The phonological domains of final lengthening

Mary E. Beckman
Ohio State University

Jan Edwards
Hunter College, CUNY

1. The phonological framework

A syllable at the end of a phrase is considerably longer than it would be phrase-internally. Similarly, a stressed monosyllable is often observed to be longer than any segmentally identical syllable that is separated from the word's edge by one or more unstressed syllables. This paper describes some experiments we conducted in an attempt to determine the precise domain of these effects. Relying on earlier reports of other seemingly related phenomena (e.g., Gee & Grosjean 1983), we assume that the domain is phonological; although surface syntactic constituency influences the effects, its influence is mediated by prosodic structure. By way of introduction, therefore, we first review the potentially relevant prosodic domains.

We consider the prosodic structure of an utterance to be a hierarchical arrangement of various prominence-lending phonological properties. This arrangement can be represented by a metrical grid with suitable bracketings at any level that also has constituents with phonologically marked edges. The grid below, for example, represents the phrase phonological structure as it might be said in isolation, with an intonation typical of citation forms. (We adopt the intonational analysis and notation of Pierrehumbert 1980.)

```
[                x    ]⌐     nuclear accent, boundary tone
     x           x           accent
  x  x           x           stress
  x x x x x      x  x        syllable
phonological structure
      |            |
      H*           H* L L%
```

The lowest level of this grid consists of seven local sonority peaks defining events called 'syllables'. Three of these syllables contain unreduced vowels, and are qualitatively longer and louder than the others, properties which define another level of events called 'stresses'. Two of these stressed syllables are autosegmentally associated to certain prominence-lending tonal configurations in the intonation contour, the two H* 'pitch accents'. The association to a pitch accent creates another level of prosodic strength, that of 'accented syllables'. The last pitch accent is followed by an unassociated L tone, the 'phrase accent'. The falling tonal pattern created by the juxtaposition of the phrase accent helps to give the syllable associated to the last pitch accent a special prominence known as 'nuclear stress' or 'sentence stress'. (The last accent itself is designated the 'nuclear accent'.) There is also a L% 'boundary tone' aligned to the edge of the phrase after the phrase accent. This boundary tone phonologically marks the end of a constituent called the 'intonational phrase'.

A fact to note about this grid is that only the highest level corresponds to any well-documented phonological constituent. Here, there is a boundary tone to mark the edges of units headed by nuclear stresses, whereas every other level only has the phonological event marking the prominence peak. An attractive hypothesis, therefore, is that phrase-final lengthening is merely the durational correlate of the boundary tone, and thus is limited to syllables at the ends of intonational phrases.

2. Intonational phrasing

We tested this hypothesis using the sets of sentences shown in Table I. The first set had a three-way contrast among pep, pepper, and peppermint, in which an identical stressed target syllable is separated from the end of the word by 0, 1, or 2 unstressed syllables. The second set had a similar two-way contrast between Pop and Poppa. It also had different verbs following the target nouns so as to keep a constant inter-stress interval length. Both corpora also contrasted pairs of sentences in which the material following the target either is or is not a kind of clause that is obligatorily set off as a separate intonational phrase. One subject read the 'pep-pepper' corpus and two subjects read the 'Poppa posed' corpus. They read the sentences from a randomized list for a total of five tokens of each type at each of three different self-selected speaking rates. The readings took place in a sound-treated recording booth, and the recorded sentences were analyzed using a digital waveform editor. (The same methods were used for the subjects in all subsequent experiments described below.)

Fig. 1 shows the overall results from the 'pep-pepper' experiment averaged over all three rates. The target syllable was nearly twice as long in pep as in the other two words, but only in the sentences where the word boundary coincided with an obligatory intonational phrase break. An analysis of variance showed significant main effects for word and for phrasing, and also a significant interaction between the two variables ($F=2,75$, $P<0.0001$). These results suggest strongly that the domain of any phrase-final lengthening is the intonational phrase.

However, one aspect of the data in fig. 1 seems to contradict this hypothesis. Although the difference was not as large as in the sentences with the obligatory intonational break, the vowel in pep was significantly longer even in the no-break condition ($F=2,16$, $P<0.0001$).

Table I. Corpora for intonational phrasing experiments

1. a. Pep, for the lack of which the party will suffer, is not to be had.
   Pepper, for the lack of which the chili will suffer, is not to be had.
   Peppermint, for the lack of which the frosting will suffer, is not ...
   b. Pep for the party is not to be had for love or money.
   Pepper for the chili is not to be had for love or money.
   Peppermint for the frosting is not to be had for love or money.

2. a. Pop opposed the question strongly, and so refused to answer it.
   Poppa posed the question strongly, and then refused to answer it.
   b. Pop, opposing the question strongly, refused to answer to it.
   Poppa, posing the question strongly, demanded an answer to it.

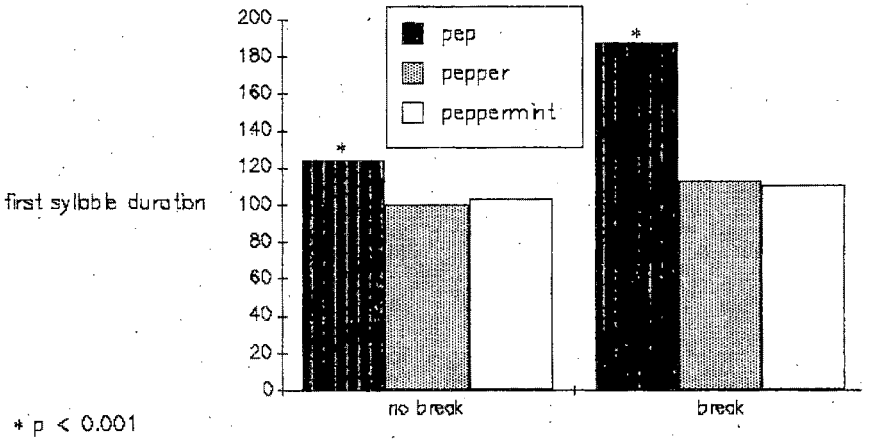

Figure 1. Mean durations in ms. for vowel in first syllable in 'pep-pepper' experiment averaged over all rates. Subject RWF.

The results for the 'Poppa posed' corpus were similar. Both subjects had a considerably longer [a] in Pop and schwa in Poppa in the sentences with an obligatory intonational phrase boundary following the target word, and both subjects also showed differences that were smaller but in the same direction for the sentences with the other syntactic structure. For subject JRE the smaller differences were significant overall, while for subject LAW they were significant only when separate ANOVA's were calculated for the different rates, and then only for the slow rate, as illustrated for the [ə] in fig. 2.

Our first thought on seeing the smaller difference in the sentences with no obligatory medial break was that the subjects must have produced optional breaks in some tokens of these sentences. The interaction with rate for subject LAW in the 'Poppa posed' corpus made this explanation seem especially likely, since speakers tend to produce more intonational phrases when they speak more slowly or more deliberately. When we listened to these utterances, and looked at their fundamental frequency patterns, however, we saw no evidence of such a drastic restructuring. We concluded that there is a real smaller effect in these sentences which is different from the substantial phrase-final lengthening at the intonational phrase boundary. We would like to think that the smaller effect is also some sort of final lengthening, perhaps for a constituent at some lower level of the grid. Since none of the other levels have independently motivated phonological constituents, however, we must first consider another possible explanation that does not involve positing a phonological phrase smaller than the intonational phrase.

English is often claimed to be stress-timed, with stressed syllables following each other at regular intervals. In strong versions of the stress-timing claim, such as that of Pike 1945, this rhythmic regularity is purportedly achieved by adjusting segment durations when different numbers of

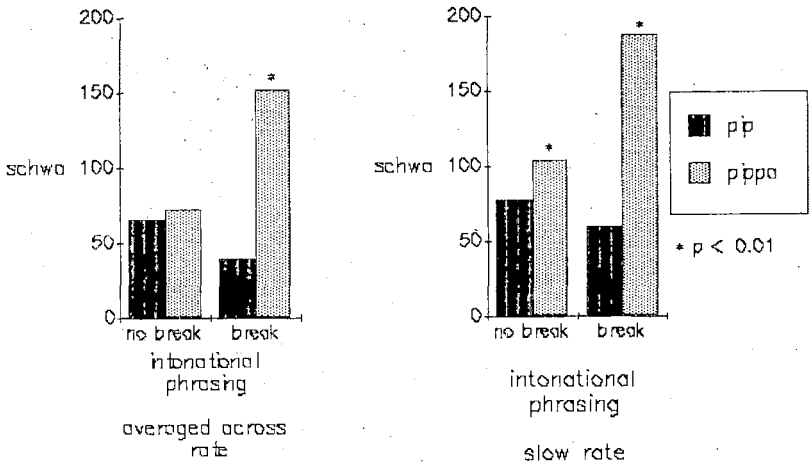

Figure 2. Mean durations for schwa in 'Poppa posed' corpus averaged over all rates (left) and for slow rate only (right). Subject LAW.

unstressed syllables intervene between stresses. Thus, the results from the 'pep-pepper' corpus could be evidence for stress-timing rather than any indication of the existence of phonological constituents smaller than the intonational phrase, as illustrated by the grids below, in which vertical lines separate the inter-stress intervals:

x x pep ..	x x x for the	x x x party	vs	x x x pepper	x x x for the	x x x chili

But stress-timing cannot explain the 'Poppa posed' corpus results, since in that corpus there was always exactly one unstressed syllable in the interval between the stress in the target noun and the stress in the following verb. Therefore, the smaller difference in the sentences where there was no medial intonational phrase break must be a final-lengthening effect and a boundary mark for some smaller phonological constituent. We labeled the effect 'word-final' (as opposed to 'phrase-final') lengthening, and did two further experiments in order to locate it more precisely in relation to the grid.

3. Accentual phrasing

The first hypothesis we considered is that word-final lengthening is a boundary mark for a constituent that is the domain of the pitch accent. This seemed a likely possibility, because accents belong to the intonation, whereas stress patterns are largely specified in the lexicon. Also, speakers may produce more pre-nuclear pitch accents in slower renditions of a given sentence, a tendency which could explain the rate effect in subject LAW's results. We therefore posited the existence of 'accentual phrases' headed by accented syllables and bounded by word-final lengthening, as shown below:

```
[              x    ]   nuclear accent/intonational phrase
[     x    ] [  x    ]   accent/accentual phrase
phonological structure
      |          |
      H*         H* L L%
```

We first tested this hypothesis with the sentences in Table II, which again contrasted the phrases Pop opposed and Poppa posed. Before saying each target sentence, the subjects read a context question, which induced contrastive focus either on the noun or on the following verb. Contrast puts nuclear stress on the item in focus. Thus, depending on the context question, there would either be nuclear accent on the verb and a medial accentual phrase boundary coinciding with the target word boundary, or there would be nuclear accent on the target noun and only one accentual phrase in the sentence. Since we assumed further that everything in an utterance belongs to some accentual phrase, we thought that the phrase containing the nuclear accent

Table II. Focus placement corpus for accentual phrasing experiment

1. a. Q. So, your dad liked the question?
   A. Pop OPPOSED the question.
   b. Q. So, your dad answered the question?
   A. Poppa POSED the question.

2. a. Q. So it was grandpa who opposed the question?
   A. POP opposed the question.
   b. Q. So it was grandpa who posed the question?
   A. POPPA posed the question.

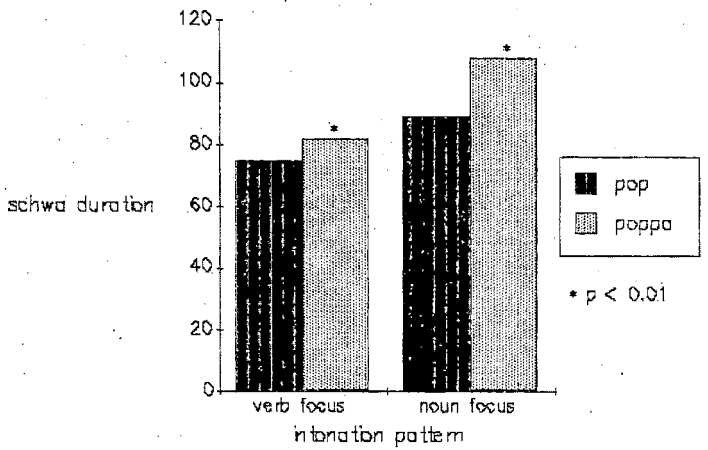

Figure 3. Mean durations for [ə] in focus experiment sentences produced at normal rate. Subject JRE.

must continue to the end of the sentence, predicting that the answers with focus on the noun would differ from those with focus on the verb by showing no word-final lengthening on the target words.

This prediction was not borne out. The same two subjects who read the 'Poppa posed' intonational phrasing corpus also read the focus corpus. Fig. 3 illustrates the results by showing the mean values for the schwa in the sentences at normal rate produced by subject JRE. The word-final schwa in Poppa posed was significantly longer than the non-final schwa in Pop opposed whether the focus was on the verb or on the noun. The results for the preceding [a] for this speaker and for both vowels for the other speaker are similar. Thus, in terms of the prediction, this experiment does not support a unit at the level of accents as the domain for word-final lengthening.

On the other hand, these results constitute evidence against the hypothesis only if everything in an utterance must belong to some accentual phrase. But if only syllables in words with accents belong to constituents at this level, the results are equivocal. The lone accentual phrase in the sentences with focus on the noun would then terminate at the end of the target word and the following material up to the end of the sentence would be unaffiliated to any accentual phrase, as illustrated below:

```
[x                        ]   nuclear accent/intonational phrase
[x  ] _____________________   accent/accentual phrase
 x     x          x          stress
 x  x  x     x    x   x      syllable
POPPA posed the question.
 |
 H*          L          L%
```

(In this grid, the underscore at the accent level highlights material that is unaffiliated to any accentual phrase.) The focus sentences thus might give evidence for the accentual phrase, but they could not disprove it.

4. Accentual phrase, stress foot, or independent prosodic word?

Table III gives sample sentences from the experiment that we designed to correct this flaw of the corpus involving contrastive focus. The target phrases in this experiment, superstition, super station, and Sioux perspective, all have the same stress pattern but different word-boundary placements. The sentences also contrasted three different intonation patterns chosen for their pitch-accent placements relative to the two stressed syllables in the target phrases. In the first pattern, the nuclear accent is on make, so that there can be no accents on either stressed syllable in the target phrase because it is in post-nuclear position. The second pattern placed 'scooped' L*+H accents on the word real preceding the target phrase and on the second stressed syllable in the target phrase, but no accents on the first syllable. The third pattern placed a pre-nuclear L* accent on the first stress and a nuclear H* on the second stress in the target phrase.

This corpus tests three hypotheses about word-final lengthening. The first is again the notion that the lengthening marks accentual phrases. The test for this hypothesis is that, since a lexical item can have more than one accent, there should be word-final lengthening internal to lexical items when

Table III. Intonation patterns for second accentual phrasing experiment

1. post-nuclear
   You may call it a superstition, but that doesn't MAKE it a superstition.
   H* L L%

2. uncertainty contour
   Q. Do you have any feigned beliefs?
   A. I have a real superstition.
   L*+H L*+H L H%

3. surprise-redundancy contour
   Don't you understand?! It's a superstition!
   L* H* L L%

---

accents are placed appropriately. Thus, superstition should pattern exactly like super station; its [u] should be shorter and its schwar longer than in Sioux perspective, but any difference among the three phrases should hold only when both stressed syllables are accented, in the surprise-redundancy contour:

```
[x  ] [ x  ]   [x  ][ x  ]   [ x ] [   x   ]   accentual phrases
 x      x       x     x       x       x
 x x    x  x    x x   x  x    x   x   x  x
super station  superstition  Sioux perspective
 |      |       |     |       |       |
 L*     H*      L*    H*      L*      H*
```

The second hypothesis is that word-final lengthening marks a 'stress foot'. If this hypothesis is correct, then there should be the durational patterns just described, but without the dependency on accent placement:

```
[x  ] [ x  ]   [x  ][ x  ]   [ x ] [   x   ]   stress feet
 x x    x  x    x x   x  x    x   x   x  x
super station  superstition  Sioux perspective
```

The third possibility is that phrasing below the intonational phrase level is independent of the prosodic hierarchy, that the word-final lengthening marks a 'prosodic word' that is not necessarily headed by any prosodic peaks such as accents or stresses. In this case, final lengthening should occur only at the edges of actual lexical items, so that the schwar in superstition should always be shorter than that in super station:

```
[   ] [    ]   [         ]   [   ] [         ]   prosodic words
super station  superstition  Sioux perspective
```

We had six subjects in this experiment, and the results showed two different patterns. For the first speaker, the [u]'s in superstition and super station were shorter than in Sioux perspective, but only in the suprise-redundancy contour, where they were accented as well as stressed (fig. 4a). The [ɚ] in superstition also patterned like that in super station (fig. 4b).

In both words, it was consistently longer than in Sioux perspective, but again, only in the suprise-redundancy intonation. The similarity of superstition to super station and the dependency on accent pattern for any difference among the words suggests that the relevant unit for word-final lengthening is an accentual phrase.

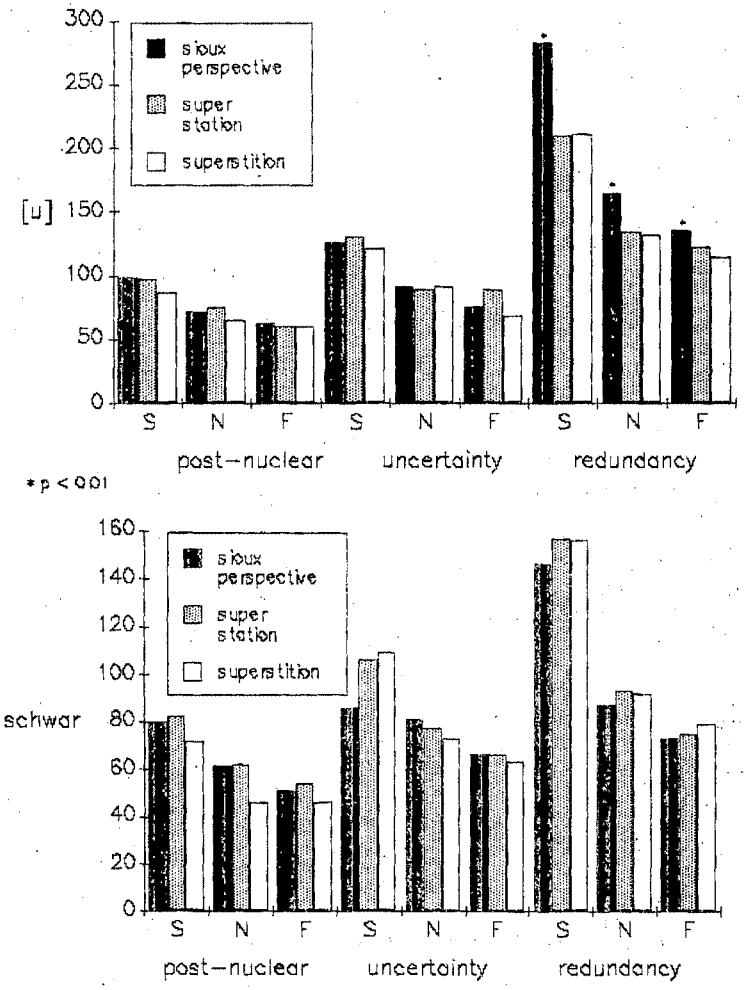

Figure 4. Mean durations for [u] (top) and [ə] (bottom) in each test phrase, averaged by rate (Slow, Normal, Fast) and by intonation pattern. Subject JRE.

The second speaker, on the other hand, showed no dependency on the accent pattern. He had a longer [u] in Sioux regardless of the intonation pattern, although it was consistently so only at the slow rate (fig. 5a). His second syllables also showed no dependency on accent (fig. 5b). The schwar in super station was longer than in Sioux perspective whatever the accent placement,

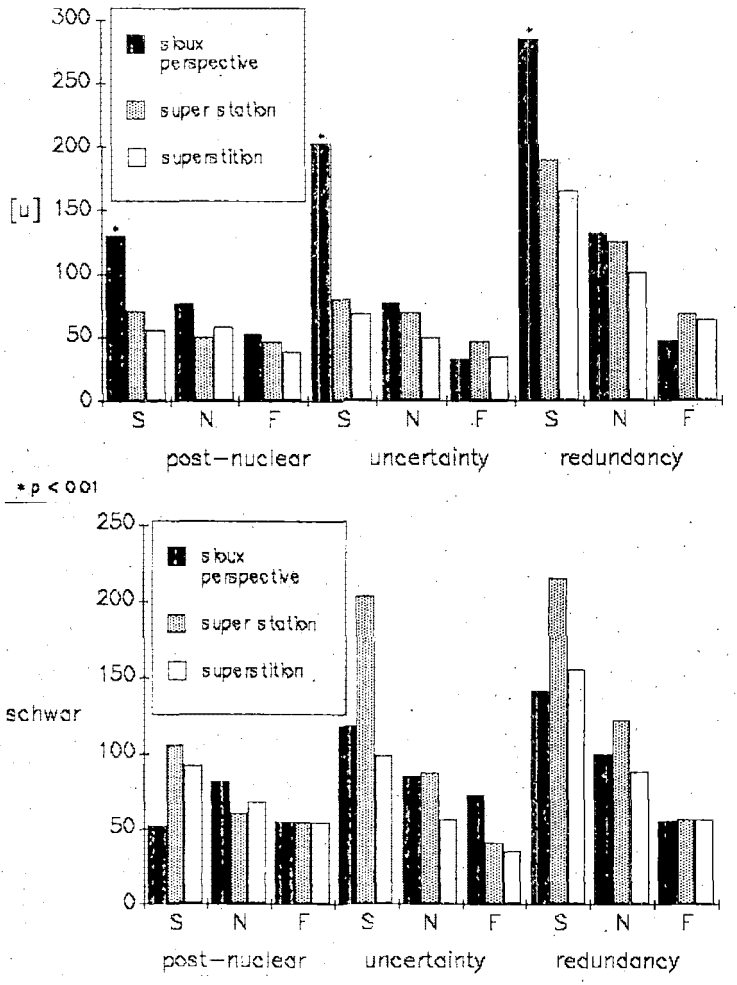

Figure 5. Mean durations for [u] (top) and [ɔ] (bottom) in each test phrase, averaged by rate (Slow, Normal, Fast) and by intonation pattern. Subject JSC.

although again only at the slow rate. Moreover, superstition did not pattern like super station. Instead, its schwar was generally shorter, like the non-final vowel in Sioux perspective. Thus, this subject's results do not support either the accentual phrase or the stress foot as the domain for word-final lengthening. They suggest rather a prosodic word that is independent from the hierarchy of stresses and accents.

Of the other subjects, two seemed to pattern like the first, showing some evidence for the accentual phrase, and two patterned more like the second, showing evidence for the prosodic word as a phrasal constituent that is independent of the prosodic hierarchy of stresses and accents. The comparisons which support these apparent patterns did not often reach significance, however. The insignificance of the differences in relation to the measure of error in the statistical analysis is perhaps inevitable given the small size of the word-final lengthening effect and the small sample sizes of the categories being compared. (Recall that each of the bars in figs. 4 and 5 represent only five tokens). Thus, since few of the crucial comparisons reached significance, these results do not argue conclusively for two possible speaker-dependent patterns in the use of word-final lengthening.

On the other hand, our experiments do sustain two important conclusions. First, they strongly suggest that there are two different final-lengthening effects: phrase-final lengthening and word-final lengthening. Phrase-final lengthening occurs at intonational-phrase boundaries, and is a large effect that is highly consistent across speakers and rates. Word-final lengthening occurs at some smaller constituent's boundaries, and is a much smaller effect that is not always discernible in experiments that have only five tokens of each type. Second, the word-final effect cannot be explained as a result of stress-timing in English and must be a true final lengthening. However, more ambitious experiments are needed to locate its domain more precisely below the intonational phrase.

Acknowledgements

We thank Lori Weiss and Younghee Chung for doing most of the duration measurements, and John Sidtis for letting us use his analysis facilities at New York Hospital (developed under NIH grants NS 18802-03 and NS 17778-04). We also thank Osamu Fujimura's department at AT&T Bell Labs for loaning an AT&T PC 6300 and A/D board to the lab at Ohio, and Joan Miller for letting us use her speech and graphics routines in a waveform editor. The experiments benefited at various stages from comments by Carol Fowler, Katherine Harris, Julia Hirschberg, Gerald Lame, and especially Janet Pierrehumbert.

References

Gee, James Paul, and Francois Grosjean. (1983). Performance structures: A psycholinguistic and linguistic appraisal. Cognitive Psychology 15.411-458.

Pierrehumbert, Janet B. (1980). The phonology and phonetics of English intonation. Unpublished Ph.D. dissertation, MIT.

Pike, Kenneth. (1945). The intonation of American English. Ann Arbor, Mich.: University of Michigan Press.