The acquisition of English stress by speakers of nonaccentual languages: lexical storage versus computation of stress*

JOHN ARCHIBALD

Abstract

In this paper I look at the acquisition of English stress by speakers of Chinese and Japanese in a longitudinal study. Based on the observed developmental patterns, and on characteristics of the first languages involved, I argue that the subjects are not computing metrical structures (as has been observed in subjects whose L1s are Spanish, Polish, and Hungarian) but rather are storing the English stress lexically.

1. Introduction

For some time now I have been investigating the acquisition of second-language stress systems in an attempt to get an idea of what the representations and processes involved in interlanguage phonology are. The study that I discuss in this paper differs in two respects from my previous studies (e.g. Archibald 1993a, 1993b), which investigated, via a cross-sectional design, how Polish, Hungarian, and Spanish speakers acquired English stress. One, it is a longitudinal study, and two, it involves different first languages (Chinese and Japanese). Ideally, it might have been better to change only one of the factors, but these are the subjects who were available to me for a longitudinal study. As a result, this paper has more to say about the influence of the first language (L1) type on second-language acquisition than about patterns of development over time, because, as we shall see, the performance didn't change very much over time. In the past, I looked at subjects whose L1s had stress, while in this study I look at subjects whose L1s do not have stress. This typological difference will prove to be significant in explaining the behavior of the Chinese and Japanese subjects.

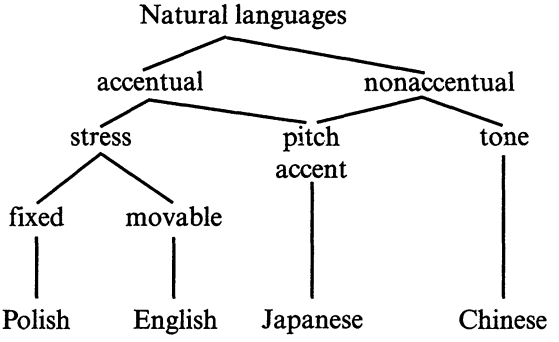
Linguistics 35 (1997), 167–181

0024–3949/97/0035–0167

2. A taxonomy of stress systems

There are several different types of stress systems found in natural languages. A taxonomy is shown in (1).

(1) A taxonomy of natural languages



There are languages like English that use pitch to signal stress accent, and languages like Chinese that use pitch phonemically. Languages like Japanese appear to have characteristics of both types. Exactly how to distinguish between stress, pitch accent, and tone languages has generated a considerable amount of discussion in the literature (Beckman 1986; Hyman 1978; Hayes 1995; Pulleyblank 1986; van der Hulst and Smith 1988). Hayes (1995: 49–50) comes out most clearly when he says, "pitch accent languages must satisfy the criterion of having INVARIANT TONAL CONTOURS on accented syllables, since tone is a lexical property. This is not so for pure stress languages, where the tonal contours of stressed syllables can vary freely" and furthermore be influenced by intonation.

In previous studies I looked at native speakers of Polish and Hungarian (essentially fixed-stress languages) acquiring English stress (Archibald 1993a). Polish usually stresses the penultimate syllable and Hungarian usually stresses the initial syllable. I also looked at native speakers of Spanish (a movable-stress language) acquiring English stress (Archibald 1993b). These studies had been conducted assuming the metrical parameters proposed by Dresher and Kaye (1990), shown in Table 1.

The differences between the languages are shown in Table 2. From these tables, it can be seen that while there may be parametric differences between the languages studied, the same kinds of representations were being constructed in the first and second languages, that is, representations of stress.

In this paper, I hope to expand my data base by looking at native speakers of a tone language (Chinese) and a pitch-accent language

Table 1. *The metrical parameters of Dresher and Kaye*

P1	The word-tree is strong on the [left/right].
P2	Feet are [binary/unbounded].
P3	Feet are built from the [left/right].
P4	Feet are strong on the [left/right].
P5	Feet are quantity-sensitive (QS) [no/yes].
P6	Feet are QS to the [rime/nucleus].
P8	There is an extrametrical syllable [no/yes].
P8A	It is extrametrical on the [left/right].

Table 2. *The parameter settings of Spanish, Polish, Hungarian, and English*

		Spanish	Polish	Hungarian	English
P1	(word tree)	right	right	left	right
P2	(foot type)	binary	binary	binary	binary
P3	(built from)	left	left	left	left
P4	(strong on)	right	right	left	right
P5	(QI/QS)	QS	QI	QS	QS
P6	(sensitive to)	rime	NA	nucleus	rime
P8	(extrametrical)	yes	no	no	yes
P8A	(extrametrical on)	right	NA	NA	right

(Japanese). As shown in (1), it has been argued that these languages are not stress languages, and therefore subjects who have these languages as their L1s may have very different kinds of representations. Furthermore, they would have to acquire a new way of representing prominence in their second language. I will argue that we see these subjects treating English stress as a lexical phenomenon.

3. Research design

The basic research design used in my earlier studies forms the basis of this study as well. Subjects perform both production and perception tasks related to stress assignment. First they had to read a list of words out loud (see Table 4). (Departing from my earlier studies, I did not have the subjects engage in any sentence-level tasks, as they were not found to perform significantly differently on these tasks.) Stress placement was transcribed on these words. Then the subjects listened to the same words they had produced as they were read out loud on a tape recorder by a native English speaker. After a training session, the subjects had to mark

which syllable they perceived stress to be on. My previous studies have shown that native speakers can do this reliably.

3.1. *The subjects*

In this study I gathered data from ten subjects in November. Only four subjects were able to be reassessed in March of the following year. The subject profiles are given in Table 3.

3.2. *The test items*

The words that the subjects had to produce and perceive are given in Table 4.

Table 5 gives the defining characteristics of the classes of words (ignoring some phonological details that are not relevant to this discussion).

One of the ways to get a picture of the interlanguage grammar is to look at the errors that the subjects make in their production and perception of the stress patterns of these word classes. In other words, native speakers of English have knowledge of these word classes, and therefore it is something that nonnative speakers have to acquire. In my previous studies, it was a robust finding that the perception tasks were significantly different (and significantly more accurate) than the production tasks.

Table 3. *Subject profiles*

	Age	L1	L2 level (out of 6)
Subject 1	19	Cantonese	5
Subject 2	19	Japanese	6
Subject 3	32	Mandarin	6
Subject 4	21	Mandarin	3

Table 4. *Test items*

Class 1	aroma	Manitoba	arena	Minnesota	horizon
Class 2	agenda	consensus	appendix	veranda	synopsis
Class 3	cinema	javelin	venison	America	cabinet
Class 7	hurricane	baritone	antelope	candidate	matador
Class 4	maintain	appear	erase	decide	achieve
Class 5	collapse	elect	observe	adapt	convince
Class 6	astonish	edit	cancel	consider	interpret

Table 5. *Phonological properties of the test items*

Class 1	noun, penultimate stress due to heavy penult (tense vowel)
Class 2	noun, penultimate stress due to heavy penult (branching rhyme)
Class 3	noun, antepenultimate stress due to lack of heavy syllables
Class 7	noun, antepenultimate stress due to stress retraction (secondary stress)
Class 4	verb, final stress due to heavy final syllable (tense vowel)
Class 5	verb, final stress due to heavy final syllable (branching rhyme)
Class 6	verb, penultimate stress due to lack of heavy syllables

That is, the subjects were better at perceiving stress accurately than they were at producing stress accurately.

4. The results

As the L1s are typologically distinct, I will separate my discussion of the Chinese subjects from that of the Japanese subject. I begin with the Chinese subjects.

4.1. *Chinese subjects*

The chart shown in Table 6 gives a profile of the numbers of errors that the subjects made (where T1 = Time 1, and T2 = Time 2). T-tests did not reveal any significant differences in the mean numbers of errors.[1] The closest to significance was between production and perception at T2.

One of the first things to note is that for all three of the subjects, the perception scores are worse than (or in one case equal to) the production scores (this is true of the means as well). We also note that overall from

Table 6. *Perception versus production (Chinese subjects)*

	Perception		Production	
	T1	T2	T1	T2
Subject 1	19	22	10	9
Subject 3	15	13	10	10
Subject 4	10	22	10	9
Total	44	57	30	28
Mean	14.7	19	10	9.3

Table 7. *Production versus perception errors by class (Chinese subjects)*

		Perception		Production	
		T1	T2	T1	T2
Subject 1	Class 1	2	3	2	2
	Class 2	3	4	1	0
	Class 3	4	4	1	0
	Class 7	2	2	2	3
	Class 4	2	1	1	1
	Class 5	4	5	2	2
	Class 6	2	3	1	1
Subject 3	Class 1	3	1	2	2
	Class 2	2	2	1	2
	Class 3	3	1	2	0
	Class 7	2	4	1	1
	Class 4	1	1	0	1
	Class 5	2	2	3	3
	Class 6	2	2	1	1
Subject 4	Class 1	2	1	2	2
	Class 2	0	4	0	1
	Class 3	4	2	0	0
	Class 7	2	5	3	1
	Class 4	1	4	2	2
	Class 5	0	3	3	2
	Class 6	1	3	0	0

Time 1 (November) to Time 2 (March) the perception scores actually got worse, though not significantly so.

If we break the above chart down into errors by class (C), the picture shown in Table 7 emerges. Clearly, the differences between word classes were minimal, as were the differences between performance at T1 and T2. In an attempt to see whether the subjects were treating different word classes differently, I combined the production and perception errors to see if class differences would emerge from this view. The result is shown in Table 8. Again, we note that for each subject, there was very little difference between word classes and very little change from T1 to T2. This can be seen more clearly when we present the data as shown in Table 9.

The differences between mean number of errors between T1 and T2 are summarized in Table 10. Obviously, the changes are very small, given that there were 35 lexical items being tested. It is also worth noting that for most word classes, the number of errors increased.

Table 8. *Combined perception and production errors (Chinese subjects)*

		T1	T2
Subject 1	Class 1	4	5
	Class 2	4	4
	Class 3	5	4
	Class 7	4	5
	Class 4	3	2
	Class 5	6	7
	Class 6	3	4
Subject 3	Class 1	5	3
	Class 2	3	4
	Class 3	5	1
	Class 7	3	5
	Class 4	1	2
	Class 5	5	5
	Class 6	3	3
Subject 4	Class 1	4	3
	Class 2	0	5
	Class 3	4	2
	Class 7	5	6
	Class 4	3	6
	Class 5	3	5
	Class 6	1	3

One of the characteristics that I had found previously in the interlanguage grammars of the Hungarian, Polish, and Spanish subjects was that they treated different grammatical categories differently when it came to stress assignment. For example, they treated English nouns and verbs differently.

Archibald (1993a) showed how Polish subjects learning English treated nouns and verbs differently when assigning stress. In a class of words represented by *horizon* (nouns with penultimate stress due to a tense vowel in the penult) the most common error made by Polish subjects on all tasks was to stress the initial syllable (i.e. *hórizon*). However, in a class of words represented by *astonish* (verbs with penultimate stress due to a lack of a heavy syllable in the penult), the most common error pattern was to stress the final syllable (i.e. *astonísh*). I argued that the learners had determined that English nouns have final rhymes that are extrametrical (if the final vowel is lax) while English verbs do not. This suggests that the learners are able to consult the notion of grammatical category when assigning stress; they were treating nouns and verbs differently.

Table 9. *Rankings by class (Chinese subjects)*

	C1	C2	C3	C7	C4	C5	C6
Time 1							
Subject 1	4	4	5	4	3	6	3
Subject 3	5	3	5	3	1	5	3
Subject 4	4	0	4	5	3	3	1
Mean	4.3	2.3	4.7	4	2.3	4.7	2.3
Ranking	C5, C3 > C1 > C7 > C2, C4, C6						
Time 2							
Subject 1	5	4	4	5	2	7	4
Subject 3	3	4	1	5	2	5	3
Subject 4	3	5	2	6	6	5	3
Mean	3.7	4.3	2.3	5.3	3.3	5.7	3.3
Ranking	C5 > C7 > C2 > C1 > C6, C4 > C3						

Table 10. *Differential performance between T1 and T2 (Chinese subjects)*

Class 1	−0.6
Class 2	+2.0
Class 3	−2.4
Class 7	+1.3
Class 4	+1.0
Class 5	+1.0
Class 6	+1.0

Similarly, there is evidence that the subjects were treating all verbs as a coherent class. The behavior of the Polish subjects was consistent with a process that could be phrased as *if it's a verb stress the final syllable.* For many of the items tested (i.e. those with a heavy final syllable) this would yield the correct result, as shown in (1):

(2) maintáin, appéar, eráse, decíde, achíeve
collápse, eléct, obsérve, adápt, convínce

But for other items (i.e. those with light final syllables) this would lead to the wrong form, as shown in (3):

(3) astonísh, edít, cancél, considér, interprét

Table 11. *Nouns versus verbs (Chinese subjects)*

	Nouns				Verbs			
	Perception		Production		Perception		Production	
	T1	T2	T1	T2	T1	T2	T1	T2
Subject 1	11	13	6	5	8	9	4	4
Subject 3	10	8	6	5	5	5	4	5
Subject 4	8	12	5	5	2	10	5	4
Total	29	33	17	15	15	24	13	13
Mean	9.7	11	5.7	5	5	8	4.3	4.3

These subjects are able to assign stress with reference to grammatical category (as native speakers do).

The Chinese subjects, though, did not seem to be consulting grammatical category when it came to stress assignment. This distribution is shown in Table 11. T-tests revealed no significant differences,[2] but the subjects were much better at perceiving stress accurately on verbs at both T1 and T2. This may be because two out of the three verbal categories had final stress (e.g. *appear, collapse*), and stress in the final position may be more salient.

The picture that is emerging from all of these results is that the subjects in this study did not seem to be acquiring the principles of English stress assignment with regard to such things as the influence of syllable structure or grammatical category on stress assignment. They seemed to be treating stress as a purely lexical phenomenon; something that has to be memorized as part of the phonological representation of a word. This analysis is supported when we look at the patterns of change from T1 to T2 by lexical item and see how many items stayed the same (whether right or wrong), how many became more nativelike, and how many got worse. This pattern is shown in Table 12.

In both perception and production, the majority of the lexical items (for all subjects) did not change their stress from T1 to T2. On the production task, more lexical items did become more nativelike (3.3) than became less nativelike (2.0). However, on the perception task, more items became less nativelike (7.7) than became more nativelike (3.7). Overall, their perception of English stress was getting worse (though the overall picture is largely influenced by subject 4; the same pattern [to a lesser degree] is found in subject 1).

At first blush, this seemed perplexing. These subjects are all native speakers of Chinese, a tone language. These subjects, I thought, should

Table 12. *Performance over time (Chinese subjects)*

	Perception			Production		
	same	better	worse	same	better	worse
Subject 1	24	4	7	32	2	1
Subject 3	29	4	1	24	5	4
Subject 4	16	3	15	30	3	1
Total (/138)	11	23	86	10	6	
Mean (/35)	23	3.7	7.7	28.7	3.3	2
Mean %	65.7	10.6	22	82	9.4	5.7

be very sensitive to differences in pitch (and, therefore, be able to perceive English stress), as pitch is phonemic in their first language. But I don't think this is necessarily the case. If we turn it around, then we should argue that, because English speakers have movable stress in their L1, and because English stress is manifested (partially) by an increase in pitch, English speakers should be sensitive to differences in tone in Chinese. Anecdotally at least (though see Juffs 1989; Leather 1990) English speakers have a hard time learning to perceive different Chinese tones. It seems likely that the difference between linguistic versus nonlinguistic processing is crucial. Obviously, English speakers have the ability to distinguish differences in pitch when they are presented as nonlinguistic stimuli (e.g. musical notes), but they are not accustomed to doing this when processing linguistic forms. Conversely, my initial expectation that Chinese speakers should be good at perceiving pitch differences in English was probably not taking into account the linguistic processing of English forms. The subjects would probably do quite well on nonlinguistic tests of pitch discrimination. But the fact that stress has not been triggered in their L1 may affect their interlanguage grammars.

The fact that pitch is phonemic in the L1 may shed some light on what is going on. When we think of other aspects of a phonemic representation, say that in Japanese /l/ and /r/ are not phonemic, this is often something that affects cross-linguistic transfer. The learner's initial assumption is that things that are phonemic in the L1 will be phonemic in the L2. This could be what is going on with pitch in these subjects. If pitch (manifested as tone) in the L1 is stored as part of the lexical entry, then the subjects may well be assuming that English pitch (manifested as stress) is also stored as part of the lexical entry.

Now, all of this may explain why their perception is bad but it doesn't say why their perception is getting WORSE. As I noted before, the difference was not significant and may have been strongly influenced by the perfor-

mance of subject 4, who was the subject at the lowest level of proficiency of the four subjects. Nevertheless it is somewhat counterintuitive that the perception scores got so much worse. I might speculate that if these subjects are lexically storing stress rather than computing it, they are less sensitive to the elements that the computational system refers to when building metrical structure (e.g. vowel quality, syllable weight, etc.). It would appear that these subjects are not paying attention to such facts as, reduced vowels tend not to bear stress in English, or heavy syllables do tend to bear stress. They are just storing an accent diacritic as part of the lexical entry. If this is the case, and subject 4 were really just guessing where the stress went when listening to a native speaker say the words, we might expect his performance to fluctuate unpredictably. I would make the following argument. First of all, we note that the subjects' perception of stress is not that good (for reasons I will get to). Therefore they would lexically store an incorrect stress. When listening to words, then, we would expect frequency effects in that subjects would perceive stress to be where it is in their lexical entry. They might be more accurate (but still not very accurate) in perceiving unknown or infrequent words because the perception is not interfered with by automatic lexical access. If they don't utilize the cues for stress (vowel quality, heavy syllables, etc.), then their perception will not improve, and we will see primarily the frequency effects as access to the incorrect lexical item becomes more automatic (harder to override; harder to notice the mismatch between the input and the output), and their perceptual accuracy may actually go down.

Let me turn now to a discussion of the Japanese subject.

4.2. *The Japanese subject*

Japanese is a pitch-accent language, which means that lexical items may store a diacritic mark that indicates where the pitch accent on a word goes. Once the location of the accent is known, then the pitch contour of the word is predictable. Table 13 shows the performance of the Japanese subject.

Table 13. *Perception versus production (Japanese subject)*

	Perception		Production	
	T1	T2	T1	T2
Subject 2	0	1	11	9

We note right away that this subject's perception is very accurate, in contrast with the Chinese subjects', and I will comment on that later.

If we break the above chart down into errors by class (C), the picture shown in Table 14 emerges. Clearly, the differences between word classes were minimal, as were the differences between performance at T1 and T2. As with the Chinese subjects, I combined the production and perception errors to see if class differences would emerge from this view. The result is shown in Table 15.

Again, we note that there was very little difference between word classes and very little change from T1 to T2. This is similar to the Chinese subjects. We turn now to the question of whether the Japanese subject was distinguishing between grammatical categories with reference to rules of stress assignment. The analysis of this subject, shown in Table 16, would suggest that he is not.

Table 14. *Errors by class (Japanese subject)*

Subject 2	Class 1	0	1	1	1
	Class 2	0	0	0	0
	Class 3	0	0	2	1
	Class 7	0	0	4	2
	Class 4	0	0	0	1
	Class 5	0	0	3	3
	Class 6	0	0	1	1

Table 15. *Combined perception and production errors (Japanese subject)*

Subject 2	Class 1	1	2
	Class 2	0	0
	Class 3	2	1
	Class 7	4	2
	Class 4	0	1
	Class 5	3	3
	Class 6	1	1

Table 16. *Nouns versus verbs (Japanese subject)*

	Nouns				Verbs			
	perception		production		perception		production	
	T1	T2	T1	T2	T1	T2	T1	T2
Subject 2	0	1	7	4	0	0	4	5

For this subject, both production and perception showed very little change over time. It may, of course, be that I just wasn't leaving enough time between probes. Given more time, we may see a change.

The Japanese subject too, then, seems to be treating stress as a purely lexical phenomenon. This analysis is supported when we look at the patterns of change from T1 to T2 by lexical item and see how many items stayed the same (whether right or wrong), how many became more nativelike, and how many got worse. This pattern is shown in Table 17.

The Japanese subject seems to be much more successful in his perception of English stress than the Chinese subjects. In terms of his production, he did not appear to be significantly different from the Chinese-speaking subjects. I suspect that the reason this subject is better at perceiving English stress has something to do with the structure of Japanese. Japanese is a mora-counting language that makes the distinction between light and heavy syllables. Furthermore, bimoraic syllables have greater duration than monomoraic syllables. As a result, the Japanese subject would be sensitive to these kinds of durational cues. He may well be utilizing these skills to perceive the longer duration of an English stressed vowel. The Chinese subjects, on the other hand, do not have length contrasts in their L1 and, as a result, may not be sensitive to perceiving the length distinctions that can cue stress in English.

5. Conclusion

The results of this paper need to be considered in the wider context of what other nonnative speakers have been shown to do when acquiring English stress. The Chinese and Japanese subjects didn't greatly change their stress patterns over time and didn't appear to be basing their stress assignment on things like grammatical category or syllable weight. When we contrast this study with the studies that have been done on native speakers of Polish, Spanish, and Hungarian (other stress-accent languages), we can see that we are actually learning something about the influence of the L1 when it's a nonaccentual language. The subjects in

Table 17. *Performance over time (Japanese subject)*

	Perception			Production		
	same	better	worse	same	better	worse
Subject 2	34	0	1	27	5	3
Mean %	97.1	0	2.9	77.1	14.3	8.6

this case appear to be treating stress as a lexical phenomenon. Subjects whose L1s are accentual languages were transferring their L1 principles and parameters of metrical structure (e.g. quantity-sensitivity, extrametricality). Subjects whose L1s were nonaccentual languages appear to be transferring quite different things.

This seems to be analogous to Carroll's (1989) study of the acquisition of gender by French-immersion students. She argued that nonnative speakers of French were representing gender in a manner that was fundamentally different from native speakers. At times this could produce behavior that was almost indistinguishable from native speakers, but the representation was thought to be different. In Archibald (i.p.) I argued that this suggests that adult L2 learners can reset existing parameters but may not be able to trigger new structures (see also Archibald et al. i.p.). This appears to be parallel to what learners from a nonaccentual language are doing when they are trying to learn an accentual language. Often they are getting the stress correct on the English words, but they seem to be doing it in a way that is very different from native speakers.

Received 21 November 1995
Revised version received
28 June 1996

University of Calgary

Notes

* I would like to thank Mount Royal College for allowing me access to their students. I would also like to thank an anonymous *Linguistics* reviewer for helpful comments. Correspondence address: Department of Linguistics, University of Calgary, Calgary, Alberta T2N 1N4, Canada.

1. T1 Prod/T1 Perc: ..2149; T2 Perc/T2 Prod: ..1012; T1 Perc/T2 Perc: ..401; T1 Prod/T2 Prod: ..1835.
2. N T1 Perc/V T1 Perc: .0339; N T2 Perc/V T2 Perc: .0351; N T1 Prod/V T1 Prod: .1835; N T2 Prod/V T2 Prod: .1835.

References

Archibald, John (1993a). *Language Learnability and L2 Phonology: The Acquisition of Metrical Parameters*. Dordrecht: Kluwer.

—(1993b). The learnability of metrical parameters by adult speakers of Spanish. *International Review of Applied Linguistics* 31, 129–142.

—(i.p.). Metrical parameters and lexical dependency. In *Generative Approaches to Second Language Acquisition*, S. Flynn, G. Martohardjono and W. O'Neil (eds.). Hillsdale, NJ: Erlbaum.

—; Guilfoyle, Eithne; and Ritter, Elizabeth, (i.p.). Functional categories in L2 acquisition: evidence of presence is not necessarily presence of evidence: a response to Epstein, Flynn and Martohardjono. *Behavioral and Brain Sciences.*

Beckman, Mary (1986). *Stress and Non-Stress Accent.* Dordrecht: Foris.

Carroll, Susanne (1989). Second language acquisition and the computational paradigm. *Language Learning* 39(4), 535–594.

Dresher, Elan; and Kaye, Jonathan (1990). A computational learning model for metrical phonology. *Cognition* 34, 137–195.

Hayes, Bruce (1995). *Metrical Stress Theory.* Chicago: University of Chicago Press.

Hyman, Larry (1978). Tone and/or accent. In *Elements of Tone, Stress, and Intonation,* Donna Napoli (ed.), 1–20. Washington, D.C.: Georgetown University Press.

Juffs, Alan (1989). Tone, syllable structure and interlanguage phonology: Chinese learners' stress errors. *International Review of Applied Linguistics* 27(2), 99–118.

Leather, Jonathan (1990). Perceptual and productive learning of Chinese lexical tone by Dutch and English learners. In *New Sounds 90,* J. Leather and A. James (eds.), 72–97. Amsterdam: University of Amsterdam Press.

Pulleyblank, Douglas (1986). *Tone in Lexical Phonology.* Dordrecht: Reidel.

van der Hulst, Harry; and Smith, Norval (eds.) (1988). *Autosegmental Studies on Pitch Accent.* Dordrecht: Foris.