

Peer Reviewed Article by Douglas N. Honorof

Reference Vowels and Lexical Sets in Accent Acquisition

1.0 Introduction

It is difficult, while reading aloud, to produce a believable *unfamiliar accent* (where the term *accent* is used to refer to the pronunciation of any variety of speech, with the term *dialect* reserved for features of the grammar and vocabulary following Hughes & Trudgill, 1996). For some, it is doubly difficult to act *past* the accent once off book. Although mastering prosody (intonation, rhythm, etc.) may provide a quick means of suspending the audience's disbelief about our ownership of a character's voice, I propose that much of the awkwardness we communicate while acting 'in accent' stems from our fear that we will become confused if we allow ourselves to stop monitoring the details of our consonants and vowels so that we can focus on our actions.

Over the past few years, I have been experimenting with a fun, actor-friendly, linguistic approach to accent training intended to reduce accent anxiety.

Within this approach, actors are encouraged to explore extensively the coordination of gestures within their own vocal tracts, thus training their ears in the process. Through exploration of their own instruments, actors make new and extremely useful discoveries that help extend their range and flexibility. They are also encouraged to memorize *cross-accent* sound-spelling correspondences (lexical sets), which diminishes confusion over sound-spelling mappings. The internalization of the lexical sets frees actors to focus on acting tasks more central to the performance moment. This two-pronged process—gestural exploration and lexical set familiarization—has fostered remarkably improved consistency (subjectively judged) and noticeably better performances, both in conservatory settings and in 'quick-fix' coaching contexts.

Some actors are intimidated by the notion of an analytical, linguistic approach to work that ultimately must become synthetic and organic. Before describing the details of a linguistic approach to accent acquisition, I offer arguments in favor of the approach, arguments that have been inspired by the naïve objections of actors who have not experienced the work. Then I define what I mean by the term *gesture* and introduce the cardinal vowels of Daniel Jones and the lexical sets of John Wells. Finally, I walk the reader through a diagnostic narrative intended for use in developing phonologically balanced speech-sample archives.

1.1 Motivations for beginning with linguistics

While it may be wise to adjust our training approach to the learning styles of individual students and actor-clients, it is not entirely reasonable for an actor, especially an inexperienced actor, to dictate the nature of training sessions from the start. As experienced teachers and coaches, we may be able to suggest techniques to actors that may work better for them than the techniques with which they are already comfortable. Therefore, in order to enlist the actor's fullest cooperation for what may sound like an unnecessary, highbrow foray into linguistic science, a bit of salesmanship is in order at the outset. In the next four subsections I outline slightly more sophisticated versions of some of the motivational arguments I have successfully employed.

1.1.1 Analytical preparation breeds confidence for later creativity

Adult actors must be well enough prepared to speak in accent extemporaneously, ideally before they are even off book. Otherwise, they find the rational



Doug Honorof has worked on-camera and off as an actor and coach (accents and dialects), serves as an Associate Editor for IDEA and has taught at the Yale School of Drama, Voice One and CAP21, the undergraduate musical theatre studio school at NYU/Tisch. He trained under theatre faculty at Yale and at several institutions in New York including the Actors Center and the Upright Citizens Brigade Theatre. Doug also holds a PhD in linguistics (phonetics and phonology) from Yale and serves as Senior Research Scientist at Haskins Laboratories in New Haven, Connecticut, where, under federal funding, he investigates imitation of speech.

mind thrust into overdrive during a performance. The actor's overtaxed 'left brain' drowns out the character's inner voice, taking the actor entirely out of the moment. The performer loses touch with even a carefully constructed imaginary world and loses all sense of *complicité* with scene partners, hearing only the voice of an inner accent-critic. Every mangled vowel, every misplaced "r," every unintended muscular tension and every inauthentic tune will resound in the actor's consciousness, preventing him or her from committing fully to living the moment in character. Fear of failure with the accent can be battled only with the confidence that naturally arises out of success in rehearsal. Hard intellectual work fosters success in rehearsal, even, perhaps, when linguistic talent is lacking. Present success breeds future courage.

1.1.2 Accents are not created *ex nihilo*

As complex as a phonological system (accent) may seem, with training and experience, there is always enough time to learn an accent well because we are never really starting from scratch. Again and again my students and actor-clients have made a small adjustment to the placement of the voice or have adopted a new feature of a target accent, and have realized on their own that the accent suddenly sounds more authentic. Mysteriously, it would seem that we know when an accent is true, even when we have never used it before. Furthermore, many of us outsiders can identify native speakers or very talented mimics when we hear them (but see Trudgill, 1983, on the limitations of passive grammatical competence). Perhaps we all have a great deal of very specific passive knowledge of accents we have heard but in which we have never spoken. I suggest that such passive perceptual experience with a range of accents can quickly be turned into active production. By way of analogy, it has been said that we need not be dead to play dead. Perhaps we can play dead because we have seen corpses and because we can draw on our own memories of rest in life. Along these lines, I suspect that we need not have produced a familiar accent to judge its authenticity because we can draw on our memory of having perceived it in the past. It may even be that actors who hone their imitative skills can sense intuitively what they would need to do to produce an accent they hold in short or long-term memory even if they lack the rhetoric to describe what is going on in their vocal tracts when they do so.

Passive competence is far from well established in the literature. The devil's advocate might argue that we know an accent sounds false simply because it is stilted or forced, or because it relies on too a narrow a range of expressive tones and rhythms. Given passive competence, however, my working hypothesis would be that the experience of having analyzed and practiced a reasonably familiar accent in advance frees the actor to turn

passive knowledge into active expertise during a performance as he or she dares to believe against all evidence that he or she is indeed a native speaker of the target accent. Under such a view, the accent breakdown is simply be a tool we use to discover what, at some level, we already know sympathetically. If we learn to access our passive linguistic expertise, we start with a definite "leg up." Besides, we do not really have to get the accent perfect for the character to be believable even to a native speaker; real people are rarely pure exemplars of a type. We accept a character as real if the actor believes and has done at least *most* of the homework.

1.1.3 It gets easier with each new accent

As hard as it may be for some of us to really work at our new first accent several hours a day until we have it down pat, there is hope. Accent work is like riding a bicycle. Once the accent is there, it does not take much of a refresher to re-access it even after a long hiatus. Perhaps the breakdown and embodiment of that first new accent reactivates a dormant accent-learning mechanism in us. Whatever the case, subsequent accents become easier and easier to acquire. It is as though, with each new voice we possess, we gain more hooks on which to hang new sounds. This hope alone should give courage to the fainthearted; courage sufficient to conquer the fear of a *seemingly* monumental task. Many of my more experienced clients report that, with the accent device once again thrust into full gear, it eventually becomes possible to simply commit to a simple inspiration, and to produce a new voice with no preparation at all, say, for an audition.

1.1.4 Analytical and synthetic approaches are compatible

We all fear the unknown. Linguistics is unknown to most actors. The fear of being asked to master abstract linguistic concepts causes some actors to simply refuse to audition for accent parts, while others insist on approaching accent work synthetically with recordings alone but no coach. Although shadowing audio recordings is crucial, doing so helps the actor more after the accent has been properly broken down. Although those hoping to master an accent by relying on their ears alone ultimately have the right idea, even the most talented among them may miss important details when they jump past the analytical work. Furthermore, even those who get the prosody and all the consonants and vowels right may not always know which words take which sounds. The right vowel in the wrong word can be jolting to the listener. Although hit-and-miss accent work may suffice for voicing cartoons or teenage twitch games, many film, television and legitimate theatre roles call for more than what I call an *accent-shaped object* (ASO). Granted, there may be legitimate differences in learning style. There may even be unusually gifted actors who do not need but a note or two from us. (See Markham, 1997,

for a survey of what is known about ranges of individual achievement at accent learning.) Still, the average actor who objects to analytic instruction may, in part, actually be objecting to something inherently confusing in the traditional sound-substitution training paradigm. With a simpler and more focused linguistic approach, the actor can get past the hard work quickly, and be freed to take risks with the synthetic work.

2.0 Overview: Gestures, reference vowels and lexical sets

I borrow the notion of the *gesture* from *articulatory phonology* (e.g., Browman & Goldstein, 1991) and from the *direct realist* theory of speech perception (e.g., Fowler, 1996), frameworks related to Gibson's work in *ecological psychology* (e.g., 1966). A gesturalist might say that we learn, produce and directly perceive dynamically *coordinated* movements within the vocal tract rather than acoustic features or static orthographic symbols that stand for segments (Fowler, 1995).

Actors are often delighted to hear that the analytical work I teach is based on coordinated movement. Movement is very tangible. I tell actors next to nothing about the technical specifications of the gesture—much less than I will describe below—but my gesturalist disposition toward articulation does lead me to teach actors a second linguistic construct: the cardinal vowels. These articulatorily defined reference vowels serve as a prerequisite to the third linguistic concept I believe actors need to know about: *lexical sets* (Wells, 1982). Each of these areas is described in more detail, below.

2.1 Definition: Gestures

My colleagues at Haskins Laboratories have spent over a decade presenting experimental data that are slowly nudging phoneticians and phonologists around the world away from a fixation on static, transcription-based (that is, alphabetic) units such as phonemes and the phonetic features of which they are comprised, toward the study of dynamic articulatory gestures which may or may not cohere into segmental units. Within the Haskins model, speech gestures are decomposed into synergies of participating articulators and further specified for such parameters as constriction degree, constriction location, aperture, protrusion, articulator orientation, etc. With these parameters in mind, speech may be viewed in terms of movement, or, more accurately, as a system of coordinated movements. This perspective on speech bears striking resemblance to the decomposition of dance into timing, direction and spatial relationships as will be familiar to students of eukinetics and Labanotation, though our definition of the gesture itself is based on a different set of principles. Fortunately, movement is something even the most non-analytical of actors can relate to, as well. The following subsection outlines some gesturalist assumptions about speech in fairly non-technical terms.

2.1.1 Coordination within a gesture

We adapt from Bernstein the notion of the functional organization of muscles into *coordinative structures* (1967). *Coordinative* is the operative word. Let us first consider coordination within the gesture that corresponds to the lip closure for a 'p'. (Note: We refer to the letter 'p'. More traditional linguists might refer instead to the phoneme /p/ or to the phone [p].) In defining a specific gesture, we look for evidence of coupling among a *synergy* of articulators and specify the geometry (closure, protrusion, etc.) targeted by the named articulators working in tandem. For 'p', the key articulators of the oral tract would be the upper lip, lower lip and jaw. The geometry aimed for would be a closure of the vocal tract at the lips, though the individual articulators might contribute in varying degrees to the closure. From a gesturalist perspective, all lip closure gestures are functionally equivalent no matter which articulator moves the greatest distance or with the greatest velocity. Perhaps accents and idiolects differ according to the favored weighting of the contributions of the individual articulators that make up the lip closure gesture, but even within a single talker, that weighting can vary with gestural context and performative circumstances. In fact, an early piece of evidence for the gesture was unearthed by experiments in which the jaw was tugged downward by a mechanical device during the production of a lip closure gesture (Kelso et al., 1984). The experiment revealed that talkers compensate almost immediately for the perturbation of the jaw (and, consequently, the perturbation of the lower lip riding on the jaw) by lowering the upper lip and raising the lower lip more than usual in order to achieve a closure of the lips at all costs. At a functional level, it is the task of closure itself that matters.

The Haskins group has devised a (computational) gestural model that allows the researcher to view the midline of a two-dimensional talking head on a computer screen, and to hear the synthetic speech it produces. To condense the technical specification into one sentence: The model relies on a differential equation that treats each gesture as a critically damped mass-spring system, the gesture having its own internal stiffness—a variable that affects the rate of movement towards a target, and therefore movement duration. (See Saltzman & Munhall, 1989.)

Intrinsic duration is one property that most strikingly distinguishes gestures from static, timeless phonemes. The math behind the model is beyond the scope of the present paper and way beyond the concerns of the actor learning an accent. Fortunately, one need not understand the math in order to benefit from a coordination-oriented way of thinking about speech. If the gesturalist is correct, the mapping between gesture and sound will be a more natural one than the mapping between a letter of the alphabet and sound, which may partially explain the difficulty so many actors experience in learning

to read from transcription. It may even explain the difficulty experienced by children who have reading disabilities; the gesture corresponds to linguistic objects in the real world, while the phoneme is merely an overlaid, learned concept, and one that is easier for some to learn than others (Mattingly, 1972; Read et al., 1986). Although a number of highly regarded techniques have been devised for teaching speech through transcription (e.g., Catford, 1988; Colaianni, 1994), in production coaching where I tend to have less time with an actor than one has in a conservatory program, I begin by making the actor aware of his or her own gestures as quickly as possible, and only then touch on transcription if at all—in fact, narrow transcription at that; even in transcription, I skirt the phoneme.

2.1.2 Coordination between gestures

Clearly, words involve multiple gestures, but even to produce just the sounds spelled with the letter ‘p’, we must move beyond lip closure. We must specify a velic closure gesture to channel airflow through the mouth but not the nasal passages. We must also specify a laryngeal abduction (that is, glottis-opening) gesture to keep ‘p’ from sounding like ‘b’. The degree of abduction and the relative intrinsic duration of the glottal gesture will vary with context in an accent-specific way. A word-initial ‘p’ heading a stressed syllable in many accents of English will have a much wider—and therefore, *ceteris paribus*, temporally lengthened—opening of the glottis than will a medial ‘p’ (*pot* versus *taper*); the larger opening of the glottis gives the ‘p’ in *pot* its aspiration. A specific temporal ‘point’ in the large glottal abduction gesture may be phased to a specific ‘point’ in the lip-closure gesture, but the glottal abduction gesture is not part of the ‘p’ itself. ‘P’ does not exist except as a crude orthographic index of what is really going on. The glottal abduction gesture is itself a potentially contrastive element of syllable onsets (Goldstein & Browman, 1986). There are at least three possibilities: the laryngeal abduction gesture is absent (implying a narrowed glottis for voicing as for the ‘b’ in *boy*), it is present but ‘less wide’ (for the ‘p’ in *taper*) or it is wide (for the ‘p’ in *pot*, *spot*, *plot*, *prof*). Note, when a wide glottal opening gesture in a syllable onset is followed by a stressed vowel, we hear different things depending on context: aspirated ‘p’ in *pot*, but unaspirated ‘p’ in *spot*, and unaspirated ‘p’ followed by voiceless ‘l’ or ‘r’ in *plot* or *prof*. The phasing of a small inventory of gestures across the oral, laryngeal and velic *tiers* expresses the aspiration rule very elegantly. It is much easier to tell an actor to blow a lot of air through the lips when the syllable onset is voiceless and the vowel is stressed than it is to tell them that ‘p’ is sometimes unaspirated but only between vowels or after ‘s’ before a stressed vowel, that ‘l’ and ‘r’ are sometimes voiceless, but only in after ‘s’ plus a plosive at the beginning of a stressed syllable, etc.

Now let us consider the gestural organization for ‘p’ in an old-fashioned, upper-class Southern British accent, that is, the accent often referred to as Period Received Pronunciation (RP), U-RP or Marked RP. Here, the target accent does not have a wide glottal gesture in syllable onsets, thus, we hear unaspirated plosives in *pot* and no devoicing of ‘l’ and ‘r’ in *plot* and *prof*. Again, this is more technical a definition than the actor needs. Furthermore, it may be that this feature of the accent is variable, but in simplifying the task for the actor, we need only instruct the actor to blow little air through the lips for a voiceless ‘p’ before a stressed vowel. In fact, because the relatively more or less wide glottal aperture gesture is a property of syllable onsets, the amount of air blown, as it were, is the same whether the initial consonant is ‘p’, ‘t’ or ‘k’.

In addition to the phasing across tiers (oral, velic and glottal), the *oral* gesture for an initial consonant such as ‘p’ is phased to the oral gesture for the following vowel, etc. Phasing is simply an instruction to the computer model specifying how far into a constriction to go before forming the next (overlapping) constriction. Thus it is important to bear in mind that two talkers of the same accent who begin with the same inventory of gestures can differ not only in how they typically form the gestures (that is, in how they coordinate their articulators in the formation of constrictions *within gestures*), but also in how they coordinate gestures *with each other*. In this way a gestural approach gives us a rhetoric for moving beyond generalized accent-study to idiolect/character ‘design’.

2.1.3 Summary: Gestures

Gestures allow us to work in terms of units that are more tangible than transcriptions and features. Furthermore, gestural awareness enables the trainer to give a higher proportion of articulatory notes than auditory notes—which may be very helpful given the finding of Catford and Pisoni that purely articulatory training is more effective than auditory discrimination training alone in teaching exotic speech sounds to adults (1970). The gestural approach has been discussed above in connection with consonants, but applies equally to the teaching of vowel systems as will be seen in the next section.

2.2 Definition: Cardinal vowels

When Western linguistics became increasingly concerned with less-familiar languages and when scholarly communication across borders became more frequent in the early 20th Century, phoneticians faced a challenge. Convenient sound recording equipment was still a thing of the future. How, then, might the field worker record in print the ‘exotic’ sounds of far-off lands for the benefit of scholars back home? Or how, even, might French phoneticians, for example, describe in print the vowels of local accents for those abroad? Editors of

traditional bilingual grammars tended to simply write, for example, that the Parisian word *tout* has a vowel like that in English 'too'. Of course, a Belfaster and a speaker of just about any variety of Plantation Southern (American) would have had wildly different pronunciations for the vowel in 'too', none of them quite like the Parisian. The situation would only be doubly complicated for the Russian or Swedish phonetician who would be accessing the Parisian pronunciation of *tout* through a form of English, already a second language. The lexicographer's difficulty in mapping one accent onto another persists to this day, as lexicographers have tended to maintain the tradition of relating unknown pronunciations to the pronunciations of familiar words from accents presumed to be more familiar to the user of the dictionary or of devising equally imprecise sight-spelling such as 'uh' for the vowel in 'cup' (See Bronstein, 1998).

Daniel Jones, a British student of William Tilly and Paul Passy, advanced an alternative to cross-system sound-mapping and sight-spelling. His alternative, based on suggestions by Ellis, Bell and Sweet, will not do for general purposes such as lexicography because some training is involved, but it is simple enough to be mastered quickly by professional users of speech. The technique he developed involves learning to recognize and produce eight primary *cardinal* vowels. Eight to ten secondary cardinal vowels were later added. The cardinals serve as reference vowels belonging to no accent at all, but pronounceable by anyone who has studied the system, and are organized in terms of tongue height.

The cardinal vowel space (see Figure 1) is anchored by the cardinal vowels occupying the four corners of an imaginary two dimensional connect-the-dots chart, so it is most reasonable to begin teaching the cardinal vowels by anchoring the four corners. Beginning with cardinal vowel 1 (cv1), the student may be asked to grin from ear to ear all the while thrusting the center of the tongue into an exceptionally high position in the palatal arch, constricting the passageway to the point of *nearly* causing the air flowing over the tongue to vibrate as for a fricative. Next the student is asked to form a widely open front vowel (cv4), a gaping open (i.e., in phonological terms, *low*) back vowel (cv5) and an extremely close (that is, *high*), back vowel with exceptional lip protrusion (cv8). Having developed a feel for the four extreme corners of the 'vowel space', it is easier to fill in the lines that connect these four corners in equal steps until eight points along the periphery of the vowel space have been explored, categorized, named (cv1 through cv8) and assigned phonetic symbols.

In order to test the hypothesis that the perceived height and relative backing of the vowel correspond to the absolute high-point of the tongue along its midline, Jones lay a ferrous

chain along the center of his tongue and had x-ray stills taken of his head while producing each of the cardinal vowels in turn. The stills showed the midline tongue shape in detail, allowing him to plot the highest point of the tongue for each vowel with a single dot on a two-dimensional chart; all are pure monophthongs. With the dots connected, the resulting plot looked something like an olive, soon evolving into the familiar vowel quadrilateral adopted by the International Phonetic Association, a simplification of the chart having been suggested by Jones himself in the interest of pedagogical ease (1940). In gramophone recordings—no longer pressed—Jones himself demonstrated these vowels on a held note so that the lack of tongue and lip movement during the vowel could be heard clearly.

When teaching these monophthongal reference vowels, I place a dot on the chart for each vowel, numbering it (e.g., 'cv3') to the left of the dot for the spread vowels and to the right for the rounded vowels. In addition, I usually write the corresponding IPA symbols under the cv number for the benefit of those who have learned a transcription system, though I am careful to point out that one need not learn the transcription system in order to benefit from the exercise.

Once my students or actor-clients have taken an honest stab at the primary cardinal vowels, I teach what Jones called the secondaries, that is, reference vowels that match the eight primaries in tongue position but have the opposite lip rounding configuration—lip spread for lip rounded and rounded for spread. I simply write cv and the number (9 through 16) on the opposite side of the dot as I move around the chart a second time from upper left-hand corner to upper right-hand corner. By convention, we plot lip-spread vowels on the left and lip-rounded vowels on the right of each dot on the vowel chart reproduced in Figure 2. Next, I make a game of asking students to switch between primary and secondary cardinal vowel for the same tongue position. Finally, I ask them to hold the lip position steady while sliding between front and back for the same degree of openness or between close and open for the same front/back position. All of these maneuvers require mastering an unfamiliar pattern of coordination, much like rubbing the belly while patting the head. For the novice, this exercise can be much harder than it might seem. The above-described progression does not originate with me, but represents my own structuring of several traditional exercises in articulatory phonetics handed down to me through distillation by several preceding generations of phoneticians, and may indeed even pre-date Jones himself.

Through playing with these articulatorily defined peripheral reference vowels, the actor is sensitized to vowel parameters such as height (that is, relative openness), backing, lip

rounding, etc.—a set of sensations that does not always come immediately. Having raised consciousness about vowel dimensions in this way, we demonstrate the real beauty of the reference vowels: their usefulness as landmarks on our way to the vowels of genuine accents. When ready, I ask my actor-students to simply plot on a blank vowel chart each of the vowels of a known or new accent with reference to the cardinal positions.

Rather than plotting a dot and a phonetic symbol, they plot a word containing the vowel (in fact, a *lexical set keyword*—see below), marking rounding etc. with makeshift notation. At this point, I usually immediately find myself explaining how arrows can be plotted to show the directionality of diphthongs or triphthongs, genuine accents tending not to have many pure monophthongs. For example, I might demonstrate my low, back, long, rounded monophthong in the word *law* and suggest that we plot that vowel (with a dot or in transcription) just a bit higher and more centralized than CV13, only to have a student chime in that he or she says the word with a diphthongal vowel starting almost as high as CV15 and ending somewhat left of CV6.

In order to reduce the clutter added by diacritics, I often find it helpful to hand out a sheet with the IPA vowel symbols marked on a chart accompanied by four blank charts, one labeled 'short vowels', one 'long vowels', one 'moving vowels' and another 'vowels before coda r'. (See Figure 3.) The idea was that one would plot vowels on the vowel charts while watching a speaker produce them, thus making lip configuration a known

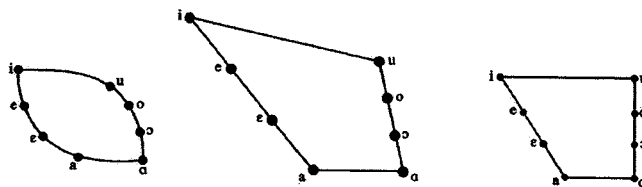


Figure 1. Cardinal vowel charts after Daniel Jones (1940:36-7) representing two-dimensional coordinates of the highest point of the tongue along its midline. The head faces left in all three charts. The chart on the left is the most accurate geometrically. The chart in the center represents "a compromise between scientific accuracy and the requirements of the practical language teacher [that the chart be easy to draw]" (1940:36). The more schematic trapezoid on the right represents a further simplification intended to ease the artistic burden on the "ordinary pupil" (1940:37fn), and is closest in shape and dimensions to the chart adopted by the International Phonetic Association (see Fig. 2).

2.2.1 Limitations of the cardinal vowels

As the reader will have noticed by now, the vowel quadrilateral is not ideal. In the best of all possible worlds, intellectual history would have given us separate vowel charts for lip-spread and lip-rounded vowels (Catford, 1981). The chart also conflates short, long and half-long vowels, forcing us to resort to messy diacritics. Its two dimensions tell us only about the tongue—and only about a single point on its midline at that, and nothing about voice quality, intrinsic fundamental fre-

quency (akin to pitch) differences among vowels, etc. Still, familiarity with these charts can increase awareness of relative vowel height and relative backing. Furthermore, (optional) familiarity with the IPA symbols may allow our students to someday take advantage of such published resources as the Longman Pronunciation Dictionary (LPD) which follows the IPA in its conventions (Wells, 2000). Whether or not our students ever learn the transcription system, and whether or not they master the cardinal vowels in one sitting (which, in all honesty, rarely happens), they will come away from the session with an awareness of

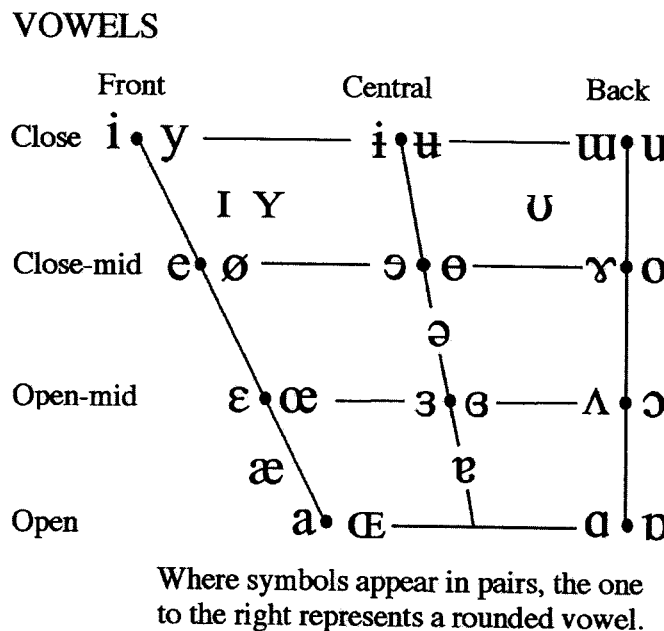
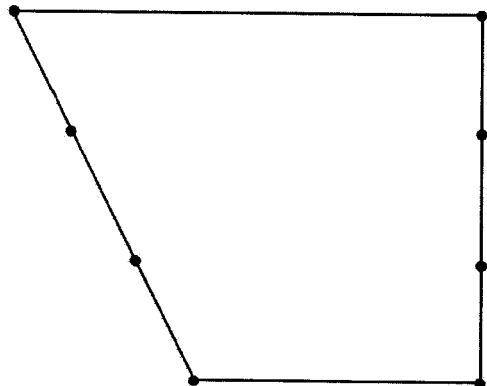


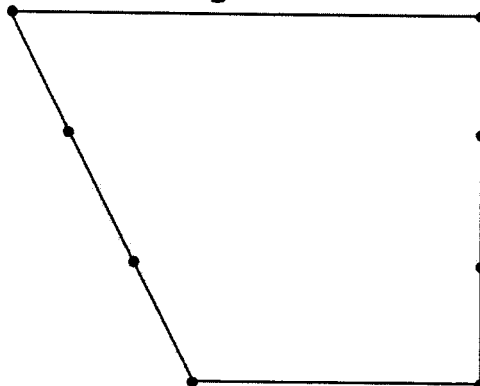
Figure 2. The vowel chart from the International Phonetic Alphabet (Revised to 1993, Updated 1996), available at no charge from the web site of the International Phonetic Association, Department of Linguistics, University of Victoria, Victoria, British Columbia, Canada (<http://www.arts.gla.ac.uk/IPA/vowels.html>).

Vowel Charts for _____

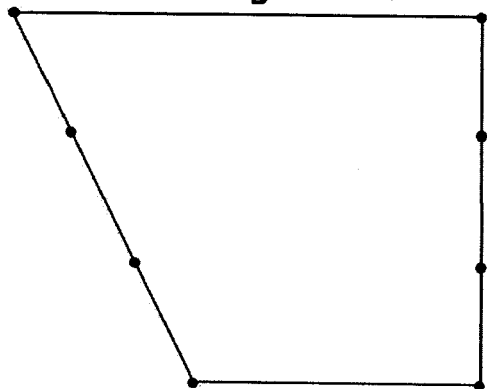
short vowels



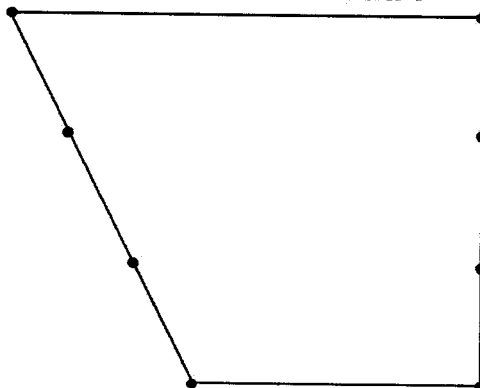
long vowels



moving vowels



vowels before coda r



lexical set keywords after J. C. Wells (1982).
Accents of English I, Cambridge Univ. Press.

- | | |
|-------------|------------|
| 1. KIT | 15. GOOSE |
| 2. DRESS | 16. PRICE |
| 3. TRAP | 17. CHOICE |
| 4. LOT | 18. MOUTH |
| 5. STRUT | 19. NEAR |
| 6. FOOT | 20. SQUARE |
| 7. BATH | 21. START |
| 8. CLOTH | 22. NORTH |
| 9. NURSE | 23. FORCE |
| 10. FLEECE | 24. CURE |
| 11. FACE | 25. happy |
| 12. PALM | 26. letter |
| 13. THOUGHT | 27. comma |
| 14. GOAT | |

vowels of the IPA (1996)

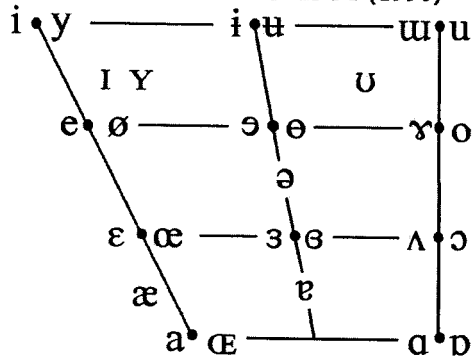


Chart courtesy International Phonetic Association, Department of Linguistics, University of Victoria, Victoria, BC, Canada.

Figure 3. Blank vowel charts, the vowel chart of the International Phonetic Association and lexical set keywords.

speech that they never had before. Below, I describe how students can then learn to apply their new gestural consciousness in learning to produce the vowels of an unfamiliar accent by plotting words rather than phonetic transcriptions on a vowel quadrilateral. The words I encourage them to plot are called *lexical set keywords*, and are motivated in the following section.

2.3 Lexical sets

2.3.1 Weaknesses of the sound-substitution approach

Traditionally, we trainers have presented adult actors with an inventory of sounds in a target accent and then asked them to substitute sound A for sound B or melody X for melody Y. Yet we have all had students complain of experiencing difficulty with these sound substitutions, even though they themselves may have implicitly requested instruction in these terms by asking us, “How do you pronounce ‘a’?” They ask such questions, because they assume that a simple list of sound changes (or unvarying letter-sound mappings if they are especially naïve) will throw them mechanically into the new accent. My job, as I see it, is to redirect the question because the sound-substitution approach creates three areas of difficulty:

- A. Speech is more than consonants and vowels. It has much to do with intonation (melody, tune), rhythm (tempo, rate), overall articulatory setting (placement), etc. This point must be obvious to our students at some level, but they may lack a rhetoric for asking about prosody and voice quality (in the broad sense of Laver, 1980). Lexical sets will not help with prosody and settings, but they do encourage a shift of focus from simple substitutions to mastery of a phonological system.
- B. Our students may not be equally familiar with the assumed base accent. Suppose we are hired to teach an old-fashioned, upper-class Southern British to a mixed group of Londoners, Glaswegians and Sydneyites. If we simply teach, for example, that the ‘y’ in the word happy is pronounced more like the vowel in ‘kit’ than like the vowel in ‘fleece’, immediately we run into a problem. The Londoner might slightly diphthongize the ‘y’, the Glaswegian might produce too open a vowel, and the Sydneyite too tense a vowel—all three missing the target vowel quality. We would have to give a different note to each actor—a nuisance when we are teaching a group.
- C. Even if we have already had the luxury of taking our students through a single uniform base (perhaps ‘standard’) system, say, in a conservatory program, there are times when the vowel system of the target accent may map in a

complex fashion onto the shared base vowel-system due to historical mergers or splits reflected in the target system. That is, sometimes a sound in the target accent does not map uniquely onto any sound in the new accent in a one-for-one substitution. We cannot simply tell our students, “A→B.” The problem of complex mappings is at the focal point of the following section.

2.3.2 Complex mappings

2.3.2.1 Mergers

In Middle English, words such as *horse* and *hoarse* were pronounced with relatively more open and relatively more close vowels, respectively, throughout the English-speaking world. Hence we find the sight-spelling *hoss* for *horse* but not for *hoarse* in rendering ‘Plantation Southern’ to name an accent that preserves the Middle English distinction between these words. However, the vowels in these words have merged so that *horse* and *hoarse* are now pronounced with the same vowel by people in many parts of the English-speaking world. One consequence of the merger is that, when teaching ‘Plantation Southern’ to a Californian, for example, we cannot simply say that the vowel in the word *horse* ‘goes to’ a lower, more spread, shorter, more nearly monophthongal vowel than is heard in the word *hoarse*. The learner will need to be told which other words pattern with *horse*, and which with *hoarse*. Here the mapping between words in the native and target accents is one-to-two, which leads us to the problem of the *lexical incidence* of a particular vowel in a particular word.

2.3.2.2 Lexical incidence and spelling

In the case of a one-to-two mapping, our students need our help if they are to extrapolate what they have learned about *horse* and *hoarse*, for instance, onto a larger set of words. That is, they need to know which words in their script rhyme with *horse* and which rhyme with *hoarse* in the target accent. We could just tell them which words are said how, but their ownership of the accent will be more profound if they have internalized the system. Fortunately, in the case of the *horse/hoarse* sets, it is relatively easy to undo the effects of the vowel merger because, with few exceptions, the conventional *spelling* of English preserves crucial information about Middle English pronunciation, and therefore about the pronunciation of these words in contemporary accents that are conservative with respect to this part of the vowel system. We can simply ask our students to memorize the following relation: words with spellings *oar*, *ore*, etc. pattern with *hoarse* while those with spellings *orC* etc., pattern with *horse*, where C stands for one or more consonants. Then, to state the matter in the simplest terms, when our students graduate from one accent that has not undergone the merger (say, ‘Plantation Southern’) to another such accent (say, conservative Hiberno-English), they

will already know which words group together. They need only learn how these vowel spellings are pronounced in the target accent. Because spelling often fossilizes vestigial information in this way, English orthography may be more helpful to us than to the average speller—a fact that I find heartening. It is perhaps a good thing for us that the spelling reforms of the 16th and 17th Centuries were not more successful. (See Dobson, 1968: 38-198.)

Sometimes lexical incidence is irregular. To pull out words that do not obey the sound-spelling rules, actors need a well-trained coach with a good ear who in turn must consult native speakers in the absence of accent-specific pronouncing dictionaries (e.g., Jones [1949 and earlier], Kenyon and Knott [1944], Wells [2000]).

2.3.2.3 Splits

Mergers as in *horsel/hoarse* are not the only possible type of challenge our students must confront when one vowel system maps untidily onto another. Sometimes spelling is less helpful. In the New York City area, many strongly localizable talkers pronounce words like *man* with a mid-open vowel or diphthong so that it sounds more like *men* than it would for, say, Londoners. This is not to say that native English-speakers from London could not hear the difference between the two words as pronounced by the stereotypical New Yorker. Make no mistake, the New York vowels have not merged by any measure, it is just that the vowel in *man* is produced with a vowel that is higher (closer) than the vowel in the word *mat*, for example, while the vowel in *men* may even be lower (more open) and more central than that in *man*. That is, to simplify matters, the set of closed syllables whose vowels are spelled ‘a’ (*mat*, *man*, *mash*, etc.), have split over time in these speech communities. By way of contrast, for Americans in certain other parts of the country, *man* and *mat* are both said with an identical vowel and thus remain together as a set as they did in Middle English. The specific vowel may be different from place to place, but the entire set of words hangs together. For example, in most of the American Far West and Lower Midlands, the vowel in *both* words is fairly open (as in the New York City pronunciation of *mat*). In the cities of the Inland North (Chicago, Detroit, Buffalo, etc.), *both* words have a vowel that is relatively close, that is, high (e.g., Labov, 1991), but perhaps more diphthongal and nasal for some talkers.

Unlike the case of vowel mergers discussed above where we saw that the spelling of the vowel can help us guess correctly which words will fall together, in the case of the *man/mat* split, London actors learning New York can virtually always guess the vowel correctly only if they observe that these ‘a’ words always have a higher vowel when followed by ‘n’ in the same syllable than when followed by ‘t’ in the same syllable. In

fact, one can list for students all the coda consonants that have the same effect on the preceding vowel as does ‘n’ and all the coda consonants that have the same effect on the preceding vowel as does ‘t’. Indeed, there really is no way to learn which word belongs with which set other than to commit to memory a list of conditioning consonants. Having to memorize how each phonetic environment conditions the split still beats having to memorize a long list of words. If we want to practice with word lists (or ask our native speaker consultants to read them for us), we can do so, but doing so need only serve as a way of assuring ourselves that the generalization about the effect of the coda consonants on the vowel is largely without exception.

Let us consider a case in point. As several Metro New Yorkers including phonetician Alice Faber have pointed out to me and as discussed on the American Dialect Society Mailing List (<http://listserv.linguistlist.org/archives/ads-1.html>), they can tell that Dennis Franz of ABC’s *NYPD Blue* has not mastered the ‘a’ split, thus *backhand* has two instances of the same vowel for him (Chicago-style, if you will) rather than a lower vowel in *back* followed by a higher vowel in *hand*. His character is supposed to be from Metro New York. He has got the right social register for his character, but the wrong regional pattern. This failure on the part of Mr. Franz’ speech coach is noticeable to locals, though, to be fair, it may be that the note was given and ignored.

2.3.3 How many lexical sets are there?

In order to explain why we need gestures and lexical sets, we must consider the alternative: *phonemes*. An average accent of English has perhaps a dozen distinct nuclear vowel phonemes in its inventory, depending how one counts. However, although two twelve-vowel systems, *Accent A* and *Accent B*, may both have /i/ and /e/, *Accent A* may use /i/ in the word *beat* where *Accent B* uses the vowel /e/.

Once we begin considering the entire inventory of the two twelve-vowel systems, we may discover that many regular diachronic sound changes in both accents have made our comparison of the two systems complex. We might have to draw a Venn diagram of the two systems in order to sort out which vowel-word correspondences intersect in A and B, and which do not. Doing so may give us more than twelve vowels, even though each accent on its own has only twelve. John Wells has laid out a scheme of such lexical sets and subsets of vowel-word (or, ideally, vowel-spelling) correspondences (1982). He bases his sets on a comparison of the vowel systems of one British and one American accent with which many of his readers are assumed to be familiar (many other accents being considered in the crafting of the lexical *subsets* as we shall see).

The lexical sets rest one leg on that accent family he regards as non-localizable within the south of Britain (Wells, 1982: 117), an accent that is arguably native to ‘up-market’ talkers from that region and acquired by elite schooling for others. That accent has passed under various names, including the opaque *Received Pronunciation* or RP, a name adopted by Daniel Jones to indicate that the speaker of this accent would be widely understood. In examining the introduction to Jones’ pronouncing dictionary (1940: x-ix), it would seem he was not advocating that social climbers learn to speak RP so that they, too, could be generally accepted, however. He was, in his own words, not a “reformer of pronunciation”. He said, “I do not regard RP as intrinsically ‘better’ or more ‘beautiful’ than any other form of pronunciation.” It seems RP’s first and most famous codifier never intended it to be regarded as a ‘standard’ (1949, §61). He recorded it because he believed that this accent of the South of England was readily understood in “most parts of the English speaking world” and because it happened to be “the only type of English pronunciation about which [he was] ... in a position to obtain full and accurate information.” Note that Jones regarded RP as a regional accent, associating it with “a majority of Londoners who have had a university education.” Historically, it has its roots in the speech of London, so its radiation from that center of influence should not be surprising. Perhaps the spread of RP among the upper and upper-middle classes throughout the south of Britain makes it only broadly localizable. As such, RP continues to thrive. Trudgill tells us that RP is alive and well, and keeping Estuary in its place as a more narrowly local non-contender (2001).

The other accent Wells has consulted in codifying his sets is so-called *General American* or *GenAm*. Perhaps referring to GenAm as a *family* of accents here will make the term marginally less objectionable to the many linguists who have marshaled evidence that there is no single, representative, non-localizable US accent of English. Here we are not concerned with the question of linguistic relativity, that is, whether there is such a thing as an unaccented talker (Esling, 1998), but with the existence of a non-localizable US accent. Aware of the debate, Wells nevertheless says, “A recognizably local accent in the United States can only come from the east or south (1982[1]: 118),” implying that GenAm finds its home in the Midwest and Far West (though Mencken goes so far as to say that GenAm extends eastwards all the way to the Connecticut river and beyond and northwards into Canada [1936: 356-67, 371]).

My every instinct as an American speech coach tells me that Wells overstates the case, so, before explaining how Wells uses GenAm in building the lexical sets that I am advocating the

reader teach, I feel compelled to distance myself from the notion of a *general* American accent. To begin with, there are many localizable accents outside of the east and south. I have heard many Westerners report the experience of having traveled outside their hometown and unexpectedly identifying by ear a hometown stranger. Perhaps these Westerners recognize familiar speech patterns by idiom or usage, or by phonetic details pertaining to stretches of speech longer than the phoneme (intonation, rhythm, voice quality, etc.). However, I suspect that even the detailed phonetics of Far Western and Midwestern vowels can be very regionally specific. In fact, speech teachers cannot even agree about matters as simple as whether words such as *cot* and *caught* are homophones in GenAm. They cannot agree because they have different GenAm archetypes in mind—those with one low back vowel and those with two.

If local Far Western and Midwestern talkers recognize ‘their own people’ by voice, then, we might think of GenAm as the family of accents that *outsiders* cannot localize. Alas, even this more limited definition fails. The cartoon and comedy industries thrive on the ability of national audience members to identify specific stereotyped Western character voices (“Fargo”, “Surfer Dude”, etc.).

Neither can we define GenAm as a highbrow social variant. In the US, the stereotype of an educated talker is not a pretty one—it is Back Bay effete or computer nerd, perhaps. In any case, although prescriptive traditions of grammar are taught in US schools, we have no analogous, consistent, unifying tradition of pronunciation training in US schools that would produce cross-regional uniformity only in scholastic achievers.

Alternatively, one hears GenAm defined as the accent of American English that we find unremarkable on the lips of a foreigner. Unfortunately, few foreigners ever reach anything close to this standard, so this hypothesis remains virtually untestable.

One also hears that many Americans who identify themselves (or are identified by others) as descendants of West African slaves may speak a non-localizable American English, or, at least, have fewer regional distinctions in pronunciation among themselves. (See Wolfram & Schilling-Estes, 1998, for a very accessible discussion of the roots and distribution of African-American Vernacular English features.) Suppose this claim of a non-localizable ‘black’ accent bears some truth (and I am *not* saying it does). Why has no one has ever thought to consider a ‘black’ accent of American English ‘general’ if it is indeed the most non-localizable native accent we have in the US? Clearly,

Lexical Sets Handout

Keywords	Typical Spellings, Simplified Subsets	Examples
1. KIT	iC	kid
2. DRESS	eC	bell, merry
3. TRAP	aC—'flat A' in RP & GenAm	tap, man, hand, marry
4. LOT	oC	God
5. STRUT	uC	cup
6. FOOT	uC, ooC, ouC	put, good, could
7. BATH	(a) aff, ath (<i>voiceless</i>), ass, aft, asp, ast, ask, augh; (b) ance, ant, anch, CVCand, ample; (c) alf, alv, an't; + many unpredictable, e.g., trans-	(a) staff, path, brass, shaft, last, ask, grasp, laugh (b) dance, plant, command, example; (c) half, halve, can't
8. CLOTH	(a) o, au + <i>fricative</i> — rhyme with THOUGHT in GenAm and Older RP; (b) o, — rhyme with THOUGHT in GenAm, never in RP; (c) or — same pattern as (b)	(a) off, soft, cross, gone, Australia; (b) moth, coffee; (c) orange, tomorrow, sorry
9. NURSE	ur, or, ir, er, earC	fur, work, fir, fern, earn
10. FLEECE	(a) e, ee, eCe (b) ea, oeC	(a) be, meet, these; (b) meat, phoenix
11. FACE	(a) aCe; (b) aiC, ay, ey, eiC, aig; (c) ea	(a) cake; (b) wait, bay, obey, rein, straight; (c) great
12. PALM	(a) native Anglo-Saxon al, a#, ah — 'broad A' in RP & GenAm; (b) a memorize foreign words	(a) calm, bra, blah; (b) Nevada
13. THOUGHT	(a) auC, ough, aw, al, alk; (b) aiC (other than 'k') — in England, rhymes with THOUGHT subset (a) or with LOT depending on region and generation	(a) caught, bought, jaw, all, talk; (b) fault, salt, also
14. GOAT	(a) o, oCe, oa; (b) ow, ol	(a) so, rode, road; (b) bowl, roll
15. GOOSE	(a) o, oo, oCe, ou (b) uCe, euC, ew, uiC, lew#, eaut	(a) do, too, tooth, move, group; (b) duke, feud, few, fruit, view, beauty
16. PRICE	iCe, i, y	write, ride, hi-fi, try, type
17. CHOICE	oy, oiC	toy, join
18. MOUTH	ouC, ow	south, cow
19. NEAR	(a; c) eer, ere, ier, ear; erV; (b) ierC, eirC	(a; c) beer, mere, pier, fear; serious; (b) fierce, weird
20. SQUARE	(a; c) are#, air#, ear#, eir#, ere#, ary; (b) arC	bare, fair, bear, their, there; Mary; scarce
21. START	(a; b) ar#, arC; (c) ar	(a; b) bar; part; (c) safari
22. NORTH	(a) or#, ar#; (b; c) orC, uar; aur	(a) or, for, nor, Thor, war (exhaustive); (b; c) horse, order, quart; Laura
23. FORCE	(a, bi, c) ore, oar#, oor#, our#; orC; orV; (bii) oarC, ourC	(a, bi, c) tore, roar, door, pour; fort; oral; (bii) hoarse, court
24. CURE	(ai, b, ci) oor#, our#; ourC; oori, ouri; (aii, cii) ure, urV, eur	(ai, b, ci) poor, tour; gourd; boorish, tourism; (aii, cii) pure, plural, Europe
25. happY	(a) y#, i#, ie#; (b) ee#, ey#, ea#	(a) city, taxi, talkie; (b) coffee, hockey, Chelsea
26. lettER	er#, or#, o(u)r#, yr#, ure#	tiger, author, harbo(u)r, martyr, figure
27. comma	a#, ia# — All non-native borrowings; Middle English had no final vowel corresponding to schwa	vodka, phobia

racist or perhaps racist-classist notions lurk within our notion of ‘generality’. Any kind of specificity seems to disqualify an accent as general unless the race and class are white and middle, respectively.

Clearly I question the existence of a non-localizable form of speech native to anyone in the US. (The nativeness criterion excludes broadcast and stage speech as contenders for the title *non-localizable*.) At best we can say that a talker may, in some contexts, use so few socially or regionally marked speech features that we tend to hear past his or her regional or social identity. Even so, GenAm speakers may indeed be localizable at a level at which most people are not attending. The perception of non-localizability, within this definition, implies that the talker’s voice simply manages to slip under the threshold of being remarkably distracting despite its localizable features. If there is no true non-localizable native speaker of GenAm, Wells has simply codified his own version of a cultural myth, albeit, out of descriptive rather than prescriptive motives. Nevertheless, even invented accents ‘exist’ at some level, so on those terms I accept Well’s GenAm family of accents as a construct on an equal footing with the RP family of accents, both of which are building blocks of the lexical sets.

2.3.4 Definition: Lexical sets

Having mapped out the lexical sets by comparing RP and GenAm vowel systems, Wells selects a *keyword* to name each of 27 major sets. The keyword is always a word that belongs to the set it names and, among other considerations, is not easily confused on hearing no matter who says the word to whom, provided each party is a native speaker of some accent of English. For example, the word FLEECE is chosen over *beat* to stand for words that are pronounced [i:] in RP and [i] in GenAm. *Beat* would be unsuitable as a keyword because some talkers may pronounce *beat* more-or-less as an RP or GenAm speaker might pronounce *bait* (Wells, 1982: 123).

The list of lexical set keywords is given in Table 1, with only the most common spellings indicated—a simplification I highly recommend for teaching actors, especially as exceptional spelling-sound mappings always arise, regardless.

Table 1. Lexical set keywords, typical spellings and examples after Wells (1982). Subset names appear in parentheses, and are from Wells. Those lexical sets whose subdivision has more relevance for historical linguistics than for present day accents of English are reorganized here, e.g., CURE (ai, b, ci) versus CURE (aii, cii). Some of Wells subsets are excluded altogether. ‘C’ stands for one or more consonants. ‘V’ stands for one or more vowels. ‘#’ stands for the end of a word (minus any grammatical endings). The last three lexical sets keywords (happy, letter, comma) are bi-syllables. We are concerned with the second syllable in each case.”

I ask my students to memorize by rote the lexical set keywords and the spellings that match each set. Where spelling is less helpful, as when two sets share a vowel spelling, I also require memorization of both sets of words along with their membership. (All sets are given in Wells, 1982, Volume 1.) For example, students familiarize themselves with the NORTH and FORCE lists if their native accents have merged the two sets. In fact, I use this example when first mentioning the lexical sets. A demonstration of the NORTH/FORCE merger in “Plantation Southern” seems to work better than a demonstration in conservative Hiberno-English for my American students. When teaching students from areas where the accents have never undergone the FOOT/STRUT split, the need for lexical sets is equally clear. As an American, I can guess with accuracy which words belong to FOOT and which to STRUT without a moment’s thought; many North Country speakers, for example, would need to learn by rote which words belong to FOOT and which to STRUT in order to acquire an accent in which the two sets are distinct. Even so, there is a lot less rote work involved than in working through the Skinner materials (1990).

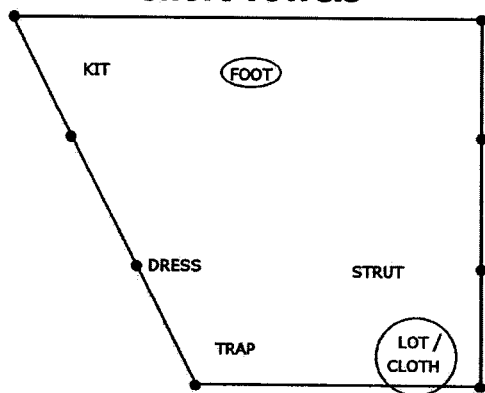
Although Wells has devised the lexical sets to codify linguistic reality, there is a necessary arbitrariness in the number and exact make-up of the sets. The Venn diagram of vowels one gets by settling on the union of two reference accents (in this case, RP and GenAm) produces a system that is not necessarily sufficient for every purpose. This is where the subsets come in. When I am teaching an accent that largely follows the Middle and Early Modern (“Elizabethan”) English practice of keeping words with *ee* spellings, for example, distinct from those with *ea* spellings (say, conservative Hiberno-English), I simply draw my students’ attention to the FLEECE(a) and FLEECE(b) subsets, respectively. Then my students can usually guess correctly whether a FLEECE word is pronounced more like *be* or *bay* in an accent with the split. However, I mention the subsets to actors as issues arise in the course of instruction. When I am asked about subsets that are not relevant to learning, say, the particular Irish accent under study, I simply say, “Don’t worry about those subsets.” I have yet to have an actor protest this directive. The subsets are discussed in detail in Wells, especially Volume I, Chapters 2 and 3 (1982).

Given the foregoing discussion, I believe the reader can imagine how learning the lexical sets and their subsets might help students internalize regular mappings between conventional orthography and the vowel system of any accent under study. This mapping is the consequence of the particular series of mergers and splits that have produced the target accent’s vowel system over a period of centuries. Whenever teaching a new accent to students who have learned the lexical sets, the trainer

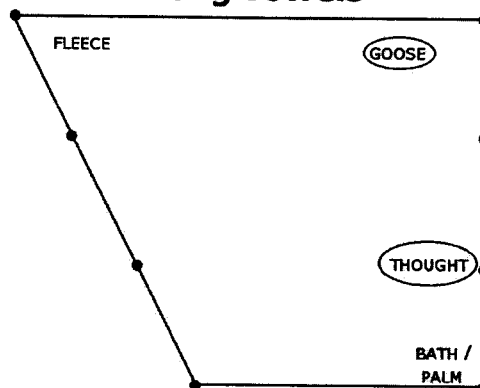
Vowel Charts for Modern RP

Impressionistic analysis © 2000-2003 Douglas N. Honorof

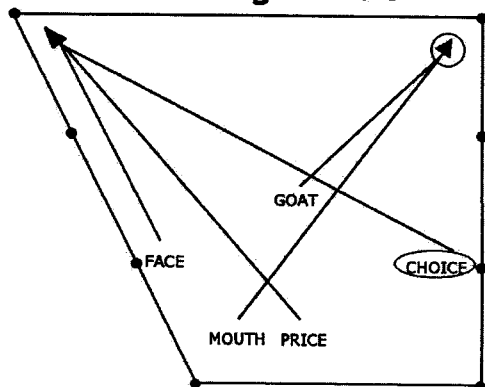
short vowels



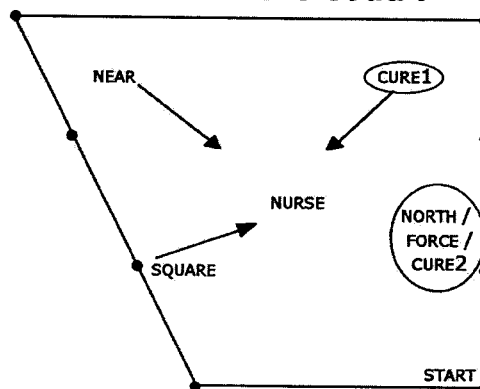
long vowels



moving vowels



vowels before coda r



lexical set keywords after J. C. Wells (1982).
Accents of English I, Cambridge Univ. Press.

- | | |
|-------------|------------|
| 1. KIT | 15. GOOSE |
| 2. DRESS | 16. PRICE |
| 3. TRAP | 17. CHOICE |
| 4. LOT | 18. MOUTH |
| 5. STRUT | 19. NEAR |
| 6. FOOT | 20. SQUARE |
| 7. BATH | 21. START |
| 8. CLOTH | 22. NORTH |
| 9. NURSE | 23. FORCE |
| 10. FLEECE | 24. CURE |
| 11. FACE | 25. happy |
| 12. PALM | 26. letter |
| 13. THOUGHT | 27. comma |
| 14. GOAT | |

vowels of the IPA (1996)

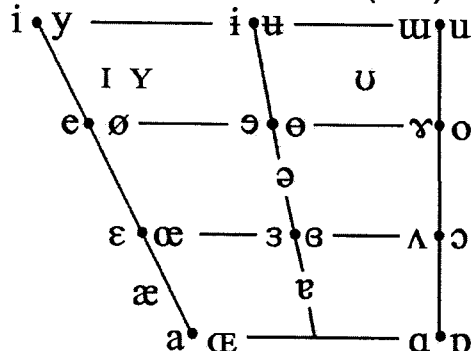


Chart courtesy International Phonetic Association, Department of Linguistics, University of Victoria, Victoria, BC, Canada.

Figure 4. The vowels of Modern Received Pronunciation. In a chronology of overtly prestigious Southern British accents, Modern RP would come between Older RP and Estuary. Lip rounding of any type is indicated with an ellipse for vowel nuclei, off-glides or entire vowels. The lexical sets happY, lettER and commA and other accent features not indicated on the charts ('smoothing', 'liquid u'/yod, consonants, rhythm, intonation, etc.) may be discussed and noted by students on the back of the sheet.

need only list the sets that are merged in the target accent, and make a note or two about specific subsets where relevant. Although the lexical sets and subsets help us remember which words rhyme with which in the target accent, the native-speaker or speech coach must still model the phonetic realization of each set in the target accent. One can imagine two accents with the same mapping of spellings and vowel categories but with different phonetic realizations. For instance, one can imagine contrasting an English accent heard in Australia with an English accent heard in New Zealand, choosing two accents, say, with essentially the same number and pattern of lexical sets and subsets, but in which the KIT set has a more close or open vowel, respectively. I state the phonetic realizations in terms of gestures, relating them to the articulatorily defined reference vowels we have already mastered. We simply plot the word KIT in different places on the two Antipodal vowel charts.

The lexical sets also help us keep track of variation among vowels that have arisen in some consonant environments but not others. For example, in RP there are many pairs of lexical sets that share a nuclear vowel, but in which one of the two sets contains words with a centring diphthong before coda 'r'. This is, in part, the reason we see 27 lexical sets for the RP/GenAm union of accents where GenAm alone has far fewer distinct vowels. In this connection, let us consider KIT versus NEAR. Wells keeps these sets distinct even though there is sometimes an 'r' following the vowel in the KIT set. *Mirror* is a KIT word, while *nearer* is a NEAR word. In RP as in many accents of English (unlike GenAm), the vowels in these words are different. The stressed syllable of *mirror* has a short monophthong in RP—roughly, [ɪ]—where *nearer* has a centring diphthong—roughly, [ɪ] + something like schwa (Wells, 1982[1]: 153). Similarly, in what may be the most famous potential merger of vowels before 'r', *marry* and *merry* belong to sets whose keywords have no coda 'r' (TRAP and DRESS, respectively), while only *Mary* has a centring diphthong in RP and therefore belongs to a set whose keyword has a coda 'r' (SQUARE). Such detailed information is provided here to help the speech trainer make sense of the lexical sets, though I rarely delve into such complexities with actors in the context of a production.

Through mastering the sound-spelling correspondences via the lexical sets, actors build reasonably reliable intuitions about

the history of English phonology. These intuitions are general enough to save them time and frustration whenever they approach a new accent of English down the road. The lexical sets can also be taught to foreigners learning to pronounce English, especially because the sound-spelling correspondence problem poses them especially great challenges. (See Flege, 1987, for a review of challenges largely peculiar to second-language accent work.) The spelling of English is irregular as charged, but not, as it turns out, totally unpredictable once the lexical sets have been mastered. However, foreign students benefit from memorizing all the word-set memberships, while native speakers of English need memorize only those where the mappings differ between their native accent and likely target accents. Indeed, in coaching native speaker actors for a production, I give out almost no paper. Traditional pronunciation guides may be helpful to the coach, but most actors treat them as noise and politely file them away in a drawer. Despite the slightly abstract nature of the linguistic approach I am advocating, I give each actor one sheet of paper introducing the lexical sets (see Table 1) and another containing blank vowel charts (see Figure 3) which I ask them fill in during the coaching session using whatever notation is meaningful to them. Actors are encouraged to make notes about consonants, rhythm, intonation, etc., on the back of the vowel-chart sheet. An example of a completed set of charts for a family of accents spoken in the south of Britain appears as Figure 4.

2.3.5 Limitations of the lexical sets

There are, of course, limitations to any set of words grouped by spelling. For example, there is some consonant-derived variation among vowels that the lexical sets help us address only indirectly. Thus we may need to advise actors that the consonant before the vowel dictates the use of yod ("liquid u") in the CURE and GOOSE lexical sets in virtually all accents that preserve historical 'yod'. For instance, in some American accents, we use a yod only and always after consonants that do not involve a tongue-tip gesture—a nice example of an instance in which a gestural description of speech is simpler and more intuitive than a list of consonants that happens to exclude sounds typically spelled with 't', 'd', 'n', 'l', 'r', 's', 'sh', 'z', 'ch', 'j' and the fricative consonant in such words as *measure*. However, the lexical subsets do help us sort out *which* CURE and GOOSE words never have yod after tongue-tip gestures in these accents. Specifically, most words that can start with a yod after certain accent-specific subsets of the tongue-tip consonants are spelled with *ew*, *eu*, *ui* and *uC* (CURE [aii, cii] and GOOSE [b]). Words spelled, for example, with *oo*, *ou* and *oCe* never had a yod historically, and so do not take yod in contemporary accents (CURE [ai, b, ci] and GOOSE [a]). Wells discusses this complex phenomenon in terms of 'yod-dropping' because he has diachronics in mind. I avoid such

terms when teaching actors because I do not want my students to misinterpret such a process-oriented term as an indication that I am assuming a single base accent as the starting point for the entire group of students. I would say, instead, “whether or not we start the vowel with a ‘y’ sound”.

The lexical sets also fail to give us any useful information whatever about irregularities. For example, we must memorize *ad hoc* whether words with the spelling *oo* belong to the GOOSE lexical set or to the FOOT lexical set. There is, unfortunately, no better solution to this problem.

In settling on RP and GenAm as the bases of the lexical sets, Wells has limited the number of lexical sets he needs to get the basic job done. Unfortunately, in working with actors, we teach many accents, including some with more conservative vowel systems than RP and GenAm. Looking at a wider sampling of accents, we find subsets that Wells discusses but does not officially list. For example, he does not distinguish between the two-to-five different spellings that encode the unmerged NURSE set in parts of Ireland and Scotland. In teaching an accent in which *fir*, *fern* and *fur* do not rhyme, for example (as they may not have rhymed for Shakespeare: Bough & Cable, 1993: 229), I draw the student’s attention to what could be thought of as supplemental subsets of the NURSE lexical set. These supplemental subsets are helpful because, unlike RP and GenAm, some accents have failed to obscure this older distinction through merger. This is one of the few cases in which the lexical sets do not provide us with sufficient grain, so I simply have my students plot the labels NURSE[ɪ], NURSE[e, ea] and NURSE[u] as mnemonics.

Note that, in *Accents of English*, Wells gives IPA transcriptions for the vowels of the lexical sets as he discusses each accent. In fact, the use of symbols may be somewhat confusing for the non-linguist because his transcriptions are sometimes broad (that is phonemic or representative, standing for a range of variant sounds) and sometimes narrow, rendering phonetic details. When Wells uses the IPA in the Longman Pronunciation Dictionary, on the other hand, he transcribes words fairly narrowly, thankfully, except for the vowel in DRESS—Wells follows Jones in using [e]—a perpetual source of confusion that has been discussed elsewhere (Barnes, 2000: 96).

The transcription-free system I use with actors who do not know the IPA does require an additional bit of work on their part. Having plotted lexical set keywords on vowel charts in place of vowel transcriptions, the actor cannot at a glance recall whether the vowel is round or spread. Because the location of the lexical set keyword on the vowel chart indicates only the highest point of the tongue, we must indicate with a

circle, color of ink or other method which vowels are produced with lip-rounding in the target accent. Information that would normally be indicated with a diacritic such as vowel duration is indicated in my four-chart system by choice of chart. The use of four charts also makes the entire system clearer and the keywords easier to read.

3.0 Comma Gets a Cure

Clearly there is more to mastering vowels than mapping spellings onto them. One needs to practice hearing and producing them in context. One could ask a native speaker to model the vowels and other features by reading aloud from Skinner’s shibboleth sentences, for example, but those sentences are stage-American-specific. To address this issue, my colleagues and I have devised a passage that includes the lexical set keywords and examples of as many subsets as practical and frequently useful. The passage also contains environments for a variety of interesting consonant phenomena (linking ‘r’, etc.), is narrative and non-technical in content, and includes grammatical constructions intended to elicit a range of vocal expression insofar as any read passage can do. The passage has come to be called *Comma Gets a Cure*. Importantly, the *Comma* passage is available for use in accent study without special permission from the authors and has been adopted as one of the texts of the International Dialects of English Archive (www.ukans.edu/~idea). An annotated version of the passage is reproduced in an appendix to the present paper. Since the first appearance of the passage on the IDEA website, I have discovered one context that is missing, namely, the environment /tn/ potentially produced with a glottal stop in place of or accompanying the tongue-tip gesture for the /t/ and with a potentially intruded vowel rather than a syllabic nasal (*Britain, mitten*). There must surely be other oversights. However, given the number of speech samples already collected, no revisions to the passage are being proposed at the present time.

4.0 Summary and future work

I have distilled into prose aspects of a process that I have found helpful in training adult actors, focusing especially on a simple way of teaching potentially complex vowel systems. The process I have described involves kinesthetic familiarization with articulatorily defined reference vowels and the mapping of those vowels onto lexical sets, that is, onto sets of historically or potentially rhyming words. I believe that, as a consequence of the specific analytical nature of the training—not simply the fact of any training at all—I see in my adult actors improved intuitions about sound-spelling correspondences (cf. Weiss, 1992). I have argued that the process I teach is quick and general enough to be learned by actors outside of conservatory settings, but also provides an excellent way to begin teaching general speech awareness to repeat clients and to

students with whom one expects to work long-term. The human ability to coordinate and perceive speech gestures has been well documented in recent years, though none of the claims made here regarding the specific approach I use in teaching accents to adult actors has been tested in any scientifically rigorous fashion. We have, however, developed a research program aimed at testing the hypothesis of passive competence that has been laid out above.

For added value, I have provided an annotation of a diagnostic passage—*Comma Gets a Cure*—a passage built around lexical sets, subsets and other accent-features. The passage has been adopted as one of two reading passages by the *International Dialects of English Archive* (IDEA) database which is searchable online at www.ukans.edu/~idea. The speech of many talkers has already been archived for IDEA using this passage, and many more samples are likely to become available. In addition to using the passage in making primary source recordings, there are no doubt many as yet unexplored ways of using the passage itself in teaching and coaching.

Acknowledgements

The author received support from National Institutes of Health Grant DC-03782 to Haskins Laboratories while writing the present paper. He thanks Barbara Somerville for providing inspiration and Alice Faber for comments on an earlier draft of the paper, but assumes all responsibility for its contents. He is also grateful to Rocco Dal Vera and Louis Colaianne for their encouragement. Please address correspondence to the author at DialectDoug@yahoo.com.

References

- Barnes, M. J. (2000). A critique of phonetic transcription in American actor training. *Voice & Speech Review* 1: 89-102.
- Baugh, A. C. & T. Cable (1993). *A History of the English Language*, Fourth Edition. Englewood Cliffs, NJ: Prentice Hall.
- Bernstein, N. (1967) *The Co-ordination and Regulation of Movements*. Oxford, Pergamon Press Ltd.
- Bronstein, A. J. (Ed.) (1998). *Conference Papers on American English and the International Phonetic Alphabet*. Publication of the American Dialect Society 80. Tuscaloosa, AL: University of Alabama Press.
- Browman, C. & L. Goldstein (1991). Gestural structures: distinctiveness, phonological processes, and historical change, in Ignatius G. Mattingly & Michael Studdert-Kennedy, Eds., *Modularity and the Motor Theory of Speech Perception* (pp. 313-338). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Catford, J.C. (1968). *A Practical Introduction to Phonetics*. Oxford, England: Oxford University Press.
- Catford, J.C. (1981). Observations on the recent history of vowel classification. In R. E. Asher & Eugénie J. A. Henderson (Eds.), *Towards a History of Phonetics* (pp. 19-31). Edinburgh: Edinburgh University Press.
- Catford, J. C. & David B. Pisoni (1970). Auditory vs. articulatory training in exotic sounds. *The Modern Language Journal* LV(7):477-481.
- Colaianne, L. (1994). *The Joy of Phonetics and Accents*. New York, NY: Drama Book Publishers.
- Dobson, E. J. (1968). *English Pronunciation 1500-1700*, Second Edition. Volume I, Survey of the Sources. Oxford, England: Oxford University Press.
- Esling, J. H. (1998). Everyone has an accent except me. In L. Bauer & P. Trudgill (Eds.) *Language Myths* (pp. 169-75). London: Penguin.
- Flege, J. E. (1987). A critical period for learning to produce foreign languages? *Applied Linguistics* 8:162-177.
- Fowler, C. A. (1996). Listeners do hear sounds, not tongues. *Journal of the Acoustical Society of America* 99 (3):1730-1741.
- Fowler, C. A. (1995). A realist perspective on some relations among speaking, listening and speech learning. In K. Elenius & P. Branderud (Eds.), *Proceedings of the XIIIth International congress of Phonetic Sciences*, Volume 1 (1993.2, pp. 470-477). Stockholm: Stockholm University.
- Gibson, J. J. (1966). *The Senses Considered as Perceptual Systems*. Boston: Houghton Mifflin.
- Goldstein, L. & C. P. Browman (1986). Representation of voicing contrasts using articulatory gestures. *Journal of Phonetics* 14, 339-342.
- Hughes, A. & P. Trudgill (1996). *English Accents and Dialects: An Introduction to*

- Social and Regional Varieties of English in the British Isles*, 3rd Edition. London: Arnold.
- International Phonetic Association (1999). *Handbook of the International Phonetic Association: A Guide to the Use of the International Phonetic Alphabet*. Cambridge: Cambridge University Press.
- Jones, D. (1949). *English Pronouncing Dictionary*, Eleventh Edition. (First edition published 1917). New York: E. P. Dutton & Co.
- Jones, D. (1940). *An Outline of English Phonetics*, Sixth Edition. New York: E. P. Dutton & Co., Inc. Originally published by Teubner, Leipzig, 1918.
- Kelso, J. A. S., B. Tuller, E. Vatikiotis-Bateson, and C. A. Fowler (1984). Functionally specific articulatory cooperations following jaw perturbations during speech: Evidence for coordinative structures. *Journal of Experimental Psychology: Human Perception and Performance* 10: 812-832.
- Kenyon, J. S. & T. A. Knott (1953). *A Pronouncing Dictionary of American English*. Merriam-Webster.
- Labov, W. (1991). The Three Dialects of English. In Penelope Eckert (Ed.), *New Ways of Analyzing Sound Change* (pp. 1-44). San Diego: Academic Press, Inc.
- Laver, J. (1980). *The Phonetic Description of Voice Quality*. New York: Cambridge University Press.
- Markham, D. (1997). *Phonetic Imitation, Accent, and the Learner*. Travaux de l'Institut de Linguistique de Lund, 33.
- Mattingly, I. G. (1972). Reading, the linguistic process, and linguistic awareness. In J. F. Kavanagh & I. G. Mattingly (Eds.), *Language By Ear and by Eye: The Relationships Between Speech and Reading* (pp. 33-147). Cambridge, ma: MIT Press.
- Mencken, H. L. (1936, reprinted 1999). *The American Language: An Inquiry into the Development of English in the United States*, Fourth Edition. New York: Alfred A. Knopf.
- Read, C., Y.-F. Zhang, H.-Y. Nie and B. Q. Ding (1986). The ability to manipulate speech sounds depends on knowing alphabetic writing. *Cognition* 24: 31-34.
- Saltzman, E. L. & K. G. Munhall (1989). A dynamical approach to gestural patterning in speech production. *Ecological Psychology* 1: 333-382.
- Skinner, E. (1990). *Speak with Distinction*. L. Mansell & T. Monich (Eds.). New York: Applause.
- Trudgill, P. (2002). Sociolinguistic Variation and Change, Ch. 16: *The sociolinguistics of modern RP*. Washington, DC: Georgetown University Press. Viewable on-line at www.phon.ucl.ac.uk/home/estuary/trudgill.htm.
- Trudgill, P. (1983). *On Dialect* (pp. 8-30). New York: New York University Press.
- Weiss, W. (1992). *Perception and Production in Accent Training*. *Revue de Phonétique Appliquée* 102: 69-82.
- Wells, J. C. (2000). *Longman Pronunciation Dictionary*, NEW Edition. Pearson Education.
- Wells, J. C. (1982). *Accents of English 1: An Introduction* (pp. 117-183). Cambridge: Cambridge University Press.
- Wolfram, W. & N. Schilling-Estes (1998). *American English: Dialects & Variation*. Cambridge, MA: Blackwell.

Appendix: Annotated Version of *Comma Gets a Cure*

Under the present formatting of our diagnostic passage, the first instance of each of Wells' lexical set keywords is set in small caps. The lexical sets help us classify the talker's vowels. Other vowel-related phenomena and potentially diagnostic consonant issues are underscored, though not necessarily the first time they occur and often only once. Care has been taken to include in the text the full range of consonants in all potentially interesting environments. Notes are offered sparingly on the right.

Comma Gets a Cure

by Jill McCullough and Barbara Somerville

edited by Douglas N. Honorof

© 2000 Douglas N. Honorof, Jill McCullough & Barbara Somerville. All rights reserved.

Passage

Well, here's a story for you: Sarah Perry was a veterinary NURSE who had been working daily at an old zoo in a deserted district of the territory, so she was very happy to START a new job at a superb private practice in NORTH SQUARE near the Duke Street Tower. That area was much nearer for her and more to her liking. Even so, on her first morning, she felt stressed. She ate a bowl of porridge, checked herself in the mirror and washed her FACE in a hurry. Then she put on a plain yellow DRESS and a FLEECE jacket, picked up her KIT and headed for work. When she got there, there was a woman with a GOOSE waiting for her. The woman gave Sarah an official letter from the vet. The letter implied that the animal could be suffering from a rare form of FOOT and MOUTH disease, which was surprising, because normally you would only expect to see it in a dog or a GOAT. Sarah was sentimental, so this made her feel sorry for the beautiful bird.

Before long, that itchy goose began to STRUT around the office like a lunatic, which made an unsanitary mess. The goose's owner, Mary Harrison, kept calling, "Comma, Comma," which Sarah THOUGHT was an odd CHOICE for a name. Comma was strong and huge, so it would take some FORCE to TRAP her, but Sarah had a different idea. First she tried gently stroking the goose's lower back with her PALM, then singing a tune to her. Finally, she administered ether. Her efforts were not futile. In no time, the goose began to tire, so Sarah was able to hold onto Comma and give her a relaxing BATH.

Once Sarah had managed to bathe the goose, she wiped her off with a CLOTH and laid her on her right side. Then Sarah confirmed the vet's diagnosis. Almost immediately, she remembered an effective treatment that required her to measure out a LOT of medicine. Sarah warned that this course of treatment might be expensive—either five or six times the cost of penicillin. I can't imagine paying so much, but Mrs. Harrison—a millionaire lawyer—thought it was a fair PRICE for a CURE.

Random notes on items of special interest

'I' as vowel, h dropping, SQUARE/DRESS before 'r'
unstressed 'been'; FACE before 'I'; GOAT before 'I'
stress placement (2); cluster simplification
quality of 'r', if any; yod after 'n, s'
yod after 'd'; str-smoothing
intrusive 'r' before C; NEARER with schwa
'I' as vowel; DRESS height before coda [st]
'I' in 'If'
KIT+'r'; aspirated plosive /_stressed V
't' as/with glottal stop
locative vs. existential 'there'; unstressed 'was'

linking 'r'

reduction to a syllabic nasal
'w'/'wh'
open vowel quality before g

flapping

stressed demonstrative 'that'
yod after 'I'; stress placement
SQUARE/TRAP before 'r'
vocative intonation
'ng'; 'h' before yod

'nt' with glottal stop
smoothing; 'ng' before '-ing'; yod after 't'
labiodental
vowel in suffix; smoothing
intrusive 'r' before a vowel
labiodental

labiodental
PRICE height/_ 't'
PRICE height/_ 'd'; stress on con-; smoothing

DRESS-KIT /_ nasal
'I' in 'I' + yod