The speech of nonnative, late English language learners usually exhibits an accent that results from transferring the phonological rules and segmental or suprasegmental features from their first language (L1) into their English speech. Second language (L2) learners may also create novel pronunciations of English sounds that are not part of their L1 (e.g., Best, 1995; Flege, 1995). Native-like attainment in L2 pronunciation by adult language learners is the exception rather than the rule (Bongearts, 1999), yet most adult learners wish to speak the L2 without an accent.

Advertising with the slogan “Speak like a native speaker – understand fast speech,” the authors of the web program SpeechInAction claim to make native-like attainment—in terms of speech perception and production—possible, a goal that likely appeals to a large audience of adult foreign language (FL) learners. The target audience of SpeechInAction is adult learners of English who have intermediate or higher proficiency. The overall content, speech samples, and exercises are appropriate for such learners. The program is designed to familiarize FL learners with a range of English accents so that they are better able to understand a range of everyday fast, spontaneous English speech. The authors cover a vast variety of dialects of the English language, including accents found in the US, the UK, Ireland, Canada, and New Zealand.1 SpeechInAction features five main modules, each containing 10 chapters. Completion of an entire module would take about 10 to 20 hours. The five modules are entitled: (a) “Listening and Pronunciation (British and Irish Voices)”; (b) “Listening and Pronunciation (US and Canadian Voices)”; (c) “Listening to Accents of the British Isles”; (d) “Listening to Accents of the USA”; and (e) “Dictations: Intensive Listening” (covering all major accents, including New Zealand). Each chapter is further subdivided into various units that target specific dialects.

One major asset of SpeechInAction is its logical and sequential organization. Overall, the program is organized by proficiency level, as some activities target intermediate learners and others target intermediate to advanced learners. Typically, the subsections move from more comprehensive listening activities (e.g., listening to longer dialogues and completing multiple-choice questions) to more specific exercises such as recognizing, distinguishing, and comparing dialects. The sequence of the exercises moves from perception to production, where the user typically compares his/her production with that of native English speakers. Each dialect unit closes with a review that displays a summary of the student’s work.

SpeechInAction has a variety of strengths and weaknesses. Strengths include its presentation of a wide variety of authentic speech samples and natural discourse. These samples provide a great
deal of practice for students who aim to determine, distinguish, and ultimately comprehend different dialects in the English language. The dialect units first present a stretch of discourse (either a monologue or a dialogue). If a learner is unable to comprehend certain passages of the discourse, he/she can access both a prose transcript as well as a detailed transcript of each speech unit which can be played separately. The discourse is then followed by multiple-choice exercises for either general or selective listening. In order to answer the multiple-choice questions correctly, the listener must employ both bottom-up and top-down processing strategies. That is, the listener must make use of contextual knowledge as well as information on linguistic features in order to sufficiently comprehend the message. The context is usually given at the beginning of each unit: “In this chapter we work with Catherine, a university lecturer who was born and brought up in Houston, Texas, but has lived and worked in France and England for twenty years. She was ‘blown away’ by French” (Module 2: Listening to Accents of the USA; Accent 1: Blown Away by French). Some exercises require less contextual information, as when the task is to notice the missing and linked sounds. According to Petersen (1991), pedagogical focus on both top-down and bottom-up processes improves L2 listening comprehension.

Whenever an answer to a question is entered, correction is immediately provided on the next page and the speech sample is then segmented into phrases and into single words. Thus, the learner may develop listening skills such as identifying words, phrases, or clause boundaries to derive meaning from the stream of L2 speech. The learner can work with the accompanying exercises, in which he/she can first record him/herself and then compare his/her pronunciation with two different dialects. The authors provide distinctive features for which the learner should listen, for example, the differences in the vowels and the consonants /t/ and /r/. The learner can further practice his/her ability to recognize different accents by listening to brief excerpts and matching each to its corresponding accent.

A weakness of SpeechInAction is that the listening experience is nonreciprocal or non-communicative, and it provides no visual input. A number of research studies by Hardison (e.g., 2003, 2005a, 2005b) indicated that the presentation of audio and visual cues is beneficial to speech comprehension accuracy by L2 learners. Hardison’s recent comprehensive overview (2007) of research studies on the visual element in phonological processing strongly emphasizes the value of facial cues in L2 processing. In addition to facial cues, gestural cues (e.g., hand gestures) appear to facilitate L2 processing. Specifically, L2 learners’ listening comprehension appears to be most accurate when the learners can make use of both hand gestures and facial cues (Sueyoshi & Hardison, 2005). Unfortunately, SpeechInAction provides auditory input only, and it neglects the substantial role of visual input which is naturally available in face-to-face communication. The program could therefore be improved by including video clips of the speakers.

SpeechInAction encourages its users to record themselves when producing given words or phrases to practice their pronunciation and to enhance their ability to detect and distinguish various dialects. Furthermore, learners are encouraged to self-assess their production on a scale of 1-3 (Yes, No, Try Again) for questions such as “I have pronounced the short vowels accurately/on time” or “I have produced the speech units at original speed.” The
advantages of self-assessment coupled with the goal of producing speech at fast, native-like speed are not quite apparent to me. If learners attempt to produce rapid speech, their spoken language may lose intelligibility. Rather than testing whether they have produced rapid speech, learners could work with a partner to judge whether their language is intelligible, which could be of a greater value to their conversational use of English with native speakers. In fact, research studies have demonstrated that L2 learners perceive faster speech to be less accented, even if the speech was less intelligible overall (e.g., Munro & Derwing, 2001). I would suggest that instead of, or in addition to, having learners self-assess their speech production, the software program could allow learners to see their speech on a spectrogram, which would allow for various displays (vowel formants, primary stress, segmental and suprasegmental features, and prosody). My suggestion is partly based on the findings of Hardison (2004, 2005a), which demonstrated that computer assisted prosody training2 reliably improves L2 learners’ pronunciation. Though the developers do not claim that SpeechInAction will enhance learners’ L2 prosody, the inclusion of visual prosody displays would serve as a beneficial accessory to enhance learners’ ability to detect the variation in the prosody and discourse intonation of different English dialects.

On several occasions, the authors claim that “the better you can imitate an accent or dialect, the better you will be able to understand both the accent itself and the wide range of pronunciations of other accents” (Introduction page). There is no empirical evidence to support this claim, and speakers might be able to comprehend and distinguish a variety of accents while not necessarily being able to imitate the various dialects.

As mentioned above, SpeechInAction is advertised with the slogan “Speak like a native speaker – understand fast speech.” Although adult learners will most likely not end up speaking exactly like native speakers (Bongearts, 1999; Flege, Munro, & MacKay, 1995; Munro & Derwing, 2002), training with this software may improve their pronunciation and refine their listening comprehension skills. The developers of SpeechInAction also emphasize the benefits of unscripted speech, thereby putting students of English in touch with real speech and a variety of accents.

In addition to the philosophy of motivating and rewarding the learner, the authors strongly emphasize the variety and authenticity of L2 speech samples they use. Although the web program is not interactive, and does not encompass a communicative approach, it does take an inductively and deductively balanced teaching approach. In some subsections, the authors provide an explicit explanation of the differences between two dialects, whereas other subsections direct learners in formulating their own “rules” and perceptual categories based on authentic speech samples. For example, learners have to notice the difference between sounds or cross out letters which are absent or nonaudible in spoken language. The developers seem to value analytical and metacognitive skills, as users have to analyze (among other features) vowel qualities, articulation of consonants, and segmental stress (the sample chapter does not display any analysis of suprasegmental features). Although the analysis of phonetic features may raise learners’ awareness of those in perception and production, overanalyzing language by thinking too much about its forms and consciously lingering on the phonetic domains tend to impede graduation to automaticity (DeKeyser, 1997).

Furthermore, whenever the learner is
recording his/her own language production, he/she is encouraged to compare this production to that of (an) English native speaker(s). The authors obviously value the “native speaker model.” Yet, whether the native speaker’s speech is the appropriate model in language teaching has been a matter of some debate (Davies, 1991; Medgyes, 1992). As I have pointed out on several occasions, native-like speech production might not be a realistic goal for adult language learners, and it is doubtful that the imitation of native dialects helps learners to become accentless English users.

By using *SpeechInAction*, the learner gains experience in listening to a variety of speakers who speak different dialects in various speech situations. This exposes the learner to the reality of language variation. The overt emphasis on native-like perception and production seems to be unrealistic and, in most contexts, inappropriate for adult L2 learners and/or L2 teaching. A better objective of the program would be to concentrate on enhancing the learners’ intelligibility and comprehensibility rather than focusing primarily on accentedness and “native-like speech rate.”
References


Endnotes

1Speech in Action does not include Australian or South African dialects.
2Training in which learners can follow along and compare their pitch contours with that of native speakers.

(Submission received 15 May 2009)
(Revision received 27 July 2009)
(Revision accepted 30 July 2009)