Serious Gaming and Vocabulary Growth
Research into the Effectiveness of a Serious Game
upon Receptive Vocabulary

Uriël Schuurs

Cito (Central Institute for Educational Measurement), Arnhem, The Netherlands
uriel.schuurs@cito.nl

Serious games claim to be both entertaining and instructive [17]. However, research into the actual effects of serious games is rare. This paper presents an effect study that was conducted in 2010 in order to evaluate the learning effects of the serious game Mijn naam is Haas (My name is Haas) upon vocabulary growth in children in grade 1 and grade 2 (age 4 – 7).

1   Introduction

The educational computer game Mijn Naam is Haas (My name is Haas) provides a playful learning environment for children aged 3 to 7. Within the online serious game children create their own interactive story by drawing the world of character Haas. Sometimes they have to move this character around and the child can also draw solutions to problems and add elements, like trees, birds, mushrooms and insects, by drawing them right into the virtual landscape. These elements then become part of the story and may even change the plot. In addition to the online game there are CD-ROM's and picture books, some of them with circular stories. Currently products are used at home and in the classroom in the Netherlands. A distinction has been made between formal and informal learning [16] and the conclusion seems to be that the main difference between the two is in the environment: at school or at home. One might argue that due to serious games like Mijn naam is Haas the threshold between formal and informal learning is diminishing.

In 2009 the developers of the serious game My name is Haas asked Radboud university Nijmegen for advice on the language use within the game to be developed. We especially focused upon the selection of target words, that is: words to be captured by the children. We decided not to make use of word frequency as a selection criterion as this does not seem to be a fruitful approach, neither in instructional design nor in estimating productive vocabulary size [11]. Instead use was made of the Streeflijst Woordenschat [15]: this list reflects teachers´ judgments upon the necessity of a receptive knowledge of the words involved by children in grade 1 up to grade 4. From this Streeflijst we assigned words from different levels to be used in the serious game Mijn naam is Haas. Deliberately words were chosen in such a way that the misunderstanding of a particular word would not hamper the understanding of the game adventure as a whole.

One of the assumptions of the designers of My name is Haas is that the game will increase young children's vocabulary. Beside that the game is supposed to stimulate children's story comprehension and to challenge their problem solving skills. The latter is triggered by means of confronting the child with situations that need some kind of solution. For instance, when Haas is walking through the woods he arrives at a river. The question is how to get to the other side. It's the child's role to come up with an answer, for instance by drawing a bridge or a little boat. The object that the child has drawn determines the way the story continues, so the game adapts itself to the input of the current user. In this way the learning environment is really interactive and it stimulates the child to come up with some kind of solution. Largest effects however were expected within vocabulary. In order to explain this I will highlight some major implications derived from international literature on Vocabulary Acquisition.

2   Theoretical Background

The question arises: How do children learn the meaning of words? When arriving at grade 1 children bring in a vocabulary of about 4.000 words. Apparently the principles that guide the vocabulary growth in the younger years are very effective. In this acquisition process three main principles can be distinguished [1, 13]: labeling, categorizing and network building.

Labeling seems to be a natural behavior for parents as they keep on talking to their children during the day while going out shopping with the children, feeding them, bathing them and putting them to bed. There is a one to one-relationship between the words, the labels they are using and the persons, objects, activities that are in the context of this language use. By labeling the adult language users make clear the relationship between words and objects. At the same time a *categorization* of objects and labels is necessary in order to get grip on the world. While comparing objects such as a couch at home, a wooden chair at Kindergarten or maybe even a desk chair, a child may notice that despite big differences there is one important similarity: all three objects are meant to sit on. A categorization of objects and activities in the world surrounding and in the words we use for labeling them, will lead to the *building of a network*. Children notice there is a relationship between *a pan* where French fries are fried in and a *plate* where the food will be eaten from. Distinctive features are used in order to decide whether a word belongs to one category or another, for instance which piece of furniture belongs to the category home, school or office, and to decide upon the relationship between (groups of) words such as *a dog*, *a duck* or maybe *an egg*.

Discovery of these relationships, both in content and in the syntactic field - that is: network building - is considered to be necessary for acquiring deep word knowledge. Just labeling an object will result in a shallow understanding of the concept [2, 9]. All kinds of criteria are being used by language learners to build up these vocabulary networks: shape, color, function and other distinctive features of the object that is labeled as well as syntactic and semantic characteristics of the word (see [4], Chapter 6 for an overview of studies). We do not know on what basis this labeling is most effective in what situation. However we do know that categorizing and network building is a powerful reinforcement to vocabulary growth.

On the basis of these theoretical insights five basic principles are generally considered to be essential for the stimulation of solid vocabulary growth [6, 21, 22]:

1. The use of a rich learning environment that is challenging for the language learner to learn in a functional and meaningful context.
2. Repeated input of target words: in order to be captured a new target word should be presented several times and in different contexts.
3. In order to build a network several distinctive features of a new target word should be presented, both syntactic possibilities, semantic aspects and the context in which the word is being used.
4. Target words should be presented in an interactive context: just like young children learn the meaning of words while playing and interacting in an verbal way with each other, elementary school pupils profit most from language input in a social environment with a lot of interaction going on. Within a serious game it's possible to reconstruct such an interactive environment.
5. Last but not least active participation: giving language input is necessary but stimulating the language learner to use the new words makes the acquisition process much more effective. In a serious game this is possible without creating anxiety or fear as is the case with more introvert children in classroom situations.

In general one can conclude that serious gaming is suitable in helping building a semantic network, especially when stories are used in which the child is participating actively [12, 17, 23]. Nevertheless, there are few empirical studies. For this reason it is stated that "Nothing much can be said about the effectiveness and efficiency of using educational games" in classroom settings [3]. Focusing on Mijn naam is Haas again it is easily demonstrated that within the game important conditions for effective vocabulary growth [4, 7] are met, including:

- target words that are carefully selected from Streeflijst Woordenschat [15];
- presentation of words in a functional context offering different cues in order to derive word meaning;
- multi-channel presentation: children can hear how the specific target words are being used and they can see the context in which the target words are used;
- target words are presented several times within one play session;
- as stated earlier: the child can participate actively within the context in which the word is being used.

Within the game these principles are operationalized in a nondescript way to the user: the player's attention will be with the game itself, not with the target words that are presented. On this basis the following hypothesis can be formulated: *vocabulary growth will be faster in classrooms where Mijn naam is Haas is used than in classrooms where no use is made of a serious game.*

3   Method

Using funds of the M&ICT program in The Netherlands research could be conducted into the effects of this serious game upon vocabulary growth within 4- 6 years old children. This research was carried out in the autumn of 2010; 12 schools at primary level were involved and test results of 412 children were taken into analysis.

In each of the 12 primary schools one class was allowed to play the game twice a week during 6 weeks (the experimental group). Two themes of the game were presented to the children using the online environment (theme *Creativity* and theme *Food and Drinks*). In the background use was made of the educational method Schatkist; during the experimental phase the pupils were not allowed to use the computer version of Schatkist.

The control group consisted of 12 parallel classes at the same elementary schools. The control group did make use of Schatkist, an exhaustive curriculum that incorporates 16 themes including stories, learning activities and computer games. In order to guarantee that the target words were presented in the control groups as well as in the experimental condition, illustrated short stories containing the target words were presented twice a week in the control group. Moreover, teachers of the control group were instructed as how to present these stories to the children in an interactive way. In order to optimise language input in the control group the interactive reading-instruction was based upon insights in the best practices in reading aloud to young children [6, 18, 22], including the repetition of words, paraphrasing sentences, asking open questions to the children, modeling and elaborating the children's responses and giving positive feedback to reactions of the children.

In both conditions teachers took note of the classroom activities during the experiment. Specifically for this experiment an intervention-related receptive vocabulary test was designed containing 28 multiple choice items related to target words derived from the two themes used in Mijn naam is Haas. In this receptive vocabulary test each test item consists of three picture alternatives while the target word is presented through audio within a sentence. The test was constructed after having pretested several items in June 2010 (mean p-score = 70.5 ; coefficient Alpha = .79). This intervention-related vocabulary test was presented to all children involved before and after the experimental intervention. As an external criterion a second receptive vocabulary test was used, derived from a standardized test set [20]. In fact we used a pretest posttest design that allows for within subject analysis. Data of children that missed one of the test sessions were deleted pairwise.

4   Results

Scores on pretest and posttest are represented in Table 1:

Table 1. Test scores per condition on pretest and posttest

	pretest		posttest		
Condition	mean	St.dev.	mean	St.dev.	n
Control group	20,53	3,695	21,73	3,658	205
Experimental group	21,24	3,755	23,04	3,236	207
Total	20,89	3,738	22,39	3,510	412

The log books of the group of teachers showed that in the experimental condition per child 12 to 20 minutes a week was spent playing the serious game My name is Haas. Within the control group in 4 out of 12 schools a double amount of time was spent reading aloud the stories to the children, discussing word meanings and doing related vocabulary exercises. Nevertheless, using the SPSS Repeated Measurement General Linear Model, the difference in vocabulary growth was significant in favour of the experimental condition. Within this model pretest scores on the TAK test [20] were used as a covariate in order to neutralize small, non-significant group differences at the pretest. Interaction between the variables Vocabulary and Condition turns out to be significant (F = 8,541, df = 1, p = .00), indicating that differences in vocabulary growth were highly dependent on the condition a child was in. As gains in the experimental group were higher the hypothesis is confirmed.

Effect size in vocabulary growth per condition was calculated using formula (1)

$$E = \frac{\mu_2 - \mu_1}{\sigma_1} \tag{1}$$

As a result effect sizes per group could be calculated (see Table 2):

Table 2. Effect sizes per condition

Condition	Means and standard deviations	Effect size
Posttest / pretest for the control group	21,73 – 20,53 / 3,695	0,32
Posttest / pretest for the experimental	23,04 – 21,24 / 3,755	0,48
Posttest data for both groups	23,04 – 21,73 / 3,658	0,36

The results show that vocabulary growth differs significantly in both groups. Table 2 indicates that growth was larger in the experimental group, effect size being 150% of the effect size in the control group.

5   Discussion and Conclusions

What can be concluded on the basis of the results reported? First of all, vocabulary growth was observed in both conditions. The conclusion might be that education - that is: our specific treatment - does at least have some effect. Secondly, there was a

significant interaction between Condition and Vocabulary with a positive result in the experimental group. Our hypothesis has been confirmed.

Some further remarks need to be made. The first conclusion was that education does have effect. In fact, we are not sure about this. We didn't make use of a null-condition without any treatment: especially at a younger age informal language input, from outside of the school is a dominant factor in language acquisition. Though the target words used in our experiment were not very frequent, it might be possible that these words were captured in informal settings we didn´t control.

The reported effect sizes are moderate [5]: although effective, the results of the treatment were not very spectacular. However, with regard to the double amount of time spent in the control group in comparison with the experimental group, the serious game might have generated much more positive results. As ceiling effects may play a role further data analysis could be carried out on the basis of a differentiation in starting level (e.g. low, medium, high).

In this experiment we did not test deep word knowledge [14] and we didn´t get insight in the quality of the mental semantic network that was built by the children [cf. 8]; as a consequence we do not know much about the persistence of the learning results. In addition we did not use a vocabulary test that generates an indication or estimate of growth in receptive vocabulary as a whole (cf. [21].

To broaden the perspective: it is challenging to find out more about the characteristics in gaming that lead to positive learning effects. It might be possible to look for effects in areas that are less obvious and maybe a little more difficult to explore, for instance in the area of problem solving skills or ict-skills. Both more complicated and more interesting will be research focussing on the interaction between learning effects and characteristics of the serious game; game characteristics may be derived from available descriptive studies, cf. [10, 19]. In the long term it might be possible to create games with built-in tests: not merely tests in order to determine game levels that are best suitable, but formative tests that give insight in educational gains or even diagnostic tests that make clear what a pupil has achieved and what elements of the curriculum still need some attention - with or without a serious game.

References

1. Aitchison, J.: Words in the mind. An introduction to the mental lexicon. Blackwell Publishers, Oxford (2002)
2. Beck, I.L., McKeown, M.G., Kucan, L.: Bringing Words to Life. Robust Vocabulary Instruction. The Guilford Press, New York (2002)
3. van den Berg, M., Simons, R.J.: Wat weten we over gaming in het PO en het VO? Stichting Kennisnet. Den Haag (2008)
4. Bloom, P.: How children learn the meaning of words. MIT Press, Cambridge (2002)

5. Cohen, J.: Statistical Power Analysis for the Behavioral Sciences. Academic Press, New York (1969)

6. Damhuis, R., Litjens, P.: Mondelinge Communicatie. Expertisecentrum Nederlands, Nijmegen (2003)

7. Duerings, J., van der Linden, B., Schuurs, U., Strating, H.: Op Woordenjacht. Creatief en effectief werken aan woordenschatuitbreiding, Uitgeverij Garant, Antwerpen, Apeldoorn (2011)

8. Greidanus, T., Nienhuis, L.: Testing the Quality of Word Knowledge in a Second Language by Means of Word Associations: Types of Distractors and Types of Associations. The Modern Language Journal 85(4), 567–577 (2001)

9. Haastrup, K., Henriksen, B.: Vocabulary acquisition: acquiring depth of knowledge through network building. International Journal of Applied Linguistics 10(2), 221–240 (2000)

10. Johnson, W.L., Vilhjalmsson, H., Marsella, S.: Serious Games for Language Learning: How Much Game, How Much AI? In: Proceedings of the 2005 Conference on Artificial Intelligence in Education: Supporting Learning through Intelligent and Socially Informed Technology (2005)

11. Meara, P.: Lexical Frequency Profiles: A Monte Carlo Analysis. Applied Linguistics 26(1), 32–47 (2005)

12. Nagy, W.E.: Why vocabulary instruction needs to be long-term and comprehensive. In: Hiebert, E.H., Kamil, M.L. (eds.) Teaching and Learning Vocabulary: Bringing Research to Practice, pp. 27–44. Erlbaum, New York (2005)

13. Nelson, K.: Language in cognitive development. Cambridge University Press, Cambridge (1996)

14. Read, J.: The development of a new measure of L2 vocabulary knowledge. Language Testing 10(3), 355–371 (1993)

15. Schaerlaekens, A., Kohnstamm, D., Lejaegere, M., de Vries, A., Peeters, L., Zink, I.: Streeflijst woordenschat voor 6-jarigen. Swets & Zeitlinger, Lisse (1999)

16. Selfton-Green, J.: Literature review in informal learning with technology outside school: A report for NESTA futurelab. NESTA Futurelab, Bristol (2006)

17. Sørensen, B.H., Meyer, B.: Serious games in language learning and teaching – a theoretical perspective. In: Situated Play, Proceedings of DiGRA Conference (2007)

18. Stoep, J.: Interactief voorlezen: een goed begin. Onderwijskunde in theorie en praktijk, red. L. Verhoeven en R. Voeten, Zwijssen, Tilburg, pp. 94–104 (2004)

19. Thorne, S.L., Black, R.W.: Language and literacy development in computer-mediated contexts and communities. Annual Review of Applied Linguistics 27, 1–28 (2007)

20. Verhoeven, L., Vermeer, A.: Taaltoets Alle Kinderen. Cito, Arnhem (2001)

21. Vermeer, A.: Breadth and depth of vocabulary in relation to L1/L2 acquisition and frequency of input. Applied Linguistics 22(2), 217–234 (2001)

22. Wasik, B.A., Bond, M.A.: Beyond the pages of a book: Interactive book reading and language development in preschool classrooms. Journal of Developmental Psychology 93(2), 243–250 (2001)

23. Wong, W.L., Shen, C., Nocera, L., Carriazo, E., Tang, F., Bugga, S., Narayanan, H., Wang, H., Ritterfeld, U.: Serious Video Game Effectiveness. In: Proceeding ACE 2007 Proceedings of the International Conference on Advances in Computer Entertainment Technology. ACM, New York (2007)

24. Yusoff, A., Crowder, R., Gilbert, L., Wills, G.: A conceptual framework for serious games. In: 2009 Ninth IEEE International Conference on Advanced Learning Technologies, Riga, Latvia, July 15 - 17 (2009)