

# Two Multimodal Interfaces to Military Simulations

Kenneth Wauchope

Navy Center for Applied Research in Artificial Intelligence  
Naval Research Laboratory, Code 5512  
4555 Overlook Avenue, S.W.  
Washington, DC 20375-5337, USA  
wauchope@aic.nrl.navy.mil

## 1 Introduction

Eucalyptus and InterLACE are two projects integrating spoken natural language interaction with the already-developed graphical user interfaces of currently existing military simulation systems. We take this approach (rather than developing new, fully integrated systems from scratch) to facilitate the development of Navy demonstrations of NL technology applied to their own tools, as well as to explore the relative strengths and weaknesses of graphical vs. NL interfaces and how the two can be used in combination to yield more powerful interaction techniques. The natural language processing capability of the two systems is described in an accompanying Proceedings technical note (Wauchope et al. 1997); here we go into a bit more detail on the issues of multimodal integration with graphical interfaces.

## 2 Eucalyptus

Our first venture into multimodal interface development was Eucalyptus, a spoken natural language interface integrated with the graphical user interface of the KOALAS AEW Test Planning Tool, a simulation-based mockup of a hypothetical Naval air combat command and control system. Our objective was to have the functionality of the NL interface duplicate and parallel that of the GUI, allowing comparison of language-based and graphical approaches to predication and reference. We also sought for the two modalities to interact in a number of ways, such as deictic reference (combined pointing and speaking) and NL interaction with graphical dialog windows.

### 2.1 Guidelines

We adhered to the following guidelines and approaches in developing Eucalyptus.

- Spoken inputs are normative English utterances only, not just arbitrary individual words or

phrases. Such utterances can include certain elliptical or fragmentary forms, but only those that people might normally use with each other under similar circumstances. As a result Eucalyptus is a spoken NL system, not just a voice command system. This is not to say that voice commands are undesirable or that people prefer to use conversational language when talking to a machine, but just represents our particular research interests.

- We address NL input only, not output. The NAUTILUS NLP processor does try to generate brief and helpful English sentence fragments in response to user queries, but this is deliberately underdeveloped since we also have ongoing work in adapting a full-fledged NL generation system for full-sentence query responses (integrated into our second demo system InterLACE, to be discussed later).
- We are careful not to provide the NL interface with any new domain functionality not already available in some way from the graphical interface. This ensures that any comparisons between the two will be based strictly on syntax and convenience of use.
- We disallow “metalinguistic” NL references to the graphical interface itself (e.g. *Open the advice explanation window, Push the 'Move Fighter' button*) in favor of direct references to the task domain (*Show the current advice, Move a fighter*). This avoids the user treating the graphical interface as primary and speech as a subservient adjunct, a kind of “vocal mouse” as it were. Again this is not to deny that speech control of a GUI might be useful for handicapped individuals or in extremely hands-busy situations, but just encourages the user’s perception of the two interfaces as parallel and equal partners.

# Report Documentation Page

Form Approved  
OMB No. 0704-0188

Public reporting burden for the collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden, to Washington Headquarters Services, Directorate for Information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington VA 22202-4302. Respondents should be aware that notwithstanding any other provision of law, no person shall be subject to a penalty for failing to comply with a collection of information if it does not display a currently valid OMB control number.

1. REPORT DATE <b>1997</b>		2. REPORT TYPE		3. DATES COVERED <b>00-00-1997 to 00-00-1997</b>	
4. TITLE AND SUBTITLE <b>Two Multimodal Interfaces to Military Simulations</b>				5a. CONTRACT NUMBER	
				5b. GRANT NUMBER	
				5c. PROGRAM ELEMENT NUMBER	
6. AUTHOR(S)				5d. PROJECT NUMBER	
				5e. TASK NUMBER	
				5f. WORK UNIT NUMBER	
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) <b>Naval Research Laboratory, Navy Ctr for Applied Research in Artificial Intelligence, 4555 Overlook Avenue SW, Washington, DC, 20375-5337</b>				8. PERFORMING ORGANIZATION REPORT NUMBER	
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)				10. SPONSOR/MONITOR'S ACRONYM(S)	
				11. SPONSOR/MONITOR'S REPORT NUMBER(S)	
12. DISTRIBUTION/AVAILABILITY STATEMENT <b>Approved for public release; distribution unlimited</b>					
13. SUPPLEMENTARY NOTES					
14. ABSTRACT					
15. SUBJECT TERMS					
16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES <b>2</b>	19a. NAME OF RESPONSIBLE PERSON
a. REPORT <b>unclassified</b>	b. ABSTRACT <b>unclassified</b>	c. THIS PAGE <b>unclassified</b>			

## 2.2 Multimodal Integration

To issue an order to a simulated fighter aircraft using the KOALAS graphical interface, the user first clicks a button labelled with the order type (e.g. Move, Recall) which pops up a dialog window requesting that the user select the desired objects by mousing them on the radar screen, echoing the objects' names in text fields as they are selected, and finally requesting an accept/quit decision. The Move Fighter dialog box, for example, prompts first for the fighter and then for the new patrol station (optional parameters like speed are not prompted for but can be adjusted at any time).

This rather specialized dialog box behavior is distinctly language-like in two ways: the resulting graphical syntax (Move, fighter, station) closely resembles NL imperative syntax (*Move this fighter to this station*), and the prompting sequence resembles a natural language dialog. As a result we elected to use KOALAS dialog boxes both for accept/quit confirmation of completely specified NL commands (*Move fighter 1 to station 4*), as well as to prompt for remaining unspecified arguments (*Move fighter 1*). In either case, speech can also be used to dismiss the dialog (*Okay, Cancel that*) as well as to provide remaining arguments to the dialog (*one of the fighters holding station 4*).

The second form of multimodal integration in Eucalyptus is through deictic reference: phrases like *this fighter* accompanied by a mouse click. Since mouse clicks in the KOALAS graphical interface do not create a "current selected object" but rather report the closest object of the type being prompted for by the current dialog box prompt, this behavior was easy to modify for deictic reference: the mouse click identifies a set of possible objects and the accompanying NL context (*this fighter, Move this here*) provides the semantic presuppositions to disambiguate the denoted object.

Finally, we allow purely graphical interactions to have elliptical or anaphoric NL followups. For example after completing a fighter move using the GUI, the NL inputs *Same with fighter 2* or *Move fighter 3 there also* will be properly understood.

## 3 InterLACE

InterLACE is a multimodal interface to the Air Force's LACE land/air combat simulation system, containing an extensive real-world cartographic database of central Germany. Unlike KOALAS, LACE came to us with no graphical interface, although all the "hooks" were in place to implement a map display with animation of simulated objects.

Since with LACE there was no graphical command interface to mirror in natural language, we opted instead to focus on database query (also included in Eucalyptus) and the issuing of verbal on-road route instructions to a simulated tank unit. Deictic reference was implemented similarly to Eucalyptus, the difference being that we chose to implement the conventional notion of "current selected object": objects clicked by the mouse are highlighted and remain so until deselected. The user can also select a set of closely adjacent objects with a double-click and then, as in Eucalyptus, use verbal context (*this town, What's the population here?*) to resolve the reference.

### 3.1 Database Query

Because of the large size of the database (over twelve thousand objects) and the need to reduce the search space during information retrieval, we elected to constrain the domain of NL discourse to only those objects currently visible in the map display, representing about one-eighth of the entire database at any one time. Hence *Does this road cross any small towns?* only reports back towns that are currently visible in the display, based on the assumption that this is the user's current focus of attention (this constraint applies only to quantified NPs and not proper names, i.e. *Does this road cross Berlin?* does not require that Berlin be visible).

### 3.2 Tank Commands

LACE provides for the instantiation of simulated ground vehicles (such as tank units and SAM missile carriers) and includes a number of routines for on-road navigation, including a route planner for computing paths between towns. Verbal instructions to the tank giving just a destination (*Go to the nearest small town*) invoke the route planner, but since the planner does not accept specification of means for getting to the destination, commands like *Head north on road E2 for 5 km* instead invoke an InterLACE function called the "stretch selector" that tries to find a single stretch of road (a sequence of road segments between two adjacent linkage points) that reconciles every component of the means. The tank understands both cardinal (*northeast*) and relative directions (*left*), and in ambiguous situations chooses the stretch that bears the most closely in the specified direction.