Social Semiotic Analysis of the Design Space of Augmented Reality

Brad Cameron[*]
Discipline of Linguistics
University of Adelaide

Christian Sandor[#]
Magic Vision Lab
University of South Australia

Peter Mickan[^]
Discipline of Linguistics
University of Adelaide

ABSTRACT

Text is one of the multimodal components used in many augmented reality (AR) applications. Fulfilling the design goals of an application depends, to a large extent, on the successful design of these components. This paper examines the role of text in AR user interfaces from a social semiotic and linguistic perspective. Since analysis of the text in AR applications has mostly focused on aspects such as legibility, placement, occlusion-related issues and exploration of methods to overcome other visual challenges, we suggest that such a social semiotic and linguistic analysis of the textual content can provide insights into user interface and application design for AR researchers and developers. Our analysis suggests that the role of text in AR can differ, and depends, at least in part, on the type of text-image interaction of which it is a part.

KEYWORDS: Augmented Reality, Social Semiotics, Multimodality, User Interface Design

INDEX TERMS: H.5.1 [Information Interfaces and Presentation]: Multimedia Information Systems—Artificial, augmented and virtual realities; H.5.2 [User Interfaces]: User-centered design; J.5 [Arts and Humanities]: Linguistics

1 INTRODUCTION

The relatively recent uptake of smartphones by the public in many countries, and the acceleration of processing power and increase in screen and video resolution makes the study of emerging applications for these devices in the context of human social interaction and multimodal semiotic resources essential. Social Semiotics is a theoretical framework for analysis of language, culture and communication and is informed by the work of Halliday [1] and more recently many others including Kress [2]. Augmented Reality (AR) research typically focuses on science and technology, although some, such as Pedersen [3] have suggested important functions for the humanities in this field.

This paper approaches the design space of AR from a social semiotic and linguistic perspective in order to analyse the role of text in terms of the text-image relationship in four AR research prototypes.

Our findings suggest that the role of text in AR prototypes varies depending on its relationship with the image and the purpose for which it is used. For example, the role of labeling an object with text is contributing corresponding content, but is not re-expressing the image in a new way or adding additional data.

We suggest that by examining the text-image relationship it is possible to understand and gain insight into the role and purpose of these modes as part of a *modal ensemble*. Importantly, this analysis allows us to see which roles the text-image

[*] e-mail: bradley.d.cameron@student.adelaide.edu.au
[#] e-mail: chris.sandor@gmail.com
[^] e-mail: peter.mickan@adelaide.edu.au

relationship is not fulfilling. This is significant because the design of AR applications depends in part on which modes are chosen and for what purpose.

2 BACKGROUND

Multimodality in computer science refers to the modes of input and output of data between a human and a computer user interface. Multimodal user interfaces are a research area in human-computer interaction (HCI) and the goals of designing multimodal rather than unimodal systems typically include ease of use, transparency, flexibility and efficiency [4]. Multimodality in social semiotics and linguistics is, in a sense, more fundamental, and describes the modes of interaction of humans with their surroundings. One example is a highway sign (Figure 1). Kress suggests that the modes of *writing* and *image* and *colour* in such signs lend themselves to doing different kinds of semiotic work, with each having distinct potentials for meaning [2].
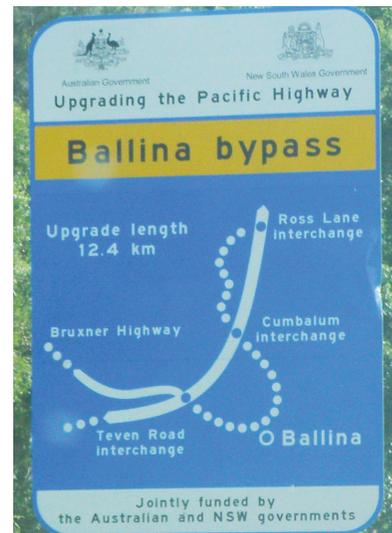

Figure 1: Australian highway sign.

We will apply the analysis of the semiotic role of modes from this perspective to AR research. Focusing on the relationship of the text to the image and the user, we can use the image-text relations framework outlined by van Leeuwen [5] and Unsworth and Chan, [6]. This framework suggests that there are two basic types of image-text interaction: *elaboration* and *extension*. In the first, one mode elaborates on the other by either adding corresponding content across modes so there is some redundancy (*equivalence*), or by a re-expression or reformulation of meaning in an alternate mode (*exposition*).

The second basic type of image-text relationship, *extension*, involves either a text or an image adding new meaning to the other (*augmentation*), or a juxtaposition of image and text to 'jointly construct activity sequences' [6], (*distribution*).

To examine this theory, we look at four AR examples: Schall et al [7]: 3d visualisation of underground infrastructure, Henderson and Feiner [8]: benefits of AR for task localization in

maintenance, Bell et al [9] : an annotated situation-awareness aid, and Bichlmeier et al [10]: evaluation of a virtual mirror as a navigational aid for AR driven minimally invasive procedures.

3 ANALYSIS

The text in the four prototypes was examined, and placed into one of four columns in Figure 2. Black text denotes that which can be observed while red text indicates types of image-text relationship that were not observed, but which could be applied.

	Elaboration		Extension	
	Equivalence	Exposition	Augmentation	Distribution
Schall et al	GIS data (labels, name of objects)	Step-by-step instructions (eg. If maintenance is required, accompanied by a visual AR component).	GIS data (additional information about the properties or status of the objects)	Introduce unrelated text to the image, or unrelated image to the text.
Henderson and Feiner	floating labels (name of parts)	Step-by-step task-related instructions	Extra data about parts (eg temperature, electrical charge, physical composition)	
Bell et al	Labels (name of objects)	Navigational instructions (eg. directions accompanied by a visual AR component).	Extra data (eg. The CPU and RAM of the computer).	
Bichlmeier et al		Textual task related instructions to reduce time needed to get used to the task.		

Figure 2. Text-image interaction in AR prototypes.

Schall et al produced an example of a maintenance prototype allowing a user to see a 3d visualization of underground water infrastructure in order to assist in on-site applications for geohydrological and environmental problems of urban waters [7]. The observable text in this prototype is fulfilling potentially essential maintenance-related roles by providing labels (*equivalence*) and additional information (*augmentation*) about objects. We did not observe any *expositional* text such step-by-step task-related instructions, which could be used in a maintenance application of this system. The database of GIS (Geographic Information System) data, from which the data for the system originates, could also provide an opportunity for gaming or new types of user experience. For example, the introduction of new and unrelated text to the image, or new and unrelated image to the text (*distribution*).

Henderson and Feiner, adapted an IETM (Interactive Electronic Technical Manual) into a working AR prototype for military mechanics maintenance in order to examine and document how AR can assist in reducing time and effort in navigating larger sequences of maintenance tasks in such complex systems [8]. The prototype included essential text outlining step-by-step instructions on a required task, in addition to labels indicating the names of objects. Observations showed that there was no text in the role of *augmentation*. One example of potentially useful text of this type for a maintenance application could be the display of temperature of a mechanical component, or its electrical charge.

Bell et al produced a situational awareness aid allowing users to see a mini map of the space they are located in complete with labels of objects and options to view extra data about the properties of the objects [9]. This concept could be extended to provide textual and visual instructions (*exposition*) to enable a navigational or task related function.

Bichlmeier et al, is an evaluation of a virtual mirror as a navigational aid in improving the precision of an interactive tool in a medical AR scenario [10]. There was no text observable in this prototype. They suggested that subjects needed time to get used to the system and the task, and therefore it is possible that re-expressing the meaning in text (*exposition*) may have reduced this time.

4 CONCLUSIONS

The results of the research indicate that the role of text varies in AR research and that there are different types of text-image relationships. Modal ensembles are based on *designs*, that is, on selections and arrangements of resources for making a specific message about a particular issue for a particular audience [2]. The design of an application depends on the choice of semiotic resources, and which meanings to assign to particular modes.

By analyzing text in a text-image relationship from a social semiotic and linguistic perspective, we have been able to show that such analysis can provide insights into the role of text in AR, and therefore increase awareness of the meaning making process in the design of applications. More investigation is needed.

REFERENCES

[1] M.A.K Halliday. (1978), Language as Social Semiotic: The Social Interpretation of Language and Meaning. Baltimore, MD: University Park Press.
[2] Kress, G (2010). Multimodality: A social semiotic approach to contemporary communication. Routledge, London.
[3] Pedersen, I (2009). Radiating Centers: Augmented Reality and Human-Centric Design, IEEE International Syposium on Mixed and Augmented Reality (ISMAR 2009).
[4] Oviatt, S (1997), Multimodal Interactive Maps: Designing for Human performance, Human-Computer Interaction, 1997, Volume 12, pp 93-129, USA.
[5] van Leeuwen, T (2005), Introducing Social Semiotics, Routledge, London.
[6] Unsworth, Len and Chan, Eveline. Bridging Multimodal Literacies and National Assessment Programs in Literacy [online]. Australian Journal of Language and Literacy, The, Vol. 32, No. 3, Oct 2009: 245-257.
[7] Schall, G, Schmalstieg, D & Junghanns, S (2010), VIDENTE - 3D Visualisation of Underground Infrastructure using Handheld Augmented Reality, CRC Press/ Taylor and Francis# Publisher: CRC; 1 edition (February 15, 2010).
[8] Henderson, S & Feiner, S (2009), Evaluating the Benefits of Augmented Reality for Task Localization in Maintenance of an Armored Personnel Carrier Turret, IEEE International Syposium on Mixed and Augmented Reality (ISMAR 2009).
[9] Bell, B, Feiner, S & Hollerer (2002), An Annotated Situation-Awareness Aid for Augmented Reality, UIST 2002, Paris, France.
[10] Bichlmeier, C (2010), Evaluation of the Virtual Mirror as a Navigational Aid for Augmented Reality Driven Minimally Invasive Procedures, IEEE International Syposium on Mixed and Augmented Reality (ISMAR 2010).