grammar dominance in phonological variation should be considered a hypothesis rather than a statement of fact. Since most phonological research on variation over the past two decades has been conducted in "grammar only" models, there has been very little consideration of how grammatical and non-grammatical factors interact.

With more research, and as more data become available, we may find that "grammar dominance" is not absolute but there are specific situations in which grammar may be trumped by various non-grammatical factors. We know, for instance, that linguistic variation if often coopted to perform linguistic identity, to do work in the social sphere. It may therefore be possible that a speaker (or a speech community) can, for the sake of establishing a community identity, opt to use a specific variant more often even if that were to result in a system that counters grammatically based generalizations such as those discussed above.

A potential example of an instance where usage patterns seem to result in a variation pattern that counters the linguistic generalizations comes from word-final *t*/*d*-deletion in Los Angeles English, as reported by Bybee (2002:275). For this corpus, Bybee shows that there is one class of words, words ending on /-nt/, that unexpectedly shows a higher rate of deletion before vowels (69%) than before consonants (44%). She hints at the possibility that usage patterns of these words may explain this deviation from the typical pattern. If these words are used more often before vowel-initial words, then it may be that pre-vocalic position is the position where the words are most predictable, so that pre-vocalic position will be the context most conducive to reduction for these words.

Whatever the reason for deviations from the general cross-linguistic patterns, given that it may be possible for non-grammatical factors to sometimes trump grammatical factors in within-language variation, the relationship between within-language variation and typology may sometimes be broken. It therefore behooves us both to do more research into the relationship between grammatical and non-grammatical factors in within-language variation, and also to be cautious in the conclusions that we draw about typology based on within-language variation, and vice versa.

References

Anttila, A. (1997). Deriving variation from grammar. In F. Hinskens, R. van Hout, & L. Wetzels (Eds.), *Variation, Change and Phonological Theory* (pp. 35-68). Amsterdam: John Benjamins.

Anttila, A. (2002). Variation and phonological theory. In J. K. Chambers, P. Trudgill, & N. Schilling-Estes (Eds.), *Handbook of Language Variation and Change* (pp. 206-243). Oxford: Blackwell.

Anttila, A. (2007). Variation and optionality. In P. de Lacy (Ed.), *The Cambridge Handbook of Phonology* (pp. 519-536). Cambridge: Cambridge University Press.

Baayen, R. H., Piepenbrock, R., & Gulikers, L. (1995). *The CELEX Lexical Database (CD-ROM)*. Philadelphia: Linguistic Data Consortium.

Bayley, R. (2002). The quantitative paradigm. In J. K. Chambers, P. Trudgill, & N. Schilling-Estes (Eds.), *The Handbook of Language Variation and Change* (pp. 117-141). Oxford: Blackwell.

Boersma, P. (1997). How we learn variation, optionality, and probability. *University of Amsterdam Institute of Ponetic Sciences Proceedings, 21*, 43-58.

Boersma, P., & Hayes, B. (2001). Empirical tests of the Gradual Learning Algorithm. *Linguistic Inquiry, 32*, 45-86.

Bybee, J. L. (2002). Word frequency and context of use in lexical diffusion of phonetically conditioned sound change. *Language Variation and Change, 14*, 261-290.

Chomsky, N. (1965). *Aspects of the Theory of Syntax*. Cambridge, MA: MIT Press.

Chomsky, N., & Halle, M. (1968). *The Sound Pattern of English*. New York: Harper & Row.

Coetzee, A. W. (2004). *What It Means to be a Loser: Non-Optimal Candidates in Optimality Theory.* (Ph.D. Dissertation), University of Massachusetts, Amherst.

Coetzee, A. W. (2006). Variation as accessing "non-optimal" candidates. *Phonology, 23*(3), 337-385.

Coetzee, A. W. (2009a). An integrated grammatical/non-grammatical model of phonological variation. In Y.-S. Kang, J.-Y. Yoon, H. Yo, S.-W. Tang, Y.-S. Kang, Y. Jang, C. Kim, K.-A. Kim, & H.-K. Kang (Eds.), *Current Issues in Linguistic Interfaces. Volume 2* (pp. 267-294). Seoul: Hankookmunhwasa.

Coetzee, A. W. (2009b). Phonological variation and lexical frequency. In A. Schardl, M. Walkow, & M. Abdurrahman (Eds.), *NELS 38. Volume 1* (pp. 189-202). Amherst: GLSA.

Coetzee, A. W. (2012). Variation: where laboratory and theoretical phonology meet. In A. C. Cohn, C. Fougeron, & M. Huffman (Eds.), *The Oxford Handbook of Laboratory Phonology* (pp. 62-75). Ofxord: Oxford University Press.

Coetzee, A. W. (2016). A comprehensive model of phonological variation: grammatical and non-grammatical factors in variable nasal place assimilation. *Phonology, 33*, 211-246.

Coetzee, A. W., & Kawahara, S. (2013). Frequency biases in phonological variation. *Natural Language and Linguistic Theory, 31*, 47-89.

Coetzee, A. W., & Pater, J. (2011). The place of variation in phonological theory. In J. Goldsmith, J. Riggle, & A. Yu (Eds.), *Handbook of Phonological Theory. 2nd Edition.* (pp. 401-434). Cambridge: Blackwell.

de Lacy, P. (2002). *The Formal Expression of Markedness.* (Ph.D.), University of Massachusetts, Amherst.

de Lacy, P. (2004). Markedness and conflation in Optimality Theory. *Phonology, 21*(2), 145-199.

Guy, G. R. (1991). Explanation in variable phonology: an exponential model of morphological constraints. *Language Variation and Change, 3*, 1-22.

Jesney, K. (2007). *The locus of variation in weighted constraint grammars*. Poster presented at the Workshop on Variation, Gradience and Frequency in Phonology. Stanford University, July 2007. [Downloaded on December 27, 2007 from http://people.umass.edu/kjesney/papers.html.].

Jesney, K., & Tessier, A.-M. (2011). Biases in Harmonic Grammar: the road to restrictive learning. *Natural Language and Linguistic Theory, 29*(251-290).

Kiparsky, P. (1993). *An OT perspective on phonological variation*. Handout from Rutgers Optimality Workshop 1993, also presented at NWAV 1994, Stanford University. Available at https://web.stanford.edu/~kiparsky/Papers/nwave94.pdf.

Kroch, A. (1989). Reflexes of grammar in patterns of language change. *Language Variation and Change, 1*(3), 199-244.

Kroch, A. (1994). Morphosyntactic variation. In K. Beals, J. Denton, R. Knippen, L. Melnar, H. Suzuki, & E. Zeinfeld (Eds.), *CLS 30: Papers from the 30th Regional Meeting of the Chicago Linguistic Society. Volume 2* (pp. 180-201). Chicago: Chicago Linguistic Soceity.

Labov, W. (1966). *The Social Stratification of English in New York City*. Washington, DC: Center for Applied Linguistics.

Newmeyer, F. J. (2003). Grammar is grammar and usage is usage. *Language, 79*, 682-707.

Pitt, M. A., Dilley, L. C., Johnson, K., Kieling, S., Raymond, W. D., Hume, E., & Fosler-Lussier, E. (2007). *Buckeye Corpus of Conversational Speech. (2nd Release). [*http://www.buckeyecorpus.osu.edu/*]*. Columbus: Department of Psychology, Ohio State University.

Postal, P. (1966). Review of "Elements of General Linguistics" by André Martinet. *Foundations of Language, 2*, 151-186.

Postal, P. (1968). *Apects of Phonological Theory*. New York: Harper & Row.

Prince, A., & Smolensky, P. (1993). *Optimality Theory: Constraint Interaction in Generative Grammar*. New Brunswick, Rutgers University. [Revised version appeared as Prince & Smolensky 2004.].

Reynolds, B. (1994). *Variation and Phonological Theory.* Ph.D. Dissertation, University of Pennsylvania.

van Oostendorp, M. (1997). Style levels in conflict resolution. In F. Hinskens, R. van Hout, & L. Wetzels (Eds.), *Variation, Change and Phonological Theory* (pp. 207-229). Amsterdam: John Benjamins.