Conference on Precedence Relations
CUNY Phonology Forum
Jan. 25, 26, 2007

The Interaction of Internal and External Factors in Metathesis and Deletion

Beth Hume
Ohio State University
ehume@ling.osu.edu

(1) Assumption: An individual's phonological system is dynamic.
One source of evidence comes from the observation that the distribution of sounds may change due to, for example, social contact with speakers of other varieties of the language.
e.g. When I moved to the United States from Canada in 1986, my vowels [aw, aj] and [ʌw, ʌj] were in complementary distribution:

- [aw, aj] occurred before voiced sounds (except flap) and word-finally, e.g. 'ride' [rajd] 'bow' [baw]
- [ʌw, ʌj] occurred before voiceless consonants and flap, e.g. 'lout' [lʌwt], 'write' [rʌjt], writer [rʌjɾr̩].

Today, twenty years later, my vowels [aw, aj] are in free variation with [ʌw, ʌj]; that is, the context in which they can occur has extended to before voiceless consonants.

(2) The goal of my research program is to understand:
a. what the factors are that shape a person's cognitive phonological system;
b. the extent to which these factors interact;
c. how we can model them mathematically in order to predict sound patterns;
d. how we can represent the interactions in a cognitive model of language.

In this talk, I will focus primarily on (2a)-(c).

With regards to (2a): What are the relevant factors involved in shaping a person's cognitive phonological system?

(3) To answer this question, we need only turn to the vast literature from many areas of study in linguistics: historical, sociolinguistic, psycholinguistic, phonological, phonetic. Based on this literature, we know that there are many factors that influence the shape of an individual's phonological system. These include (at least):
a. the person's existing phonological system, including information about prosodic structure, sound and feature categories, relations among sounds (allophonic, contrastive), etc.
b. perceptual factors: quality of the acoustic/auditory cues to the identification of a sound or sequence of sounds influenced by acoustic/auditory similarity between the sounds involved, etc.
c. production factors: the amount of precision required to produce a given sound or sequence of sounds, influenced by, e.g. the complexity of the articulation, similarity among sounds in a sequence
d. contextual probability: the probability that a linguistic element (feature, sound, contrast, morpheme, word, syllable, etc.) will occur in a particular context, as a function of frequency (type, token), and predictability (transitional probabilities)
e. lexical factors: word probability, neighbourhood effects;
f. cognitive factors: generalization, expectation;
g. social factors: the amount of social value accorded a particular sound or sequence.

Acknowledgements: For their comments and suggestions, I am very grateful to members of the Phonies and Changelings discussion groups in the Ohio State University Dept. of Linguistics, and to audience members at the CUNY Phonology Forum on Precedence in Phonology, in particular, Chuck Cairns, Greg Guy, Paul de Lacy, Bill Idsardi, Andrew Nevins and Andrew Yu.

Regarding (2b): To what extent do these factors interact?

(4) Less research has been done on the interaction of the factors in (3). Some thoughts on why:
- practical reasons: understanding human language is a huge endeavor, hence the need to break it down into tractable areas of research
- depth before breadth: we need to understand which factors are relevant before we can start talking about how they interact

Case studies: Research into metathesis and consonant deletion provide good examples illustrating the interaction of various factors in each process.

Case study 1: Metathesis

(5) a. 'Metathesis' is used here as a descriptive term to refer to a language sound pattern in which a sequence of sounds appears in one order in one context but in the opposite order in a related context. The contexts being compared may come from within a single synchronic system, from two related dialects, or from different synchronic states of a language.
b. I am concerned in this discussion only with cases of metathesis in which the differing orders arose by **precedence reversals**, i.e. xy > yx, where {x, y} are speech sounds. (Blevins & Garrett 1998 suggest that differences in the ordering of sounds can also arise though subsequent stages of deletion and insertion.)

(6) Factors influencing metathesis include (at least): perception; production; structural conditions (e.g. syllable structure, segment quality); contextual probability; word frequency; generalization; expectation (Blevins & Garrett 1998, Grammont 1933, Hock 1985, Hume 2004, Wanner 1989, inter alia)

(7) Hume (2004):

- Metathesis necessarily involves two sounds (e.g. *x, y*) which occur as *xy* in one context, and *yx* in another. From the perspective of language change, we may consider one order as original (the input), and the other order as the result of metathesis (the output).
- *Predicting the input to metathesis*: For a sequence of sounds to be subject to a reversal in precedence (i.e. metathesis input), there must be indeterminacy in the acoustic/auditory signal regarding the onset and offset of the speech sounds involved, i.e. with respect the order of the sounds. Indeterminacy is defined as a function of:
  - the listener's experience with the elements involved (e.g. sounds, sound sequences, morphemes, words, etc.), and
  - the quality of information in the signal as determined by the types of sounds involved, the context in which they appear, the phonetic cues present, etc. For example, similarity between sounds favors metathesis (same place or manner).
- *Predicting the output of metathesis*: The sequence of sounds or structure (e.g. syllable type) that results from metathesis is more frequent in the given language than the input order of sounds or structure. Thus, for two sounds, e.g. *xy*, to undergo metathesis, a) the sound sequence *yx* must be attested in the language, *and/or* the prosodic structure, e.g. open syllable, which includes *yx* must be attested (in this latter case then, a novel order of sounds may occur as the result of metathesis provided that the prosodic sequence already occurs in the language); and b) *yx* or the relevant prosodic structure must be more frequent than that of the input.
- This means that a person's knowledge of the language's sound patterns and how frequently they are used influences how the signal is processed and the order in which the sounds are parsed.

- Factors involved in metathesis thus include (at least):
  - perception
  - person's knowledge of sound structure
  - contextual probabilities of sound sequences or categories
  - (production, social and lexical factors can also be shown to play a role)

Case study 2: Deletion

(8) Word medial t/d deletion (Raymond et al. 2006)

- Study focuses on alveolar stops /t/ and /d/ when they occur word-internally, e.g. as in *s<u>t</u>op*, *be<u>tt</u>er*, *a<u>d</u>vice*, or *i<u>t</u>'s*; that is, in any position other than the beginning or end of an orthographic word.
- Data comes from the *Buckeye* corpus of spontaneous speech (Pitt, Johnson, Hume, Kiesling, & Raymond, 2004). The Buckeye corpus contains over 300,000 words of speech from forty individual speaker interviews. For access to the Buckeye Corpus, go to http//vic.psy.ohio-state.edu/.
- The subset of the corpus used for this study consisted of about 100,000 words from 14 of the corpus speakers. Interviews from 7 male (2 older and 5 younger) and 7 female (2 older and 5 younger) speakers were included in the subset.

(9) Some results (see Raymond et al. 2006 for discussion of all results)

a. <u>Syllable position matters</u> (not surprising): t/d in coda was more likely to delete that t/d in onset.
Possible consonant-vowel combinations in onsets and codas (for reference)

Syllable position	Syllable structure	/t/ Examples	/d/ Examples
Onset	V.__V	be*tt*er, poli*t*ics	di*d*n't, somebo*d*y, la*d*y
	V.__C =/wry/	be*t*ween, ma*tt*ress	a*d*rift, Broa*d*way
	./C=s/__V	s*t*ill, ins*t*ead	–
	.C=/s/__C=/r/	s*t*rike, ins*t*ruct	–
	C.__V	in*t*o, ac*t*ive	har*d*er, hol*d*ing
	C.__C=/wry/	elec*t*ric, en*t*wine	hun*d*red, pen*d*ulum
Coda	V__C=/sz/.	tha*t*'s, tro*t*skyist	ki*d*s, a*d*s
	V__.C≠/wr/	discree*t*ly, swee*t*ness	a*d*mit, ma*d*ness
	C__C=/sz/	prin*t*s, ac*t*s	han*d*s, hol*d*s
	C__.C≠/wr/	par*t*ner, exac*t*ly	gran*d*parents, fon*d*ling

b. <u>Segmental context matters</u> (again, not surprising):

- The likelihood of **onset** /t,d/ deletion by preceding consonant class

  Most likely homorganic sonorant (flapping context)
  ↓
  Least likely non-homorganic non-sonorant

- The likelihood of **coda** /t,d/ deletion by preceding consonant class (consonant cluster simplification)

  Most likely non-homorganic non-approximant
  ↓
  homorganic non-approximant
  ↓
  Least likely approximant

c. Contextual probabilities: Only significant for t/d in onsets

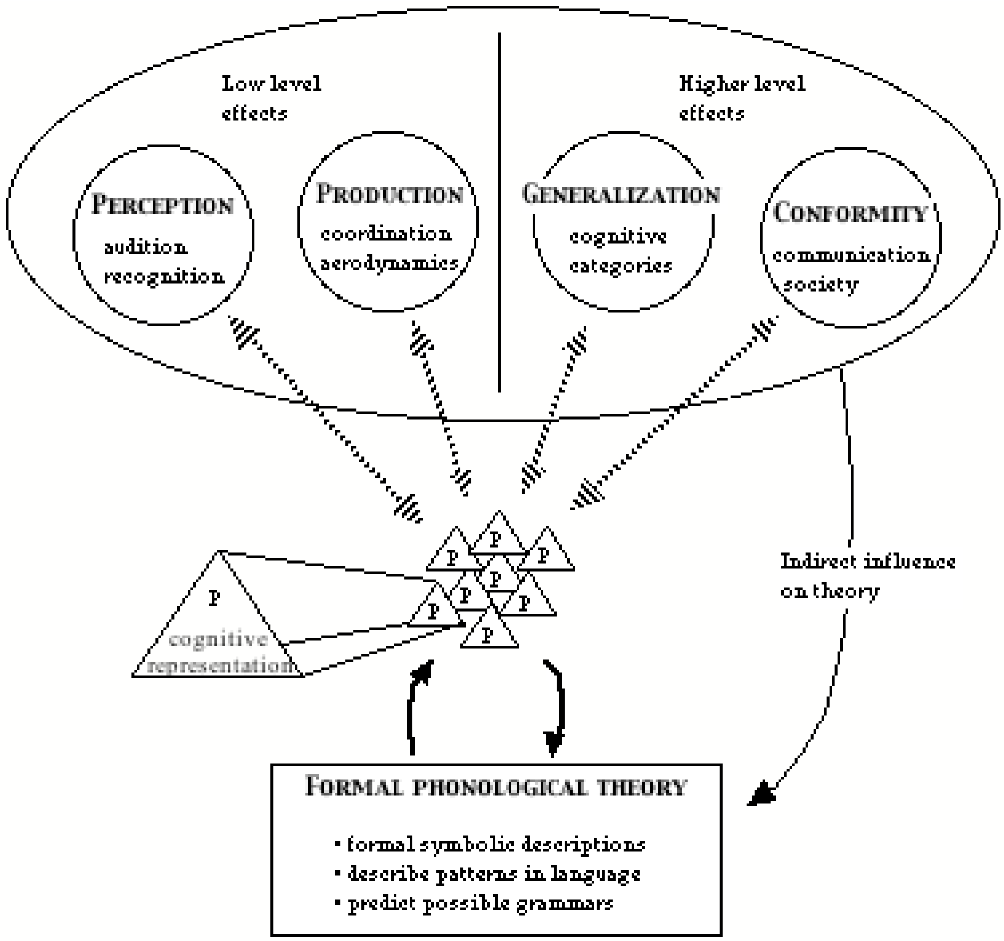
- *Word predictability*
  Onset t/d in flapping context, e.g. be*tt*er: t/d in words that are more predictable from the word that follows them are more likely to be deleted than t/d in less predictable words. The effect is perhaps because words that are more expected in context will be produced more fluently, resulting in more reduction in words containing stops in onsets through lenition (assumes a model of articulatory planning – see Raymond et al. 2006 for discussion).
  e.g. "it's *better than* nothing" vs. "it's *better for* everyone"
  If "better" is more predictable before "than", than before "for", the /t/ in "better than" will be more likely to delete than the /t/ in "better for".

- *Bigram frequency*
  Onset t/d: higher frequency vowel–stop bigrams promote deletion of /t,d/ tokens in onsets, but not in codas.

d. Production was also a predictor of deletion in that:

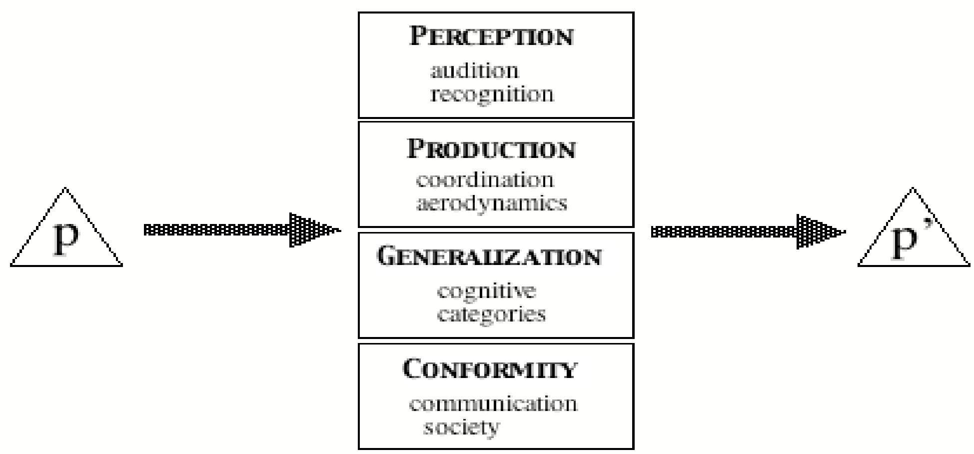
- Production demands for coda t/d can be assumed to be higher than for onset t/d, resulting in greater likelihood of deletion of the stops when in coda position. Part of the reason for this:
- Deletion of coda t/d was most frequent when flanked between consonants, putting greater production demands on the stop than when followed by a vowel. Further, when the preceding consonant was a non-approximant, cluster simplification resulting in deletion was more likely when the consonant was non-homorganic, also increasing the articulatory difficulty of the cluster.

(8) One approach to explaining these interactions is proposed in Hume & Johnson (2001)
a. A general model of the interplay of "external factors" and phonology, broadly defined.



- *p*: the cognitive symbolic representation of a person's phonological knowledge, embodied in the person's brain. We may assume that *p* is a component of *l*, the cognitive symbolic representation of a language.
- factors can both influence the sound system of language as well as be influenced by one's language, hence the bi-directional arrows in the diagram between these factors and *p*.

(9) The interaction of external factors with phonological knowledge (*p*, *p'*). External factors as filters.



- Interactions of opposing tendencies occur in cycling *p > p'(> p...)* where the interval between cycles is very short.

- A change that reduces cost on one function may produce increased cost on another and so be quickly reversed.

E.g., the sound pattern [nt] may be changed to [nd] in order to achieve lower articulatory cost (avoiding the modulation of voicing). In the next cycle, [nd] may be changed back to [nt] because [nd] conflicts with conformity (e.g. [nd] diverges too much from the socially accepted pronunciation norm).

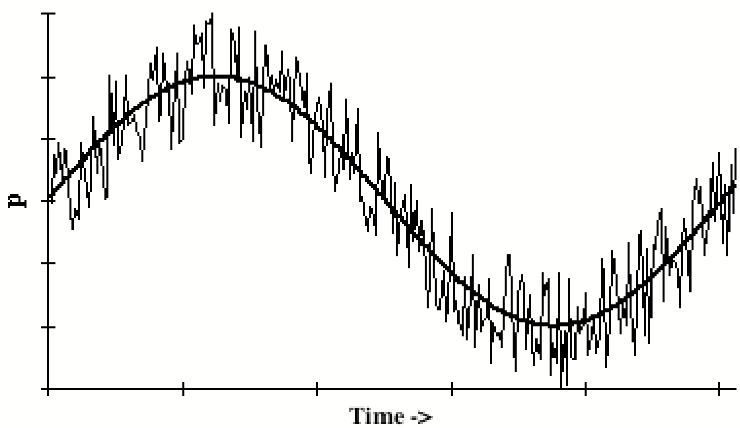
(10)

a. Unlike the traditional view that sound change occurs over hundreds or thousands of years (a course-grained time scale), our model handles interactions by using a fine-grained scale, as shown at the top of page 6.

b. The slowly changing line shows general tendencies, while the fine-grained time scale shows rapidly fluctuating change. Time in this illustration is on the horizontal axis, and the vertical axis is meant to show, in an abstract one dimensional projection, the location of sound system *p* of a language in the space of possible languages.

c. We then see the development of *p* over time, with local noise overlaying global stability. Through the sequential interaction of forces, it is a self-organizing system that is nonetheless in constant flux.

d. This view is consistent with Labov's (various) and Joseph & Janda's (1988) view that language change occurs in the present.



(11) This approach lays the groundwork for understanding the interaction of factors. Yet, the model says nothing about the contributions of the various factors and how these work together to predict a particular sound pattern.

Recall, for example, that bigram frequency is relevant for onset t/d but not for coda t/d; both perception and frequency are relevant for metathesis.

Regarding (2c): How can we model interactions mathematically in order to predict sound patterns?

A probabilistic model of factor interaction:

(12) a. We can express the likelihood of a particular phonological process occurring in terms of probability. For example, let's assume that *x*, *y* are speech sounds in some language system S, and that there is an

80% chance of the sequence *xy* changing to *yx* (and a 20% chance of it not changing). This can be expressed as:

$$p(\text{xy} \text{ -> } \text{yx}|\text{S}) = .8$$

b. But how do we arrive at the value .8 (or whatever the relevant value is)? This is the tricky part and the part where our knowledge as linguists comes into play.

c. We can assume that each factor (perception, production, etc.) has a probability distribution. For example, the probability of a precedence reversal occurring in (12a) can be determined by considering the probability that each factor will contribute to metathesis.

- <u>production factors</u>: the probability of *xy* being pronounced incorrectly, i.e. as something other than *xy*.

  $$1 - p(\text{xy}|[\text{xy}])_{\text{prod}}$$

  The greater the probability of the sequence being mispronounced, the greater the probability that metathesis will occur.

- <u>perceptual factors</u>: the probability of *xy* being misperceived, i.e. not perceived as *xy* (this relates to the requirement that the perceptual cues to identifying the sequence undergoing metathesis are acoustically/auditorily indeterminant) .

  $$1 - p(\text{xy}|[\text{xy}])_{\text{perc}}$$

  The greater the probability of the sequence being misperceived, the greater the probability that metathesis will occur.

- <u>contextual probability</u>: the probability that some linguistic element occurring in a given context. For metathesis, recall that the output needs to be (a) an attested phonological structure in the language and (b) more frequent than the corresponding input structure. We can formulate this as:
  - the probability of (output) YX occurring in the language, minus the probability of (input) XY occurring in the language, where YX represents some phonological structure (feature, segment, contrast, syllable, etc.) which includes the segment sequence *yx*.

  $$p(\text{YX})_{\text{cp}} - p(\text{XY})_{\text{cp}}$$

  For metathesis to occur, the probability of the output, $p(\text{YX})_{\text{cp}}$ must be greater than $p(\text{XY})_{\text{cp}}$.

Note that in this analysis, the factor in (3) referring to "existing phonological structure" is included in the factor category, "contextual probability." Thus, any phonological structure (e.g. syllable, feature, segment) can be measured in terms of its frequency of occurrence. Well-formedness is then assumed to be probabilistic rather than categorical, an assumption shown to be necessary for predicting metathesis patterns (Hume 2004).

c. The probability of a precedence reversal affecting the sequence *xy* in language S can be expressed as below. For simplicity, I include only three factors though all factors would figure into the complete model. Also, I assume that factors are independent and equally weighted, assumptions that will most likely need to be modified.

$$p(\text{xy -> yx}|S) = 1 - p(\text{xy}|[\text{xy}])_{perc} \cdot 1 - p(\text{xy}|[\text{xy}])_{prod} \cdot p(p(\text{YX})_{cp} - p(\text{XY})_{cp})_{cp}$$

(12) Hypothetical case 1:
Consider a language where there is recent innovation with the intervocalic sequence [nl] and it is now sometimes pronounced as [ln] in language S. Other sequences such as [nt] do not show this variation. I assume in this illustration that the crucial difference in the two cases relates to **perceptual factors**: there is greater indeterminacy with respect to precedence in [nl] than with respect to [nt]. The influence of all other factors remains constant.

$$\begin{aligned} p(\text{nl -> ln}|S) &= 1 - p(\text{nl}|[\text{nl}])_{perc} \cdot 1 - p(\text{nl}|[\text{nl}])_{prod} \cdot p(\text{ln})_{cp} - p(\text{nl})_{cp} \\ &= \mathbf{.8} \cdot .7 \cdot (.6 - .2) \\ &= .8 \cdot .7 \cdot .4 \\ &= .224 \end{aligned}$$

$$\begin{aligned} \text{p}(\text{nt -> tn}|S) &= 1 - p(\text{nt}|[\text{nt}])_{perc} \cdot 1 - p(\text{nt}|[\text{tn}])_{prod} \cdot p(\text{tn})_{cp} - p(\text{nt})_{cp} \\ &= \mathbf{.2} \cdot .7 \cdot (.6 - .2) \\ &= .2 \cdot .7 \cdot .4 \\ &= .056 \end{aligned}$$

Prediction: given the differing contributions of perceptual factors, intervocalic [nl] is more likely than [nt] to undergo a precedence reversal.

(13) Hypothetical case 2:
Consider another language, S'. In this case, the intervocalic sequence [mk] is being pronounced as [km]. There is no variation in the pronunciation of [nk], however. In both cases, the perceptual cues of nasal + stop are better than those of stop + nasal. In terms of production, they are similar as well. The crucial difference relates to the frequency of the various sequences: [mk] is less frequent than [km], while [nk] is more frequent than [kn].

$$\begin{aligned} p(\text{mk -> km}|S') &= 1 - p(\text{mk}|[\text{mk}])_{perc} \cdot 1 - p(\text{mk}|[\text{mk}])_{prod} \cdot p(\text{km})_{cp} - p(\text{mk})_{cp} \\ &= .2 \cdot .7 \cdot (.6 - .2) \\ &= .2 \cdot .7 \cdot .4 \\ &= .056 \end{aligned}$$

$$\begin{aligned} \text{p}(\text{nk -> kn}|S') &= 1 - p(\text{nk}|[\text{nk}])_{perc} \cdot 1 - p(\text{nk}|[\text{nk}])_{prod} \cdot p(\text{kn})_{cp} - p(\text{nk})_{cp} \\ &= .2 \cdot .7 \cdot (.2 - .5) \\ &= < 0, \text{ thus undefined} \end{aligned}$$

Prediction: Given the differing contributions of contextual probability, intervocalic [mk] is more likely than [nk] to undergo a precedence reversal.

(14) Deletion (and any other phonological process) can be modeled in a similar way.

(15) Conclusion:
Advantages of the model:
- mathematically rigourous
- integrates knowledge from a range of fields in order to elucidate language sound systems; encourages collaborative research across areas of specialization
- can make predictions about sound patterns within and across language varieties

Challenges:

- o contextual probabilities: we need more corpora of languages, coded for relevant phonological structure
- o perceptual factors: to calculate probability values, one source of data is confusion matrices, but more are needed since they do not exist for all languages, or all segment types in a given language
- o production factors: measures for calculating relative articulatory complexity?
- o user-friendly tools for searching databases

References

Blevins, Juliette and Andrew Garrett. 1998. The origin of consonant-vowel metathesis. *Language*. 74. 3. 508-55.

Grammont, Maurice. 1933. *Traité de phonètique*. Paris: Librairie Delagrave.

Hock, Hans Henrich. 1985. Regular metathesis. *Linguistics* 23. 529-46.

Hume, Elizabeth, 1997. Metathesis in phonological theory: The case of Leti. *Lingua* 104. 147-86.

Hume, Elizabeth. 2004. The Indeterminacy/Attestation Model of Metathesis. *Language* 80(2).

Hume, Elizabeth & Keith Johnson. 2001. A model of the interplay of speech perception and phonology. In E. Hume & Keith Johnson (eds.), *The role of perception in phonology*. New York: Academic Press. 3-26.

Joseph, B. & R. Janda (1988) The how and why of diachronic morphologization and demorphologization. In M. Hammond & M. Noonan (eds.), *Theoretical Morphology*. New York: Academic Press.

Pitt, Mark, Keith Johnson, Elizabeth Hume, Scott Kiesling, and William Raymond. 2004. The Buckeye Corpus of Conversational Speech: Labeling Conventions and a Test of Transcriber Reliability. *Speech Communication*.

William, Raymond, Robin Dautricourt & Elizabeth Hume. (2006). Word-medial /t,d/ deletion in spontaneous speech: Modeling the effects of extra-linguistic, lexical, and phonological factors. *Language Variation and Change,* 18. 55-97.

Wanner, Dieter. 1989. On metathesis in diachrony. In Wiltshire, Caroline, R. Eraczy & B. Music (eds.), *CLS 25: Papers from the 25th annual regional meeting of the Chicago Linguistic Society*. Chicago: CLS.