Effects of Position of Real-Time Translation on AR Glasses

Anonymous Author(s)*

ABSTRACT

Augmented reality (AR) provides users with contextually relevant multimedia content by overlaying it on real-world objects. However, overlaying virtual content on real-world objects can cause occlusion. Especially for learning use-cases, the occlusion might result in missing real-world information important for learning gain. Therefore, it is important to understand how virtual content should be positioned relative to the related real-world information without negatively affecting the learning experience. Thus, we conducted a study with 12 participants using AR glasses to investigate the position of virtual content using a vocabulary learning task. Participants learned foreign words shown in the surrounding while viewing translations using AR glasses as an overlay, on the right or below the foreign word. We found that showing virtual translations on top of foreign words significantly decreases comprehension and increase users' task load. Insights from our study inform the design of applications for AR glasses supporting vocabulary learning.

CCS CONCEPTS

• **Human-centered computing** → **Mixed / augmented reality**; **User studies**.

KEYWORDS

Augmented reality, AR, overlay, vocabulary learning

ACM Reference Format:

Anonymous Author(s). 2020. Effects of Position of Real-Time Translation on AR Glasses. In *MuC '20: Mensch und Computer, September 06–09, 2020, Magdeburg*. ACM, New York, NY, USA, 8 pages. https://doi.org/10.1145/1122445.1122456

ACM ISBN 978-1-4503-9999-9/18/06…$15.00

https://doi.org/10.1145/1122445.1122456

1 INTRODUCTION

Augmented Reality (AR) allows enriching real-world physical environment by layering virtual information over the physical world [7]. By seamlessly integrating virtual content within the user's environment, AR provides users with access to interactive and contextually relevant virtual information that seemingly coexists in the real world [1, 4, 6]. The augmented information may include various multimedia content that can be perceived and interacted within the real-world environment.

AR is applied in many application areas, such as navigation, training, manufacturing, gaming, and particularly education and teaching [25]. Previous work investigates the potential of using AR for enhancing learning experiences [5, 30, 32]. It is used to support learning in different domains, including learning history [27], chemistry [10], mathematics [24], and foreign languages [33, 35]. It has been shown that AR can enhance problem-solving and collaboration among students [3, 42], increase their learning interest and motivation [38], and contribute to the overall learning outcome [41].

Previous work argued that embodied cognition and interactivity are the advantages that AR provides for learning [34]. Another advantage of using AR for learning is displaying the explicit relationship of virtual information and real-world objects in the users' environment. For example, contextually relevant virtual content might overlay a physical object to provide additional information about it. It was argued that virtual content presented together with contextually related real-world information might act as a memory cue and facilitate the memorability of the real-world content [15, 16]. However, since AR systems seamlessly layer virtual content on top of the environment, crucial real-world information can be occluded. Especially in a learning context, this might have adverse consequences, such as missing or not paying attention to significant learning material.

Some research-related and commercial AR applications that support language learning activities [8], such as TranslatAR [14] or Google Translate [1], overlay foreign words with translations. In these cases, the real-world information is completely hidden from the user while using such applications. As the virtual content can also be moved or displayed around real-world target objects, it is currently unknown

[1] https://translate.google.com

how the position of the virtual content affects the learning experience.

In this paper, we investigate the effect of position of augmented information using binocular see-through AR glasses during a learning task. As the learning use case, we use a vocabulary learning task and a dictionary strategy [20]. Vocabulary learning tasks are frequently used for AR studies by previous work [19, 33, 34]. Through an AR user study, we compared three positions for displaying translations for the vocabulary. Translations were overlaid on top, displayed below, or shown on the right of unfamiliar foreign words. We selected the below and right positions due to the reading direction in the country of the study (left to right and top to bottom). We conducted a study with an application we developed for the Microsoft HoloLens. Participants could interact with the virtual translations by turning them on or off using the HoloLens clicker. During the study, participants learned vocabulary in the foreign language while reading the words in the real world and viewing the translation on the above-mentioned positions using AR glasses. By comparing the positions of AR translation during the vocabulary learning task, we show that overlaying translations on top of unfamiliar foreign words using AR glasses decreases comprehension and increases task load.

2 RELATED WORK

Our work is based on previous research investigating AR for language learning and the effects of text placement on AR smart glasses that we discuss in the following.

AR for Language Learning

To create interactive learning environments, AR allows embedding educational experiences within the real-world environment by overlaying virtual content on top of the physical world. This enables lasting connections within users' knowledge base [5]. A recent meta-analysis of using AR for learning by Garzón and Acevedo [18] identified that AR has a medium effect on the learning gains of students. Interestingly, 82% of the articles collected to conduct their meta-analysis were intended to enhance language learning experiences. It was mentioned that learning gains and motivation are the two most indicated advantages of using AR for language learning.

Previous work investigated the potential of AR to support vocabulary learning. Ibrahim *et al.* [22] conducted an experiment with a system that augmented physical objects around the participants with translations in an unfamiliar foreign language. They found that vocabulary learning with AR is both more effective and more enjoyable compared to learning using the flashcard method. Participants who learned using AR remembered more vocabulary after a four day delayed post-test. Previous work also investigated applying a game-based learning approach using AR technology to support vocabulary learning [9, 26]. The findings indicate that using these kinds of AR vocabulary learning systems improves learners' motivation and facilitates their language learning.

Previous work also investigated real-time translation tools using AR. However, these works were mainly focused on text recognition and translation techniques without considering the language learning experience or the usability of these tools. TranslatAR [14] is a handheld AR application that translates text in real-time and displays it as an overlay. Meda and Kumar [28] presented an application that allows augmenting Telugu text on top of English text in real-time. Toyama *et al.* [40] proposed using eye-tracking to detect the users' attention area on a document in a foreign language and overlay the translation in that area.

Text Placement on AR Smart Glasses

Previous work investigated various aspects of text placement on AR smart glasses. In a study, Orlosky *et al.* [29] found that users prefer placing text in the environment than on the screen of AR smart glasses. Furthermore, they described a system that uses a camera to track dark and uniform areas within the user's field of view to display text in real-time using AR smart glasses to improve readability. Similarly, by evaluating the images of the scenes behind the smart glasses taken from a camera attached to it, Tanaka *et al.* [37] suggested a method to determine the area to display text while walking. Chua *et al.* [11] investigated nine different display positions of monocular smart glasses for showing information while performing a visually intensive primary task. Their findings suggest using middle-center and bottom-center positions for dual-task scenarios when the information is urgent. However, they recommend using middle-right, top-center, and top-right positions for dual-task scenarios when the information is not urgent as the center of the vision is needed for primary tasks, especially when the smart glasses have to be used for a long duration. Rzayev *et al.* [31] compared top-right, center, and bottom-center text positions while walking and sitting. They found that independent of the mobility displaying text in the top-right position of smart glasses increases subjective workload and reduces comprehension.

Summary

In summary, vocabulary learning is an essential part of language learning. Previous work showed that AR can facilitate language learning experiences and increase users' motivation. Especially, using AR smart glasses to display real-time translations in the environment is a promising approach for various real-world scenarios. For example, Google Translate, a commercial real-time translation application with AR functionality, is widely used for language learning [8]. Previous

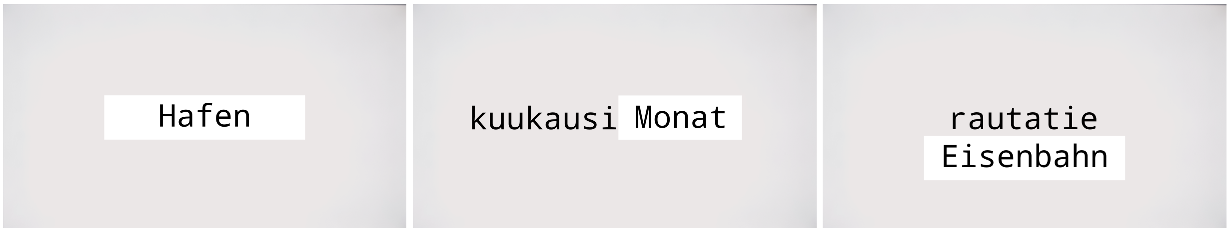

Figure 1: Screenshot views from HoloLens displaying *overlay, right* and *below* POSITIONS. Words in the Finnish language were displayed on the screen of a 46" TV. The translation into German was presented using the AR glasses.

work investigated the effect of different positions for textual content using AR smart glasses. Furthermore, previous work on AR-based translation used different positions for translations. While Fragoso *et al.* [14] overlayed the words in a foreign language with the translation, several works [28, 36] showed the translations below the word in the environment using AR. However, several commercial real-time translation applications (*e.g.*, Zoi Meet application for Vuzix Blade smart glasses [2] or the Word Lens application for Google Glass [3]) display translations not in the user's environment but on the smart glasses' screen. Consequently, it is not clear how the position for real-time translations using AR smart glasses support the learning experience. Therefore, insights on the position of translated text in AR glasses are missing.

3 METHOD

We conducted a study to investigate how the position of a virtual translation using AR glasses affects the learning experience. In the study, we compared three Positions for the translations around or over the words in a foreign language. The translation was layered on top of the foreign word *overlay* similar to the mobile Google Translate application, displayed on the *right* or *below* the foreign word, as used in previous work on real-time translation using AR [13, 39] (see Figure 1). We selected the right and below positions due to the reading direction (left to right and top to down) in the country of the study. The unfamiliar foreign words were translated into German, the official language in the country of the study. As the foreign language, we used Finnish since it is not commonly used in the country of the study and belongs to a different language family (*i.e.*, Uralic languages) than the German.

Study Design

We employed a repeated-measures design with Position as the only independent variable resulting in three conditions: 1) the translation was displayed on top of (*overlay*), 2) on the right of (*right*) or 3) below the foreign word (*below*). The order of positions was counterbalanced to avoid sequence effects. As dependent variables, we measured the *comprehension* of learned words, time to learn words (task completion time – *TCT*), subjective task load assessed by the Raw TLX (*RTLX*) questionnaire [21], and usability using the System Usability Scale (*SUS*) questionnaire [2] after each condition. Finally, participants provided qualitative feedback for each condition.

Apparatus

To conduct the study, we implemented an application for the Microsoft HoloLens that enables detecting text in the foreign language in the environment and displaying translation for these words using three positions. The HoloLens app was developed with Unity 2018.3.11[4]. Since text recognition was not the focus of our study, and software development kits, such as Vuforia[5] did not support non-English characters, we used Vuforia's image recognition feature to detect foreign words. Using image recognition, we created mappings between pictures of all Finnish words selected for the study and their translations. The HoloLens application recognizes the predefined Finnish words as they are in the field of view of the AR glasses. When a word in the foreign language is recognized, the HoloLens's clicker can be used to present the translation of the word relative to the environment. The translation can be activated and switched off by pressing the button on the clicker. The translations are displayed using black sans-serif text as frequently used in studies on text readability in AR [17, 23]. Based on the condition, the app covers the original word with the translation (*overlay*) or places the translation below (*below*) or to the right (*right*) of the original word. To cover the Finnish word in the environment and increase the contrast for the overlay condition, we used a white background for the translations.

For the study, we selected 30 words in Finnish. To prevent their meaning from becoming accessible to participants

[2] https://ir.vuzix.com/press-releases/detail/1724/vuzix-blade-ar-smart-glasses-now-support-popular
[3] http://questvisual.com

[4] https://unity.com
[5] https://vuforia.com

Figure 2: A participant translating a unfamiliar Finnish word using the HoloLens app.

through their knowledge of other languages, we did not include words that resemble the words in German and English languages. The selected words were common words that a language student might learn at the beginning of their studies, or a tourist might learn before visiting a foreign country, such as a railroad or a restaurant.

We displayed Finnish words on a 46" TV to be able to switch between them during the study easily. To avoid any occlusion effect and to provide participants with the possibility to learn a single word at a time, we showed the Finnish words one-by-one on the TV. To show these words, we prepared PDF files displaying each word on a separate page with a white background in a randomized order.

Procedure

After introducing the purpose of the study, we asked participants to sign a consent form and answer questions about their demographics and familiarity with the technology. We then introduced the Microsoft HoloLens, gave short information about AR, helped participants to wear the device, and handed them the HoloLens's clicker. We invited participants to sit in front of the TV with a distance of $2m$. We then introduced the translation app, explained its usage and the translation POSITIONS. We informed the participants that they could enable and disable the translation for each foreign word by using the clicker. To become accustomed to the app, through a training session, participants looked for the translation of several Finnish words in all POSITIONS (see Figure 2). Afterward, we asked participants to imagine the following scenario: *"You are on vacation in Finland. During your travels, you encounter unfamiliar Finnish words, for example, on street signs or restaurant menus. To learn their translation on the go, you use a real-time translation app installed on your AR glasses."* We told them that during the study, they had to learn the translation for ten words in the foreign language per condition. We informed participants about the pending *comprehension* tests and that they could spend as much time as needed to see the translations. After learning the translation for a single word, participants could tell the experimenter to switch to the next word.

As the participants were familiar with the study procedure, they started with the first condition. After viewing the translation of the first ten words, participants took off the HoloLens and received a laptop to complete the *RTLX* and the *SUS* questionnaires. Afterward, we measured *comprehension* using a multiple-choice vocabulary test. As the vocabulary test, participants had to select the correct translation among four possible answers for each learned Finnish word. The provided four possible answers contained the correct translation, two words that were learned, and a word that was not learned during the last learning session. Then, participants gave qualitative feedback about the condition. The time to learn Finnish words was manually measured by the experimenter. Afterward, participants continued with the remaining conditions. In the end, participants were asked for their final feedback about the conditions, and the most and least preferred translation positions. The study took on average 60 minutes per participant.

Participants

We recruited 12 participants (5 females, 7 males) through our university's mailing list. Their average age was $M = 25.0$ ($SD = 3.91$) years, and most had a background in IT and were university students. While six participants had no experience with AR, three participants indicated previous experience with using a HoloLens. Four participants reported using AR applications for learning and entertainment purposes. We ensured that no participant could speak the foreign language used in the study. All participants came from a culture with a left-to-right reading direction and, German was their mother tongue. Participants received course credits for participating in the study.

4 RESULTS

During the study, 12 participants completed a learning session for each of the three POSITIONS. We performed a qualitative analysis of the collected objective and subjective data. We used one-way repeated-measures ANOVAs and paired-sample t-tests with Bonferroni correction for the parametric data. For the nonparametric data, we used a Friedman test following by Connover's post hoc tests with Bonferroni correction. Table 1 summarizes the descriptive measures.

Quantitative Measurements

There was a statistically significant effect of POSITION on *comprehension* ($\chi^2(2) = 6.077, p < .05, W = 0.705$) (see Figure 3 (left)). Pairwise comparison for POSITION revealed a significant difference between the *overlay* and the *below* ($p < .05$) condition. Figure 3 (right) shows the time took in seconds to learn the words. There was no statistically

	Overlay		Right		Below	
	M	**SD**	**M**	**SD**	**M**	**SD**
Comprehension	8.83	1.99	9.17	1.03	9.58	0.79
RTLX	53.08	21.00	38.00	17.91	39.17	19.42
SUS	61.46	20.60	76.88	15.60	74.17	17.20
TCT	229.25	81.06	241.42	101.32	244.67	78.90

Table 1: Descriptive results for all conditions.

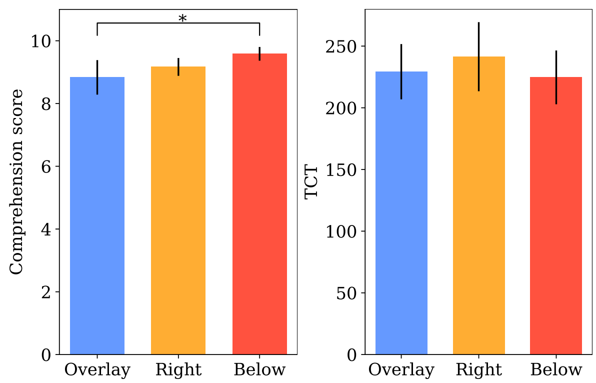

Figure 3: Average *comprehension* score (left) and *TCT* in seconds (right) for all conditions (error bars show standard error).

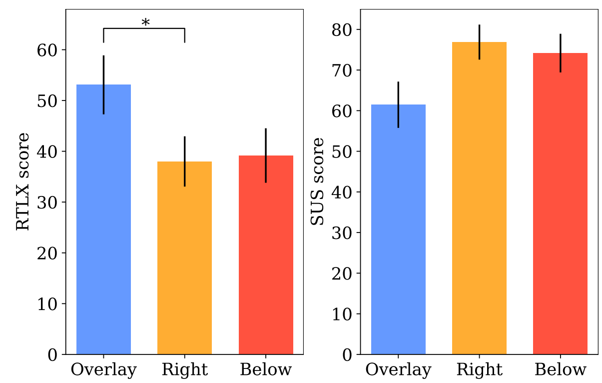

Figure 4: Average *RTLX* (left) and *SUS* (right) scores for all conditions (error bars show standard error).

significant effect of Position on *TCT* ($F_{2,22} = 0.278, p = .076, \eta^2 = 0.007$).

There was a statistically significant effect of Position on perceived task load ($F_{2,22} = 4.242, p < .05, \eta^2 = 0.278$) (see Figure 4). Post hoc test of the *RTLX* scores revealed significant difference between the *overlay* and the *right* ($p < .05$) conditions. Comparing the *SUS* scores (see Figure 4 (right)), we found a statistically significant effect of Position, $\chi^2(2) = 6.195$, $p < .05, W = 0.633$. However, post hoc test could not reveal statistically significant differences between the conditions (all $p > .066$).

Qualitative Feedback

After each session, participants provided feedback about their learning experience when using the HoloLens and the particular translation position. Furthermore, at the end of the study, each participant provided general feedback on the study. 7 (58.3%) participants preferred the *below* position the most: *"Using this position, both the word and the translation could be seen at the same time but were separate from each other. There was no need to look far away to see the translation"* (P9, P10). Furthermore, P5 commented that the *below* position was the easiest to see and learn the translation considering the flow of information. However, P3 and P8 considered seeing the foreign word and the translation under each other as *"the unfavorable arrangement of the words"*.

4 (33.3%) participants considered the *right* position as the most precise presentation of the translation of a word: *"One has the foreign word and the translation next to each other and can display or cancel the translation with a click. It feels like a vocabulary book, where one page is covered"* (P8). However, for P9, this position was the lowest preferred one.

P3 considered the *overlay* position as the best since it was possible to view the foreign word and the translation separately. However, 9 (75%) participants commented on the negative aspects of this position: *"It is not possible to see the direct connection between the foreign word and the translation of it"* (P5). *"It disturbed me that I had to switch off the translation to see the word in the foreign language"* (P1, P6, P10).

5 DISCUSSION AND LIMITATIONS

The study showed that displaying translation in the *overlay* position resulted in the lowest comprehension and the highest task load. Consequently, most of the participants did not prefer this position due to not being able to see the foreign word and the translation simultaneously. Compared to the other positions, in the *overlay* position, participants had to interact with the AR glasses each time to view either the foreign word or the translation for it. This negatively affected both the perceived task load and usability. The SUS scores for the *right* and *below* POSITIONS were higher than for the *overlay* POSITION. However, post hoc tests did not show a statistically significant difference between the POSITIONS.

The qualitative feedback revealed that most participants considered the *below* POSITION as the best one to show the translation for a foreign word. The objective data supported this finding, as with the *below* POSITION, participants' comprehension scores were the highest.

Considering the task completion time, we did not find statistically significant differences between the conditions. Thus, participants spend a similar amount of time for all POSITIONS during vocabulary learning tasks. We can conclude that the time spent on learning was not the reason that comprehension scores of the conditions significantly differed.

For this study, to compare the effect of positioning augmented information on learning experience, we used a simple vocabulary learning task. The task required paying attention to both real-world and virtual information. In general, quantitative and qualitative results revealed that the possibility to view both the real-world and related virtual information simultaneously outperforms the condition where the virtual information is occluding the real-world one. Considering other tasks that do not require viewing both the real-world and the virtual information, the effects of POSITIONS might be different. However, future research is needed to confirm these effects.

For the study, considering previous work and the reading direction in the country of the study, we selected *overlay*, *below* and *right* POSITIONS. However, the result for participants from cultures with a right-to-left reading direction might be different. Future work is needed to investigate the relation between reading direction and the preferred position. Furthermore, we used a vocabulary learning task since it has been frequently used in previous work. However, other common use cases for a real-time translation application might be the translation of short phrases or sentences. Learning short phrases or sentences might increase the cognitive load and cause longer learning time, fatigue effect, and, consequently, a longer study duration. Future work is also needed to investigate these effects.

As for the study we investigated only the position of the translation on AR glasses, we displayed the foreign words on the screen of a TV with the same size and background color. This allowed us to easily change the foreign words during the study and randomize them. However, in a real-world scenario, foreign words in the environment, such as street signs or restaurant menus, would appear in various sizes and backgrounds. Future research is needed to investigate the effect of various properties of text presentation on the vocabulary learning experience.

6 CONCLUSION

We investigated the effect of three positions (*overlay*, *below*, *right*) to show translations on a vocabulary learning task using AR glasses. We studied how the position affects vocabulary comprehension, learning duration, perceived task load, and usability. We supported our investigation with quantitative data and qualitative feedback. The results showed that simultaneously presenting a word in a foreign language and the translation for this word outperformed overlaying the translation on the original word. The latter position decreased the comprehension score and increased the perceived task load.

In our study, we presented translations in a textual form. However, previous work showed that users can learn more if the translation is presented both with a text and a picture compared to only using text [12, 43]. For a real-world use case displaying translations in multiple modalities using AR might require more space around the foreign word in the environment. We suggest that future work investigates the effect of the position of multimodal information as a translation on vocabulary learning. Moreover, we compared only three POSITIONS that are registered in 3D space. However, there are commercial real-time translation applications for AR glasses that present translations attached to the glasses as a head-up display (*e.g.*, Zoi Meet application for Vuzix Blade smart glasses). Future research is needed to compare the effect of further presentation concepts for augmented content using AR glasses on learning experiences.

REFERENCES

[1] Ronald T Azuma. 1997. A survey of augmented reality. *Presence: Teleoperators & Virtual Environments* 6, 4 (1997), 355–385.

[2] Aaron Bangor, Philip T Kortum, and James T Miller. 2008. An empirical evaluation of the system usability scale. *Intl. Journal of Human–Computer Interaction* 24, 6 (2008), 574–594.

[3] Mark Billinghurst and Andreas Duenser. 2012. Augmented reality in the classroom. *Computer* 45, 7 (2012), 56–63.

[4] Mark Billinghurst, Hirokazu Kato, and Ivan Poupyrev. 2001. The magicbook-moving seamlessly between reality and virtuality. *IEEE Computer Graphics and applications* 21, 3 (2001), 6–8.

[5] Matt Bower, Cathie Howe, Nerida McCredie, Austin Robinson, and David Grover. 2014. Augmented Reality in education–cases, places and potentials. *Educational Media International* 51, 1 (2014), 1–15.

[6] David E Breen, Eric Rose, and Ross T Whitaker. 1995. Interactive occlusion and collision of real and virtual objects in augmented reality. *Munich, Germany, European Computer Industry Research Center* (1995).

[7] Julie Carmigniani and Borko Furht. 2011. Augmented reality: an overview. In *Handbook of augmented reality*. Springer, 3–46.

[8] Nee Nee Chan and Chin-Siang Ang. 2017. Investigating the use of mobile applications in everyday language learning. *IJMLO* 11 (2017), 378–394.

[9] Shih-Yeh Chen, Chao-Yueh Hung, Yao-Chung Chang, Yu-Shan Lin, and Ying-Hsun Lai. 2018. A study on integrating augmented reality technology and game-based learning model to improve motivation and effectiveness of learning English vocabulary. In *2018 1st International Cognitive Cities Conference (IC3)*. IEEE, 24–27.

[10] Yu-Chien Chen. 2006. A study of comparing the use of augmented reality and physical models in chemistry education. In *Proceedings of the 2006 ACM international conference on Virtual reality continuum and its applications.* 369–372.

[11] Soon Hau Chua, Simon T. Perrault, Denys J. C. Matthies, and Shengdong Zhao. 2016. Positioning Glass: Investigating Display Positions of Monocular Optical See-Through Head-Mounted Display. In *Proceedings of the Fourth International Symposium on Chinese CHI (ChineseCHI2016).* Association for Computing Machinery, New York, NY, USA, Article Article 1, 6 pages. https://doi.org/10.1145/2948708.2948713

[12] Dorothy M Chun and Jan L Plass. 1996. Effects of multimedia annotations on vocabulary acquisition. *The modern language journal* 80, 2 (1996), 183–198.

[13] Jun Du, Qiang Huo, Lei Sun, and Jian Sun. 2011. Snap and translate using windows phone. In *2011 International Conference on Document Analysis and Recognition.* IEEE, 809–813.

[14] Victor Fragoso, Steffen Gauglitz, Shane Zamora, Jim Kleban, and Matthew Turk. 2011. TranslatAR: A mobile augmented reality translator. In *2011 IEEE Workshop on Applications of Computer Vision (WACV).* IEEE, 497–502.

[15] Yuichiro Fujimoto, Goshiro Yamamoto, Hirokazu Kato, and Jun Miyazaki. 2012. Relation between location of information displayed by augmented reality and user's memorization. In *Proceedings of the 3rd Augmented Human International Conference.* 1–8.

[16] Yuichiro Fujimoto, Goshiro Yamamoto, Takafumi Taketomi, Jun Miyazaki, and Hirokazu Kato. 2012. Relationship between features of augmented reality and user memorization. In *2012 IEEE International Symposium on Mixed and Augmented Reality (ISMAR).* IEEE, 279–280.

[17] Joseph L Gabbard, J Edward Swan, and Deborah Hix. 2006. The effects of text drawing styles, background textures, and natural lighting on text legibility in outdoor augmented reality. *Presence: Teleoperators & Virtual Environments* 15, 1 (2006), 16–32.

[18] Juan Garzón and Juan Acevedo. 2019. A Meta-analysis of the impact of Augmented Reality on students' learning effectiveness. *Educational Research Review* (2019).

[19] Robert Godwin-Jones. 2016. Augmented reality and language learning: From annotated vocabulary to place-based mobile games. *Language Learning & Technology* 20, 3 (2016), 9–19.

[20] Yongqi Gu and Robert Keith Johnson. 1996. Vocabulary learning strategies and language learning outcomes. *Language learning* 46, 4 (1996), 643–679.

[21] Sandra G Hart. 2006. NASA-task load index (NASA-TLX); 20 years later. In *Proceedings of the human factors and ergonomics society annual meeting*, Vol. 50. Sage Publications Sage CA: Los Angeles, CA, 904–908.

[22] Adam Ibrahim, Brandon Huynh, Jonathan Downey, Tobias Höllerer, Dorothy Chun, and John O'donovan. 2018. Arbis pictus: A study of vocabulary learning with augmented reality. *IEEE transactions on visualization and computer graphics* 24, 11 (2018), 2867–2874.

[23] Jacek Jankowski, Krystian Samp, Izabela Irzynska, Marek Jozwowicz, and Stefan Decker. 2010. Integrating text with video and 3d graphics: The effects of text drawing styles on text readability. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems.* 1321–1330.

[24] Hannes Kaufmann and Dieter Schmalstieg. 2002. Mathematics and geometry education with collaborative augmented reality. In *ACM SIGGRAPH 2002 conference abstracts and applications.* 37–41.

[25] Kangdon Lee. 2012. Augmented reality in education and training. *TechTrends* 56, 2 (2012), 13–21.

[26] Lap-Kei Lee, Chi-Pang Chau, Ki-Nam Tsoi, Nancy Ling Yang, and Nga-In Wu. 2019. A Mobile Game for Learning English Vocabulary with Augmented Reality Block Builder. In *International Conference on Technology in Education.* Springer, 116–128.

[27] Ann M McNamara. 2011. Enhancing art history education through mobile augmented reality. In *Proceedings of the 10th International Conference on Virtual Reality Continuum and Its Applications in Industry.* 507–512.

[28] Prema Meda, M Kumar, and Ramu Parupalli. 2014. Mobile augmented reality application for telugu language learning. In *2014 IEEE International Conference on MOOC, Innovation and Technology in Education (MITE).* IEEE, 183–186.

[29] Jason Orlosky, Kiyoshi Kiyokawa, and Haruo Takemura. 2014. Managing Mobile Text in Head Mounted Displays: Studies on Visual Preference and Text Placement. *SIGMOBILE Mob. Comput. Commun. Rev.* 18, 2 (June 2014), 20–31. https://doi.org/10.1145/2636242.2636246

[30] Iulian Radu. 2014. Augmented reality in education: a meta-review and cross-media analysis. *Personal and Ubiquitous Computing* 18, 6 (2014), 1533–1543.

[31] Rufat Rzayev, Paweł W. Woundefinedniak, Tilman Dingler, and Niels Henze. 2018. Reading on Smart Glasses: The Effect of Text Position, Presentation Type and Walking. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems (CHI '18).* Association for Computing Machinery, New York, NY, USA, Article Paper 45, 9 pages. https://doi.org/10.1145/3173574.3173619

[32] Hannu Salmi, Helena Thuneberg, and Mari-Pauliina Vainikainen. 2017. Making the invisible observable by Augmented Reality in informal science education context. *International Journal of Science Education, Part B* 7, 3 (2017), 253–268.

[33] Marc Ericson C Santos, Takafumi Taketomi, Goshiro Yamamoto, Ma Mercedes T Rodrigo, Christian Sandor, Kato Hirokazu, et al. 2014. Evaluating augmented reality for situated vocabulary learning. In *22nd International Conference on Computers in Education (ICCE 2014).* Asia-Pacific Society for Computers in Education, 701–710.

[34] Marc Ericson C Santos, Takafumi Taketomi, Goshiro Yamamoto, Ma Mercedes T Rodrigo, Christian Sandor, Hirokazu Kato, et al. 2016. Augmented reality as multimedia: the case for situated vocabulary learning. *Research and Practice in Technology Enhanced Learning* 11, 1 (2016), 4.

[35] Olga Scrivner, Julie Madewell, Cameron Buckley, and Nitocris Perez. 2016. Augmented reality digital technologies (ARDT) for foreign language teaching and learning. In *2016 future technologies conference (FTC).* IEEE, 395–398.

[36] Aulia Akhrian Syahidi, Herman Tolle, Ahmad Afif Supianto, and Kohei Arai. 2018. BandoAR: Real-Time Text Based Detection System Using Augmented Reality for Media Translator Banjar Language to Indonesian with Smartphone. In *2018 IEEE 5th International Conference on Engineering Technologies and Applied Sciences (ICETAS).* IEEE, 1–6.

[37] Kohei Tanaka, Yasue Kishino, Masakazu Miyamae, Tsutomu Terada, and Shojiro Nishio. 2008. An information layout method for an optical see-through head mounted display focusing on the viewability. In *2008 7th IEEE/ACM International Symposium on Mixed and Augmented Reality.* IEEE, 139–142.

[38] Wernhuar Tarng and Kuo-Liang Ou. 2012. A study of campus butterfly ecology learning system based on augmented reality and mobile learning. In *2012 IEEE Seventh International Conference on Wireless, Mobile and Ubiquitous Technology in Education.* IEEE, 62–66.

[39] Takumi Toyama, Daniel Sonntag, Andreas Dengel, Takahiro Matsuda, Masakazu Iwamura, and Koichi Kise. 2014. A Mixed Reality Head-Mounted Text Translation System Using Eye Gaze Input. In *Proceedings of the 19th International Conference on Intelligent User Interfaces (IUI '14).* Association for Computing Machinery, New York, NY, USA, 329–334. https://doi.org/10.1145/2557500.2557528

[40] Takumi Toyama, Wakana Suzuki, Andreas Dengel, and Koichi Kise. 2013. User attention oriented augmented reality on documents with

document dependent dynamic overlay. In *2013 IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*. IEEE, 299–300.

[41] Minjuan Wang, Vic Callaghan, Jodi Bernhardt, Kevin White, and Anasol Peña-Rios. 2018. Augmented reality in education and training: pedagogical approaches and illustrative case studies. *Journal of ambient intelligence and humanized computing* 9, 5 (2018), 1391–1402.

[42] Christopher Wasko. 2013. What teachers need to know about augmented reality enhanced learning environments. *TechTrends* 57, 4 (2013), 17–21.

[43] Makoto Yoshii. 2006. L1 and L2 glosses: Their effects on incidental vocabulary learning. *Language Learning & Technology* 10, 3 (2006), 85–101.