### Just in Time Learning: Implementing Principles of Multimodal Processing and Learning for Education of Children with Special Needs

Dominic W. Massaro Department of Psychology University of California Santa Cruz, CA 95064 USA massaro@ ucsc.edu

### ABSTRACT

Our language-training program, which utilizes embodied conversational agents as tutors, who guide students through a variety of exercises designed to teach vocabulary and grammar, to improve speech articulation, and to develop linguistic and phonological awareness. With our Lesson Creator, teachers, parents, and even students can build original lessons that allow concepts, vocabulary, and pictures to be easily integrated. This user-friendly application allows the composition of lessons with minimal computer experience and instruction. Although it has many options, the program has wizard-like features that direct the coach to explore and choose among the alternative implementations in the creation of a lesson. The Lesson Creator application facilitates the specialization and individualization of lessons by allowing teachers to create customized vocabulary lists Just in Time as they are needed. The Lesson Creator allows the coach to give descriptions of the concepts as well as corrective feedback, which allows errorless learning and encourages the child to think as they are learning. Finally, there is an example of lesson created in the Lesson Creator, and are playable in Timo Vocabulary.

### **1. INTRODUCTION**

The goal of this paper is to confront challenges facing computerassisted learning for language-challenged students. Its major premise, however, is that the design of optimal learning environments for these students will share features with ideal learning situations for all individuals across a variety of academic and other settings. Consider, for example, the importance of vocabulary learning for perceptual and cognitive development (LaSasso & Davey, 1987; Massaro, 2004). It might be believed that only language-challenged students require direct instruction in vocabulary whereas normally-developing students will acquire vocabulary on their own, in reading for example. However, it is now accepted that all students benefit from vocabulary instruction (Beck, McKeown, & Kucan, 2002).

### 1.1 Principles of Learning

Principles of learning can inform computer-assisted instruction even though the learning of content might be domain dependent. For example, learning Shakespeare will generalize very little to solving word problems in mathematics. However, the psychological processes involved in one domain overlap with learning processes in other domains. Psychologists and Educators have uncovered a number of principles that help us understand and optimize learning. For example, research and applications have consistently revealed the value of time on task so much so that it has become legend. From the initial stages of learning to the attainment of expertise, time spent on focused deliberate practice is essential and the more time spent the more learning (Ericsson et al., 2006). A related effective principle of learning is the value of distributed or spaced practice relative to massed practice (Pashler, Dohrer, Depeda, & Carpenter, 2007). Given an equivalent amount of time for learning, it is better to space this practice across multiple learning sessions rather than having fewer sessions of learning. Finally, comprehension and learning new material is critically dependent on what is already known about that subject. These principles and others have guided our implementations of our technology.

## 2. BALDI TECHNOLOGY AND PEDAGOGY

The value of visible speech in face-to-face communication was the primary motivation for the development of Baldi®, a computer-animated talking head. Baldi has an underlying threedimensional wireframe model that is controlled analogously to controlling a puppet on a set of strings (see Massaro, 1998, Massaro, 2004). There are several advantages to utilizing animated tutors and multisensory and multimodal learning environments. Research has shown that our perception and understanding of spoken language are influenced by a speaker's face and accompanying context and gestures, as well as the actual sound of the speech. Given the value of face-to-face interaction, our persistent goal has been to develop, evaluate, and use animated agents to teach speech and language.

Baldi<sup>®</sup> is an accurate three-dimensional agent appropriately aligned with either synthesized or natural speech. Baldi provides realistic visible speech that is almost as accurate as a natural speaker. For examples in a variety of different languages, see http://mambo.ucsc.edu/psl/international.html. Also, Baldi can be animated in real time on a commodity PC, and is able to say anything at any time in interactive applications (Massaro, 2004; Massaro, Ouni, Cohen, & Clark, 2005).

One of the first applications of Baldi as an animated tutor was implemented for deaf and hard of hearing students at the Tucker-Maxon School of Oral Education (TMOS; http://www.tmos.org/). Baldi was used in a variety of school exercises in mathematics, history, science, and social studies (Massaro, Cohen, & Beskow 1999). Eventually, a lesson creator and tutor were developed for vocabulary learning, which eventually evolved into Timo Vocabulary and the Lesson Creator described in this paper. An incentive to employing computer-controlled applications for training is the ease in which automated practice, feedback, and branching can be programmed. Another valuable component is the potential to present multiple sources of information, such as text, sound, and images in parallel. Instruction is always available to the child, 24 hours a day, 365 days a year. Several advantages of utilizing Baldi as a computer-animated agent as a tutor are clear, including the popularity of computers and embodied conversational agents. We have found that the students enjoy working with Baldi because he offers extreme patience, he does not become angry, tired, or bored, and he is in effect a perpetual teaching machine. This type of animated tutoring can be applied in science learning for not only students with special needs but for all students.

Perhaps the greatest advantage of computer-aided instruction is that instruction can be tailored exactly to the student's needs, which is best implemented in a one-on-one learning environment for the students. Other benefits include the ability to seamlessly meld spoken and written language as well as sign language and Cued Speech, and provide a semblance of a game-playing experience while actually learning. Given that education research has shown that children can be taught new word meanings by using direct instruction methods (Beck, McKeown, and Kucan, 2002), Animated Speech Corporation implemented these basic features in a commercial application to teach vocabulary and grammar. Timo Vocabulary, which is a direct descendant of Baldi's Tutor (Massaro, 2004, 2006a), includes a curriculum of 127 vocabulary-building lessons covering more than 650 words and images drawn from K-4th grade curriculums of TMOS and other schools for deaf and hard of hearing students [http://animatedspeech.com/products/products vocabulary.html].

One of the principles of learning that Timo Vocabulary exploits most is the value of multiple sources of information in perception, recognition, learning, and retention. An interactive multimedia environment is ideally suited for learning. Incorporating text and visual images of the vocabulary to be learned along with the actual definitions and sound of the vocabulary facilitates learning and improves memory for the target vocabulary and grammar. Many aspects of these lessons enhance and reinforce learning. Timo Vocabulary is derived directly from the Baldi technology and pedagogy, except that Baldi has morphed into a new animated tutor called Timo. As in the previous implementations and tests of our pedagogy (Massaro, 2006a), Timo Vocabulary guides the students to 1) Observe the words being spoken by a realistic talking interlocutor, 2) Experience the word as spoken as well as written, 3) See visual images of referents of the words, 4) Click on or point to the referent or its spelling, 5) Hear themselves say the word, followed by a correct pronunciation, 6) Spell the word by typing, and 7) Observe and respond to the word used in context. Half of the exercises involve multiple-choice testing, which boosts performance on later tests (Marsh, Roediger, Bjork & Bjork, 2007). The other half of the tests involves either spoken or written generation of the students' answers, which facilitates learning (Metcalfe & Kornell, 2007). The test exercises can be viewed as learning exercises because testing has been demonstrated to increase learning and retention (McDaniel, Roediger, & McDermott, 2007).

One of the most valuable features of this application is that the tutoring can be individualized for each student. A profile is created for each student to specify which exercise should be included, the reward selection, and whether or not captioning should be on—an important concern for deaf and hard of hearing students.

# 2.1 Potential Limitations of Computer Assisted Instruction

There are several properties of computer-assisted instruction and the lesson environment we have described that might limit the efficacy of learning. We use three-dimensional animated agents because they can say anything at any time in real time. Thus, new vocabulary can be recorded and new lessons can be made without the necessity of first video recording a real person. One potential limitation is that the visual input the student receives is necessarily on a two-dimensional screen rather than existing as an actual live three-dimensional person. Important pedagogical characteristics of a three-dimensional world might not be adequately represented in just two dimensions. A second possible limitation is that the lessons use synthetic speech and facial animation. Although these properties make our tutoring agent capable of saying anything at any time, they might compromise the instructional dialog in the tutoring situation.

### 2.1.1 Two Dimensional Displays

It is feasible that limiting the students' experience to the two dimensional world of computer monitors would constrain learning relative to a live teacher. Our computer-animated tutor, Baldi, was developed to provide informative visible speech as well as the standard auditory speech. It is important to know if the visible speech from a three-dimensional face projected onto a twodimensional surface is as informative as a live person seen in three dimensions. Surprisingly, there does not seem to be a direct comparison of these two conditions. However, we can evaluate the influence of visible speech across studies that used these two different displays. It should be noted that these studies were carried out with typical hearing participants, and noise was added to the auditory speech to make it difficult to understand completely. The results are equally relevant to hard-of-hearing persons in noise-free environments, however, because the two groups behave identically in these respective conditions (Massaro, 1998; Massaro & Cohen, 1999, 2000).

The benefit of the face is usually evaluated by comparing word identification performance when the words are presented with the face and voice relative to just the voice. Sumby and Pollack (1954) used a live talker whereas Jesse, Vrignaud, and Massaro, (2001) used a video recording of the talker. The benefit given visible speech appears to be roughly equivalent in the two cases. Comparing experimental conditions that matched as much as possible, the accuracy improved from about 30% with just auditory speech to about 57% when the two-dimensional visible speech was also present. This 27% improvement compares favorably with the 28% improvement with the three-dimensional presentation. Thus, we can expect that the benefit of a visible tutor exists for two-dimensional viewing as well as well as for a live presentation.

The success of two-dimensional media such as the television and the internet, however, is a real-world experimental proof of the sufficiency of two dimensions for learning. Furthermore, the use of video-based sign language interpreting in the classroom does not seem to present more of a challenge for deaf students than live classroom presentations (Marschark, Pelz, Convertino, Sapere, Arndt, & Seewagen, 2005). Thus, tutoring on two-dimensional surfaces appears to be as effective as live tutoring.

## 2.1.2 The Use of Synthetic Speech and Facial Animation

In a few instances, individuals have reacted negatively to the use of synthetic auditory speech in our applications. Not only did they claim it sounded relatively robotic (in some cases, people thought there was a resemblance to our California governor in his previous life as a terminator), they were worried that children may learn incorrect pronunciation or intonation patterns from this speech. However, this worry appears to be unnecessary. In agreement with the positive outcomes of direct experimental evaluations described below, Baldi has been used in many different pedagogical applications at the Tucker-Maxon School of Oral Education (http://www.tmos.org/), where Baldi tutored quite successfully with about 16 hard of hearing children who were about 8 to 14 years of age (Barker, 2003; Soland, 2007). The students had either hearing aids or cochlear implants, and were tutored by Baldi an average of about 20 minutes per day. Baldi taught these children receptive vocabulary directly, and also was used in various applications reinforcing the school's curriculum.

As part of the vocabulary tutor, there were recorded speech tasks in which these students imitated and elicited words prompted by Baldi's synthetic speech models. The teachers' impressions were that these children did use Baldi's synthetic speech to produce fairly intelligible words (Soland, 2007). These students had severe-to-profound hearing losses (90 dB HL or greater) with varying degrees of speech intelligibility and delayed vocabulary skills. But their productions of these new words spoken by Baldi seemed to be no better or no worse than their normal articulatory patterns. The teachers thought these production tasks were beneficial to the students. In addition, the teachers were able to correct the speech synthesizer's pronunciation of a word when it was initially mispronounced by modifying the text input. This was necessary because they noticed that when Baldi mispronounced a word or gave it inappropriate accenting, students were likely to pronounce or intone the word in a similar manner.

A number of these students, also described in Barker (2003), who began using Baldi and synthetic speech 7 years ago now have graduated from high school (Connors Fortier, 2007). Obviously, they were still able to achieve academically despite regular exposure to synthetic speech at a fairly young age (Barker, 2003). It should be noted that the primary goal was to improve deficit language bases among deaf and hard of hearing children, which was believed to be much more critical to academic achievement than perfect pronunciation. For example, a student could read and write quality assignments even though some of the words would be mispronounced. But, in fact, many of the children's receptive vocabulary work with the tutor carried over into intelligible expressive vocabularies (Connors Fortier, 2007).

In addition to these observations, experimental tests demonstrated that hard of hearing children improved their pronunciations of words as a direct result of Baldi's tutoring (see Massaro, 2006a, for a review). In vocabulary lessons, the children not only improved in their receptive vocabulary but also in their productions of these words (Massaro & Light, 2004a). In speech production tutoring on specific speech segments such as /s/, /z/, /t/, and /d/, the application was successful in teaching correct pronunciation of the target words and also generalized to the segments in novel words (Massaro & Light, 2004b). This is gratifying because the value of synthetic speech like our animated visible speech tutor is that anything can be said at any time by simply entering the appropriate written text. Natural speech would

require that the content be prerecorded by voice talent. This constraint would negate the just in time feature of creating lessons. Finally, notwithstanding these justifications, synthetic auditory speech has improved considerably and the synthetic voice of the newer Timo is much more natural sounding than Baldi's original voice.

Analogous arguments exist for facial animation. We have shown that Baldi can be speechread almost as accurately as a real person. In the Jesse et al. (2000/2001) study described earlier, one of 65 auditory sentences was randomly presented in noise on each trial, and the hearing participants were asked to watch and listen to each sentence and to type in as many words as they could for each sentence. There were three presentation conditions: auditory, auditory paired with the face of the original talker, and auditory paired with the face of Baldi. Pairing the original talker with the auditory speech improved performance by 54% whereas pairing Baldi with the auditory speech gave a 47% improvement. Thus, the large and similar improvement in the two conditions demonstrates that Baldi provides respectable visible speech even though he is synthetic. Although Timo is based on Baldi, research is in progress to test whether Timo's visible speech is as effective as Baldi's. Given this foundation in educational practice, we now turn to a short review of studies using this pedagogy in vocabulary learning ..

### 2.2 Evaluation of Effectiveness

The multisensory approach with a computer-animated agent to vocabulary learning has been tested in several experiments studies. A detailed review of these tests with deaf and hard of hearing students in the learning of speech and language is given in Massaro (2006a). Several evaluation experiments showed that both hard-of-hearing and autistic children learned many new words, grammatical constructions and concepts (Bosseler & Massaro, 2003; Massaro & Light, 2004a), proving that the application provided an effective learning environment for these children. The research strategy insured that any learning was due to the intervention itself rather than from outside of the lesson environment. Students learned all of the items that they were specifically tutored on and not the items that were only tested. In addition, a delayed test given more than 30 days after the learning sessions took place showed that the children retained over 85% of the words that they learned. This learning and retention of new vocabulary, grammar, and language use is a significant accomplishment for these children.

In a recent experimental test, Massaro (2006b) used the same multisensory approach with a computer-animated agent to evaluate the effectiveness of teaching vocabulary to beginning elementary students learning English as a second language. Children, whose native language was Spanish, were tutored by Timo (http://animatedspeech.com/), a new animated character based on Baldi, and tested on English words they did not know. The children were pretested on lessons in the application in order to find three lessons with vocabulary that was unknown to the children. A session on a given day included a series of three test lessons, and on training days, a training lesson on one of the three sets of words. Different lessons were necessarily chosen for the different children because of their differences in vocabulary knowledge. As shown in Figure 1, the test session involved the presentation of the images of a given lesson on the screen with Timo's request to click on one of the items, e.g., Please click on the oven. No feedback was given to the child. Each item was



Figure. 1: The left panel shows a typical screen shot from the Timo Vocabulary application. The right panel gives the average percentage of correct identifications for each of the three sets of words. The three vertical bars indicate when training was initiated for each of the three sets. Performance improved on each set of words after training on that set was initiated.

tested once in two separate blocks to give 2 observations on each item. Three different lessons were tested, corresponding to the three sets of items used in the multiple baseline design. A training session on a given day consisted of just a single lesson in which the child was now given feedback on their response. Thus, if Timo requested the child to click on the dishwasher and the child clicked on the spice rack, Timo would say, "I asked for the dishwasher, you clicked on the spice rack. This is the dishwasher. The training session also included the Elicitation and Imitation sections in which the child was asked to repeat the word when it was highlighted and Timo said it, and the child was asked to say the item that was highlighted. Several days of pretesting were required to find lessons with unknown vocabulary. Once the 3 lessons were determined, the pretesting period was followed by the training days. As can be seen in Figure 1, the children learned the words when they were tutored but not words that were simply tested. This result replicates the previous studies carried out on hard of hearing and autistic children with Baldi as the animated conversational tutor. In other experiments, we have also observed that Baldi's unique characteristics allow a novel approach to tutoring speech production to both children with hearing loss (Massaro & Light, 2004b) and adults learning a new language (Massaro & Light, 2003).

Given the success of the tutoring program, it is important to assess whether the facial animation is a significant influence on learning vocabulary. To evaluate this question, an experiment compared to what extent the face facilitated vocabulary learning relative to the voice alone (Massaro & Bosseler, 2006). The vocabulary learning consisted of both the receptive identification of pictures and the production of spoken words. Five autistic children were tutored in vocabulary with and without the face. Each child continuously learned to criterion two sets of words with the face and voice and two sets with just the voice and without the face. The rate of learning was significantly faster and the retention was better with than without the face. Although two of the children did not show a large advantage with the face, the research indicates that at least some autistic children benefit from the face. The better learning and retention with the face was most likely due to the additional information provided by the face but it is still possible that the face was more engaging and motivating which in turn would

benefit performance. In either case, it shows the value of animated tutors in the teaching of vocabulary.

## **2.3 Montessori's Principles of Educational Practice**

Given the success of Montessori education (Stoll-Lillard, 2005), it is challenging to evaluate how direct multisensory instruction relates to their principles. Montessori's Principle 1 claims that motor behavior and cognition are closely intertwined and that physical movement can enhance thinking and learning (Stoll-Lilliard, 2005, Stigler, 1984). At first glance, this principle seems the antithesis of direct computer-aided instruction with an animated tutor. However, we have learned that our nervous systems appear to be wired in a way that observations of actions activate neural mechanisms involved with the actual performance of those actions. The so-called mirror neurons (Rizzolatti & Craighero, 2004) involved in performing an action are activated when that action is observed. One possibility, therefore, would be to implement lessons on Nintendo's Wii [http://wii.com] to allow the child to have larger physical movements. Another would be to have animated movies as well as pictures for learning.

Montessori's Principle 2 states that choice and perceived control promote children's concentration and contentment in the learning process. As is currently exists, direct instruction does not appear to allow much choice. On the other hand, the child can be given a library of lessons and she can choose the lesson to study. A precocious child might even be able to create a lesson of her choosing.

Principle 3 assumes that personal interest enhances learning in a context where interests build on prior knowledge and the children's own questions. For example, a deaf French child used the Lesson Creator to document her travel and holiday pictures in a set of English vocabulary lessons. Thus, learning a new language was facilitated by involving her direct experience and interests with a normally tedious task.

Principle 4 indicates that extrinsic rewards negatively impact long-term motivation and learning. Rewards and feedback can be controlled exactly in computer-assisted learning. Directed feedback can allow errorless learning without focusing on rewarding the child for correct answers and punishing the child for incorrect answers.

According to Principle 5, collaborative (child–child) arrangements are conducive to learning. Although most automated instruction is one-on-one and precludes collaborative learning, this principle can be instantiated in several different ways. First, the animated agent can be a child who works along with the child (Tartaro & Cassell, 2008). Second, children can work together on a lesson or on creating lessons, and can even distribute the required learning and thereby achieve the benefits of the Jigsaw Classroom (http:// www.jigsaw.org/).

Principle 6 assumes that learning situated in and connected to meaningful contexts is more effective than learning in abstracted contexts (Gee, 2003). Although most automated instruction can be considered relatively unsituated and not connected to a meaningful context, the Lesson Creator allows the immediate creation of lessons on subjects that are currently taught: Just In Time learning. Thus, the child sees the value and appropriate context of the lesson when it is connected to her appropriate interest and cognitive level.

Principle 7 claims that sensitive and responsive (nurturing) teaching is associated with more optimal outcomes. Tutors can be created and programmed to be highly nurturing. For example, the difficulty of the lessons can be controlled to meet the child's preferred difficulty level, and errorless feedback can be provided.

Principle 8 assumes that order in the environment promotes and establishes mental order and is beneficial to the child. Direct instruction is highly orderly in its functioning, which adheres to this principle.

Given this background, the Lesson Creator, which further optimizes the effectiveness of the learning process, is described next.

### 3. LESSON CREATOR

The Lesson Creator, which is a direct descendant of the Vocabulary Wizard (Massaro, 2004, 2006a), adds flexibility and many new pedagogical features to Timo Vocabulary, and allows the easy creation of new lessons. Because vocabulary is essentially infinite in number, it is difficult to anticipate all of the vocabulary that a student will need. The Lesson Creator solves this problem by Just in Time learning. Teachers, parents, and even students can build original lessons that meet unique and specialized conditions. New lessons can be made to allow personalized concepts, vocabulary and pictures to be easily integrated. This user-friendly application allows the composition of lessons with minimal computer experience and instruction. Although slightly more complex than for example your typical installation wizards because of the many more options, the program has wizard-like features that direct the coach to explore and choose among the alternative implementations in the creation of a lesson.

Eight student exercises can be appropriately modified within the Lesson Creator. The evolving design of this lesson pedagogy is based on educational principles to optimize learning, which are not always intuitive. The Lesson Creator allows the coach to tailor the lesson to the needs of the student, to seamlessly meld spoken and written language, bypass repetitive training when student's responses indicate that material is mastered, provide a semblance of an interactive and engaging experience while actually learning, and to lead the child along a growth path that always bridges his or her current "zone of proximal development."

The Lesson Creator allows more effective cognition and learning by adding a significant pedagogical feature to Timo Vocabulary by allowing the coach (whether parent, teacher, or peer) to give descriptions of the nouns as well as corrective feedback, which allows errorless learning and encourages the child to think as they are learning vocabulary. For example, Timo might ask the child to click on elephant. If the child clicks on elephant, Timo would say "Right on, you clicked on elephant. An elephant is an animal with a long trunk" On the other hand, if the child clicked on giraffe instead of elephant, Timo would say, "You clicked on giraffe. A giraffe is an animal with a long neck. An elephant is an animal with a long trunk. Can you click on elephant?" With this type of supportive and corrective feedback, the child learns about both animals, and is encouraged to think about their differences. This kind of language interaction is denied to many deaf children due to a lack of fully effective communication with their hearing parents. This interaction could be further facilitated and enhanced by including multimodal embellishments in the communication setting.

The Lesson Creator assists in the specialization and individualization of vocabulary and grammar lessons by allowing teachers to create customized vocabulary lists from words already in the application or with new words. If a teacher is taking her class on a field trip to the local Aquarium, for example, she will be able to create lessons about the marine animals the children will see. A parent could prepare lessons with words in the child's current reading, or names of her relatives, schoolmates, and teachers. Lessons can also be easily created for the child's current interests. Most importantly, given that vocabulary is essentially infinite in number, it is most efficient to instruct vocabulary Just in Time as it is needed.

### 3.1 Conclusion

The Lesson Creator is highly functional. For example, I have also used the Lesson Creator to implement science lessons. Although relatively mellow by video game standards, it has enough engaging interactive features to engage the student in mastering the lesson. Some of these properties include 1) providing information when needed, 2) operating at the outer edge of the student's competence because of the value of learning at the zone of proximal development, 3) rewarding commitment of the learner's self, and 4) encouraging the student to think about the relationship among the things being learned. The resulting lessons encompass and instantiate the developments in the pedagogy of how content is learned, remembered and used. We have found that mimicking various aspects of video games enhance learning effectiveness (Gee, 2003). Finally, as described earlier, we have aimed to be faithful to the valuable principles of educational practice that were proposed by Montessori about a century ago.

### 4. ACKNOWLEDGEMENT

Baldi<sup>®</sup> is a trademark of Dominic W. Massaro. The research described in this paper and writing of the paper are either currently supported or have had previous support from the Special Hope Foundation, the National Science Foundation, the Public Health Service, a Cure Autism Now Foundation Innovative Technology Award, the National Alliance for Autism Research, and the University of California, Santa Cruz. The author thanks Joseph Bunnett and Alexandra Jesse for their comments.

### 5. REFERENCES

- Barker, L. J. (2003). Computer-assisted vocabulary acquisition: The CSLU vocabulary tutor in oral-deaf education. Journal of Deaf Studies and Deaf Education, 8, 187-198.
- [2] Beck, I. L., McKeown, M. G., & Kucan, L. (2002). Bringing words to life: Robust Vocabulary Instruction. New York: The Guilford Press.
- [3] Bosseler, A. & Massaro, D.W. (2003). Development and Evaluation of a Computer-Animated Tutor for Vocabulary and Language Learning for Children with Autism. Journal of Autism and Developmental Disorders, 33, 653-672.
- [4] Connors Fortier, Pamela (July 5, 2007). Teaching Staff, Tucker-Maxon School of Oral Education. Report by email correspondence.
- [5] Ericsson, A.; Charness, N.; Feltovich, P. J.; & Hoffman, R. R. (Eds.) (2006). The Cambridge Handbook of Expertise and Expert Performance. Cambridge: Cambridge University Press, 2006.
- [6] Gee, J.P. (2003). What video games have to teach us about learning and literacy. New York: Palgrave/Macmillan.
- [7] Jesse, A., Vrignaud, N., & Massaro, D. W. (2000/01). The processing of information from multiple sources in simultaneous interpreting. Interpreting 5, 95-115.
- [8] LaSasso, C. & Davey, B. (1987). The relationship between lexical knowledge and reading comprehension for prelingually, profoundly hearing impaired students, The Volta Review, 89, 211-220.
- [9] Marschark, M., Pelz, J. B., Convertino, C., Sapere, P., Arndt, M. E., & Seewagen, R. (2005). Classroom Interpreting and Visual Information Processing in Mainstream Education for Deaf Students: Live or Memorex®? Journal of American Educational Research, 42, 727-761).
- [10] Marsh, E.J., Roediger, H.L., Bjork, R.A. &Bjork, E.L. (2007). The Memorial Consequences of Multiple-Choice Testing. Psychonomic Bulletin & Review, 14, 194-199.
- [11] Massaro, D. W. (1998). Perceiving Talking Faces: From Speech Perception to a Behavioral Principle. MIT Press: Cambridge, MA.
- [12] Massaro, D. W. (2004). Symbiotic Value of an Embodied Agent in Language Learning. In R.H. Sprague, Jr.(Ed.), Proceedings of 37th Annual Hawaii International Conference on System Sciences, (HICCS'04) (CD-ROM,10 pages). Los Alimitos, CA: IEEE Computer Society Press. Best paper in Emerging Technologies.
- [13] Massaro, D.W. (2006a). A computer-animated tutor for language learning: Research and applications. In P.E. Spencer,& M. Marshark (Eds.), Advances in the spoken language development of deaf and hard-of-hearing children (pp. 212-243). New York, NY: Oxford University Press.
- [14] Massaro, D.W. (2006b). Embodied Agents in Language Learning for Children with Language Challenges. In K. Miesenberger, J. Klaus, W. Zagler, & A. Karshmer, (Eds.), Proceedings of the 10th International Conference on Computers Helping People with Special Needs, ICCHP 2006 (pp.809-816). University of Linz, Austria. Berlin, Germany: Springer.

- [15] Massaro, D.W. & Bosseler, A. (2006). Read my lips: The importance of the face in a computer-animated tutor for autistic children learning language. Autism: The International Journal of Research and Practice, 10, 495-510.
- [16] Massaro, D.W., Cohen, M.M., & Beskow, J. (1999). From theory to practice: rewards and challenges. In Proceedings of the International Conference of Phonetic Sciences (pp. 1289-1292). San Francisco, CA. Regents of the University of California.
- [17] Massaro, D. W., & Light (2003, September). Read my tongue movements: bimodal learning to perceive and produce non-native speech /r/ and /l/. Eurospeech 2003-Switzerland (Interspeech). 8th European Conference on Speech Communication and Technology, Geneva, Switzerland.
- [18] Massaro, D.W., & Light, J. (2004a). Improving the vocabulary of children with hearing loss. Volta Review, 104(3), 141-174.
- [19] Massaro, D.W., & Light, J. (2004b). Using visible speech for training perception and production of speech for hard of hearing individuals. Journal of Speech, Language, and Hearing Research, 47(2), 304-320.
- [20] Massaro, D.W., Ouni, S., Cohen, M.M., & Clark, R. (2005). A Multilingual Embodied Conversational Agent. In R.H. Sprague (Ed.), Proceedings of 38th Annual Hawaii International Conference on System Sciences, (HICCS'05) (CD-ROM,10 pages). Los Alimitos, CA: IEEE Computer Society Press.
- [21] McDaniel, M.A., Roediger, H.L. & McDermott, B. (2007). The Memorial Consequences of Multiple-Choice Testing. Psychonomic Bulletin & Review, 14, 200-206.
- [22] Metcalfe, J. & Kornell, N. (2007). Principles of Cognitive Science in Education: The Effects of Generation, Errors, and Feedback. Psychonomic Bulletin & Review, 14, 225-229.
- [23] Pashler, H., Rohrer, D., Cepeda, N.J & Carpenter, S. K. (2007). Enhancing Learning and Retarding Forgetting: Choices and Consequences. Psychonomic Bulletin & Review, 14, 187-193.
- [24] Rizzolatti, G., & Craighero, L. (2004). The mirror-neuron system. Annual Review of Neuroscience, 27, 169-192.
- [25] Soland, Christine (June 15, 2007). Speech Language Pathologist, Tucker-Maxon School of Oral Education. Report by email correspondence.
- [26] Stigler, J. W. (1984). "Mental abacus": The effect of abacus training on Chinese children's mental calculation. Cognitive Psychology, 16, 145–176.
- [27] Stoll-Lillard, A. (2005). Montessori: The Science Behind the Genius. Oxford: Oxford University Press.
- [28] Sumby, W. H., & Pollack, I., 1954. Visual contribution to speech intelligibility in noise. Journal of the Acoustical Society of America, 26, 212-215.
- [29] Tartaro, A. & Cassell, J. (2008). "Playing with Virtual Peers: Bootstrapping Contingent Discourse in Children with Autism" Proceedings of International Conference of the Learning Sciences (ICLS), June 24-28, Utrecht, Netherlands.