Listen-Communicate-Show (LCS): Spoken Language Command of Agent-based Remote Information Access

Jody J. Daniels and Benjamin Bell
Lockheed Martin Advanced Technology Laboratories
1 Federal Street, A&E 3W
Camden, NJ 08102

{jdaniels, bbell@atl.lmco.com}

ABSTRACT

Listen-Communicate-Show (LCS) is a new paradigm for human interaction with data sources. We integrate a spoken language understanding system with intelligent mobile agents that mediate between users and information sources. We have built and will demonstrate an application of this approach called LCS-Marine. Using LCS-Marine, tactical personnel can converse with their logistics system to place a supply or information request. The request is passed to a mobile, intelligent agent for execution at the appropriate database. Requestors can also instruct the system to notify them when the status of a request changes or when a request is complete. We have demonstrated this capability in several field exercises with the Marines and are currently developing applications of this technology in new domains.

Keywords

Spoken language understanding, agents, dialogue management.

1. INTRODUCTION

An LCS system listens for information requests, communicates both with the user and networked information resources, and shows a tailored visualization to the individual user. The LCS-Marine system employs a spoken language understanding system (SLS) for assisting the user in placing a request and mobile, intelligent agents for information access to implement the LCS paradigm. The SLS converses with the user to generate a request or to check status, amend, or cancel an existing request. Once sufficient information is obtained from the user, the SLS launches an agent to accomplish the requested task. The agent accesses the appropriate databases via whatever network services are available (including existing tactical communications networks). Once the agent's tasks are complete, it returns to the SLS, which generates an appropriate response to the user. The response may be visual, verbal, or a combination, depending on the available devices.

2. SYSTEM OVERVIEW

The LCS-Marine system consists of four major components: an SLS, a collection of agents for information access, real-world operational databases, and communications networks to connect the user to the SLS and the agents to the databases. The underlying architecture for the system is the MIT Galaxy II conversational architecture [3]. It is a distributed, component-based middleware product designed to be "plug and play". Specialized servers handle specific tasks, such as translating audio data to text. All Galaxy II-compliant servers communicate with each other through a central server known as the Hub. The Hub manages flow control, handles traffic among distributed servers, and provides state maintenance.

In the SLS, speech is sent from the Audio I/O server to the Recognizer. The top *n* recognitions are then parsed, prior context added, and processed using the Natural Language (NL) servers (Frame Construction and Context Tracking) to verify the new input's validity and context. The Turn Manager (TM) determines how to proceed with the conversations and generates a response. NL (Language Generation) converts it to text and the Synthesis server generates the verbal response. The audio server then speaks the waveform file to the user. We customize the various servers to work with domain specific issues and application-specific information and training. Figure 1 shows our LCS architecture.
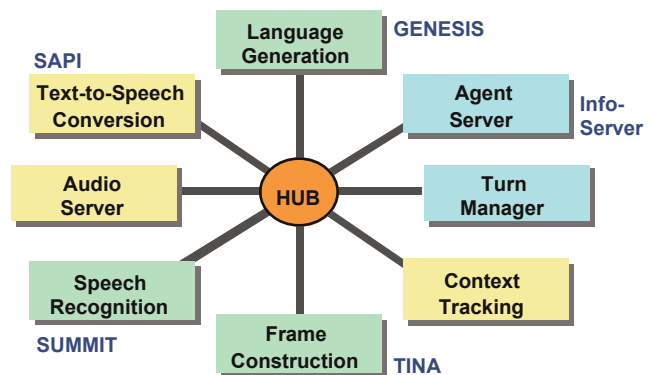

Figure 1. The LCS-Marine architecture.

We have integrated an additional server into the architecture to support information access—an Agent server. The Agent server manages a collection of agents that can be tasked to accomplish a variety of missions, including migration to

distant machines with possibly different operating systems to gather information or to monitor and report events [2].

Typically, the Agent server receives its tasking from the TM and supplies the TM with information from the data source(s). For persistent tasks, the Agent server becomes the initiator of a dialogue to inform the user of specific events by passing agent reports to the TM. When a visual display is present, the Agent server will dispatch an agent to pass the updated information to the display machine.

For the LCS-Marine application our agents had to interact with a logistics database that could be between one to one hundred miles away. We later describe how our agents were able to reach this live database over the tactical communication links available.

Users interact with the LCS-Marine system using the voice capture device appropriate to their organization (telephone, cell phone, tactical radios, computer headsets, etc.).

3. MARINE COMBAT SERVICE SUPPORT PROBLEM

Marines work in a dynamic, fluid environment where requirements and priorities are constantly subject to change. Under current operations, it might take up to 72 hours before a Marine in a Combat Service Support Operations Center (CSSOC) can confirm with a requesting unit that their order is in the logistics system. This is due to a lack of resources available to the tactical units as well as a difficulty in turning logistics data into information to enable timely analysis and decision making. For Marines conducting tactical operations, these restrictions and limited visibility into the supply chain hamper logistics planning, decision, execution, and assessment. Figure 2 shows the various echelons involved in tactical Marine logistics operations. It is noteworthy that tactical units have no organic means of accessing the logistical databases other than via radio contact with personnel at the CSSOC.

The focus of the LCS-Marine project is to provide Marines in the field with this missing visibility into the supply chain. By using standard radio protocols and a common form, Marines can now converse with a system that understands their task and end goal and can assist them in getting both the information and supplies they need. Figure 3 shows a sample of the Rapid Request form, used when placing an order.

Supporting the LCS-Marine domain required understanding and using proper radio protocols to communicate. It required the system to understand call signs, military times, grid coordinates, and special ordinance nomenclature. Additionally, to fully support the dynamic environment, LCS-Marine needed the ability to understand and translate usages of the military phonetic alphabet. This alphabet is used to spell difficult or unusual words. For example, to give the point of contact for the request as Sergeant Frew, the user could say: " P O C is Sergeant I spell Foxtrot Romeo Echo Whiskey over." LCS-Marine would convert the phonetic words to the proper letter combination. This way the vocabulary is potentially much larger than that used for system training.

Supporting the dynamic aspects of the Marine environment, the system is speaker independent. This is critical in applications where the user may change and there is no additional time for training the system for a new operator.

The recognizer is trained on the domain vocabulary, but not on individual operator voices. The system also fully supports natural, conversational dialogue, i.e., the recognizer expects utterances at a normal rate of speech and the speaker does not need to enunciate each syllable.

It is important to note that the amount of time spent training personnel to use the LCS-Marine system is generally less than 10 minutes. After a short introduction, the user is shown a sample dialogue for familiarization. The user is also given information about meta-instructions – how to start over or to clear their previous statement – before they begin operation.

4. OPERATIONAL EVALUATION

To measure the effectiveness of the LCS paradigm under operational conditions—real users placing real requests, accessing a live database, and using existing communications links—we conducted a series of Integrated Feasibility Experiments (IFE). The IFEs ranged from a pilot study that featured scripted dialogue, replicated databases, and testing in the lab with prior military personnel, to field experiments where active duty Marines used the system operationally over a series of days as their sole means of interaction with the logistics system for rapid requests. We tested the system's support of placing and checking on requests for ammunition (Class V), fuels (Class III), and subsistence (Class I) supplies. More on the experimentation protocols can be found in [1] and [4].
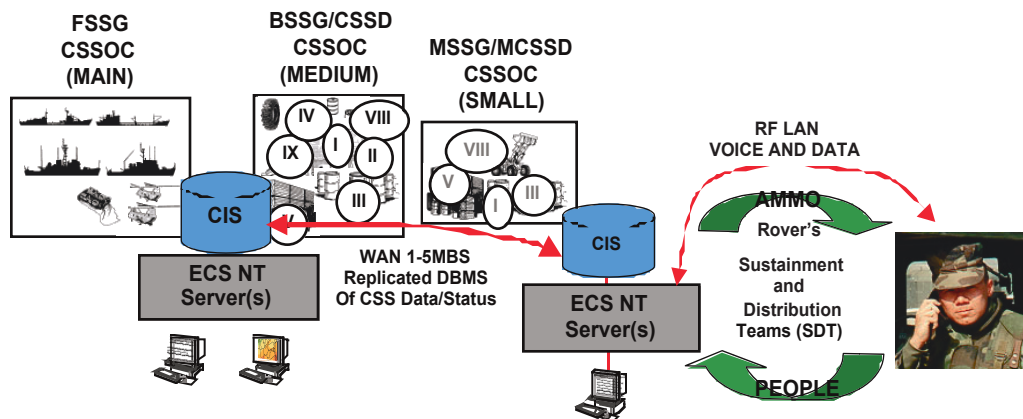

Figure 2. The Marine logistics ordering chain.

Figure 3. Partially Complete Rapid Request Form along with a portion of the database.

Over the course of the IFE process we were able to experiment with differing server configurations as well as varying communications linkages between servers. The most recent IFE (December 2000) used the server layout shown in Figure 4.
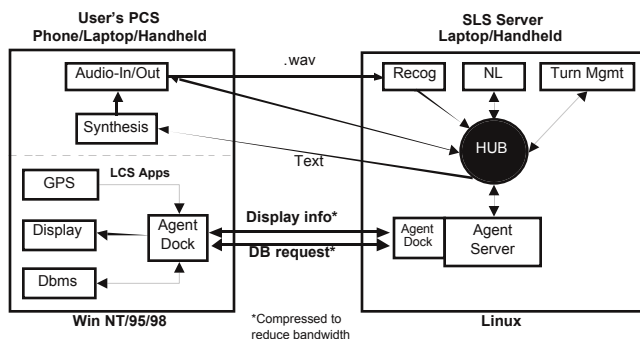

Figure 4. The physical LCS-Marine server layout.

The ideal configuration of the system would have a Marine using their organic communications system calling in to a remote location and communicating with the SLS there. This would not add any additional cost or hardware to the existing Marine infrastructure. This operational layout is depicted in Figure 5. Unfortunately, the current tactical radio, the Single Channel Ground and Airborne Radio System (SINCGARS), can create a large amount of channel noise, which alters or distorts the acoustic signal. Current recognizers can not yet compensate for this distortion, although there is active research into solving this problem.

We used a second operational layout to test the system and get operator feedback on using a spoken language understanding interface. This layout is depicted in Figure 6. In this layout, we required the user to beat the same location as the entire SLS system and the agents migrated over the SINCGARS data link
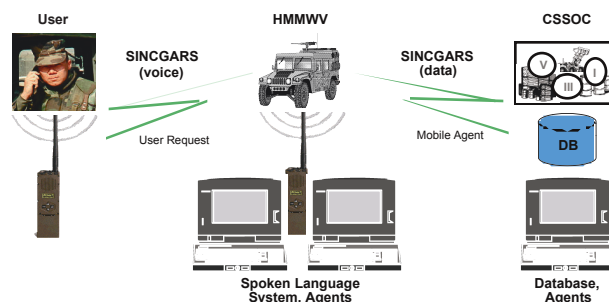

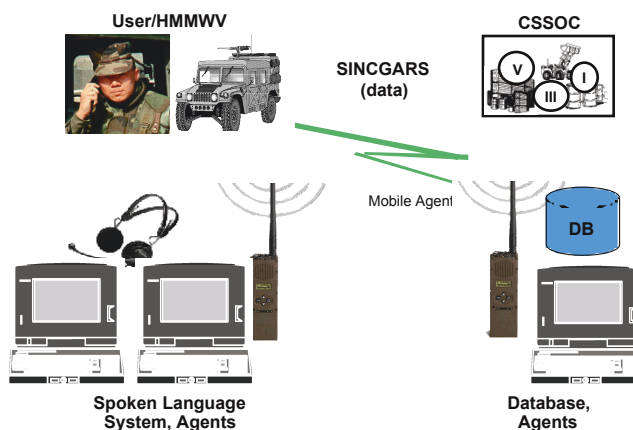**Figure 5. The ideal LCS-Marine operational layout.**

Figure 6. The LCS-Marine actual operational layout.

to reach the logistics database. The recognizer still had to contend with the issue of a noisy and dynamic background, but the acoustic distortion was eliminated.

5. CONCLUSION

We have built a system that integrates a spoken language understanding system with a mobile, intelligent agent system that allows users in a hostile acoustic environment to place and access data requests via a conversational interface. LCS-Marine is speaker independent and requires little training. The time to accomplish a task is significantly lower than the manual input method it seeks to enhance, but it can still be improved. Being able to rapidly access, insert, modify, and delete requests gives the users greater visibility into the supply system.

6. ACKNOWLEDGMENTS

7. REFERENCES

[1] Daniels, J. Integrating a Spoken Language System with Agents for Operational Information Access. In *Proc.. of Innovative Applications of Artificial Intelligence (IAAI-2000)*, August, 2000, Austin, TX.

[2] McGrath, S., Chacón, D., and Whitebread, K. Intelligent Mobile Agents in the Military Domain. In *Proc.. Of Autonomous Agents 2000 Workshop on Agents in Industry*. Barcelona, Spain.

[3] Seneff, S., Lau, R., and Polifroni, J. 1999. Organization, Communication, and Control in the GALAXY-II Conversational System. In *Proc.. of Eurospeech '98*. Budapest, Hungary.

[4] Stibler, K., and Denny, J. A Three-tiered Evaluation Approach for Interactive Spoken Dialogue Systems. In *Proc.. of the Human Language Technology Conference HLT-2001*, Mar, 2001, San Diego, CA.