

December 2020

Evaluating a Speech Training Software Program Called NativeAccent Based on Empirical Studies

Farideh Nekoobahr

University of Houston, friedanekoo@gmail.com

Jacqueline Hawkins

University of Houston, jhawkins@uh.edu

Kristi L. Santi

University of Houston, klsanti@central.uh.edu

Janeen R. S. Antonelli

University of Houston, jrantone@central.uh.edu

 Part of the [Adult and Continuing Education Commons](#), [Bilingual, Multilingual, and Multicultural Education Commons](#), [Curriculum and Instruction Commons](#), [Educational Assessment, Evaluation, and Research Commons](#), [Educational Methods Commons](#), [Educational Technology Commons](#), [Language and Literacy Education Commons](#), [Online and Distance Education Commons](#), and the [Teacher Education and Professional Development Commons](#)

Johanna Leigh Thorpe

University of Houston, thorpe.johanna@gmail.com

Find similar works at: <https://stars.library.ucf.edu/jele>

University of Central Florida Libraries <http://library.ucf.edu>

This Article is brought to you for free and open access by STARS. It has been accepted for inclusion in Journal of English Learner Education by an authorized editor of STARS. For more information, please contact STARS@ucf.edu.

Recommended Citation

Nekoobahr, Farideh; Hawkins, Jacqueline; Santi, Kristi L.; Antonelli, Janeen R. S.; and Thorpe, Johanna Leigh (2020) "Evaluating a Speech Training Software Program Called NativeAccent Based on Empirical Studies," *Journal of English Learner Education*: Vol. 11 : Iss. 1 , Article 7.

Available at: <https://stars.library.ucf.edu/jele/vol11/iss1/7>

INTRODUCTION

Unlike a few years ago, when utilizing digital media in English language classes was optional, or when many English language programs were skeptical about the effectiveness of technology-supported pedagogy (McClanahan, 2014), multimedia and online technology sources are now indispensable instructional tools for teaching English. The main reason for such integration is the close relationship between globalization of English and digitization of education (Lotherington & Jenson, 2011). In other words, while English was only an international language in the pre-digital era, it is now also the language of technology which makes English more globalized than before. By the same token, applying technology to teaching English, especially in recent years, is due to a number of factors such as “availability of technology, convenience,” [and] consumers’ need” (Lai, 2017, p. 21). While technology is becoming more available and more convenient than in previous decades, it is now considered not only a tool but a foundation in many aspects of life, including education and language learning.

Given this significance of technology in language learning, it should be noted that “in most high quality language courses beginner students will learn to pronounce the phones of the target language that are not present in his native language” (Eskenazi, 1996, p. 1465). Proficiency in listening, speaking, and pronunciation can directly enhance English Language Learners’ (ELLs) scores on speaking and listening sections of standardized tests and improve their overall literacy. By contrast, deficiency in those skills may not only impede ELLs’ academic achievements, but it may also damage their self-esteem and intensify their acculturative stress as they are not able to efficiently communicate and interact with others. Among all language areas, pronunciation is the most marginalized in the curriculum of

English language programs (Caballero & Thompson, 2016; Gilbert, 2018). To fill this gap, English language programs should provide training for their faculty to teach all language areas systematically by applying research-based pedagogical approaches and by allocating appropriate time for each specific skill. It is hoped that technology can facilitate this process. For that reason, using a Computer Assisted Language Learning (CALL) platform that is research-based can be a rational solution. The present article introduces and evaluates **NativeAccent**, an innovative and systematically designed language-training software program that focuses on pronunciation. It is also worth noting that while there is a wide range of language-training tools on the market, it is not easy to find the one that is time-effective, cost-effective, and designed based on empirical studies. The purpose of evaluating this software is to bring awareness to educators and ELLs of this platform that meets their various needs.

BACKGROUND

What Is NativeAccent, and How Does It Work?

NativeAccent is an online English language assessment and training software program for ELLs produced by Carnegie Speech Company at the Language Technologies Institute of Carnegie Mellon University (Carnegie Speech, 2018c). NativeAccent was originally devised as the Fluency Project in 1996 and was developed into a language-learning technology tool or a “pronunciation training system” in 2001 (Pelton, 2012a, para. 2). Since then, it has been modified and improved based on the needs of its users and in line with technological advances. It should be noted that this article focuses on NativeAccent version 3 (v.3). This section delineates how NativeAccent v.3 is designed based on the following highly tested instructional strategies: an *initial assessment*, *training sessions*, a *final reassessment*, *repeated measurement*,

graphic presentation and systematic and rigorous interpretive guidelines, and individualized dosage.

The Initial Assessment

Before taking the initial assessment, the Carnegie Speech website provides users with a step-by-step tutorial and a roadmap to using the software, including instructions on creating an account, adjusting the microphone for noise reduction, and taking the assessment. This tutorial is an integral part of the instruction since the subsequent personalized curriculum materials are built upon this information. For instance, the system needs to know the language of users when they create their accounts so that it can provide more featured exercises for specific languages. After completing the instructions and configuring the system, users take the initial assessment, which contains 89 items on phonemes, fluency, and grammar. They answer each item on the assessment section by clicking on the record button and reading the sentences into the microphone. Then they stop recording and proceed to the next item. For grammar questions, users need to first choose the right answer and then record it into the microphone. Thus, users' pronunciation, grammar, and fluency skills are evaluated in the initial assessment.

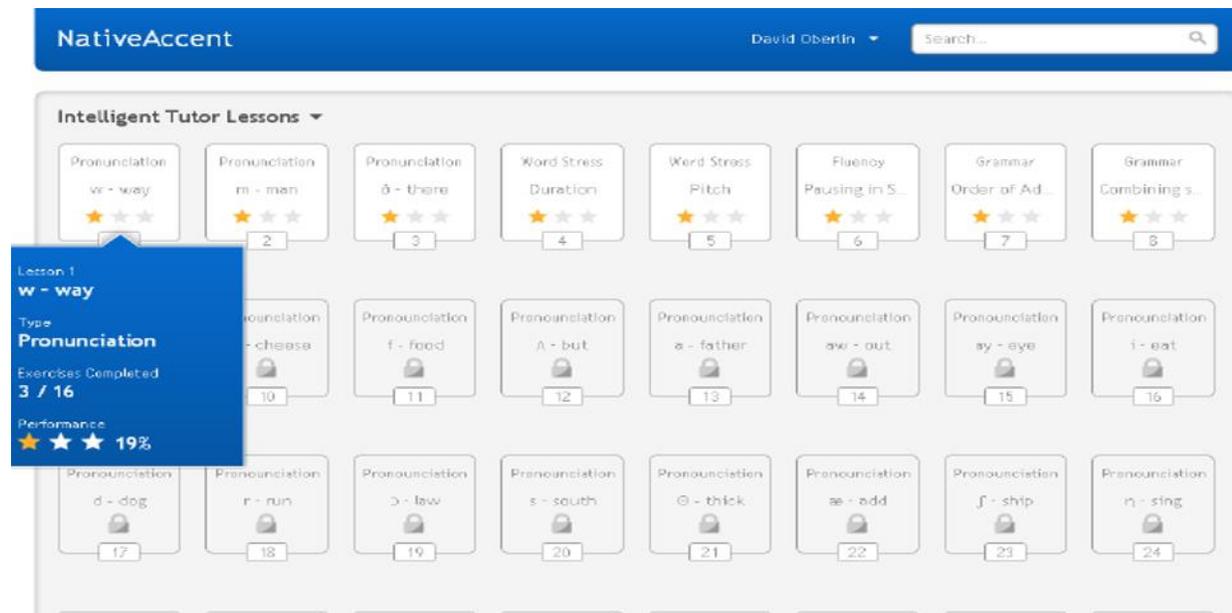
Training Sessions

Intelligent Tutor

One of the most important features of NativeAccent v.3 is its Intelligent Tutor which is designed based on the Intelligent Tutoring System (ITS). Through this system, each learner is provided with a personalized curriculum based on the data gained from the initial assessment, the learner's native language, as well as his or her current English language skills. In this sense, it focuses on learners' knowledge by being knowledge-centered (IRIS Center, 2017).

Automatic Speech Recognition

Another important feature of NativeAccent is its Automatic Speech Recognition (ASR), which pinpoints users' errors on phonemic pronunciation, fluency, and spoken grammar skills based on their recordings. Thus, the Intelligent Tutor receives the data that are detected and analyzed by ASR and accordingly creates personalized training sessions based on the errors found in the initial assessment. The training sessions include phonemic pronunciation, fluency, and grammar exercises presented in boxes as shown in Figure 1. The boxes are prioritized by the Intelligent Tutor based on the importance of the errors that need to be corrected by learners. Each box itself contains a number of exercises that take approximately 15 minutes for users to answer. Once the exercises are completed successfully in one box, another box is unlocked and ready for practice. It is worth mentioning that only a few boxes are available for users at a time, and the other ones are locked to guide users to practice them in order. Having a plan ahead of them and being aware of the amount of work done, as well as the amount of work to be accomplished, can guide learners to be goal oriented and to "stay on track" (Carnegie Speech, 2018c). Having such a roadmap can be a motivating and engaging factor that encourages learners to be focused and be responsible for their own learning (Mawson, 2017). It also creates a gamelike condition enticing learners to challenge themselves.

Figure 1*Intelligent Tutor Lessons*

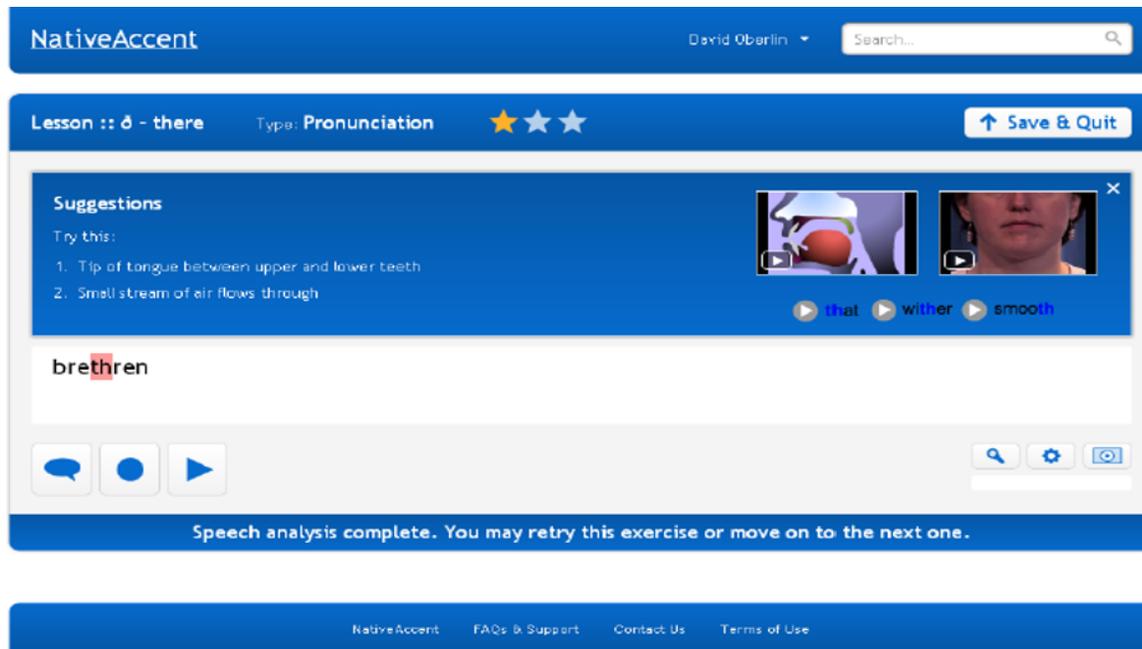
(From “Mining pronunciation data for consonant cluster problems,” by Pelton, 2012)

Throughout the training sessions, the ASR detects pronunciation errors and where they are made by using its “patented pinpointing technology” (Carnegie Speech, 2018a) which “compares a user’s pronunciation to a statistical model of native speakers” (Pelton, 2012a, para. 2). Users can see their errors highlighted as depicted in Figure 2. The ASR and the Intelligent Tutor analyze the errors and provide immediate feedback and detailed instructions on how to correct them. The feedback is “multi-modal” (Pelton, 2012b, p. 10), that is, it comes in different forms such as text (shown as *suggestions* in Figure 2), graphics of speech organs accompanied with audio, an audiovisual aid of a model speaker illustrating how the sounds are produced, as well as written and audio examples highlighting the specific errors made by learners. Users can also listen to their recordings and compare them to their corrected errors made by model speakers as many times as they wish. Then they are guided to correct their

errors, record them again, and compare their recording to the model speaker. Any recording is automatically saved in the system and is available to users and trainers.

Figure 2

Pinpointing Feedback in NativeAccent



(From “Mining pronunciation data for consonant cluster problems,” by Pelton, 2012b)

The Intelligent Tutor in NativeAccent v.3 also provides learners with corrective feedback through “a reporting mechanism” (Eskenazi et al., 2007, p. 124) such as progress charts and bar graphs that are automatically created for all the recordings made by learners. Figure 3, for instance, demonstrates a summary assessment report for two applicants: Applicant One showed *needs training* on Word Stress, *failed* on Pronunciation & Clarity and Fluency, and *passed* the Grammar section, whereas Applicant Two *passed* all concepts. The progress charts, which also show the amount of work done by users, can be created for individual exercises or in the form of a summary report on all exercises. The reporting mechanism not only helps the trainees to be

aware of students’ progress, but it also can engage and motivate them to remove their errors. Furthermore, the adaptable reporting mechanism helps teachers/trainers to be mindful of their students’ needs. It also gives them the ability to create formative and summative reports for one learner or a group and to compare students’ progress levels to each other.

Figure 3

Summary Assessment Report

Name	Date/Time	Length of Assessment	Word Stress	Pronunciation & Clarity	Fluency	Grammar
Applicant, One	Tuesday, December 8, 2009, 01:37 PM	17 mins 43 secs	2/3 (Train)	14/38 (Fail)	1/3 (Fail)	11/13 (Pass)
			Pitch-rising	Below Train Threshold	Below Train Threshold	Verbs-Present Perfect vs. Past Perfect
Applicant, Two	Friday, November 27, 2009, 02:19 PM	19 mins 22 secs	3/3 (Pass)	35/38 (Pass)	3/3 (Pass)	12/13 (Pass)
				i - leaf u - shoe ū - shq uld		Prepositions-Multi-Word Prepositions

(From “Mining pronunciation data for consonant cluster problems,” by Pelton, 2012b)

Although providing corrective feedback can facilitate the learning process, it is not sufficient; feedback is more effective if it is immediate, “understandable,” “actionable,” accompanied with models for the learners to follow, and systematically organized in the form of a curriculum based on the personal needs of users (Carnegie Speech, 2018b, para. 6).

Additionally, learning takes place more effectively if after providing corrective feedback, the correct response is elicited from the learner (Ellis & Bogart, 2007). Through its ASR technology as well as the ITS, NativeAccent v.3 provides its users with immediate, corrective feedback. The ASR detects the errors through its *error detection system* and then guides the learner in a “self-discovery” (Eskenazi et al., 2007, p. 125) process to correct them. Through the *self-discovery*

process, learners are trained in how to listen and how to correct their errors (elicitation) through explicit illustrations, which can result in improvement in both listening and pronunciation skills.

In addition to providing learners with immediate, corrective feedback and model speakers, NativeAccent is self-paced. Namely, it gives users the opportunity to manage the personalized training curriculum based on their abilities as well as the amount of time they allocate to practicing. The button labeled *Save & Quit* in Figure 2 (on the right-hand side, above the model speaker), for instance, shows that learners can stop the process, save their information, and resume later. Therefore, it is not expected that all learners spend the same amount of time and have the same speed during the process since the training curriculum is adaptable and customizable. Customer support is always available, making it possible for individual users to practice independently; however, the best results are produced when the presence of a trainer or instructor can facilitate and accelerate the whole process. These options also lend themselves to a single- case research design that is described below.

The Final Reassessment

Following the training sessions, individuals or groups are given the final reassessment to evaluate the extent of their progress compared to the results of the initial assessment. This process can provide a summative pre-test and post-test option for users and/or instructors.

NativeAccent and Single-Case Research Design

Single-case research design (SCRD) is “useful for investigating nuances to intervention effects” (Sheridan, 2014, p. 306). Since the major focus of NativeAccent is on improving pronunciation, it provides a variety of exercises and guidelines on the intricacies of English pronunciation rules and on many of those sounds that might not exist in users’ native languages. The important features of SCRD that are applicable to NativeAccent are as follows:

(a) *repeated measurement* (Sheridan, 2014), (b) “graphic presentation [and] systematic and rigorous interpretive guidelines” (Sheridan, 2014, p. 306), and (c) “individualized dosage” (Kaiser, 2014, p. 316).

Repeated Measurement

NativeAccent repeatedly measures its learners’ progress through the Intelligent Tutor. This is carried out through ABAB design. In phase A (initial assessment), the ASR assesses each user’s phonemic pronunciation, fluency, and grammar skills and pinpoints the user’s problem areas. Based on the errors made in the initial assessment, the user’s current level of spoken language, and his or her native language, the Intelligent Tutor provides training exercises presented in boxes. In phase B, for every error made in each exercise, the user receives an intervention in the form of corrective feedback and instructions in how to correct them, which paves the way for the third phase (phase A) where the intervention is withdrawn, and the user practices independently. The reappearance of the intervention (phase B) depends on the user’s performance, and this cycle is repeated until errors are removed.

Graphic Presentation and Systematic and Rigorous Interpretive Guidelines

“Intensive graphic presentation and visual inspection, including the use of systematic and rigorous interpretive guidelines” (Sheridan, 2014, p. 306) is considered another important feature of SCRD. The Intelligent Tutor in NativeAccent constantly provides guidelines for students to follow through its “multi-modal” (i.e., “aural, written, and pictorial”) feedback system (Pelton, 2012b, p. 10). The progress charts and graphs also show “accurate measurement” (Kaiser, 2014, p. 316) by providing detailed information on users’ performance on assessments, all exercises, type of exercises completed, and amount of time spent on each exercise.

Individualized Dosage

Based on users' performance on the initial assessment, different exercises are systematically assigned to each learner and matched with his or her specific needs and speech patterns. The online Intelligent Tutor also provides models for users and monitors their progress continuously. In doing so, it focuses on learner-centeredness by providing a different curriculum path for its different users, and accordingly, provides each learner with individualized feedback.

Conclusion

As can be seen, NativeAccent v.3 is designed based on a highly systematic structure (an *initial assessment*, training sessions, a *final reassessment*, *repeated measurement*, *graphic presentation and systematic and rigorous interpretive guidelines*, and *individualized dosage*), pinpointing technology, and speech recognition technology, all through which it presents a customized curriculum to each learner.

ISSUES, CONTROVERSIES, PROBLEMS

This section evaluates NativeAccent v.3 based on the four empirically-supported theories of **Universal Design for Learning (UDL)**, **Learning Science (LS)**, **Intelligent Tutoring System (ITS)**, and **Automatic Speech Recognition (ASR)**.

NativeAccent and Universal Design for Learning

While the principal goal of Universal Design for Learning (UDL) is providing “equal opportunities” and “effective instruction to all learners” (National Center on Universal Design for Learning [UDL], 2014), the question is whether NativeAccent can achieve this goal, and if so, how it is done. The following section elaborates on how the three UDL principles of *multiple means of representation*, *action and expression*, and *engagement* as well as the four UDL curricular components of *goals*, *materials*, *methods*, and *assessment* (IRIS Center, 2017;

National Center on UDL, 2014) are applied to NativeAccent.

NativeAccent and UDL Principles

Based on the first UDL principle of *multiple means of representation*, information is presented in various ways to be available for all learners (IRIS Center, 2017; National Center on UDL, 2014). NativeAccent presents information through feedback based on the specific errors made by users. In other words, after each exercise in the training sessions, the online Intelligent Tutor provides its various users with multi-modal feedback/information such as text, graphics, audio, and video in the form of a model speaker and thus follows this principle. The second UDL principle focuses on providing *multiple means of action and expression*, meaning that students can express what they learn in different ways. While NativeAccent provides models for each new lesson by applying scaffolded instruction, and students practice their pronunciation through individual sounds, words, or sentences, they have to correct their mistakes through speaking only, as NativeAccent is a pronunciation training tool.

The third UDL principle aims to provide *multiple means of engagement* in order to maintain different learners' engagement and motivation. The Intelligent Tutor in NativeAccent can motivate and maintain learners' engagement by providing corrective and immediate feedback in multiple ways, such as text, graphics, audio, model speakers, and progress charts. The Intelligent Tutor also guides learners to be goal-oriented by showing them a roadmap in the form of boxes of exercises (as shown in Figure 1) to answer. This helps users to be engaged in the learning process and responsible for their own learning.

NativeAccent and UDL-Based Curriculum Components

A UDL-based curriculum hinges on four major components: *goals, materials, methods, and assessment*.

Goals

The first component is setting *clear and measurable learning goals* so that all students can achieve them (IRIS Center, 2017). Setting goals in NativeAccent is done through the initial assessment. After the assessment, each learner receives a personalized curriculum. The curriculum path is clearly presented through the training sessions, which are arranged based on their importance for each learner. Additionally, the goals are measurable, meaning that users are reminded of their errors by getting immediate feedback for each exercise and are assisted by the Intelligent Tutor in how to correct them. Therefore, learners are aware of the specific steps as well as the major goal of the training. By providing a clear individualized roadmap for each learner, NativeAccent acts like a navigator facilitating and accelerating a student's language learning process.

Materials

The second component of a UDL-based curriculum is providing “flexible” and “well-designed” instructional *materials* to meet the needs of every individual learner (IRIS Center, 2017; National Center on UDL, 2014). The ITS of NativeAccent presents corrective feedback and instructions in the forms of text, audio, and video, and each learner can access the personalized curriculum through a series of boxes as depicted in Figure 1. The materials in the training sessions are designed based on the results of the initial assessment—they are flexible and vary based on each learner's specific needs. Materials are also prioritized, and therefore *well-designed*, according to the importance of the errors that need to be corrected by learners.

The instructions and models are personalized based on the specific mistakes made by each learner; therefore, all training sessions, as well as instructions, are unique for each learner. For instance, for one learner the order of pronunciation exercises might be /b/ as in *bag*, /tʃ/ as in

chair, /ð/ as in *they*, and /dʒ/ as in *job*; for another student, the training exercises might be completely different such as /ʊ/ as in *could*, /u/ as in *shoe*, /au/ as in *out*, and /d/ as in *desk*.

Methods

The third component of a UDL-based curriculum is applying various instructional *methods* to meet the needs of all learners who have different abilities and backgrounds (National Center on UDL, 2014). In addition to text, the Intelligent Tutor in NativeAccent presents its training through audio and video. It also provides numerous examples and highlights users' errors to make them aware of problem areas. Additionally, the Intelligent Tutor activates learners' prior knowledge by guiding them to do the exercises before providing them with answers.

Assessment

The fourth component of a UDL-based curriculum is applying multiple methods of *assessment*. The ITS in NativeAccent does this by continuously giving formative assessments during and after each lesson, as well as summative assessments at the beginning and the end of the course. The Intelligent Tutor also provides learners with formative and summative progress charts that help them to be conscious of their progress. Different colors on the charts signify different meanings (Figure 3). For instance, the color *green* shows that the learner has passed the assessment and needs minimal training; *yellow* shows the learner has passed the assessment but needs moderate training; *red* shows the learner has failed the assessment and needs intensive training; and *gray* shows the assessment is not completed. Learners can also receive a summative report on all their activities at any stage of their progress. However, as mentioned above, since NativeAccent is a pronunciation learning tool, it assesses learners' pronunciation through spoken language; consequently, learners are not free to choose the way they prefer to be

assessed.

NativeAccent and Learning Science Theory

The **Learning Science** (LS) strategies that this article applies to NativeAccent v.3 are *retrieval practice*, *variable training* and Vygotsky's *zone of proximal development (ZPD)*.

Retrieval Practice

Through *retrieval practice*, the brain remembers what it has learned by applying different techniques such as *elaboration*, *reflection*, *testing effect*, *spacing out*, and *generation*. Through *elaboration* learners restate what they have just learned and try to use it in various contexts. *Reflection* can make learning more durable by practicing a new skill/idea repeatedly until it becomes reflexive and automatic (Brown et al., 2014). *Testing* per se is not sufficient; effective and lasting learning can take place if testing is effortful and systematic, i.e., *testing effect*. By *spacing out*, a learner leaves out some space between learning and practicing instead of massed practicing (Brown et al., 2014). Another technique that can make learning more lasting is *generation*, through which learners use their prior knowledge and make an effort to guess what the new idea means (Bransford et al., 2000; Brown et al., 2014).

Variable Training

Variable training provides the opportunity for learners to practice a new idea/skill in different ways. *Interleaved practice*, as another form of *variable training*, takes place when different concepts are presented and practiced at one time (Brown et al., 2014).

Vygotsky's Zone of Proximal Development

Vygotsky's *ZPD*, another LS strategy, refers to the level at which a student can perform a task by imitating a more competent individual or a teacher (Chaiklin, 2003). The important elements in *ZPD* are the *possibility of developing a skill*, *learner's potential to develop*,

learner's imitation, and the assistance of a more competent individual.

How NativeAccent Applies Retrieval Practice, Variable Training and Vygotsky's ZPD

Unlike conventional instructional methods that usually begin with direct instruction prior to assessment, NativeAccent presents its lessons after assessing learners' potential knowledge in order to provide them with appropriate and specific guidance. Namely, the course begins with the initial assessment (*testing*), the result of which determines what specific lessons should be included in the curriculum/training sessions for each learner. In other words, each learner receives a personalized curriculum based on the responses given in the initial assessment, his or her native language, and current spoken English skills.

The initial assessment also provides the *possibility of developing a skill*, which matches with Vygotsky's *ZPD* (Chaiklin, 2003) in a way that no training exercises are assigned without assessment. Therefore, to practice a new concept, prior knowledge is required. Another component of *ZPD* is *assessing learner's potential to develop* (Chaiklin, 2003). In NativeAccent, this is done through not only the initial assessment but also all the training exercises. More simply, in order to move on to the new set of exercises, learners should be able to accomplish the previous material.

Additionally, since the initial assessment measures the intricacies in the English language such as 38 phonemes, 28 grammar skills, and 16 micro fluency questions (Mawson, 2017; Pelton, 2012b), the training sessions are designed based on the errors made by learners and those areas that should be improved. Even so, the training exercises cover pronunciation, grammar, and fluency, and in some exercises learners should pay attention to all three aspects of spoken language skills simultaneously. This characteristic pertains to *interleaved practice* in LS through which learners should be able to organize their knowledge by differentiating

similarities and differences in multiple concepts (Bransford et al., 2000; Brown et al., 2014).

Unlike conventional instructional methods, the training sessions also do not begin with direct instruction but with *generative training* or *generation*, i.e., exercises through which learners link the new material to their former knowledge by making an effort to provide correct answers. This trial- and-error practice to generate responses (as a type of *testing*) can lead to effective learning by strengthening memory (Bransford et al., 2000; Brown et al., 2014). Providing correct responses means that learners are ready for the next set of exercises; however, if learners make errors, corrective feedback is immediately provided in different forms. In LS this strategy is called *varied practice* or *variable training* (Bransford et al., 2000; Brown et al., 2014).

The multi-modal feedback system in NativeAccent v.3, which includes text, audio, and a model speaker (Pelton, 2012b), provides learners with detailed instructions such as where the errors are and how they should be corrected. Learners should imitate the model speaker by repeating after it and correcting their errors. This strategy matches with *imitation* in *ZPD*. Additionally, all the training exercises are presented through the Intelligent Tutor, which acts like the *more competent individual* in *ZPD* (Chaiklin, 2003).

As noted above, instruction in NativeAccent does not come first and is not the same for all types of learners. In NativeAccent instruction is carried out after *pinpointing* students' errors, and it varies based on types of errors. After providing feedback, NativeAccent relates what has been taught to learners' prior knowledge by providing more exercises until all errors are corrected. In LS this strategy, as a type of *retrieval practice*, matches with *elaboration* (Brown et al., 2014). Additionally, detailed exercises on phonemes, grammar skills, and micro fluency skills are repeated in different contexts throughout the personalized curriculum until

they become automatic and reflexive (Brown et al., 2014). This strategy corresponds to *reflection* as a form of *repeated retrieval* in LS. The reappearance of the new concepts throughout the curriculum gives learners the chance to leave some space between each practice. As another learning strategy in *retrieval practice*, this is called *spacing out* which can consolidate learning the new material (Brown et al., 2014).

While each learner begins the course with a pretest in the initial assessment, he or she also takes a posttest or a final reassessment at the end of the course. As can be seen, *testing*, as a *retrieval practice*, is applied in multiple ways throughout the course such as the initial assessment (a summative test), all training exercises (formative tests), and a final reassessment (a summative test). *Testing* is also an indispensable pedagogical strategy in NativeAccent that can facilitate the process of learning by slowing down the process of forgetting (Bransford et al., 2000; Brown et al., 2014).

NativeAccent and the Intelligent Tutoring System (ITS)

Intelligent tutors facilitate “learning in a meaningful and effective manner by using a variety of computing technologies” (Carnegie Speech, 2012, p. 20). In general, intelligent tutors are designed based on *model tracing*, i.e., creating individualized instructions for each learner by following cognitive models (Koedinger et al., 1997). Intelligent tutors also provide learners with personalized feedback, which allows them to “practice at their own speed” (McCrocklin, 2016, p. 27) and manage their time and pace according to their needs and skills (Koedinger et al., 1997).

The Intelligent Tutor in NativeAccent does *model tracing* by prioritizing and personalizing “English pronunciation training exercises based on each user's unique pronunciation needs” (Carnegie Speech, 2018d) and by providing model speakers. The

personalized curriculum is based on the data gained from the initial assessment and is in response to the errors made in the training session. The Intelligent Tutoring System (ITS) in NativeAccent also does *knowledge tracing* through assessing learners' knowledge and identifying their problems (Koedinger et al., 1997). *Knowledge tracing* is done not only through the initial and final assessments but also in all the training exercises throughout the course of study. Intelligent tutors can also be engaging and motivating by providing learners with one-on-one, "pertinent," (Eskenazi, 1999, p. 449) "timely," (Koedinger et al., 1997, p. 35; Pitler et al., 2012, p. 38) and immediate feedback.

Since the ITS in NativeAccent provides learners with a roadmap as well as feedback for all their training exercises, it can motivate them and save them time by guiding them to be goal-oriented and engaged in their own learning process. One of the advantages of intelligent tutors over human tutors is their accuracy in identifying problems and giving corrective feedback (Le & Pinkwart, 2012). NativeAccent v.3 is designed to identify problems and give corrective feedback on 38 phonemes, 28 grammar skills, and 16 micro fluency skills; it does not, however, detect or correct word stress and sentence stress errors, a factor to be considered in its later versions.

The exercises in the training sessions of NativeAccent are created based on learners' mistakes and current language skill level in the initial assessment. The first principle asserts that the components of the desired skill should be decomposed (Anderson et al., 1995). The ITS in NativeAccent, which supports learners' spoken skill, *decomposes the components of the desired skill* by assigning exercises on 38 phonemes, 28 grammar skills, and 16 micro fluency skills and in response to the specific needs of its users. Nevertheless, since word stress and sentence stress errors are not detected, they are not included in the training sessions either.

Training exercises which are presented in boxes, like a roadmap, show learners the goals so that they can achieve mastery in a skill. This point matches with the second principle of the ITS, which is communicating “the goal structure underlying the problem solving” (Anderson et al., 1995, p. 179).

In NativeAccent, the ITS provides specific instructions on how errors should be corrected, as well as corrective, immediate feedback on all errors made in the training exercises except for word stress and sentence stress errors. The feedback is always accompanied with visual aids or scaffolded support (Eskenazi, 1996) such as a model speaker as illustrated in Figure 2. Depending on the type of error, the feedback can be on a specific point or a more general rule. In order to avoid confusion, the Intelligent Tutor leaves only a few related boxes unlocked and available for practice while all the other boxes are locked. After the learner successfully practices the exercises in one box, another box becomes unlocked and ready for practice.

NativeAccent and Automatic Speech Recognition (ASR)

Automatic Speech Recognition (ASR) is a technology used for *identifying* (Kim, 2006) and *decoding* spoken words and transcribing them into text (Voicebox Technologies Corporation, 2017). This section applies Witt and Young’s (1998) four ASR modules of (a) *instruction and modeling*, (b) *monitoring*, (c) *evaluation*, and (d) *assessment*, as well as Eskenazi’s (2009) two levels of speech error detection: *specific* and *general errors*.

Based on the first module of *instruction and modeling*, the ASR provides clear instruction and modeling for pronunciation lessons (Witt & Young, 1998). In NativeAccent, instruction and modeling are provided through feedback such as text, audio, graphics, and model speakers for all pronunciation lessons. The second ASR module refers to *monitoring*

through *error detection* (Witt & Young, 1998). The ASR system in NativeAccent detects and analyzes pronunciation errors not only in the initial assessment but also throughout all training exercises. Additionally, recognizing and detecting errors is carried out at two levels: *specific errors*, such as phonemes, and *general pronunciation errors*, such as overall fluency (Eskenazi, 2009). The ASR in NativeAccent detects 38 phonemes, 28 grammar skills, and 16 micro fluency questions (Pelton, 2012b), such as pausing, reduction, palatalization, and linkage (Mawson, 2017). According to the third module of *evaluation*, the ASR evaluates learners through error analysis, error correction, and feedback (Witt & Young, 1998). The ASR in NativeAccent detects and analyzes errors, based on which the Intelligent Tutor provides feedback and instructions on how those errors should be corrected. The fourth module refers to *assessment*, meaning the ASR continuously assesses learners' progress (Witt & Young, 1998).

CONCLUSION

The purpose of this article was to demonstrate how a CALL software program called NativeAccent v.3 can be systematically evaluated by applying four research-based theories. By matching the components of these theories (UDL, LS, ITS, and ASR) to NativeAccent, it was concluded that this software (a) is systematically structured based on the components of UDL and LS; (b) can be personalized based on the needs of its users and their native language speech patterns; (c) identifies and detects errors (except for word stress and sentence stress) through its ASR feature; (d) provides timely and immediate feedback through its Intelligent Tutor; and (e) supports pronunciation, fluency, and grammar skills. The result of this research can guide ELLs, educators, and administrators who are seeking ways to enhance the development of language learning process in choosing the most appropriate technology tool. While NativeAccent v.3 is a pronunciation training program, the guidelines presented in this

article can be applied to other language training platforms as well.

This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

References

- Anderson, J. R., Corbett, A. T., Koedinger, K. R., & Pelletier, R. (1995). Cognitive tutors: Lessons learned. *The Journal of the Learning Sciences*, 4(2), 167–207.
<http://www.jstor.org.ezproxy.lib.uh.edu/stable/1466690>
- Bransford, J. D., Brown, A. L., & Cocking, R. R. (Eds.). (2000). *How people learn: Brain, mind, experience and school* (Expanded ed.). <https://ebookcentral.proquest.com>
- Brown, P. C., Roediger, H.L., III, & McDaniel, M.A. (2014). *Make it stick: The science of successful learning*. Belknap Press.
- Caballero, K. T., & Thompson, S. (2016). *The COLOR VOWEL® Approach: Resources for connecting pronunciation to vocabulary, reading, and spelling*. English Language Training Solutions, LLC. doi:10.3998/mpub.8307407
- Carnegie Speech. (2012, November). *NativeAccent student manual*. Pittsburgh: Carnegie Speech.
<http://edurep.com/na3asiatraining/students/pgcc%20docs/Student%20docs/NativeAccent%20v.3%20Student%20Manual.pdf>
- Carnegie Speech. (2018a). *Carnegie Speech Assessment*.
<https://www.carnegiespeech.com/sales/index.php/vstore/csa>
- Carnegie Speech. (2018b). *Education*. <http://www.carnegiespeech.com/education/education.php>
- Carnegie Speech. (2018c). *Enterprise*. <http://www.carnegiespeech.com/industries/bpo.php>
- Carnegie Speech. (2018d). *Training overview*.
<http://www.carnegiespeech.com/products/nativeaccent.php>
- Chaiklin, S. (2003). The zone of proximal development in Vygotsky's analysis of learning and instruction. In A. Kozulin, B. Gindis, V.S. Ageyev, & S.M. Miller (Eds.), *Vygotsky's educational theory in cultural context* (pp. 1–21). Cambridge University Press.

- Ellis, N. C., & Bogart, P. S. H. (2007, October). Speech and language technology in education: The perspective from SLA research and practice. *Proceedings of the SLaTE Workshop on Speech and Language Technology in Education, Farmington, PA, USA*, 1–8.
- Eskenazi, M. (1996). Detection of foreign speakers' pronunciation errors for second language training - preliminary results. *Proceedings of the Fourth International Conference on Spoken Language Processing, Philadelphia, PA, 3*, 1465–1468.
- Eskenazi, M. (1999). Using a computer in foreign language pronunciation training: What advantages? *CALICO Journal*, 16(3), 447–469.
- Eskenazi, M., Kennedy, A., Ketchum, C., Olszewski, R., & Pelton, G. (2007). The NativeAccent pronunciation tutor: Measuring success in the real world. In *SLaTE-2007*, 124–127.
- Eskenazi, M. (2009). An overview of spoken language technology for education. *Speech Communication*, 51(10), 832–844.
- Gilbert, J. B. (2018). Issues in teaching pronunciation: Prosody, intonation, and vowels. In J. I. Lontas & M. DelliCarpini (Eds), *The TESOL encyclopedia of English language teaching*. <https://doi.org/10.1002/9781118784235.eelt0241>
- IRIS Center. (2017). *Universal design for learning: Creating a learning environment that challenges and engages all students*.
<http://iris.peabody.vanderbilt.edu/module/udl/cwrap/#content>
- Kaiser, A. P. (2014). Using Single-Case Research Designs in Programs of Research. In T. R. Kratochwill & J. R. Levin (Eds.), *School Psychology Series. Single-case intervention research: Methodological and statistical advances* (pp. 309–323).
- Kim, I. (2006). Automatic speech recognition: Reliability and pedagogical implications for

- teaching pronunciation. *Journal of Educational Technology & Society*, 9(1), 322–334.
- Koedinger, K., Anderson, J. R., Hadley, W. H., & Mark, M. A. (1997). Intelligent tutoring goes to school in the big city. *International Journal of Artificial Intelligence in Education*, 8, 30–43.
- Lai, P. (2017). The literature review of technology adoption models and theories for the novelty technology. *Journal of Information Systems and Technology Management*, 14(1). doi:10.4301/s1807-17752017000100002
- Le, N., & Pinkwart, N. (2012). Can soft computing techniques enhance the error diagnosis accuracy for intelligent tutors? *Intelligent Tutoring Systems Lecture Notes in Computer Science*, 7315, 320–329. doi:10.1007/978-3-642-30950-2_42
- Lotherington, H., & Jenson, J. (2011). Teaching multimodal and digital literacy in L2 settings: New literacies, new basics, new pedagogies. *Annual Review of Applied Linguistics*, 31, 226–246. doi:10.1017/s0267190511000110
- Mawson, C. (2017). Carnegie Speech Company. NativeAccent. English speech training. *Journal of Second Language Pronunciation*, 3(1), 140–143.
<https://benjamins.com/#catalog/persons/290104843>
- McClanahan, L. (2014). Training using technology in the adult ESL classroom. *Journal of Adult Education*, 43(1), 22–27. <https://www.learntechlib.org/p/155499/>
- McCrocklin, S. M. (2016). Pronunciation learner autonomy: The potential of Automatic Speech Recognition. *System*, 57, 25–42.
<http://www.sciencedirect.com/science/article/pii/S0346251X15001980>.
- National Center on Universal Design for Learning. (2014). *What is Universal Design*

for Learning. <http://www.udlcenter.org/aboutudl/whatisudl>

Pelton, G. (2012a). Improving intelligent tutoring of pronunciation consonant cluster problems.

In G. S. Lim & E. D. Galaczi (Convenors), *Cambridge English Centenary Symposium on Speaking Assessment*. Symposium conducted at the Guildhall of Cambridge, England. <http://events.cambridgeenglish.org/cambridge-english-centenary-symposium/files/pelton-improving-intelligent-tutoring.pdf>

Pelton, G. (2012b). *Mining pronunciation data for consonant cluster problems*.

http://www.academia.edu/8568681/Mining_pronunciation_data_for_Consonant_cluster_problems

Pitler, H., Hubbell, E. R., & Kuhn, M. (2012). *Using technology with classroom instruction that works*. <https://ebookcentral.proquest.com>

Sheridan, S. M. (Ed.). (2014). Single-case designs and large N studies: The best of both worlds.

In T. R. Kratochwill & J. R. Levin (Eds.), *Single-case intervention research: Methodological and statistical advances* (pp. 299–308). American Psychological Association.

Voicebox Technologies Corporation. (2017). Automatic Speech Recognition (ASR).

How ASR works and how Voicebox overcomes real-world challenges to ASR.

<http://www.voicebox.com>

Witt, S., & Young, S. (1998). Computer-assisted pronunciation teaching based on

automatic speech recognition. In S. Jager, J. Nerbonne, & A. van Essen (Eds.), *Language Teaching & Language Technology* (pp. 25–35). Swets & Zeitlinger.