Frequency effects and Optimality Theory

James Myers and Yingshing Li
National Chung Cheng University
First Theoretical Linguistics Conference, National Cheng Chi University, Taiwan
May 21, 2005

[removed Chinese font and reformatted for US letter paper March 4, 2006]

Introduction

(1) Frequency affects lenition

a. English vowel deletion (Hooper 1976)

higher frequency:	év[e]ry	[∅]
mid frequency:	mém[o]ry	[ə] ~ [∅]
lower frequency:	mámm[a]ry	[ə]

b. Southern Min syllable contraction

higher frequency:	"you see"	/li + k^{h}ũã/ → [nĩã]	(Chung 1996, Hsu 2003)
mid frequency:	"time"	/si + kan/ → [ɕi kan] ~ [ɕian]	(Li, in prep)
lower frequency:	"period"	/ki + kan/ → [ki kan]	(Li, in prep)

(2) The problem

❶ The patterns are clearly "postlexical", yet are sensitive to lexical information. In fact, they are "phoneticky" (variable across utterance tokens and articulatorily gradient), suggesting that lexical representations are themselves phonetically detailed.
❷ The phonological/lexical link is systematic, not arbitrary as such links should be.

(3) Our solution

❶ We encode gradient lexical representations using phonetically based Optimality Theory (e.g. Boersma 1998, Hayes et al. 2004).
❷ We reject the functionalism of phonetically based OT in favor of diachronic explanations for "naturalness" (e.g. Hale and Reiss 2000, Blevins 2004). Thus synchronically, frequency effects are indeed arbitrary, just as lexical effects should be.

(4) Overview

- Phonetically gradient frequency effects in lenition
- A detailed example: Southern Min syllable contraction
- Gradience and variability in Optimality Theory
- Lexical effects in Optimality Theory
- Alternative OT analyses

<u>Phonetically gradient frequency effects in lenition</u>

(5) Generative phonology should study frequency effects, given their systematicity, ubiquitous role in cognitive science, and relation to lexical factors long recognized by phonologists (syntactic category, morphology, lexical exceptions).

(6) Many kinds of frequency effects have been observed in phonological research. Here we are only concerned with the last kind in the following list.

- <u>Type frequency effects</u>. The number of words that conform to a pattern vs. the number that violate it can be used by linguists (and perhaps learners) as a diagnostic of productivity (e.g. Bybee 2001, but cf. Duanmu 2004).
- <u>Negative token frequency effects in the lexicon</u>. Token frequency counts how often words are used (e.g. *it* has higher token frequency than *id*). Phillips (1984, 1998, 1999, 2001) observes that a certain type of diachronic change targets low-frequency words, giving a negative correlation between frequency and patterning in the lexicon (e.g. emergence of the N/V stress contrast in pairs like *áddress/addréss*). Change involves phonological contrasts, isn't phonetically motivated, and is not very systematic (e.g. Yaeger-Dror and Kemp 1992 describe a diachronic contrastive vowel substitution that is influenced by lexical semantics rather than lexical frequency).
- <u>Negative token frequency effects in acquisition</u>. Children lenite less (thus better mimicking "marked" adult targets) in words they hear/use more often (e.g. Tyler and Edwards 1993, Gierut et al. 1999).
- <u>U-shaped token frequency effects on judgments</u>. Bailey and Hahn (2001) report that English nonsense syllables are rated as more English-like if they are similar to either high-frequency words or (regular) low-frequency words, with the lowest ratings for syllables similar to mid-frequency words; we haven't replicated this in Chinese, though.
- <u>Positive token frequency effects on lenition</u>. Lenition is stronger in common words.

(7) A speculation on the origins of negative vs. positive token frequency effects in adults

- Negative effects involve the frequency of *lemmas* (morphosyntactic words), which compete with each other, encode much nonphonological information, and have only semi-systematic effects in phonological priming (Levelt et al. 1999).
- Positive effects involve the frequency of *lexemes* (word forms), which facilitate rather than compete with each other and encode only phonophonetic information; Jurafsky at al. (2002) address evidence for this claim.

(8) Why focus on positive token frequency effects in lenition?

- Most systematic
- Most problematic (others easier to dismiss as "performance" or nonsystematic)
- Most phonological: patterns are "natural" and have been widely recognized by generative phonologists as "genuine phonology"

(9) More examples

- English vowel reduction (Fidelholz 1975, Hammond 2004, Jurafsky et al. 2002)
- English flapping (Patterson and Connine 2001)
- English t/d deletion (Myers and Guy 1997, Bybee 2000, 2002)
- English rhythm rule (Hammond 1999)
- Mandarin syllable contraction (Tseng 2005)
- Probably every phonetically leniting process, at least to some degree

(10) These processes are both variable and phonetically gradient.

- Obviously these processes are variable, since they are all optional.
- Evidence for phonetic gradience is more ambiguous, since most studies use categorical measures (e.g. judgments or corpus-derived counts), which is consistent with frequency merely affecting the selection rate among two categorical allomorphs (lenited vs. nonlenited) rather than truly affecting gradient phonetics.
- However, some of the above studies used continuous measures which also showed frequency effects (e.g. duration in Jurafsky et al. 2002).
- As shown below, Li (in preparation) provides relatively clear evidence for the phonetic gradience of frequency effects in Southern Min syllable contraction.

(11) What forces give rise to positive frequency effects in lenition?

- <u>Ease of articulation</u>. Gestures involving smaller movements require less energy.
- <u>Predictability</u>. Frequency increases predictability, but so can discourse context; predictive contexts also increase lenition (Charles-Luce 1997, Jurafsky et al. 2001). It's not clear if predictability is a listener-oriented or speaker-internal factor (e.g. priming).

(12) Consequences for lexical representations (Pierrehumbert 2002; see also Bybee 2002)

a. The standard phonological production model is modular and feed-forward: Categorical representations in the lexicon are passed to articulation, where gradient effects appear.
b. But frequency information is also lexical: It is predictive of lenition, but is not itself predictable, and so cannot emerge "online" like discourse context effects.
c. Therefore the effect of frequency on lenition must occur prelexically, meaning that lexical representations are themselves gradiently lenited, not categorical.
d. Rather than choosing between two allomorphs, the production system chooses among a very large number of phonetically detailed exemplars, thereby deriving variability.
e. Pierrehumbert (2002) calls this <u>word-specific phonetics</u>.

(13) Consequences for Optimality Theory

❶ (12b) implies that lexical representations are phonetically detailed, consistent with the "direct" version of phonetically based OT (e.g. Kirchner 1997, Boersma 1998, Myers and Tsay 2003, Zhang 2004)
❷ But (12c) implies that the systematicity of frequency effects is due to "diachronic" forces, consistent with critics of functional OT (e.g. Hale and Reiss 2000, Blevins 2004). Specifically, it arises via feedback from prior usages rather than being encoded in the grammar; synchronically, frequency is just as arbitrary as any other lexical factor.

Frequency effects in Southern Min syllable contraction

(14) Is Southern Min syllable contraction really phonology, rather than pure physics?

- Many phonologists think it's phonology (e.g. Cheng 1985, Chung 1996, Hsu 2003).
- It is systematically sensitive to well-recognized, abstract phonological factors like syllable structure and sonority, making orthodox analyses possible.
- It must be learned by speakers, since it shows language-specific (even idiolect-specific) variability: It's not the same as Mandarin syllable contraction (Chung 1997, Hsiao 2002); Taiwan Southern Min speakers don't agree on all Xiamen examples in Chen (2000); even Chung (1996) and Hsu (2003) disagree on some data.

(15) Nevertheless, it is articulatorily motivated

- It applies more readily in fast speech.
- Its phonetic effects are consistent with lenition (reduction of articulatory energy), since it reduces the number and duration of gestures.
- It applies more readily with higher frequency items (Cheng 1985 and Tseng 1996 note this explicitly; Chung 1996 and Hsu 2003 cite only high frequency items):

higher frequency / more contraction	k^hi kau	k^hi lai	tu tɕia
mid frequency / some contraction	pi kau	bi lai	tsu tɕiaʔ
lower frequency / less contraction	bi kau	i lai	bu tɕia

(16) Two of the open questions addressed by Li (in preparation)

❶ Is syllable contraction phonetically gradient?
❷ If the above is true, is frequency positively correlated with the phonetic degree of syllable contraction?

(17) Frequency measures came from the Taiwan Southern Min Adult Corpus (Myers and Tsay 2004)

- Transcriptions of approximately 40 hours of "chat radio" programs in southern Taiwan
- Approximately 400,000 word tokens
- Word frequencies can be automatically calculated

(18) The key experiment

- Items produced in isolation to eliminate discourse context effects on predictability
- 120 target items varying in frequency:

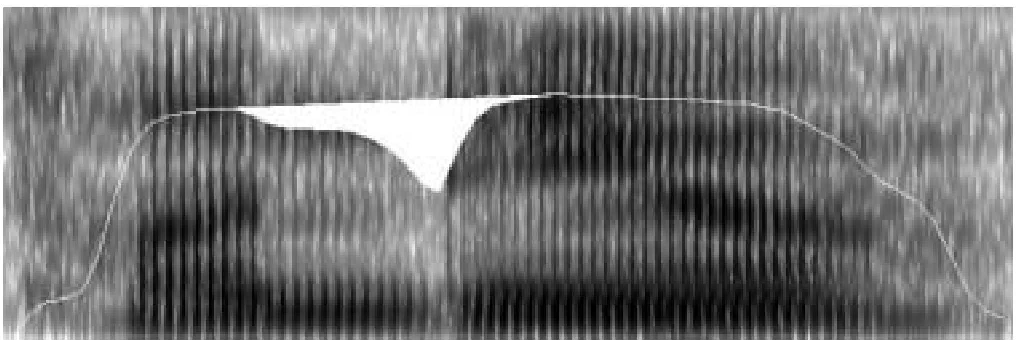
"city area"	1 token in corpus
"friend"	571 tokens in corpus

- Spoken uncontracted forms were played to 20 native speakers of Taiwan Southern Min, who were asked to repeat them back quickly but naturally; they were not explicitly told to contract them.

(19) Analysis

- Following Jurafsky et al. (2002), multiple regression was used to study the effect of frequency on contraction even when other influences were factored out, i.e.:
    - Duration, as a measure of speaking rate
    - Time to initiate speech, as a measure of overall ease of lexical access
    - Manner of first coda and second onset, which inherently vary in duration
- Key dependent measure was size of gap between the two syllables' amplitude peaks
- Question ❶ was addressed by testing whether the distribution of gap sizes was consistent with a single continuous, randomly varying process.
- Question ❷ was addressed by the multiple regression itself (conducted across speakers using the method of Lorch and Myers 1990).

(20) Sample gap measurement (partially contracted "restaurant")



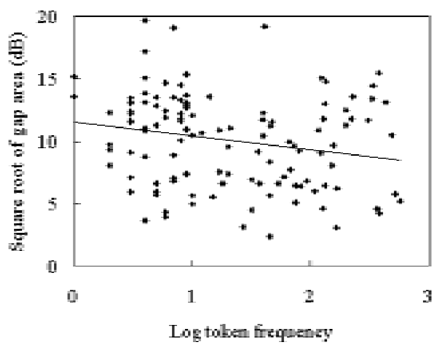
(21) Answer to question ❶

- The measure (gap area) was two-dimensional but the process is presumably linear (contraction of gestures in time).
- Consistent with this, the distribution of raw data was rather skewed (Shapiro-Wilk $W = .93$, $p = .00001$), while the distribution of square roots of the data was quite close to Gaussian normal ($W = .98$, $p = .03$; fully normal would be $W = 1$, $p > .05$).
- This implies that syllable contraction results from a phonetically gradient system, rather than a mere binary choice among allomorphs (full vs. contracted).

(22) Answer to question ❷

- The multiple regression found the following statistically significant effects:
    - Shorter durations (faster speech) meant smaller gaps (more contraction).
    - Faster reactions (less difficult) meant smaller gaps (more contraction).
    - First syllable obstruent codas meant larger gaps (less contraction).
    - Second syllable obstruent onsets meant larger gaps (less contraction).
- Most importantly, even with the above influences factored out, (log) frequency did indeed correlate with the (square root of) gap size: higher frequency items were produced with smaller gaps, an effect found in 19 out of 20 speakers ($p < 0.00001$).

- The role of frequency was nevertheless relatively minor, accounting for only about 4% of the total variability in gap size (r = -.22):



Gradience and variability in Optimality Theory

(23) Phonetically based OT already has the tools for handling variability and gradience in patterns like syllable contraction: phonetically detailed faith and structure constraint families (e.g. Kirchner 1997, Boersma 1998, Myers and Tsay 2003, Zhang 2004).

(24) In the production grammar of Boersma (1998), structure constraints are violated by expenditures of energy (e.g. *GESTURE), while faith constraints are violated by mismatches between produced forms and perceptual targets (e.g. *REPLACE).

- Since the key parameter in syllable contraction seems to be the duration between syllable peaks, we could encode this with the following constraint families:

  *GESTURE(dur_i) >> *GESTURE(dur_j), where $dur_i > dur_j$
  *REPLACE(dur_i, dur_j) >> *REPLACE(dur_i, dur_k), where $dur_j > dur_k$

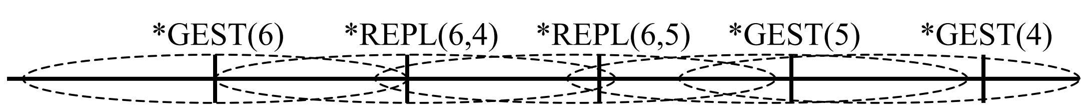
- For example, languages with contraction fixed at 5 units would have a ranking like so:

/dur = 6/	*GEST(6)	*REPL(6,4)	*REPL(6,5)	*GEST(5)	*GEST(4)
[dur = 6]	*				
☞ [dur = 5]			*	*	
[dur = 4]		*			*

(25) Variability is handled by variable constraint ranking.

- In Boersma and Hayes (2001), competence ranks constraints along a continuous ranking scale, while in performance rankings vary randomly around these fixed points.

- For example, languages in which contraction varies around 5 units would have a ranking like so, assuming a fixed performance variability range for each constraint:



• Some outputs for /dur = 6/, and the performance rankings that generate them:

[dur = 5]: *GEST(6) >> *REPL(6,4) >> *REPL(6,5) >> *GEST(5) >> *GEST(4)
[dur = 6]: *REPL(6,4) >> *REPL(6,5) >> *GEST(6) >> *GEST(5) >> *GEST(4)
[dur = 4]: *GEST(6) >> *GEST(5) >> *REPL(6,4) >> *REPL(6,5) >> *GEST(4)

(26) Advantages of this approach to gradient phonology, for our purposes

- Extends standard phonological formalism to language-specific phonetics
- Still provides room for innate or "symbolic" aspects to phonological knowledge
- Constraints seem to be able to handle gradience more naturally than rules could

(27) Claimed advantage that we will reject

- Constraint families are "projected from the phonetics" and "reflect the speaker's knowledge that a structure that is phonetically more demanding should be banned before a structure that is less so" (Zhang 2004:176; see also Myers and Tsay 2003).

<u>Lexical effects in Optimality Theory</u>

(28) The two challenges that frequency effects pose for OT

❶ Can OT *describe* word-specific phonetics at all?
❷ Can OT *explain* the positive correlation between frequency and lenition?

(29) Regarding ❶, the OT literature offers three ways to restrict phonological generalizations to a specific lexical class LC:

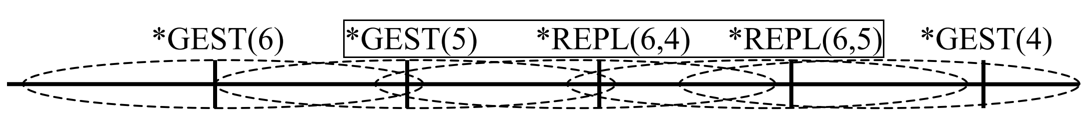
- Lexical interface constraints
  - Faith interface constraints (e.g. Alderete 2001): FAITH$_{LC}$ >> STRUC >> FAITH
  - Structure interface constraints (e.g. Pater 2000): STRUC$_{LC}$ >> FAITH >> STRUC
- Lexical cophonologies (e.g. Anttila 2002): {FAITH >> STRUC}$_{LC}$, {STRUC >> FAITH}$_{\neg LC}$
- Lexically prespecified representations (e.g. Inkelas, Orgun, and Zoll 1997): used only for isolated lexical exceptions

(30) Lexical cophonologies are overall the most elegant (Anttila 2002, Inkelas and Zoll 2003)

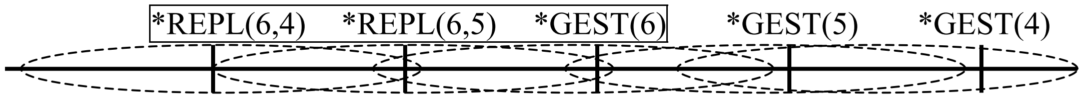
- This approach builds on the fundamental mechanics of OT
- It is capable of handling lexical classes with "leaking" borders (e.g. where most, but not all, items in a class show a given pattern)
- It is capable of handling both well-defined classes and isolated lexical exceptions

(31) Descriptive cophonologies for a syllable-contraction-type pattern

- A cophonology for a higher frequency item will tend to output [dur = 4]



- A cophonology for a lower frequency item will tend to output [dur = 6]



(32) While Inkelas, Orgun, and Zoll (1997) approve of cophonologies for handling the effect of multiword lexical classes (see also Inkelas 1998), they argue against a cophonology approach to lexical exceptions (and by extension, word-specific phonetics):

- The cophonology approach can't describe cases where a word is an exception to a pattern only in one part: e.g. in Turkish [is.tib.dat], [b] violates coda devoicing while [t] conforms to it.

- Lexicon Optimization (Prince and Smolensky 1993) makes prespecification principled, forcing inputs to match outputs unless there is reason to do otherwise.

- The cophonology approach allows for the possibility that every word has its own randomly chosen phonology: systematicity is left unexplained by the grammar.

(33) But prespecifying just part of a word is irrelevant for lenition, which seems to affect all parts of an item together.

- For example, it seems to be true that if an English word allows two opportunities for /t/ to flap (e.g. *itty-bitty*), either both will flap or neither will, depending on speaking rate and other general factors, never just one.

(34) Moreover, cophonologies can also be principled.

- Anttila (2002): lexical exceptions happen where phonological constraints are weakest.
    - As we saw above, frequency effects in syllable contraction are relatively weak; if phonological factors are not controlled (e.g. the manner of the first syllable coda or second syllable onset), they are harder to detect.

- Moreover, Golston (1996) shows that lexical representations can be described in terms of constraint violations, a notion that can be extended to representing all lexical items as cophonologies (see also Hammond 2000 for an interface constraint variant).
    - For example, the lexical representation [dur = 4] would be coded as the first ranking in (31), while [dur = 6] would be coded as the second ranking. This idea still needs some fleshing out, though, since these rankings still assume inputs.

(35) Finally, we admit that word-specific cophonologies allow for a 10,000-word language to have 10,000 different phonological systems, but it's not just our problem.

- Inkelas et al. (1997) themselves admit that their approach predicts that nonalternating patterns cannot be shown to be part of grammar, since Lexicon Optimization will always choose the surface form as the input.

- For example, in the following Lexicon Optimization tableau for a contracted surface form, the structure constraint *GESTURE has no effect:

Input	Output	*GEST(6)	*REPL(6,4)	*REPL(5,4)	*REPL(4,4)
☞ /dur = 4/	[dur = 4]				*
/dur = 5/	[dur = 4]			*	
/dur = 6/	[dur = 4]		*		

- The more general problem is the premise of question ❷: Phonological theory, as a model of the mental phonological system, is not obliged to explain patterns that are already sufficiently explained by extra-grammatical models.

(36) Substance-free (or "pure") phonology (Hale and Reiss 2000, Reiss 2003, Blevins 2004)

- Key aspects of grammar posited by theoretical phonology actually arise diachronically.
  - The OCP, geminate integrity, autosegmental phonology generally
  - Structure preservation, elsewhere condition, other pattern interaction generalizations
- (Although Hale and Reiss (2000) and Reiss (2003) also argue that phonological representations are categorical and phonological operations involve derivational rules, not constraints, these formal claims conflict with word-specific phonetics.)

(37) The logical argument for substance-free phonology

- <u>Occam's razor</u>: If articulatory and perceptual phonetics, operating diachronically, are sufficient to account for patterns, why "project" them into the grammar?
- <u>The logical irrelevance of markedness to the learner</u> (Hale and Reiss 2000): The child still has to rely on actual data to determine if a given constraint applies in her language, so why does it matter if constraints describe "naturalness" or not?
  - Note that this is even true for the learning models in phonetically based phonology. Boersma (1998) assumes that rankings like *GESTURE(dur_i) >> *GESTURE(dur_j) are fixed by phonetics, but the learning model of Boersma and Hayes (2001) makes this irrelevant. Thus if the ranking *GEST(5) >> *REPL(6,4) >> *REPL(6,5) helps predict the surface forms heard by the child, she will learn it; otherwise, she won't.

(38) The psycholinguistic argument for substance-free phonology

- Recent studies suggest that markedness is indeed irrelevant to actual children, who rely on matching surface patterns (e.g. Zamuner, Gerken and Hammond 2005).
- Markedness may also be irrelevant to the adult speaker, which is something we're currently beginning to investigate.

(39) Thus phonological knowledge consists primarily of representations, not patterns.

- These representations (at least some of them) seem to be phonetically detailed.
- Since the human brain is not an all-purpose tape-recorder, it recodes speech in terms of innate units (or at least units that humans are innately equipped to generate).
- Phonetically based OT can provide a framework for identifying the proper units in the form of specific constraint families (e.g. *GESTURE and *REPLACE).

Alternative analyses

(40) Above we explained why we prefer cophonologies over prespecification in describing frequency effects. Here's why we prefer cophonologies over interface constraints.

- Hammond (1999) proposed frequency-ranked faith interface constraints to describe why lenition is blocked in lower-frequency forms in adult speech:

  Adult ranking: $\text{FAITH}_{\text{LowFreq}}$ >> LENITION >> $\text{FAITH}_{\text{HighFreq}}$

- However, Gierut, Morrisette, and Champion (1999) independently proposed the same sort of constraints to describe why lenition is blocked in *higher*-frequency forms in child language, thus requiring the reverse ranking:

  Child ranking: $\text{FAITH}_{\text{HighFreq}}$ >> LENITION >> $\text{FAITH}_{\text{LowFreq}}$

  - Note that the reversed patterns undermine any attempt to explain the adult pattern by claiming that it emerges during language acquisition!

- Hammond (2003) himself points out another problem: the above ranking falsely predicts that zero-frequency items (e.g. neologisms, nonce forms used in an experiment, or novel phrases) should always be lenited, since they won't have any faith constraint.

(41) More arguments against an interface constraint approach to frequency effects:

- Hammond (2003) proposes structure interface constraints, thus describing lenition in adult speech as follows (note how this solves the zero-frequency problem):

  Adult ranking: $\text{LENITION}_{\text{HighFreq}}$ >> FAITH >> $\text{LENITION}_{\text{LowFreq}}$

- But applying this to child language, the zero-frequency problem reemerges, since we now falsely predict that *children* will never lenite novel forms:

  Child ranking: $\text{LENITION}_{\text{LowFreq}}$ >> FAITH >> $\text{LENITION}_{\text{HighFreq}}$

(42) Phonetically detailed representations cause more serious problems.

- Given word-specific phonetics, each *word* would have to be linked to a different phonetically detailed structure interface constraint.

  - For example, to describe different degrees of syllable contraction in different words, we would need constraint families like *GESTURE(dur_i, $freq_x$).

- But this would make ranking entirely superfluous, since each *GESTURE constraint would be linked to a different word and thus would never interact.

- We could capture the positive correlation with frequency by stipulating a principle like the following, but this would make even the *GESTURE constraint superfluous!

  $freq_x > freq_y \Rightarrow dur_i < dur_j$

(43) In any case, the formalisms of Hammond (1999, 2003) and Gierut et al. (1999) are unable to *explain* the correlations they attempt to describe (and above we saw why even a single acquisition-based explanation can't work for both). Thus even if their formalisms worked, the cause of the correlations would have to be left outside the grammar, as we argue.

(44) Two nice consequences of admitting defeat

- Frequency effects turn out to be just like all other lexical effects in phonology: they are indeed arbitrary (from the viewpoint of mental grammar).
- Our nonfunctional formalism can also accommodate all the other types of frequency effects, since the explanations for their differences are also left outside of grammar.

Conclusions

(45) Major lessons

- The vast and growing literature on phonetically based phonology (including non-generative approaches) demonstrates that mental phonological grammar must refer to phonetic detail.
- Substance-free phonology is ultimately the most valid approach to mental grammar, though it is currently less popular since it implies that most of the phonological literature isn't really about phonology (mental grammar) at all.
- These two currently warring camps should unite, and this paper provides some tips on why and how this may be done.

References

Alderete, John. 2001. Dominance effects as transderivational anti-faithfulness. *Phonology* 18:201-253.

Anttila, Arto. 2002. Morphologically conditioned phonological alternations. *Natural Language and Linguistic Theory* 20:1-42.

Bailey, T. M. and U. Hahn (2001) Determinants of wordlikeness: Phonotactic or lexical neighbourhoods? *Journal of Memory and Language* 44:568-591.

Blevins, Juliette. 2004. *Evolutionary phonology: The emergence of sound patterns*. Cambridge, UK: Cambridge University Press.

Boersma, Paul. 1998. *Functional phonology.* The Hague: Holland Academic Graphics.

Boersma, Paul, and Bruce Hayes. 2001. Empirical tests of the Gradual Learning Algorithm. *Linguistic Inquiry* 32:45-86.

Bybee, Joan L. 2000. The phonology of the lexicon: evidence from lexical diffusion. In *Usage-based models of language*, ed. Michael Barlow and Suzanne Kemmer, 65-85. Stanford: CSLI Publications.

Bybee, Joan L. 2001. *Phonology and language use*. Cambridge University Press.

Bybee, Joan. 2002. Word frequency and context of use in the lexical diffusion of phonetically conditioned sound change. *Language Variation and Change* 14:261-290.

Charles-Luce, Jan. 1997. Cognitive factors involved in preserving a phonetic contrast. *Language and Speech* 40 (3): 229-248.

Chen, Matthew Y. (2000) *Tone sandhi: Patterns across Chinese dialects*. Cambridge: Cambridge University Press.

Cheng, Robert L. 1985. Sub-syllabic morphemes in Taiwanese. *Journal of Chinese Linguistics* 13:12-43.

Chung, Raung-Fu. (1996). Syllable contraction. In R.-F. Chung (ed.), *The segmental phonology of Southern Min in Taiwan* (ch. 5). Taipei: The Crane publishing Co.

Chung, Raung-Fu. (1997). Syllable contraction in Chinese. *Chinese Languages and Linguistics III: Morphology and Lexicon*, 199-235. Taipei: Academia Sinica.

Duanmu, San. 2004. A corpus study of Chinese regulated verse: phrasal stress and the analysis of variability. *Phonology* 21:43-89.

Fidelholz, James L. 1975. Word frequency and vowel reduction in English. In *Proceedings of CLS 11*, 200-213. University of Chicago.

Gierut, Judith A., Michele L. Morrisette, and Annette Hust Champion. 1999. Lexical constraints in phonological acquisition. *Journal of Child Language* 26:261-294.

Golston, Chris. (1996). Direct Optimality Theory: Representation as pure markedness. *Language* 72 (4):713-748.

Hale, Mark, and Charles Reiss. 2000. Phonology as cognition. In N. Burton-Roberts, Phillip Carr and Gerry Docherty (eds.), *Phonological knowledge*, pp. 161-184. Oxford: Oxford University Press.

Hammond, Michael. 1999. Lexical frequency and rhythm. In M. Darnell, E. Moravcsik, F. Newmeyer, M. Noonan, and K. Wheatley (eds.), *Functionalism and formalism in linguistics, vol. 1: General papers*, pp. 329-358. Amsterdam: John Benjamins.

Hammond, Michael. 2000. There is no lexicon! Coyote Papers 10:55-77. Earlier version available as ROA-43.

Hammond, Michael. 2004. Frequency, cyclicity, and optimality. *Studies in Phonetics, Phonology, and Morphology* 10:349-364.

Hayes, Bruce, Robert Kirchner, and Donca Steriade (eds.) 2004. *Phonetically based phonology*. Cambridge, UK: Cambridge University Press.

Hooper, Joan B. 1976. Word frequency in lexical diffusion and the source of morphophonological change. In W. M. Christie, Jr. (ed.), *Current progress in historical linguistics: Proceedings of the Second International Conference on Historical Linguistics*, pp. 95-105. Amsterdam: North-Holland Publishing Company.

Hsiao, Yuchau E. (2002). Tone contraction. *Chinese Languages and Linguistics VIII,* 1-16. Taipei: Academia Sinica.

Hsu, Hui-Chuan. 2003. A sonority model of syllable contraction in Taiwanese Southern Min. *Journal of East Asian Linguistics* 12:349-377.

Inkelas, Sharon. 1998. The theoretical status of morphologically conditioned phonology: A case study from dominance. In Geert Booij and Jaap van Marle (eds.), *Yearbook of Morphology 1997*, 121-155. Dordrecht: Foris.

Inkelas, Sharon, C. Orhan Orgun, and Cheryl Zoll. 1997. The implications of lexical exceptions for the nature of grammar. In Iggy Roca (ed.) *Derivations and constraints in phonology* (pp. 393-418). Oxford: Clarendon Press.

Inkelas, Sharon, and Cheryl Zoll. 2003. Is grammar dependence real? University of California at Berkeley and MIT ms. ROA-587.

Jurafsky, Daniel, Alan Bell, and Cynthia Girand. 2002. The role of lemma in form variation. In *Laboratory phonology 7*, ed. Carlos Gussenhoven and Natasha Warner, 1-34. Berlin: Mouton de Gruyter.

Jurafsky, Daniel, Alan Bell, Michelle Gregory, and William D. Raymond. 2001. Probabilistic relations between words: Evidence from reduction in lexical production. In Joan Bybee and Paul Hopper (eds.) *Frequency and the emergence of linguistic structure*, pp. 229-253. Amsterdam: Benjamins.

Kirchner, Robert. 1997. Contrastiveness and faithfulness. *Phonology* 14: 83-111.

Levelt, W. J. M., Roelofs, A. P. A., & Meyer, A. S. (1999). A theory of lexical access in speech production. *Behavioral and Brain Sciences* 22 (1):1-37.

Li, Yingshing. In preparation. *Lexical frequency effects in Taiwan Southern Min syllable contraction*. National Chung Cheng University MA thesis.

Lorch, Jr. Robert F. & Jerome L. Myers. (1990). Regression analyses of repeated measures data in cognitive research. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 16 (1):149-157.

Myers, James, and Gregory R. Guy. 1997. Frequency effects in Variable Lexical Phonology. *University of Pennsylvania Working Papers in Linguistics* 4.1:215- 228.

Myers, James, and Jane Tsay. 2003. A formal functional model of tone. *Language and Linguistics* 4 (1):105-138.

Myers, James, and Tsay, Jane. 2004. *Phonological competence by analogy: Computer modeling of experimentally elicited judgements of Chinese syllables (II).* (Project report). National Cheng Chung University. Research Project funded by National Science Council, Taiwan. (NSC 91-2411-H-194-022XXX).

Pater, Joe. 2000. Non-uniformity in English secondary stress: The role of ranked and lexically specific constraints. *Phonology* 17:237-274.

Patterson, D., & Connine, C. M. (2001). A corpus analysis of variant frequency in American English flap production. *Phonetica* 58 (4):254-275.

Phillips, Betty S. 1984. Word frequency and the actuation of sound change. *Language* 45:9-25.

Phillips, Betty S. 1998. Word frequency and lexical diffusion in English stress shifts. In *Historical Linguistics 1995: Volume 2: Germanic Linguistics*, ed. Richard Hogg and Linda van Bergen, 223-232. Amsterdam: John Benjamins.

Phillips, Betty S. 1999. The mental lexicon: evidence from lexical diffusion. *Brain and Language* 68: 104-109.

Phillips, Betty S. 2001. Lexical diffusion, lexical frequency, and lexical analysis. In *Frequency and the emergence of linguistic structure*, ed. Joan L. Bybee & Paul Hopper, 123-136. Amsterdam: John Benjamins.

Pierrehumbert, Janet B. (2002). Word-specific phonetics. In Carlos Gussenhoven & Natasha Warner (Eds.), *Laboratory phonology, 7* (pp. 101-139). Berlin: Mouton de Gruyter.

Prince, Alan, and Paul Smolensky. 1993. *Optimality Theory: Constraint interaction in generative grammar*. Rutgers University and University of Colorado ms. ROA-537, http://roa.rutgers.edu/. Published 2004 by Blackwell.

Reiss, Charles. 2003. Quantification in structural descriptions: Attested and unattested patterns. *The Linguistic Review* 20: 305-338.

Tseng, Chin-Chin. (1996). Contraction in Taiwanese: Synchronic analysis and its connection with diachronic change. *Proceedings of the fifth international symposium on Chinese languages and linguistics, V: Interactions in languages* (pp. 205-232). Taipei: Academia Sinica.

Tseng, Shu-Chuan. (2005). Contracted syllables in Mandarin: Evidence from spontaneous conversations. *Language and linguistics*, 6 (1), 153-180.

Tyler, Ann A., and Mary Louise Edwards. 1993. Lexical acquisition and acquisition of initial voiceless stops. *Journal of Child Language* 20: 253-273.

Yaeger-Dror, Malcah, and William Kemp. 1992. Lexical classes in Montreal French: the case of (ɛ:). *Language and Speech* 35:251-293.

Zamuner, Tania S., Louann Gerken and Michael Hammond (2005). The acquisition of phonology based on input: A closer look at the relation of cross-linguistic and child language data. *Lingua* 115 (10):1403-1426.

Zhang, Jie. 2004. The role of contrast-specific and language-specific phonetics in contour tone distribution. In Bruce Hayes, Robert Kirchner, and Donca Steriade (eds.) *Phonetically-based phonology*, pp. 157-190. Cambridge, UK: Cambridge University Press.