Paper for 15th International Congress of Phonetic Sciences, Barcelona, Spain, 3-9 August, 2003

Development and Evaluation of a Computer-Animated Tutor for Language and Vocabulary Learning

Dominic W. Massaro†, Alexis Bosseler† and Joanna Light†

† Perceptual Science Laboratory, Department of Psychology University of California, Santa Cruz

Santa Cruz, CA. 95064 U.S.A.

E-mail: massaro@fuzzy.ucsc.edu, abosseler@yahoo.com, joannalight@hotmail.com

ABSTRACT

After making the case for the value of the direct teaching of vocabulary and computer-aided instruction, we review two studies using a Language Wizard/Player for teaching new vocabulary items to hard of hearing and autistic children. To insure that the program itself was responsible for the learning, we used a within student multiple baseline design where certain words were continuously being tested while other words were being tested and trained. Knowledge of the words remained negligible without training and learning occurred fairly quickly for all words once training began Finally, knowledge of the trained words did not degrade after training, generalized to new images, and was retained at least 4 weeks after training ended.

1. INTRODUCTION

The Essential Role for Vocabulary Knowledge

Although there is no consensus on the best way to teach or to learn language, there are important areas of agreement. One is the central importance of vocabulary knowledge for understanding the world and for language competence in both spoken language and in reading (Gupta & MacWhinney, 1997). There is empirical evidence that very young children more easily form conceptual categories when category labels are available than when they are not (Waxman, 2002). Increasing the pervasiveness and effectiveness of vocabulary learning offers a huge opportunity for improving conceptual knowledge and language competence for all individuals, whether or not they are disadvantaged because of sensory limitations, learning disabilities, or social condition. Finally, vocabulary knowledge is positively correlated with both listening and reading comprehension (Anderson & Freebody, 1981), and predicts overall success in school (Vermeer, 2001).

Validity of the Direct Learning of Vocabulary

These are important reasons to justify the need for direct teaching of vocabulary. Although there is little emphasis on the acquisition of vocabulary in typical school curricula, research demonstrates that some direct teaching of vocabulary is essential for good language development (Beck et al., 2002). Contrary to a common belief that learning vocabulary is a necessary outcome of reading in which new words are experienced in a meaningful context, context seldom disambiguates the meaning of a word sufficiently. Finally, knowing a word is not an all-or-none proposition. Acquiring semantic representations appears to be a gradual process that can extend across several years (McGregor et al, 2002). Thus, it is important to overtrain vocabulary, and to present the items in a variety of contexts in order to develop rich representations. Picture naming and picture drawing are techniques that can be used to probe and reinforce these representations.

Effectiveness of Computer-Based Instruction

Computer-based instruction is emerging as a prevalent method to train and develop vocabulary knowledge for both native and second-language learners (Druin & Hendler, 2000; Wood, 2001) and individuals with special needs (Barker, in press). 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 (Chun & Plass, 1996). 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. Dubois and Vial (2000), for example, found an increase in recall of second-language vocabulary when training consisted of combined presentations of spoken words, images, written words and text relative to only a subset of these.

"Baldi" and Potential for Student Engagement

Computer-based instruction makes it possible to include embodied conversational agents rather than simply text or disembodied voices in lessons. There are several reasons why the use of auditory and visual information from a talking head is so successful, and why it holds so much promise for language tutoring (Massaro, 1998). These include a) the information in visible speech, b) the robustness of visual speech, c) the complementarity of auditory and visual speech, and d) the optimal integration of these two sources of information (Massaro, 1998).

We have also witnessed that the student's engagement is enhanced by face-to-face interaction with Baldi, our 3-D computer-animated talking head. Baldi provides realistic visible speech that is almost as accurate as a natural speaker (Massaro, 1998, Chapter 13). The quality and intelligibility of Baldi’s visible speech has been repeatedly modified and evaluated to accurately simulate naturally talking humans (Massaro, 1998). Baldi’s visible speech can be appropriately aligned with either synthesized or natural auditory speech. Baldi also has teeth, tongue and palate to simulate articulation inside of the mouth, and the tongue movements have been trained to mimic natural tongue movements. This technology has the potential to help individuals with various language challenges and those learning a new language.

Language Wizard and Player

Our Language Wizard is a user-friendly application that allows the composition of lessons with minimal computer experience. The Language Player engages the student in lesson (Bosseler & Massaro, in press). They encompass and instantiate the developments in the pedagogy of how language is learned, remembered and used. Education research has shown that children can be taught new word meanings by using drill and practice methods. It has also been convincingly demonstrated that direct teaching of vocabulary by computer software is possible and that an interactive multimedia environment is ideally suited for this learning (Berninger & Richards, 2002; Wood, 2001). Following this logic, many aspects of our lessons enhance and reinforce learning. For example, the existing program makes it possible for the students to 1) Observe the words being spoken by a realistic talking interlocutor (Baldi), 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, and 6) Spell the word by typing, and 7) Observe and respond to the word used in context.

Other benefits of our program include the ability to seamlessly meld spoken and written language, provide a semblance of a game-playing 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 Wizard allows the coach to exploit this zone with individualized lessons, and with lessons that can bypass repetitive training when student responses indicate that material is mastered.

2. EXPERIMENTAL EVALUATION

Effectiveness for Hard of Hearing Children

The Language Wizard/Player has also been in use at the Tucker Maxon Oral School in Portland, Oregon, and Barker (in press) examined if training with the animated tutor software would result in vocabulary acquisition and retention among hard of hearing and hearing children. Students were given cameras to photograph objects and surroundings at home. The pictures of these objects were then incorporated as items in the lessons, using the Language Wizard. A given lesson had between 10 and 15 items. Students worked on the items about 10 minutes a day until they reached 100% on the posttest. They then moved on to another lesson. About one month after each successful (100%) posttest, they were retested on the same items. Ten girls and nine boys participated in the applications. There were six hard of hearing children and one hearing child between 8 and 10 years of age in the “lower school.” Ten hard of hearing and two hearing children, between ages 11 and 14, participated from the “upper school.”

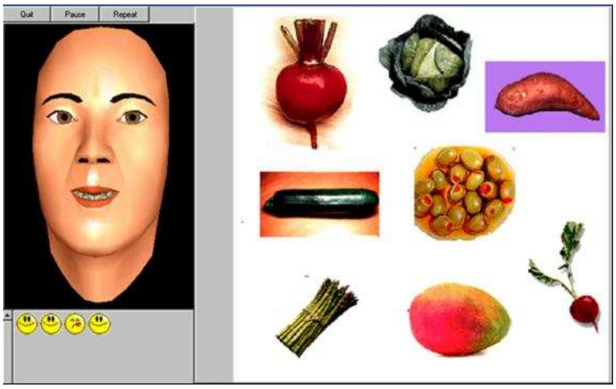

Figure 1. A computer screen from a vocabulary lesson on fruits and vegetables, illustrating the format of the Language Player during one of the exercises. Each lesson contains Baldi, the vocabulary items and written text (not present in this exercise), and “stickers”. In this application the students learn to identify fruits and vegetables. For example, Baldi says "Click on the beet". The student clicks on the appropriate region and feedback in the form of Baldi’s spoken reaction and stickers (e.g., happy and disgusted faces) are given for each response.

Similar results occurred for the two groups. Students knew some of the items without any learning, they successfully learned the other items, and retained about one-half of the newly learned items when retested 30 days later.

The results of the Barker (in press) evaluation indicated that hard of hearing children learned a significant number of new words, and retained about half of them a month after training ended. No control groups were used in that evaluation, however, and it is possible the children were learning the words outside of the Language Player environment. Furthermore, the time course of learning with the Language Player was not evaluated. It is of interest how quickly words can be learned with the Language Player to give some idea of how this learning environment would compare to a real teacher. Finally, both identification and production of the words should be assessed given that only identification was measured previously.

To address these issues, Massaro and Light (2003) carried out an experiment based on a within student multiple baseline design (Baer et al., 1968) where certain words were continuously being tested while other words were being tested and trained. Although the student's instructors and speech therapists agreed not to teach or use these words during our investigation, it is still possible that the words could be learned outside of the Language Player environment. The single student multiple baseline design monitors this possibility by providing a continuous measure of the knowledge of words that are not being trained. Thus, any significant differences in performance on the trained words and untrained words can be attributed to the Language Player training program itself rather than some other factor.

Eight hard of hearing children, 2 male ages 6 and 7, 6 female ages 9 and 10, were recruited from The Jackson Hearing Center in Los Altos, California and were given parental consent to participate. The male students were in grade 1 and the female students in grade 4 respectively and all students needed help with their vocabulary building skills as suggested by their regular day teachers. One child had a cochlear implant and the seven other children had hearing aids in both ears except for one child with an aid in just a single ear. Using the Language Wizard, the experimenter developed a set of lessons with a collection of vocabulary items that was individually tailored for each student. Each collection of items was comprised of 24 items, broken down into 3 categories of 8 items each. Three lessons with 8 items each were made for each child.

Images of the vocabulary items were presented on the screen next to Baldi as he spoke, as illustrated in Figure 1. Some of the exercises required the child to respond to Baldi's instructions such as "click on the cabbage", or "show me the yam", by clicking on the highlighted area or by moving the computer mouse over the appropriate image until an item was highlighted and then clicking on it. Two other exercises asked the child to recognize the written word and to type the word, respectively. The production exercises asked the child to repeat after Baldi once he named the highlighted image or to name the highlighted image on their own, followed by Baldi's naming of the image.

Figure 2 gives the results of identification and production for one of the eight students. The results were highly consistent across the eight students. As expected, identification accuracy (mean = .72) was always higher than production accuracy (mean = .64). This result is not unexpected because a student could know the name of an item without being able to pronounce it correctly. There was little knowledge of the test items without training, even though these items were repeatedly tested for many days. Once training began on a set of items, performance improved fairly quickly until asymptotic knowledge was obtained. This knowledge did not degrade after training on these words ended and training on other words took place. In addition, a reassessment test given about 4 weeks after completion of the experiment revealed that the students retained the items that were learned.

The average number of trials required to reach criterion was 5, 4.3, and 3.4 for mastering the first, second, and third sets of categories. Given that the word lists were randomized across participants, this significant difference indicates that inherent differences in the difficulty of the word sets are

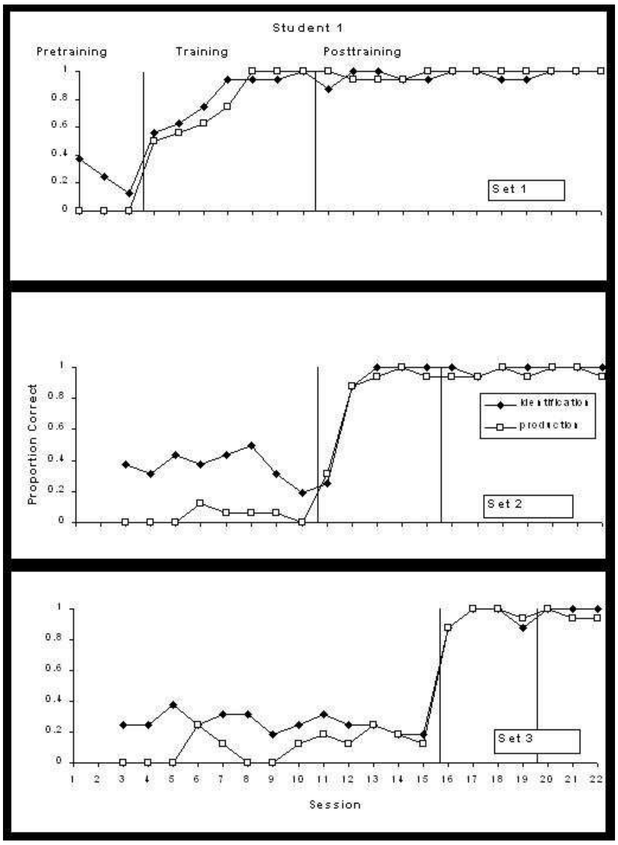

Figure 2. Proportion of correctly identified (solid black triangles) and correctly produced (empty white squares) items across the testing sessions for student 1. The training occurred between the two vertical bars. The figure illustrates that once training was implemented identification performance increased dramatically, and remained accurate without further training.

probably not responsible and that there was a learning-to-learn process within the context of the training procedure. Not surprisingly, the students may have become increasingly more familiar with Baldi and the training.

Effectiveness for Children with Autism

The Language Wizard/Player has also been used in evaluating vocabulary acquisition, retention and generalization in children with autism (Bosseler & Massaro, in press). This study consisted of two phases. Phase 1 measured vocabulary acquisition and retention. Phase 2 tested whether vocabulary acquisition was due to the Language Player or outside sources and whether the acquired words could be generalized to new images and outside of the Language Player environment. Vocabulary lessons were constructed, consisting of over vocabulary items selected from the curriculum of two schools (Bosseler & Massaro, in press). The participants were eight children diagnosed with autism, ranging in age from 7-11 years. All of the students exhibit delays in all areas of academics, particularly in the areas of language and adaptive functioning. Seven of the eight children were capable of speech.

The average results indicated that the children learned many new words, grammatical constructions and concepts, proving that the Language Player provided a valuable learning environment for these children. In addition, a delayed test given more than 30 days after the learning sessions took place showed that the children retained the words that they learned. This learning and retention of new vocabulary, grammar, and language use is a significant accomplishment for autistic children.

Although all of the children demonstrated learning from initial assessment to final reassessment, it is possible that the children were learning the words outside of our learning program (for example, from speech therapists or in their school curriculum). Furthermore, it is important to know whether the vocabulary knowledge would generalize to new pictorial instances of the words. To address these questions, a second investigation used the single subject multiple probe design, as was done in Massaro and Light (2003) study. Once a student achieved 100% correct, generalization tests and training were carried out with novel images. The placement of the images relative to one another was also random in each lesson. Assessment and training continued until the student was able to accurately identify at least 5 out of 6 vocabulary items across four unique sets of images. The students identified significantly more words following implementation of training compared to pre-training performance, showing that the program was responsible for learning. Learning also generalized to new images in random locations, and to interactions outside of the Language Player. These results show that our learning program is effective for children with autism, as it is for children who are hard of hearing.

ACKNOWLEDGEMENTS

The research and writing of the paper were supported by the National Science Foundation (Grant No. CDA-9726363, Grant No. BCS-9905176, Grant No. IIS-0086107), Public Health Service (Grant No. PHS R01 DC00236), a Cure Autism Now Foundation Innovative Technology Award, and the University of California, Santa Cruz.

REFERENCES

[1] Anderson, R. C., & Freebody, P. (1981). Vocabulary knowledge. In J. T. Guthrie (Ed.), Comprehension and teaching: Research perspectives (pp. 71-117). Newark, DE: International Reading Association.

[2] Baer, D, M., Wolf, M. M., Risley, T. R. (1968). Some current dimenstions of applied behavior analysis. *Journal of Applied Behavior Analysis, 1*, 91-97.

[3] Barker, L. J. (in press). Computer-assisted vocabulary acquisition: The CSLU vocabulary tutor in oral-deaf education. *Journal of Deaf Studies and Deaf Education,* in press.

[4] Beck, I. L., McKeown, M. G., & Kucan, L. (2002). *Bringing words to life: Robust Vocabulary Instruction.* New York: The Guilford Press.

[5] Berninger, V. W., & Richards, T. L. (2002). *Brain literacy for educators and psychologists.* San Diego: Academic Press.

[6] Bosseler, A. and Massaro, D.W. (in press). Development and Evaluation of a Computer-Animated Tutor for Vocabulary and Language Learning for Children with Autism. *Journal of Autism and Developmental Disorders*. http://mambo.ucsc.edu/pdf/autism.pdf

[7] Chun, D. M., & Plass, J. L. (1996). Effects of multimedia annotations on vocabulary acquisition. *Modern Language Journal. 80,* 183-198

[8] Druin, A., & Hendler, J. (Eds.) (2000). Robots for Kids: *Exploring new technologies for learning.* San Francisco: Morgan Kaufmann.

[9] Dubois, M., & Vial, I. (2000). Multimedia design: the effects of relating multimodal information. Journal of *Computer Assisted Learning, 16*, 157-165.

[10] Gupta, P., & MacWhinney, B. (1997). Vocabulary acquisition and verbal short-term memory: computation and neural bases. *Brain and Language, 59*, 267-333.

[11] Massaro, D. W. (1998). *Perceiving Talking Faces: From Speech Perception to a Behavioral Principle.* MIT Press: Cambridge, MA.

[12] Massaro, D.W., & Light, J. (2003). Improving the vocabulary of children with hearing loss, submitted.

[13] McGregor, K. K., Friedman, R. M., Reilly, R. M., Newman, R. M. (2002). Semantic representation and naming in young children. *Journal of Speech, Language, and Hearing Research, 45*, 332-346.

[14] Qian D.D. (2002). Investigating the Relationship Between Vocabulary Knowledge and Academic Reading Performance: An Assessment Perspective. *Language Learning, 52*, 513-536.

[15] Vermeer, A. (2001). Breadth and depth of vocabulary in relation to L1/L2 acquisition and frequency of input. *Applied Psycholinguistics, 22,* 217-234.

[16] Waxman, S. R. (2002). Early word-learning and conceptual development: Everything had a name, and each name gave birth to a new thought. In U. Goswami (Ed.) Blackwell *Handbook of childhood cognitive development* (pp. 102-126). Malden, MA: Blackwell publishing.

[17] Wood, J. (2001). Can software support children's vocabulary development? *Language Learning & Technology, 5*, 166-201.