4   Phonology of AAE

Focal point The different sound patterns of AAE are reflected in the combination of sounds that are allowed in the variety. Words in AAE and general American English that have the same meanings may have different pronunciations due to constraints on sounds. For example, in AAE the *th* sound in general American English (and other varieties of English) *bath* is produced as *f* (i.e., *baf*), and the *th* sound in general American English *bathe* is produced as *v* (i.e., *bav*). Also, as explained in the preceding chapter, yes-no questions can be produced without an initial auxiliary. Does this affect the rhythmic patterns of sentences in AAE such that they are different from those patterns in other varieties of English? This chapter addresses sound and rhythmic patterns in AAE.

It's sofer tissue!

[facial tissues television commercial]

4.1   Introduction

The unique pronunciation related to stress patterns of aspectual markers *BIN* and *dən* has been mentioned in the preceding chapters on the lexicon and syntax of AAE. Especially in the case of *BIN*, the stress pattern affects meaning. While it would seem that these patterns are significant in the study of AAE, they have not been the focus of major studies on the sound system.

The most thorough studies of phonology or sound patterns in AAE were conducted in the 1960s and 1970s (Wolfram 1969, Wolfram and Fasold 1974, Luelsdorff 1973). These studies discuss the production of word final sounds (e.g., *test* and *kind* are pronounced as *tes* and *kin*, respectively), the production of *f* in environments in which other varieties of American English produce *th* (e.g. *baf/bath*) and the vocalization of *r* and *l*, the process which results in the homophony of *court* and *coat* in some regions in the United States. A substantial amount of this research was cast in the framework of Variation Theory, in which accounts were given of the inherent variability in language.[1] In his study of the speech of African Americans in Detroit, Wolfram (1969) showed the importance of the correlation between the use of sound patterns such as

106

the ones mentioned above and extralinguistic factors such as age, socioeconomic class and sex.

It has been noted that in different varieties of American English (including registers of standard American English), word final consonant groups such as *nd* and *st* are reduced to single consonants *n* and *s*, respectively. That is, some speakers of varieties of English may produce *kind* as *kin* (rhymes with *pine*) in certain environments, so both the AAE speaker and the speaker of another variety of general American English may at some time or another produce *kin*. Speakers may differ by using *kin* in varying environments, and they may also differ in the extent to which they reduce final consonant groups to single consonant sounds. That is, the percentage rate of reduction may be greater for some speakers than for others. It is probably this feature of consonants that has received the most attention in the phonological studies of AAE, and it has been used in comparing AAE to other varieties of American English and to West African languages. On the one hand, studies have been used to argue that the patterns of final consonants in AAE parallel those in other varieties of English. On the other hand, it has been argued that AAE may resemble West African languages that do not have final consonant clusters.

This chapter presents a descriptive summary of the system of sounds in AAE. The description will focus on the systematic nature of sound combinations and the systematic way in which sounds are produced in certain environments. Along these lines, one of the goals is to show how AAE follows rules in sound combinations. In addition to considering single segments such as consonant sounds, this chapter reports on rhythmic patterns of words and sentences in AAE. Here a general overview of research in this area, as well as further directions that related studies can take, will be presented. Generally speaking, only the surface of research on rhythmic patterns in AAE has been scratched, so study in this area offers great possibilities for contributions to the description of the language system.

4.2   Final consonant sounds

Analyses of final consonant groups or clusters in AAE have been used as evidence that the language variety is systematic and governed by rules, and they have also been used to support claims about the historical origin of AAE. Two explanations for cluster data in AAE can be summarized as follows. One explanation is that pronunciations such as *tes* ('test'), *des* ('desk') and *han* ('hand') in which the final consonant clusters *st*, *sk* and *nd* are pronounced as *s*, *s* and *n*, respectively, are a result of a process called consonant cluster reduction. A second explanation is that words such as *test*, *desk* and *hand* are pronounced as *tes*, *des* and *han*, respectively, because AAE, like African languages from which it descended, does not have final consonant clusters. In presenting a general description of the process operating on final clusters, I comment on these explanations.

Consonant cluster reduction is a process in which the final consonant group or cluster, composed of two consonant sounds, is reduced to a single consonant sound. This is illustrated in (1) below, where 'C' stands for consonant.

(1)   CC→C
      || |
      nd→n
      kind→kin

In (1) the consonant cluster *nd* (CC) is reduced to *n* (C), so the word *kind* is produced without the final *d* sound (i.e., *kin*). Under this analysis, the clusters *st* in *test* and *sk* in *desk* are reduced to the consonant sound *s*. As a result of the consonant cluster process, the words *tes* ('test') and *des* ('desk') rhyme, and are minimally different in that they contrast only in the initial *t* and *d* sounds. One property of the consonant cluster reduction analysis is that it assumes, at some level in AAE, final clusters are intact, so *tes* actually starts out as *test*, and the final *t* is deleted under certain conditions. This analysis treats a word such as *test* in AAE as being identical to *test* in mainstream English. If there are differences in the pronunciations, such that in mainstream English the word is *test* in specific environments and in AAE it is pronounced as *tes*, these different pronunciations are the result of phonological processes, and not the result of different representations of the words in the two varieties. For example, the word in AAE is *test*, but it is pronounced as *tes* because a rule, as shown in (1), reduces the final *st* cluster to *s* by deleting the final *t*. AAE, then, is just like mainstream English in this respect; the only difference is that the reduction rule applies more often to words in AAE in specific environments when certain conditions are met.

The second account of the production of words such as *tes* ('test'), *kin* ('kind') and *contac* ('contact') is an analysis which leans toward West African origins and postulates that speakers have such pronunciations, not because the final consonant sound is deleted in some environments, but because the languages from which AAE originated do not have final consonant clusters (so there are no clusters like *nd* and *st*). Ernie Smith (1998), who takes the Africanist approach, makes the following remarks:

For example, the scholars who view African-American speech as a dialect of English describe the absent final consonant cluster as being "lost," "weakened," "simplified," "deleted," or "omitted" consonant phoneme. But viewed as an Africanist Language system that has adopted European words, African-American speech is described by Africologists as having retained the canonical form, or shape, of the syllable structure of the Niger-Congo African languages. Thus, in Ebonics homogeneous consonant clusters tend not to occur. This is not because the final phoneme has been "lost," "weakened," "simplified," "deleted," or "omitted," but because *they never existed in the first place*. Hence it is by relexification (that is, 'the replacement of a vocabulary item in a language with a word from another, without a change in the grammar,' – see Dillard 1972) that in Ebonics English words such as *west*, *best*, *test*, *last* and *fast* become *wes*, *bes*, *las* and *fas*; the words *land*, *band*, *sand* and *hand* become *lan*, *ban*, *san* and *han*; the words *left*, *lift*, *drift* and *swift* become *lef*, *lif*, *drif* and *swif* – and so forth.

[p. 56]

Under the African origin view, words such as *test* start out with single final consonant sounds, not with the final consonant cluster (e.g., The form *tes* does not start off with a final *st*). So according to this view, when speakers actually produce *test* in certain environments, such as that preceding a word that begins with a vowel, they actually add

the final *t* to the word. Let's illustrate with an example. In the sentence *I will test out of math*, it is argued that the word *test* is otherwise pronounced as *tes*, but the *t* is added because the word precedes *out*, which begins with a vowel. The difference is that in Smith's view, in AAE, words such as *test* and *kind* start out with single final consonant sounds (i.e., *tes* and *kin*), while in general American or mainstream English, the word starts off with final consonant clusters (i.e., *test*, *kind*). This view presents AAE as patterning phonologically with West African languages, as summarized in Smith's quote.

Having presented a general overview of two accounts of final consonant sounds, I want to give a description of consonant cluster phenomena which focuses on the environments in which the reduction process is said to occur and a summary of the types of analyses that have been offered to account for the reduction process. Pronunciations such as the following surface in AAE with greater or less frequency in certain environments. Greater or lesser frequency refers to cases such that some speakers may produce the words as they are given in (2) – without final clusters – more often in certain environments and less often in others. Some speakers may always use the pronunciations in (2), however. In all of the words below, *s* is pronounced instead of the clusters *st*, *sp* and *sk*, *f* is pronounced instead of the cluster *ft*, *p* instead of *pt* and *k* instead of *kt*. Finally, *n* is pronounced instead of *nd*, and *l* is pronounced instead of *ld*:

(2) | **AAE** | **Phonetic transcription** | |
|---|---|---|
| a. pos | [pos] | 'post' |
| b. was | [wɔs], [was] | 'wasp' |
| c. mas | [mæs] | 'mask' |
| d. gif | [gɪf] | 'gift' |
| e. adop | [ədap] | 'adopt' |
| f. conduc | [kandʌk] | 'conduct' |
| g. ban | [bæn] | 'band' |
| h. bol | [bol] | 'bold' |

Key: The phonetic transcription is given for each word as a means of providing a more accurate representation of the pronunciation.[2] The vowel sounds used by AAE speakers in different regions of the United States may vary. These vowel sounds and the ones presented throughout this chapter are primarily based on the vowel sounds I have heard all my life in southwestern Louisiana.

In studying final consonant sounds, we cannot just consider words in isolation; they should also be analyzed in different environments: whether they are produced in careful or casual speech, whether the final consonant precedes a suffix that begins with a consonant or vowel and whether the final consonant precedes a word that starts with a consonant or vowel. These environments are important because speakers might be expected to retain clusters more often when they are using more careful speech. Labov (1972) discusses the effect of such conditioning factors. Also, Guy (1991) presents a detailed analysis of conditioning factors for consonant cluster reduction.

In addition to the pronunciations in (2), those in (3) also occur with varying degrees of frequency. These words, unlike those in (2) are composed of two minimal units of meaning or two morphemes:

(3)   **AAE**   **Phonetic transcription**

    a. push   [pʊʃ]   'pushed'

    b. page   [pedʒ]   'paged'

    c. raise   [rez]   'raised'

    d. pick   [pɪk]   'picked'

    e. jump   [dʒʌmp]   'jumped'

    f. miss   [mɪs]   'missed'

In the general American standard variety, the morphemes in (3d) are *pick* and *ed*. In the AAE example in the appropriate context, *pick* is interpreted as if it is in the past tense. It would be understood as past in a sentence such as *Yesterday, she pick me to be on the team*. The adverb *yesterday* indicates past, so that information does not have to be redundantly stated on the verb *pick*. But in the sentence *They pick me*, we would need further contextual information to determine whether the past or non-past interpretation is intended because there are no cues within the sentence. Labov (1969b) reports that some speakers do not perceive the final *-ed* (pronounced as *t* in *picked*), in sentences such as the following:[3]

(4)   a. He pick me.

    b. I've pass my test.

    c. Last week I kick Donald in the mouth, so the teacher threw me out the class. (p. 57).

In classroom correction tests in which AAE speakers were given sentences in (4), Labov notes that the speakers were not able to change the sentences into classroom English by adding *-ed* to *pick*, *pass* and *kick*. This result led him to conclude that his speakers "have little ability to detect the absence of *-ed* as a grammatical element to be corrected" (p. 57).

If you think about the general English pronunciation of the word *missed* (3f), you will realize that it sounds just like *mist*, as if it ends in *t* (although the ending is spelled with *-ed*). The final *st* sound in *missed* forms the cluster that is reduced to *s* (as in *miss* in AAE). As a result of the cluster reduction process, *missed* and *mist* will be homophonous; they will sound the same (as *miss*).

Generalizations about the data in (2) and (3) will be helpful in explaining word final consonant patterns in the phonology of AAE, and they will also be helpful in characterizing groups of sounds that are involved in the process. The cluster reduction phenomenon can be captured by a formal description that will explain where it occurs and what happens as a result of it. One particular analysis that has been used to account for the patterns in the words above can be called the *voicing generalization*. It states that a final consonant cluster (e.g., *-st* as in *mist* or the pronunciation of the final sound in *missed*) is reduced by deleting the final member if the two consonants forming the cluster have the same voicing value, in which both are [+voice] (voiced) or both are [−voice] (voiceless). Voicing is an articulatory description that is used in characterizing consonant sounds. The voicing feature is used to characterize the state of the glottis, the opening between the vocal folds. When the vocal folds – flaps covering the

glottis – are apart, the air flows freely through, producing voiceless sounds (such as *p*, *t*, *k*, *s* and *f* sounds). On the other hand, when the vocal folds are together, the air forces its way through, causing vibration and subsequently producing voiced sounds (such as sounds *b*, *d*, *g*, *z*, *v*, *l*, *m*, *n*, and ŋ [representation for the *ng* sound at the end of *sing*]). (See the diagram below from McManis, Stollenwerk and Zheng-Sheng 1989.) Reprinted with permission of Ohio State University Press.

Larynx: state of the glottis

According to the voicing generalization, clusters such as *st* in *fast* and *nd* in *pound* are predicted to be reduced to *s* and *n*, respectively, because the consonants forming the clusters have the same voicing value, whereas clusters such as *nt* in *paint* and *mp* in *jump* are predicted to remain intact because the consonants forming the clusters have different voicing values. This is shown in (5), where [−v] is an abbreviation for [−voice] (voiceless) and [+v] is an abbreviation for [+voice] (voiced):

(5) a. f   a   s   t   →   fas
              | |
             [−v] [−v]
     Both *s* and *t* are [−voice].
     b. j   u   m   p   →   jump
              | |
             [+v] [−v]
     m is [+voice] and *p* is [−voice].

Although they did not call it the *voicing generalization* and give it a formal characterization, Wolfram and Fasold (1974) noted this pattern of final consonant clusters in AAE.[4]

This analysis assumes that in AAE there are underlying consonant clusters, so a word such as *fast* starts off with the cluster (*st*) intact, but the cluster is reduced by deleting the final consonant (*t*) of the cluster because it has the same voicing value as the preceding consonant (*s*). This point was made above in the description of the voicing

generalization (5a). Some arguments have been offered to show that this is a valid way of looking at the data and that AAE should be analyzed as having underlying consonant clusters (i.e., words in AAE actually end with final consonant clusters or groups) that are reduced when the proper conditions are met. One argument in support of this view is carried out along the following lines: We know that AAE has full final consonant clusters because when suffixes that begin with vowels (e.g. *-able*) are added to words such as *accept* to yield *acceptable*, the full cluster (*pt*) is pronounced. For some, this means that in AAE, these words must have final consonant clusters in the first place. How else would speakers pronounce the *pt* cluster? The argument goes on to say that the interaction between suffixes and consonant clusters is important. The generalization is that the cluster remains intact preceding suffixes that begin with vowels, but when clusters precede suffixes that begin with consonants (e.g., *-ly* in *friendly*), they are more likely to be reduced (*frienly*). The argument, then, is that AAE has 'underlying' clusters, but whether or not the clusters are pronounced depends on a number of factors. One is whether the cluster precedes a vowel initial or consonant initial suffix.

Now consider the examples below, which illustrate what happens when suffixes are added to words:

(6) **Consonant cluster preceding vowel initial suffix**   **Word final**
    a. accept+able→accept**able**   accept→accep
    b. expect+able→expect**able**   expect→expec

(7) **Consonant cluster preceding consonant initial suffix**   **Word final**
    a. friend+ly→frien**ly**   friend→frien
    b. soft+ness→sof**ness**   soft→sof

The words in (7) are included to provide representative examples of clusters preceding consonant initial suffixes. In (7) the *nd* cluster is reduced preceding a consonant initial suffix and at the end of the word. The major focus is on vowel initial suffixes and consonant clusters, so let's return to (6). In (6a) the *pt* cluster is retained when it precedes the vowel initial suffix (*-able*) but reduced word finally (*accep*), and the same pattern occurs with the *ct* (but formally represented as *kt*) cluster in (6b). Data collected in (Green 1991) provide evidence that the analysis is slightly more complicated in that not all vowel initial suffixes have the same effect on preceding clusters. That is, not all vowel initial suffixes are equal or behave the same. In some cases, the consonant cluster preceding vowel initial suffixes *-er* (8a) and *-ing* (8b) can also be reduced:

(8)   a. cold+er→col+er→coler
        b. spend+ing→spen+ing→spening

In the example in (6), the generalization is that the final cluster is more likely to be retained when it precedes a vowel initial suffix (such as *-able*) than when it precedes a consonant or ends a word, but the cluster is not necessarily retained when it precedes a vowel initial suffix. In fact, it may very well be reduced if the vowel initial suffix is *-er* or *-ing*, yielding the results in (8). The clusters preceding *-er* and *-ing* may also be

retained, as in *colder* and *spending*. This means that both pronunciations *colder*, *coler* and *spending*, *spening* occur in AAE. It is not uncommon to hear speakers use the word in (8b) in the following way: *She don't mind spening a lot of money on a good pair of shoes*. The pronunciation, *spening*, is virtually indistinguishable from *spinning*, the word that is used to describe the action of a top or wheel in motion. On the one hand, *spening* is like *spinning* because the *-nd* cluster does not obligatorily occur before *-ing*. On the other, the vowel sounds in the words are pronounced the same; just as some speakers do not distinguish the vowel sounds in *pen* and *pin*, they do not distinguish them in *spend* and *spin*.

Now consider the quotation at the beginning of this chapter, in which the speaker in a television commercial of a famous brand of facial tissues presents a clear example of cluster reduction preceding the vowel initial suffix (*-er*) when he says, "It's sofer tissue." He does not produce the *ft* cluster preceding the vowel initial suffix *-er* in *softer*; only the *f* of the cluster is pronounced. Contrary to views in earlier research, the behavior of consonant clusters preceding vowel initial suffixes does not provide hard and fast evidence that words in AAE have clusters such as those in general American English. This is related to the point that Smith was trying to make about the relation between AAE and Niger-Congo languages syllable structure although he did not address suffixes. Different types of suffixes have different effects on consonant clusters.

There is also evidence that this type of reduction occurs with the *ct* (i.e., *kt*) clusters preceding the vowel initial suffix *-ing* in *acting*. When the word *acting* means performing in the strict sense, as in the case of someone on television or in a play, AAE speakers can say of the person: "She sure did a good job acting her part" (where the *ct* cluster preceding *-ing* remains intact). On the other hand, if *acting* is used to refer to behavior, it can be produced with a reduced cluster: "Stop acking like that" or "Stop acking silly," (in which the *ct* cluster preceding the *-ing* suffix is reduced).

The variation in occurrence of consonant clusters preceding suffixes *-er* and *-ing* and the invariability of retention preceding *-able* seem to be due, in part, to the type of suffix. It has been noted, in some studies in phonology, that *-er* and *-ing* are one type of suffix, and *-able* is defined as a different type. A discussion of suffix types goes beyond the scope of this chapter, but the discussion and references in Selkirk (1982) address related issues in English.

In summary, clusters preceding vowel initial suffixes *-er* and *-ing* may be reduced, but clusters preceding the vowel initial *-able* are more likely to be retained. So far, a number of pieces fit together in the explanation of consonant cluster reduction. Let's look at them as we explain *spening*. We start off with the word *spending* and given the voicing generalization, the cluster (*nd*) is reduced (to *n*) although it precedes a vowel initial suffix (*-ing*). It is interesting to note that whatever the difference between vowel initial suffixes *-ing* and *-er* on the one hand and *-able* on the other, it is reflected with consonant clusters.
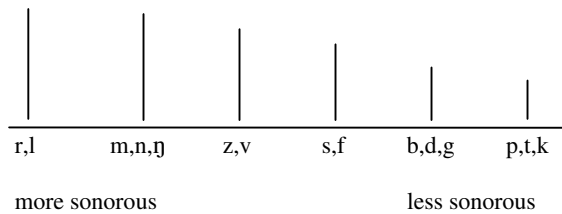
To some, cluster reduction in words such as *pos* ('post'), *was* ('wasp') and *mas* ('mask') yields words that are only consistent with unintelligent speech, and to them,

the pronunciation of the plural forms of these words (*posts*, *wasps*, *masks*) makes even less sense: *poses* [posɪz], *wases* [wɔɪz]/[wasɪz] and *mases* [mæsɪz], respectively. (The pronunciation of the vowel sound in *wases* may vary from region to region. The phonetic transcriptions are also given, so refer to the key in note 2.) But these pronunciations are systematic; the singular forms can be explained by reduction and the voicing generalization, and the plural forms are also governed by rules. It is true that the singular/plural forms *pos*/*poses*, *was*/*wases* and *mas*/*mases* are not the mainstream English pronunciations; nevertheless, these singular/plural pairs are based on clear patterns of sound combinations. I am not aware of current studies on the regular production of these plurals in other varieties of American English although Miller (1986) reports on the three examples of *wases* in *The Linguistic Atlas of the Middle and South Atlantic States* by whites. In order to understand the rule that describes the way the plural is formed, we have to understand that all of the singular forms end in *s* for one reason or another, either because the voicing generalization results in consonant cluster reduction or because the words were adopted into AAE without final consonant clusters, that is with just final *s*. Given that the singular forms *pos*, *was* and *mas* are pronounced as if they end in *s* instead of the final consonant clusters *st*, *sp* and *sk*, respectively, their plural endings are identical to plural endings of words that end in -*s*. That is to say that the plurals of *pos*, *was* and *mas* are formed in the same way that the plurals of *dose*, *boss* and *pass* are formed. In American English (including AAE), words that end in the *s* sound form the plural by adding -*es* (pronounced [əz] or [ɪz]). This explains why the plural for *boss* is *bosses* and *pos* ('post') is *poses*. The plural ending for both words is formed with -*es*. If the word *pos* in AAE actually ended in an *st* cluster, the plural would not likely be formed by adding -*es*.[5]

One speaker, an African American male in his 30s, was heard saying, *We having a lot of conteses at work* ('We're having a lot of contests at work'), where *conteses* [kantɛsɪz] is the plural form for general English *contests*. The point here is that, at first glance, some plural forms and other sound patterns may appear to be random, but linguistic analysis provides evidence that speakers are adhering to plural formation and other rules. The plural of words that are spelled with a final *st*, *sk* or *sp* cluster but pronounced with final -*s* (not the final cluster) is formed by adding -*es*, the same way plurals of other words that end in -*s* are formed.

The voicing generalization has been proposed to account for reduced clusters, but there is another possible explanation for the results that are shown in examples such as those in (2). Whereas the voicing generalization is based on the voicing value of word final consonants (e.g., *s* and *t* have the same voicing value), an alternative explanation is based on the intensity or loudness of sounds, a property referred to as sonority. Some sounds (such as *n* and *m*) are more sonorous or intense than other sounds (such as *d* and *b*). Briefly, the analysis that is based on sonority provides an account of the data in (2) by placing the consonants forming the clusters on a sonority scale and explaining that some clusters (such as *st* and *nd*) do not surface because the consonants forming them are too close in sonority. Consonant sounds are arranged from more sonorous to less sonorous on the sonority scale in (9):

(9)   Sonority scale

| r,l | m,n,ŋ | z,v | s,f | b,d,g | p,t,k |

more sonorous   less sonorous

I will not go into too much detail here, but I will give a general idea about how the scale works. Let's consider the final consonant clusters *st* (in *fast*) and *nd* (in *kind*), which are reduced to *s* (*fas*) and *n* (*kin*), respectively. According to the sonority scale, they are reduced because the consonants forming the clusters are too close in sonority. That is, the *s* is too close to the *t* (*st* cluster), and the *n* is too close to the *d* (*nd* cluster) in sonority. The closeness results in consonant cluster reduction. Now, consider the *nk* cluster (i.e., ŋk) as in *think* and *pink* and the *nt* cluster as in *paint*, which are not reduced. (Note: The final *nk* cluster in *pink* is pronounced as if it contains a *g*, *ngk*. If you pronounce the word slowly, you will hear *ng* (i.e., final sound in *sing*) before you produce *k*.) These clusters are not reduced because there is no conflict in sonority between ŋ and *k* and *n* and *t*; the consonants are far enough apart in sonority. A more accurate statement of the sonority analysis would also spell out what it means to be too close and far enough apart in sonority. Both the voicing generalization and the sonority analysis are ways of formally characterizing consonant cluster reduction, ways of making a formal statement about the rules that speakers use in producing combinations of sounds. Careful consideration of these consonant patterns will help to give us more insight into the AAE sound system.

In the previous discussion of the voicing generalization and sonority scale, it was noted that the *nt* cluster in words such as *paint* and *pint* remains intact. In auxiliaries, the *n't* cluster in *ain't*, *can't* and *don't* are not usually fully produced, so the words are pronounced as *ain'*, *cain'* and *don'*, without the final *t*. This is not unusual for function words such as auxiliaries. Note, also, that the initial consonants, *d* and *g*, are not obligatorily produced when *don't* and *am gonna* are used in specific environments. *I don't know* may very well be produced as *I'on know*.[6]

We have seen contexts in which the cluster is reduced such that the final *t* or *d* does not occur, for example. There are also some words in which the opposite process occurs; *t* or *d* is inserted. This is what happens with words such as *light skinned*, *big boned*, and *two faced*. In AAE some speakers produce these words as *light skinded* [skɪndɪd], *big boneded* [bondɪd], and *two faceted* [festɪd], respectively. In each word, a *t* or *d* sound occurs right before the *-ed* [ɪd] suffix, and the result is that the word is pronounced with an additional syllable. The final syllables in *skinded* in *light skinded* and *boneded* in *big boneded* sound like the last syllable in *candid*, and the last syllable in *faceted* in *two faceted* is like the last syllable in *rested*. In a sense, speakers treat the words *skin*, *bone*, and *face* as if they actually end in a *t* or *d* sound, in the way the word *bond* ends in a *d* sound and *rest* ends in a *t* sound.

When we study the words carefully, we find that speakers do not randomly insert *t* or *d* in these words; the sounds preceding the inserted *t* or *d* determine which sound (*t* or *d*) will be inserted. If the word ends in a voiced sound, as *n* in *bone* and *skin*, then *d* is inserted, but if the word ends in a voiceless sound, as the final *s* sound in *face*, *t* is inserted.[7] (Note: The end of the word *face* is spelled *ce*, but the final sound is formally or phonetically represented as *s*.) So far, we have seen that the voicing property of sounds plays a major role in the way words are pronounced. The voicing properties affect consonant cluster reduction and *t/d* insertion in some words.

4.3   Devoicing

The process of consonant devoicing, that is, making a voiced consonant voiceless, applies to some consonants at the ends of words. Words ending in the voiced *b*, *d* and *g* that undergo this process are pronounced as if they end in *p*, *t* and *k*, respectively. This process results in the pronunciation of *cab*, *feed* and *pig* as *cap*, *feet* and *pick*, respectively. In such cases, the final voiced consonant is devoiced, that is, made voiceless:

(10)   *b*   →   *p*
        voiced   voiceless
        d   →   *t*
        voiced   voiceless
        g   →   *k*
        voiced   voiceless

Baran and Seymour (1976), in early research on AAE and communication disorders, investigated children's performance in differentiating minimal word pairs in which the final consonant sound of one word was voiced and the final consonant sound of the other word in the pair was voiceless, as in *cab* and *cap*, respectively. They found that whites listening to blacks made the most errors in distinguishing the words in the pairs and that "blacks understood themselves better than they understood either other blacks or whites" (p. 470). More specifically, Baran and Seymour found that when blacks listened to these pairs produced by whites, they mistook words ending in a voiceless consonant for its counterpart ending in a voiced consonant. They explain that

Black children had little difficulty in matching the Standard English production of *pig* with its corresponding picture, but were less certain on this task for *pick*. Since the voiceless /k/ may be used in the production of both *pig* and *pick* in Black English, standard usage of /k/ by white children seemed to create more listener confusion than did /g/.

[p. 472]

The research by Baran and Seymour shows how final consonant devoicing can result in ambiguity. This type of information may be useful for educators teaching AAE speaking students who use this devoicing strategy.

The following section continues the discussion of the AAE sound system by providing a description of sound patterns that are related to *th* sounds in different parts of words.

4.4   Sound patterns and th

Another well-known pattern in the sound system of AAE is the production of *t/d* and *f/v* in environments in words in which the *th* sound occurs in general American English. The occurrence of *t/d* and f/*v* in certain environments is systematic in the sense that each sound is chosen based on phonetic properties; these sounds are not used randomly. To explain the distribution of these sounds, it is necessary to give the articulatory descriptions of them. The articulatory description of a consonant sound includes three different types of information that are used to identify or characterize them:

(a) state of the glottis in the production of the consonant (voicing, section 4.2.)
(b) place of articulation of the consonant
(c) manner of articulation of the consonant

The *t/d* and *f/v* will be described by providing the information in (a)–(c) for each sound. In defining the state of the glottis, we determine whether the sound is voiced or voiceless. The place of articulation describes the place in the mouth at which the sound is made, and the manner of articulation identifies the way the sound is made, for example, whether air is stopped at a particular point in the mouth. This information is given below for the sounds *t*, *d*, *f*, and *v*.

(11)

Sound	State of glottis	Place of articulation	Manner of articulation
t	vocal folds apart, voiceless	alveolar	stop
d	vocal folds together, voiced	alveolar	stop
f	vocal folds apart, voiceless	labiodental	fricative
v	vocal folds together, voiced	labiodental	fricative

The alveolar is the place of articulation for the *t* and *d* sounds because they are made by touching the ridge behind the top teeth (alveolar ridge) with the tip of the tongue, and labiodental is the place of articulation for *f* and *v* because they are formed by touching the bottom lip with the top teeth (i.e., labiodental, lip and teeth). The *t* and *d* are stops because the air is stopped at the alveolar ridge when they are produced, and the f and *v* are fricatives because, when these sounds are made, friction is produced, as the air escapes through the very small opening between the lip and teeth. (For an introductory discussion of the articulatory description of consonants, see Fromkin and Rodman [1998] and Ladefoged [1993].)

The *t* and *d* on the one hand and the *f* and *v* on the other differ from each other only in one respect: state of the glottis. The *t* is voiceless and the *d* is voiced, and

likewise, the *f* is voiceless and the *v* is voiced. This distinction is very important, and it will be shown that speakers of AAE actually make it, as do speakers of other varieties of English and other languages. Before this distinction is discussed, it is necessary, however, to give the description of *th*, as it has been noted that *t/d* and *f/v* occur in the same environments in which *th* sounds occur in general American English. There are actually two *th* sounds, but they are both represented by one spelling *th*. Incidentally, the dictionary distinguishes these sounds by underlining the *th* (**th**) in words such as *then* and by not underlining the *th* in words such as *thin*. Can you figure out what distinguishes the two *th* sounds? Two different phonetic symbols are used to refer to these sounds.

(12)

Sound	State of glottis	Place of articulation	Manner of articulation
th (θ) (*thin*)	voiceless	interdental	fricative
th (ð) (*then*)	voiced	interdental	fricative

The *th* sounds are interdental because they are made at the point where the tongue is between the teeth, and they are fricatives because friction is caused when the air escapes through the small opening between the tongue and teeth. The *th* sound represented by /θ/ is voiceless, and the *th* sound represented by /ð/ is voiced. Voicing (i.e., state of the glottis) is the only difference between the two sounds, but it is a very important distinction.

The articulatory descriptions in (11) and (12) will be helpful is explaining the patterns that we find in the production of *t/d* and *f/v* in AAE. Consider the list below:

(13)

	AAE	Phonetic transcription	
a.	thing	[θɪŋ]	'thing'
b.	think	[θɪŋk]	'think'
c.	dese	[diz]	'these'
d.	dat	[dæt]	'that'
e.	baf	[bæf]	'bath'
f.	wif, wit	[wɪf], [wɪt]	'with'
g.	mont, monf	[mʌnt], [mʌnf]	'month'
h.	Beflehem	[bɛfləhɪm]	'Bethlehem'
i.	bave	[bev]	'bathe'
j.	smoove	[smuv]	'smooth'
k.	mova[8]	[mʌvə]	'mother'

The words in (13) show that the *t* and *d* sounds occur at the end (*wit*, *mont*) and beginning (*dese*) of words, but the *f* and *v* sounds occur in the medial position (*Beflehem*, *mova*) and at the end of words (*baf*, *bave*). Given the data in (13), the three generalizations are as follows:

Generalization 1: These sounds (*t*, *d*, *f*, *v*) occur in positions in which the *th* (i.e., [θ], [ð]) sounds occur in general American English.

Generalization 2: The voiceless sounds *t* and *f* occur in medial and final environments of words in which the voiceless *th* sound occurs in general American English. This

means that words ending in voiceless *th* in mainstream English (e.g., *baf*) will end with a voiceless *t* or *f* in AAE. But note that the voiceless *th* sound is pronounced at the beginning of words (12a, b).

Generalization 3: The voiced sounds *d* and *v* occur in the same environments in which the voiced *th* sound occurs in general American English. This means that words ending in voiced *th* in general American English will end with voiced *d* or *v* in AAE. Also, words beginning with voiced *th* in general American English (e.g., *these*) will begin with voiced *d* in AAE.

The explanation for the types of patterns that emerge in the words in (13) is that AAE speakers use the sounds consistently and adhere to set rules and patterns. It is not that speakers make lazy substitutions in using *t/f* and *d/v*; they use these sounds in well-defined environments. The distinction between the two *th* sounds is maintained in AAE when speakers use voiceless sounds in one environment and voiced sounds in the other. AAE speakers have rules that govern the occurrence of these sounds in word initial, medial and final positions.

For the reader who might be inclined to conclude that speakers of AAE make substitutions for the *th* sounds (that occur in general American English) because they cannot produce these interdental fricatives (i.e., *th* sounds) by making the so-called difficult movement of placing the tongue between the teeth, it should be noted that AAE speakers do indeed produce *th* sounds. These same speakers who produce *t/d* and *f/v* in some environments in which *th* occurs in general American English also produce the *th* sounds in some environments. They produce the voiceless *th* sound at the beginning of words, so they say words such as *thin*, *thigh*, and *thing* as they are pronounced in general American English and other varieties of English, even though they begin with the *th* sound.[9] So speakers do not use *t* in environments in which voiceless *th* begins a word in English. I have introduced a significant amount of specialized terminology in laying out descriptions that will be useful in explaining patterns in the sound system of AAE. These descriptions should be informative for those interested in AAE, but they should also be useful for educators teaching students who use such patterns. The following section discusses the production of *r* and *l* in AAE, and it explains the pronunciation of final sounds in words such as *mother* in (13k).

4.5   *r* and *l*: liquid vocalization[10]

The data in this chapter have been used to show that the sound system in AAE operates according to set rules, so speakers do not delete and add sounds haphazardly. What may sound like ignorant and uneducated speech to those who are unfamiliar with the variety or who have some preconceived notions about the people who use this variety is actually rule-governed language use. This becomes clear once systematic inquiry is made into the sound system of AAE, and descriptions are provided.

This section continues the discussion of consonants, focusing on the liquids *r* and *l*. Some patterns associated with *r* and *l* in AAE have also been recorded for other varieties of English.[11] In some environments, *r* and *l* pattern similarly in AAE. When

these sounds follow vowels within words, they are not necessarily produced as liquids; instead they may be produced as an unstressed vowel (schwa ə or *uh* sound) if any sound is produced at all. Consider the examples below. (The vowel sounds may vary from region to region, so speakers in Southern regions may produce the vowel sounds in the words below differently than speakers in the Northern region although both groups of speakers may have liquid vocalization):

(14) **AAE**   **Phonetic transcription**
  a. cout   [kɔt]   'court'
  b. bea   [bæə]   'bear'
  c. brotha [brʌðə]   'brother'
  d. toe   [to]   'tore'

(15) **AAE**   **Phonetic transcription**
  a. bea   [bɛə]   'bell'
  b. pia   [pɪə]   'pill'
  c. coo   [ko:]   'cold'

(The *a* in the AAE representations is pronounced as the first vowel sound in *about*.)

Some generalizations can be made about the examples in (14) and (15):

Generalization 1: In all of the examples, the *r* and *l*, if they were pronounced, would follow vowel sounds.

Generalization 2: *r* and *l* may become vocalized, which means that they are pronounced more like vowel sounds than like consonant sounds, so they are represented by the unstressed vowel (ə), as in the phonetic transcription for the word *pill* (15b).

Generalization 3: *r* and *l* may become unstressed to the point to which they are not produced at all, the case with *toe* ('tore') in (14d).

    The pronunciations with respect to *r* and *l* in (14) and (15), respectively, are very similar in that the sounds do not occur in the position following a vowel. Notice that in (14) and (15), a schwa (ə) is used in some positions in which r occurs in general American English and in positions in which *l* occurs in general American English. The schwa is used to indicate that the full liquid is not pronounced. This is basically the explanation of the term vocalization, which is taken to mean that the liquids in these environments take on vowel-like qualities, so instead of a full liquid (i.e., *r* and *l*), a schwa (vowel sound) is produced.

    Different studies on dialects of English have noted that *r* vocalization occurs in different dialects or that the *r* is simply not produced in some dialects of English. Northern dialects such as those of Boston and New York City have been characterized as *r*-less varieties. In addition Southern English dialects have also been described as being *r*-less. AAE has this property in common with a number of English varieties, and the previous examples can also be used to show that the linguistic system patterns similarly with other varieties of English. It is interesting to note that the varieties are

quite similar with respect to the vocalization of *r* and *l*, but they may differ with respect to the actual vowel sounds of the vocalized element. I made this point when I introduced (14) and (15). Linguists (Labov 1991, Wolfram 1991, Edwards 2000) have noted that, in general (not specifically regarding vowel sounds in *r* and *l* vocalization), speakers of AAE have not participated in the vowel changes that have taken place in different parts of the United States.

There is a more interesting difference in the way the vocalization process operates in AAE and in other varieties of English. In some varieties of English, there is a restriction on the vocalization of *r*, such that the sound is vocalized in the environment following a vowel, provided that the *r* is not between two vowel sounds. (The *r* in *tore* is actually not between two vowel sounds because the *o* is pronounced, but the *e* is silent.) Another way of stating this is by saying that in some varieties of English, the *r* sound cannot be vocalized when it is between two vowel sounds. In these varieties, the *r* in the word *Carol* is not vocalized because it is in the environment between two vowel sounds. On the other hand, in AAE (and perhaps in some Southern English varieties) vocalization can occur in this environment, so 'Carol' can be pronounced as *Ca'ol* [kæəl]. Wolfram and Fasold (1974) make the following observation:

> In Vernacular Black English and in some white Southern varieties, it is possible for *r* to be absent after a vowel and before another vowel within the same word. The result is pronunciation like *sto'y*, *ma'y* and *te'ific* for *story*, *marry* and *terrific*, respectively. If the vowel directly before the *r* belongs to a prefix and *r* belongs to the base, then *r* cannot be deleted, so that *be'eave* and *re'un* are not possible pronunciations for *bereave* and *rerun*
>
> [p. 140].

In vocalization cases, the *r* is not produced, but there is some schwa sound or vowel lengthening as a reflex of *r* vocalization. The point about lengthening is important in that if the liquid is not produced as an unstressed vowel, then a vowel that is already present may become longer. This is what happens in the word *coo* 'cold' (14c), as indicated by the two *o*'s and the ':' (*oo*:) in the phonetic transcription.

As has been noted above, the liquids *r* and *l* pattern similarly with respect to liquid vocalization. They do not, however, share the feature of being able to vocalize between two vowels in that the *l* must be present in this environment.

The following section summarizes some additional phonological patterns in AAE for which there is a smaller body of literature.

4.6 Additional phonological patterns

4.6.1 *-in*

The sound *ng* (ŋ) in the *-ing* suffix is realized as *n* in most contexts, so the following words end in an *n* sound:

(16)   **AAE**
    a. walkin   'walking'
    b. runnin   'running'
    c. spenin[12]   'spending'
    d. thinkin   'thinking'
    e. listnin   'listening'
    f. openin   'opening'

This pattern is restricted to the suffix -*ing*, that is to words with more than one syllable, so it never occurs in the -*ing* in words with one syllable, such as *sing* and *ring*, to yield **sin* and **rin*, respectively. The property of -*ing* (ɪŋ) as being pronounced as -*in* (ɪn) is not restricted to AAE. This pattern is actually found in nonstandard varieties of English as well as in general American English in unstressed syllables. As a result, when the final syllables of *nothing* and *something* are unstressed, the words become *nothin* and *somethin* (also *sumpm*, see Wolfram and Fasold 1974).

4.6.2   *skr*

Another phonological pattern that is mentioned from time to time, often by speech pathologists, is the occurrence of *skr* in syllable initial position where *str* occurs in general American English and other varieties of English. Presumably speech pathologists were concerned about this use of *skr* because it was not clear whether the combination of sounds was an indication of a disorder or a dialectal pattern. Still the *skr* feature has not been observed or recorded in the literature nearly as often as other sound patterns. There are three possible reasons for this: (1) One is that because *skr* only occurs in positions where *str* can occur in general American English, there will be limited opportunity to produce the sound. (2) Secondly, *skr* may be viewed as a feature of the speech of young AAE speakers that is not maintained in adult AAE. (3) Thirdly, *skr* may be associated with AAE spoken in certain regions of the United States.

    Common words in which the sequence *skr* occurs are given below:

(17)   **AAE**   **Phonetic transcription**
    a. skreet   [skrit]   'street'
    b. skrawberry   [skrɔbɛri]   'strawberry'
    c. skretch   [skrɛtʃ]   'stretch'
    d. skraight   [skret]   'straight'

In summarizing her research on the cluster, Dandy (1991) notes that the form is found in Gullah and in the speech of some African Americans born in the South. She explains that the cluster is a highly stigmatized feature and that many of the students in her study who used it were referred to speech pathologists. She goes on to note the following about her research: "I also found a continuum that may indicate sound change in progress. If children said *skretch* for *stretch*, they probably used the *skr* alternation in other words that contained the feature: *skreet* for *street*, *skrong* for *strong*, *skrike* for *strike*, *skranger*/*deskroy* for *stranger* and *destroy*. There were some who said *skreet* but

did not make the alternation on other words with that sound" (p. 44). Also, although Dandy does not make this point, it is important to note that the students' use of *skr* may have been affected by the training they were getting from speech pathologists. I will return to this feature in chapter 8 in the discussion of the interaction between a student who uses *skr* in *skreet* 'street' and a teacher who tries to 'correct' the pronunciation.

4.6.3   *oɪ* and other vowel sounds

The diphthong [oɪ], a two-part vowel, that occurs in some environments in which *oa* as in *coach* occurs in general American English is also used in AAE. This feature has not received very much attention, and it is similar to *skr* in that respect as well as in other ways: based on my observations, this two part vowel is used by older speakers in some Southern states. So like *skr*, it is used by AAE speakers in a certain age group, and it may be restricted to specific regions in the United States. In fact, it is not clear that it is used by children at all. It is possible that the [oɪ] is an example of an older pattern in AAE, but judgment must be reserved here until further historical research is completed in this area. Older speakers who use this pattern use the diphthong [oɪ] in words such as *coach*, *road* and *roach*. The first part of the diphthong is the *o* [o] sound in *coach*, and the second part is the *i* [ɪ] sound at the beginning of *itch*. Speakers who use this diphthong have the following pronunciations: *coach* [koɪtʃ], *road* [roɪd], *approach* [əproɪtʃ] and *roach* [roɪtʃ]. The examples here are from speakers who either live in or grew up in the South (in particular, Georgia, North Carolina and Louisiana), but it would be interesting to determine whether AAE speakers actually use this diphthong if they live in other parts of the United States and do not have close ties with the South. The most common examples of [oɪ] are in the environment of *oa* spelling in general English, but the [oɪ] sound also occurs in *porch* (i.e., [poɪtʃ]) for some speakers. Note here that *r* vocalization also occurs in this pronunciation of *porch*.

Burling (1973), in characterizing the pronunciations of words in AAE, lists different vowel patterns in the language system. He notes that in some cases, black and white speakers in different regions in the United States maintain the contrasts between the vowel sounds in different words such as *pin/pen*, *for/four* and *horse/hoarse*. In different varieties of English, the [ɪ] sound in *pick* is used in both *pin* and *pen*; there is no distinction between the vowels preceding nasal sounds *n*, *m* and *ng* (ŋ). (Recall the *spin/spend* example.) An important point is that even if contrasts are lost or maintained, they are done so according to systematic rules affecting sounds in certain environments. The processes of losing and maintaining vowel contrasts apply in the speech of all speakers including mainstream English speakers, not just in the speech of AAE speakers.

Another pattern that occurs in some regions in which AAE is spoken is the lowering of the [ɛr] sound in words such as *prepare*, *care* and *hair*. Due to lowering, the second syllable in *prepare* is affected, so that syllable almost sounds like *par*. This pattern is produced by speakers of AAE in all age groups in cities in central and northern Texas. Karen Pollock reported on her observations of this type of lowering by speakers in Memphis, Tennessee, at the American Speech-Language and Hearing Association annual meeting in San Antonio, Texas (November 1998).

4.7   Prosodic features: stress and intonation

So far in this chapter, I have focused primarily on individual segments, consonant and vowel sounds, within words. This section will consider patterns that go beyond individual consonant and vowel sounds within words and look at patterns affecting syllables, entire words, phrases and sentences. The features that will be discussed here are related to pitch of the voice and rhythm of speech, and they are called prosodic or suprasegmental features. Some specific prosodic features are stress (accentuation or emphasis placed on syllables or words) and intonation (modulation of the voice or tonal inflection).

The prosodic features of AAE are very important for a number of reasons. One reason is that research in this area may be useful in defining, at least in part, what is meant by 'sounding black,' if anything at all, and the extent to which this judgment is based on certain types of prosodic patterns. Rickford (1972) raised questions about the issue of blacks 'sounding black' and whites 'sounding white.' According to his data, listeners who heard speech samples were able to identify speakers' ethnicity with some degree of accuracy. In the study, it was suggested that a number of features such as stress patterns, pronunciation and tone of the voice are indicators of the ethnicity of speakers. Labov, *et al.* (1968) also conducted a study to determine the extent to which listeners could identify a speaker's ethnicity based on speech. They concluded their data showed that the listeners had the "ability to recognize the clear cases of dialects" they know "rather than any general ability to identify the ethnic background of speakers" (p. 285). Rickford (1972) and Labov *et al.* (1968) disagree about whether listeners can detect a speaker's ethnicity by prosodic features, and Wolfram and Fasold (1974) explain that intonation "appears to be one of the main reasons why some standard-speaking blacks may be identified ethnically" (p. 147). In effect, their claim is that even when African American speakers use mainstream English syntactic and other patterns of speaking, their intonational patterns identify them as black. Wolfram and Fasold's conclusion is strong, but it is not clear what type of evidence (e.g., experiments based on listener judgments) they have used to reach it. If Wolfram and Fasold are right, what are the features of intonation that set African American speakers apart from other speakers?

Certainly other listeners believe that it is possible to recognize ethnicity through speech patterns and voice. This must be the view that led one of the attorneys in a widely publicized court case to suggest that a speaker's race could be determined from his voice and speech. The *New York Times's* report by David Margolick the following day about the attorney's comment: "Simply by suggesting that someone's race can be gleaned from the sound and timbre of his voice, Mr. Darden opened up once more the volcanic issue of race . . ." (Thursday, July 13, 1995). Baugh (1999) raises questions about sensitivity to differences between AAE and general American English. Although he does not discuss suprasegmental features, it may be that he is also concerned about the role that such features play in helping listeners detect differences between types of speech.

Another reason that research in this area is important is that it can add to our understanding of the role that prosodic features play in the meaning and interpretation

of phrases and sentences in AAE. In the discussion of the lexicon and syntax of AAE, it was noted that *BIN* and *dən* are produced with certain stress patterns, which are related to the way the markers are interpreted. That is to say that if *BIN* is not stressed, then it will not indicate remote past meaning.

From the discussion in the preceding chapters, we learn that the different meanings and uses of lexical items in AAE and general American English may lead to misunderstanding. In the same way, different intonational patterns, rhythm and pitch may also contribute to misunderstanding among speakers of different varieties of American English. It may be that speakers of different varieties of American English have these tones, rhythms and pitch in mind when they say that the speech of African Americans indicates that the speakers have an attitude or are angry, confrontational and rude, as well as when they say that African Americans have more expressive and soulful speech. Finally research on intonation and other prosodic features is important in that it can provide insight into what some researchers have called standard AAE, which adheres to the 'standard' rules of general American English, but incorporates intonational patterns and other types of expressive language use of AAE. That is, it is suggested that speakers of standard AAE may not use the type of syntactic patterns in chapters 2 and 3 and the sound patterns (e.g., *f/v* in *th* environments) described earlier in this chapter, but they will use intonational and rhythmic patterns associated with AAE. Standard AAE is also addressed in chapter 8 (note 13), but briefly the concept is an interesting one, and impressionistically speaking, there is no doubt that it exists. The problem in characterizing it is that a substantial amount of research has not been completed in that area. (Standard AAE has been addressed in Taylor 1971, Lewis 1981 and Spears 2000, 2001).

Comments about the way some African Americans use language and 'talk' are made in general conversation and research on AAE. For example, Kochman (1972) describes speakers as having "a fluent and lively way of talking" (p. 242). In the following passage from Fordham's (1996) ethnographic study of African Americans in a high school in the United States, a female comments on the language of her mother. In expressing her attitude toward her mother's 'telephone speech,' is the student making some implicit comment about her mother's intonational patterns and pitch?

She just talks like that on the telephone, I'll put it like that. When she talks, she puts on airs, you know, sounds White . . . so you can't tell whether she's White or Black. But when she's around the house, she talks, you know, regular; but when she's out around other people, anywhere out besides the house, she talks in a proper way.

[p. 114]

If we could force the daughter to expound on the way her mother "puts on airs," "sounds White," "talks regular" and "talks in a proper way," which properties would she describe? What cues lead listeners to assess speech as being "lively" and "regular?" How can these labels be translated into a more formal description of ways of talking?

Tarone (1973) asserts that "investigators have consistently recorded the impression that these suprasegmental features are probably just as distinctive in black English

as the other, more carefully described, features of syntax and segmental phonology" (p. 29). Almost thirty years later, we are still relying on impression about intonation, but we have begun to make progress in this area. Still stress and intonation have not been well studied, so the unique patterns associated with them have not been defined, and they have not been formally linked to the tonal patterns of West African languages, nor have they been linked to patterns in the United States. As the debates on the origin of AAE continue, intonation and other prosodic features may prove to be useful areas of consideration. Research in this area may also shed light on the tonal properties of (stressed) *BIN* and (unstressed) *dən*. A point of departure for discussing intonational patterns in AAE and West African languages is Gullah, which has been observed as having tonal qualities that are similar to those of West African languages (Turner 1949).

Tarone's (1972, 1973) studies of intonation in AAE, are based on speech samples she collected from African American adolescents in Seattle, Washington, while they were engaged in conversation in a social setting. According to Tarone (1973), "Intonation, in its function as communicator of attitude, might therefore be expected to reflect this difference in attitude between the participants in black and white middle-class speech events" (p. 30). Some recurring patterns that she found in the speech of African American adolescents are given below:

(1)   A wider pitch range, extending into higher pitch levels than in English or formal black English, and often shifting into a falsetto register
(2)   More level and rising final pitch contours on all sentence types in an informal situation were used
(3)   Apparent greater use of falling final pitch contours with *yes-no* questions in formal, threatening situations, but level and rising final contours in informal, familiar situations
(4)   The use of non-final intonation contours, without the use of the lexical item *if*, to mark the dependent clause of some conditional sentences

Tarone gives the following as an example of the characteristic in (4): ^{2}You ^{3}a^{2}ble to ^{3}do ^{2}it, just do it (p. 34). The dependent clause (*You able to do it* 'If you're able to do it') is represented as ending in a non-final intonational contour by having 3-2 (high to medium pitch contour) assigned to it. The non-final contour simply indicates that something will follow; we are not at the end of the sentence.

Tarone suggests that "the wider pitch range may be directly traceable to the competitive nature of the black speech events" and that "the element of aggressiveness and competitiveness was reflected in a wider range of intonation" (p. 32). In other words, it may be that listeners will translate the wider range in intonation as meaning that the speakers have an attitude or are angry and confrontational. Tarone linked wider pitch range in the speech of African American adolescents in her study to the speech events in their conversations although it certainly must not be the case that wider pitch range is reflected only in speech events. In earlier and more current studies on AAE, researchers have associated aggressiveness, indignation and negativity with speech events and the use of certain lexical items. Obviously, these attitudes can be expressed in speech; however, it

would also be a good idea to determine what else is associated with wider pitch range. It is not always clear to me, at least, that the listener's assessment of aggressiveness, indignation and negativity is shared by the speaker or captures what the speaker is trying to convey. Furthermore, it is not clear how insightful linking the wider pitch range to aggressiveness is, especially if wider pitch range is manifested in other environments.

The summary of characteristics in Tarone's four points are more formal descriptions of the impressionistic statements about 'sounding black.' If there is something distinctive about the rhythm and modulations of speech used by some African Americans, then some of the distinctions may be captured by the pitch range and falling and rising patterns.

Intonational patterns used by one African American are recorded in a more recent unpublished study by Scott Meredith and Caroline Henton (no date). Whereas the Tarone study, which is much earlier, tries to understand the meanings that are associated with certain contours as well as the differences between the AAE intonational patterns and those of other varieties of American English, Meredith and Henton are interested in determining the extent to which the system they have chosen to use to transcribe intonational contours can be extended to account for patterns in their African American English research. They base their data on speech samples from a 55-year-old African American female, who is "a (presumed) speaker of Black Vernacular English and who, at the time of the study lived in San Francisco, but lived in New Orleans for about 50 years." In their brief discussion, they point out that their data led them to conclude that the speaker has an extremely wide pitch range, an observation that has been made about AAE intonation in independent studies (see the summary of Tarone's research above).

The preliminary study in Green (1990) is similar to Tarone's study in that one of its goals was to determine whether certain characteristics and intonational patterns were associated with utterances in AAE. Also, like the Meredith and Henton study, the Green (1990) pilot study was designed to address questions about ways of formalizing statements about intonation. The scope of the study was narrow; it focused specifically on questions in AAE. Two goals were to determine (1) whether a special type of intonation is associated with questions (*yes-no* and *wh*-questions) in AAE and (2) whether these contours could be characterized by using a formal system. In short, the aim of the study was to determine whether unique intonational patterns occur in AAE and if they do whether they are associated with certain sentence types.

The data for this study were taken from speech samples produced by teenage and adult speakers who live in a small community in southwest Louisiana. Although the findings are preliminary, they match a significant finding in Tarone (1972), and they are in accord with preliminary findings by Tempii Champion, Ph.D., and Elisabeth Selkirk, Ph.D., in preliminary research conducted at the University of Massachusetts, Jun and Foreman (1996) and Foreman (1999).

In the Green study, the most striking pattern that recurred in *yes-no* questions was the final level tone. Think about the question *Do you want to read*? In asking the question, it is natural for some speakers to produce the word *read* with a final rise in intonation, but the yes-no questions in the data that I collected did not always have

the final rise. This pattern was interesting because, for the most part, final rises have been associated with yes-no questions in English. That is not to say that questions are only produced with final rises. Cruttenden (1986) warns that "there is no such thing as 'question intonation' although some tones may be more common on questions than others" (p. 59). He goes on to suggest that high-rise and low-rise tones are more common on questions. There were a number of cases in which the final tone on the *yes-no* questions in Green (1990) was level. At the end of the question, the speaker produced neither a rising nor falling pattern; the final tone was level. The claim is not that level tones are the only contours that are associated with yes-no questions in AAE but that it is often the case that these contours are used in such sentence types.

The sentences in the study were transcribed using the model in Pierrehumbert (1980). The basic tonal units that are employed in the model are H (high) and L (low), and the tonal entities that are used in the system are the pitch accent, phrase accent and boundary tone. The pitch accent (H*, L*) is associated with the stressed syllable; the phrase accent marks the end of the phrase, and the boundary tone is realized with respect to the syllables at the edge of the domain. (I will not go into detail in summarizing Pierrehumbert's study because it requires some background in phonology.)

The ends of the questions in Green (1990) are marked by HHL, in which case the final L (low tone) is used to show that the level contour is maintained throughout the end. These final patterns were observed in sentences such as the following:

(18)   a. You can get good grades and sleep? (attested)
            'Can you get good grades and sleep?'
        b. You traded your other one in? (attested)
            'Did you trade your other one in?'
        c. You want some water too? (attested)
            'Do you want some water, too?'
        d. And you let him drink your coffee like that? (attested)
   'And do you let him drink your coffee like that?'

The questions in (18) are similar to some of the ones that were discussed in chapter 2, in which the auxiliary does not precede the subject. Instead they begin with the subject; the auxiliary is present in (18a), but subject-auxiliary inversion does not occur in any of the sentences in (18). Questions like this may be rhetorical in nature in that they will not require an answer. That is to say that the questions may be asked with an answer in mind, so all the speaker is really asking is that the listener agree with or affirm what the speaker expects to be the case. A rhetorical question might be *You took the trash out already*? after the observer notices an empty waste basket, in which case the observer strongly suspects that the listener has performed that chore. Such questions are not true *yes-no* questions in that they are not genuine requests for new information; however, it is clear that the questions that are the focus of discussion in Green (1990) are indeed 'true' yes-no questions in that they are requests for yes or no responses. The types of final level patterns that I am concerned with are represented in the sentences in (19):

(19)   Yes-no question pitch tracks (Green 1990)
    a. You traded your other one in?
        'Did you trade your other one in?'

You   traded   your other   one   in

0.000 s   0.200 sec./tic   0–500 Hz Time Domain f0 plot
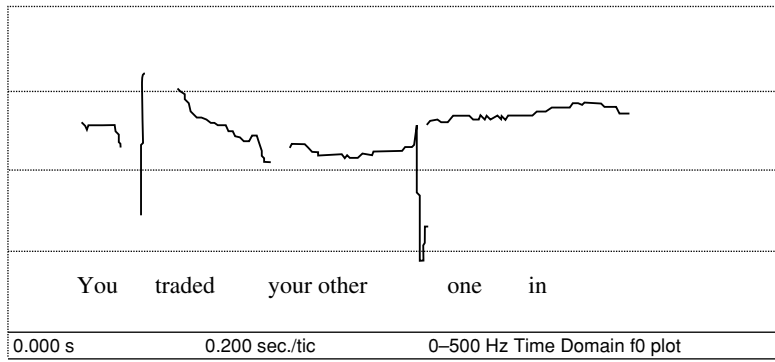
    b. Well Reverend— called y'all all one up at a time?
        'Well did Reverend— call y'all all one up at a time?'

Well Reverend ⎯ called y'all all   one   up   at a   time

1.000 s   0.200 sec./tic   0–500 Hz Time Domain f0 Plot   3.0
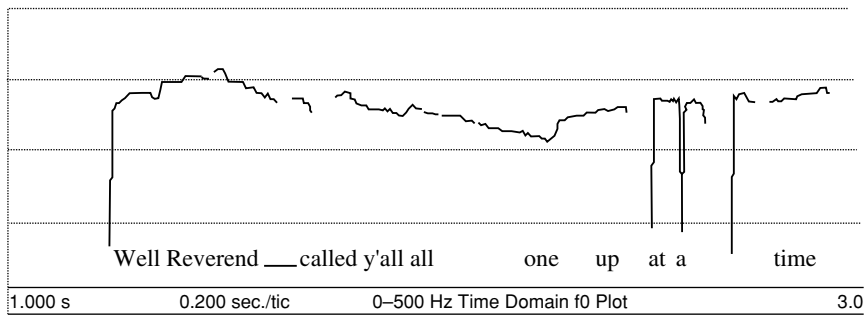
The questions in (19a) and (19b) end in a level or flat tone. In the discussion of questions in the preceding chapter, it was noted that AAE is not unique in using question intonation, but as more research is conducted in this area, we can explore issues about whether AAE uses unique intonational patterns, especially in marking questions that are not formed by subject-auxiliary inversion. The studies on question intonation in AAE strongly suggest that studying yes-no questions can yield important information about AAE intonational contours. As noted, one of the same patterns associated with the yes-no questions in Tarone's study was also found in Green's study. In addition, Foreman (1999) notes that the AAE speakers in her study did use the standard rise. However, she also found that they often used a high level or flat tone and falling tones in their yes-no questions, which are associated with declaratives. The level and falling tones are noted in Tarone (1972), and the level pattern is also associated with yes-no questions in Green (1990). The level or flat contour is used in questions and in other types of constructions in AAE. Perhaps that tone is a salient property in AAE which

can be used by listeners to form opinions about the speech of AAE speakers and might be used in distinguishing the variety from other varieties of English.

The preliminary analysis of *wh*-questions in Green (1990) shows that the contours for those questions are similar to the contours for *wh*-questions in general English. Cruttenden (1986) and Pierrehumbert and Hirschberg (1987) agree that the tunes used with *wh*-questions in general American English are basically the same as those used with simple declaratives. The lowering pattern which has been called downstepping, is associated with these types of questions. This lowering occurs consistently throughout the statement, such that following low tones are lower than previous ones. This is shown in the diagrams below:
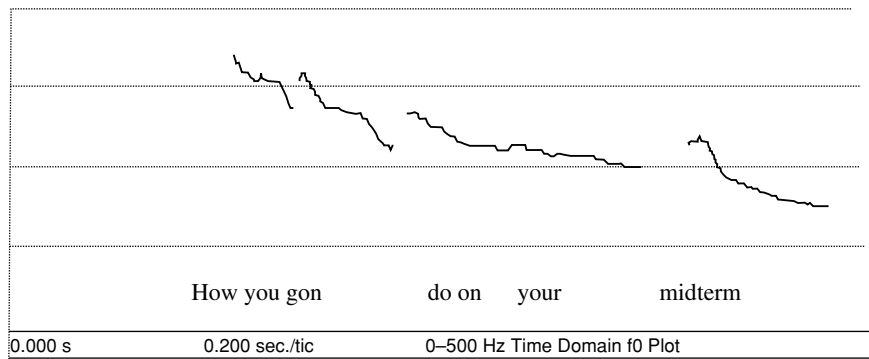
(20)   *Wh*-question pitch tracks (Green 1990)
   a. What they say?
      'What do they say?'

What they   say
0.000 s   0.200 sec./tic   0–500 Hz Time Domainn f0 Plot
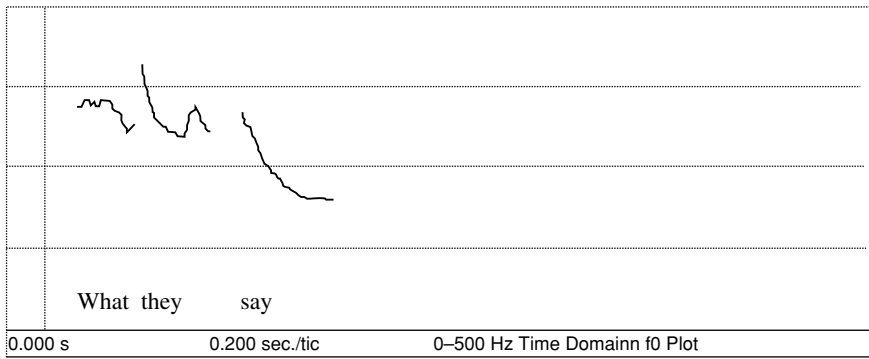
   b. How you gon do on your midterm?
      'How are you going to do on your midterm?'

How you gon   do on   your   midterm
0.000 s   0.200 sec./tic   0–500 Hz Time Domain f0 Plot

Note that the *wh*-questions are not formed with overt auxiliaries as in *What do they say*? and *How are you gon do on your midterm*?.

Research on intonational patterns may ultimately provide a body of information that can be used to give more substance to the impressionistic observations that are made about the rhythmic patterns in the variety.

In continuing the discussion of suprasegmentals, I now turn to stress patterns within the word in AAE. Investigators who have worked on AAE have commented on the stress in initial syllables of specific words in the language system. Smitherman (1977) describes this phenomenon of placing primary stress (indicated by capital letters) on the initial syllable as front shifting, and gives the following examples (p. 18):

(21) a. PO-lice   'police'
      b. DE-troit   'Detroit'

Interestingly enough, different researchers list the syllable initial stress pattern as a common feature of AAE, and they all present the same words as examples, which suggests that the process only occurs within a limited set of words.[13] According to Baugh (1983a), this process "is limited to a portion of bisyllabic words" (p. 63). He goes on to give examples of sentences in which words (*pólice*, *défine*, *próduce*, *révise*, *pólite*) that bear initial stress occur:

(22) a. If the pólice catch em, then he'll be sorry.
      b. But that's cause you let the man défine your problem.
      c. They can't lock . . . you . . . they better próduce some evidence first.
      d. He say he ain't coming back to work less they révise the schedule.
      e. He gots to be pólite around his moms – an like it too.

Baugh identifies this pattern as the forestressing of bisyllabic words and represents it in the following syllable structure, where stress in AAE is transferred to the first syllable. The diagram in (23) gives a general description of the differences between forestressing in bisyllabic words (note the consonants [C] and vowels [V]) in AAE and the second syllable stress in these words in general American English:

(23)

AAE		General American English	
CV	CVC	CV	CVC
PO	l i c e	po	LIC E
syllable 1 (stressed)	syllable 2	syllable 1	syllable 2 (stressed)

Wolfram and Fasold (1974) do not go into detail, but they note that this first syllable stress pattern is found in AAE with words such as *police*, *hotel*, *July* and *Detroit*, the same subset of words given in other studies (cf. Smitherman 1977 and Baugh 1983a). They go on to note that the stress pattern occurs in "some Southern white varieties" (p. 147), but do not go on to suggest that there are similarities between the intonational patterns in AAE and in Southern white varieties of English. The point is that while comparison between AAE and Southern white varieties have been made on a number of levels, not much has been said in relation to intonation. However, Feagin (1997) does suggest that there may be some similarities between the wide pitch ranges used by African American males and Southern

white females, but to my knowledge there has not been any research on which to base this claim. These similarities, as she points out, may have "Caribbean roots as well" (p. 132).

A final note on the topic is that this stress pattern also occurs in words with more than two syllables. In data from a sermon that I recorded in 1990, the word *protector* is produced by the minister as *PROtector*, in which the first of three syllables is stressed. (A portion of that sermon is reproduced in chapter 5, section 5.5.)

Summary

In addressing sound patterns in AAE, I considered segmental and suprasegmental features. In the discussion of segmental features, I considered initial, medial and final consonant sounds and vowel sounds in certain environments. It is easy for listeners to evaluate the sounds and sound combinations that have been discussed as lazy speech, but the descriptions show the way in which combinations are very systematic and based on defined rules that make reference to specific environments in which sounds occur. The voicing value of consonant sounds plays a major role in the production of sounds and the specific sound that is used in a certain environment. For example, *f* and *v* occur in environments in which voiceless *th* and voiced *th*, respectively, occur in general American English. Data on the production of consonant sounds have been used in arguments about the origin of AAE. For instance, it has been argued that some clusters are less likely to occur due to the nature of languages from which AAE descended. In some of the discussion, technical terminology has been introduced in giving explanations for why sounds occur or do not occur in some environments. In general, one of the goals was to show that speakers do not haphazardly delete and insert sounds because they have 'lazy tongues.' The patterns in the sound system of AAE are completely regular.

Suprasegmental features such as stress and intonation have not been well studied, but they are often discussed in relation to unique features of AAE. The research in this area raises questions about the types of intonational contours that are associated with constructions in AAE. The studies that have been conducted on yes-no questions in AAE converge on the point that both level (or flat) and falling tones are used in these constructions. It may be that the types of intonational contours that speakers use mark the speech as being unique. Segmental features have been used to argue that AAE has retained features of West African languages, and they have also been used to argue that AAE is similar to other varieties of American English; however, suprasegmental features have not been brought squarely into the debate.

Exercises

1. The *ks* cluster at the end of *saks* (as in Saks Fifth Avenue) is not reduced. That is, *ks* does not become *k*, so speakers do not say *Sak Fifth Avenue. Does the *voicing generalization* make the correct prediction about *ks* clusters? Explain your response.

2. The cluster *ld* as in *build* is also reduced, so *build* in general American English can be pronounced as *buil* in AAE. As you will see, the sonority scale does not give the correct results, that is, it does not show that the *ld* cluster should be reduced. Explain how it fails to show this.

3. Now consider the verb *building* in the sentence *They building another church over there* ('They are building another church over there'). After reviewing the section on consonant cluster reduction and your response in the preceding question (number 2 above), explain how the verb *building* would be pronounced in the sentence. Also, say whether you think it is possible to get two different pronunciations. If so, what are the pronunciations? Be specific about the information you used in reaching your conclusion.

4. Consider the medial *th* in the word *birthday* in general American English. What do you think the *th* sound would become in AAE? How would the word be pronounced in AAE?

5. Think of African Americans who are in the media (e.g., Kweisi Mfume, Oprah Winfrey, Michael Jordan, etc.). Of those who come to mind, which ones (if any) would you describe as speaking what has been referred to as standard AAE? Explain your answer, and give examples to illustrate your points.