WordPlay: A Table-Top Interface for Collaborative Brainstorming and Decision Making

Seth Hunter, Pattie Maes

Massachusetts Institute of Technology

Ambient Intelligence Group, Media Laboratory

hunters/pattie@media.mit.edu

Abstract

In this paper, we present "WordPlay", a collaborative tabletop interface for generating, organizing, and exploring ideas. The concepts, ideas and arguments discussed in a meeting around the table are visualized and can be organized and manipulated in natural ways by the participants. Users add ideas to the table by speaking, manipulate the properties of words with hand gestures, and explore related concepts by tapping them. WordPlay combines Speech Recognition software, Natural Language Processing, a semantic knowledge network, and multi-touch capability to provide a public forum for brainstorming and support group decision making.

1. Introduction: Augmenting Brainstorming and Decision Making

With the recent introduction of affordable multi-touch technologies to work environments [7, 11] we see an opportunity to design a more fluid platform for the support, engagement, and contribution of computing to collaborative discussions.

Much of the research in computer supported collaborative meeting interfaces [17, 21] enables visualization [8, 5], shared displays [1], document sharing [14, 16], organization of content [3, 9], and group decision making [6]. We designed *WordPlay* to support these functions in a multi-touch environment and extend these modalities by including the computer as a participant in the conversation. The increased accuracy of speech recognition systems [4] and the development of a common sense knowledge databases [2] enable the system to provide associative suggestions that may trigger a novel branch of thought during brainstorming and decision making scenarios.

2. User Experience: Generating, Organizing, and Exploring Ideas

So far we have experimented with and built support for two types of meetings: deciding between a set of alternatives and idea generation. Our group has gathered to brainstorm about possible domain names, product usage scenarios, and to try and think of new idea for game development. During demos of the table we have asked people for their thoughts about presidential candidates or for creative suggestions regarding a contemporary political decision.

Before a meeting, a facilitator prepares the table and chooses a background that corresponds the subject at hand. Participants gather at the table and contribute content by speaking into the microphone, or selecting a scalable, multi-touch keyboard. They can arrange ideas by sorting, expanding, and deleting them on the surface of the table. Throughout the session, any user can tap on ideas to request associations and suggestions from the system.

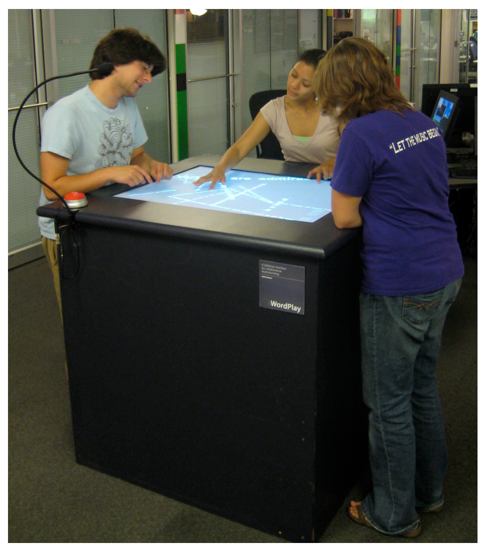

Figure 1. Participants at the WordPlay Table

2.1 Generating Ideas with Voice and Touch

In order to facilitate more natural and seamless input to the table we implemented a speaker independent voice engine using Dragon NaturallySpeaking9 [4]. We found that by using a noise cancelation microphone, recognition worked moderately well even during gallery openings and events.

To compensate for a 10% error rate in the speech recognition and errors related to users with a nonenglish accents, phrases can be deleted and a multi-touch keyboard is always available for the generation of words not in the speech dictation dictionary.

2.2 Organization for Decision Making

An important feature of *WordPlay* is the ability to toggle a background that defines a matrix for positioning factors in decision making scenarios. Users can position their opinions and thoughts within a specific context. During our public events we used a matrix with two candidates for political nomination on the left and right axis, and pros and cons on the top and bottom. (see figure 2 below)

Using multi-touch rotation algorithms, the software enables users to modify the properties of ideas or the entire canvas. We decided to let size correspond to the importance of the idea, rotation indicate which user had placed the idea, and position be relative to the way the user felt about that idea related to each axis of decision matrix background.

2.3 Exploring Computational Associations

One of the "playful" aspects of *WordPlay* is exploring the extensive databases linked to each of the words on the table. Etymologies, definitions, and relationships from the ConceptNet 3 semantic knowledge database, part of the OpenMind Commons [2] expand from words selected by the user.

We found that the relationships like "Is a", "Has the property of", "Desires", "Motivated By Goal", and "Used For" were particularly useful for tasks like naming a product or free associating about possible product usage scenarios.

Like the Visual Thesaurus [22], new branches expand recursively. Within two or three touches the landscape of the table can be visually and contextually transformed.

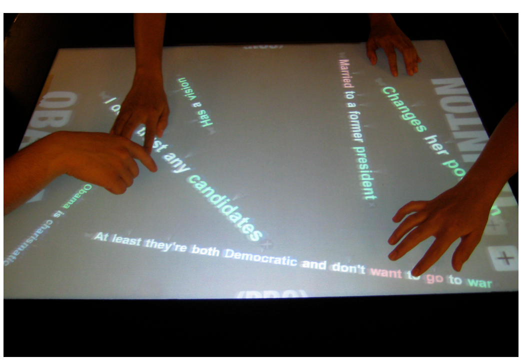

Figure 2. A decision making scenario between two candidates

3. UI Implementation: Design Principles

3.1 Typographic Animation

Ideas visualized on the table have animated behaviors [23] in order to address orientation, collision, and content generation. To orient equally around all four sides of the table they follow Lissajous curves [10], a system of parametric equations describing complex harmonic motion. Phrases also have a lifetime of two minutes, their transparency fading slowly if the user does not select the phrase for positioning on the table. This reduces the complexity and emphasizes only intentional layout decisions.

3.2 Aesthetic Organization

We used WordNet [22] to separate an speech phrase into its parts (adjectives, adverbs, nouns, and verbs) and identify concepts. These attributes were used to determine color and size properties of the words as well as to negate articles of speech and commonly used words.

3.3 Gestural Vocabulary

Like existing taxonomies of gestural touch efforts [15] [12], we attempted to incorporate a "stroke" or "glide", a directional touch, as a method of triggering new associations to grow from an idea. Without immediate visual feedback, this gesture was not intuitive or transparent, so we added an animated icon in front of the finger which follows the gesture path and acts as a confirmation of intent. The development of a grammar and vocabulary for gestural feedback is still an underdeveloped research area.

4. Platform Specifications

4.1 Hardware

The prototype is a closed table approximately 40 inches cubed. The display is an FTIR display [7] with a 1 mm silicon compliant surface [18] and a Mylar Projection surface of 32 by 24 inches. It is equipped with a low cost vision camera, a mirror, and a current sink to provide headroom to the LED arrays embedded in the acrylic surface. In addition, there is a 24 inch Buddy Gooseneck noise cancelation microphone and a large press button mounted on the corner of the table for speech input. (see Figure 1)

4.2 Software

The architecture and software tools used in *WordPlay* are outlined in the diagram below. Data flows from the microphone to a series of software programs which attempt to parse the audio into text and find appropriate associations. We used the Natural Language Processing Toolkit [13], WordNet [22], and ConceptNet3 [2] with a software written in the Python scripting language.

Camera events are handled by a modified version of the open source TouchLib [19] libraries and are passed to the Graphics Engine (see section 4). The end users see only the visual output of the graphics engine. We used the OSC protocol for communication between engines.

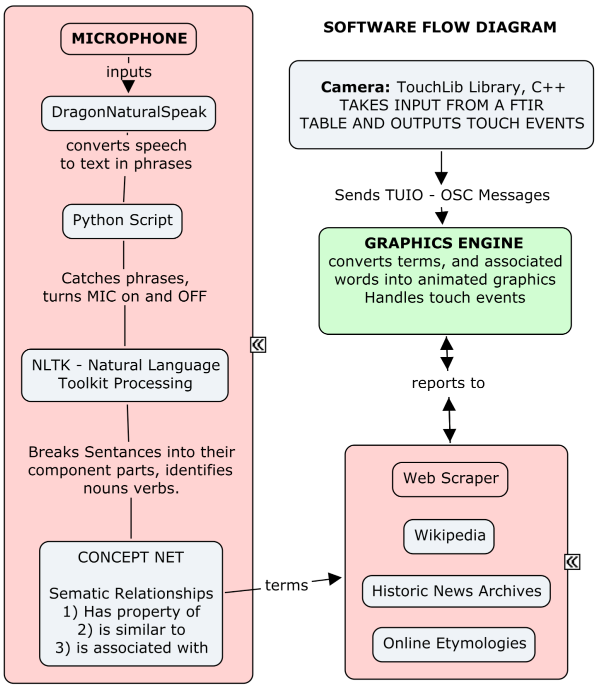

Figure 3. Software flow diagram for WordPlay

5. Usage Observations

Presently, evaluations are based on usage over three months of lab demos, museum events, and our group's internal usage of the table. Some of the lessons we learned during usage are described below.


Figure 4. Users in a Museum

5.1 Platform Design

The physical build of a table should be as sturdy as possible without compromising the design. We found that for surfaces larger than 32 inches, 3/8 inch acrylic will bow in the middle up to 1/4 inch in response to pressure. The top of the table should account for significant weight of participants in group situations, especially for builds at counter height 36-40 inches. Participants lean on the frame with one hand and touch the screen with the other about 40% of the time they are using the table. Even for tall tables knee room is an important factor. Knees and legs tend to rock against the sides of the table, influencing the stability of the projector.

5.2 Meeting Application Software

We found that when multiple people are using the interface, scaling the entire canvas can create confusion between users. Canvas scaling should be toggled for individual control. For brainstorming applications a larger, more elongated surface is preferable to reduce overlapping of phrases and increase the clarity of the arrangements.

Feedback animations attached to gestures were the most successful way for users to learn the capabilities of the system. We found that people would orient by experimenting quietly for a minute or two before fully engaging with the applications.

6. Conclusion and Future Work

Multi-touch computing seems especially suited to augmenting collaborative discussions by enabling users to visualize, modify, and expand on their ideas in a fluid manner. Applications like *WordPlay* create space to explore the associative properties of words together in a context of social engagement where the computer acts as an agent in the discussion.

We are excited by the design possibilities at the intersection of social conversation spaces and multi-touch computing platforms. *WordPlay* is just one facet in a larger domain of social interfaces that can provide a set of tools to augment, visualize, associate, and archive our creative group discussions.

Future work will incorporate pre and post session tools for preparing media, and accessing archival information. Our next tabletop display will have a larger surface area, a network for proximity sensing, RFID identification of participants, and drawing capability. Layering streams of audio and video into the design of the system may also help provide a more compelling space to gather, design, brainstorm, and make decisions.

7. References

[1] Borchers, J., Ringel, M. Tyler, J., Fox, A., Stanford interactive workspaces: IRoom, a framework for physical and graphical user interface prototyping, IEEE Wireless Communications, 2002 Vol 9, pp 64-69

[2] ConceptNet 3, A Semantic Network Representation of the Open Mind Common Sense Project http://conceptnet.media.mit.edu/

[3] Crabtree, A., Roden, T. & Mariani, J. Collaborating around Collections: informing the continued development of photoware. In Proc. CSCW 2004, ACM Press (2004), 396-405

[4] Dragon Naturally Speaking Speech Recognition Software, Nuance, 2007, http://www.nuance.com/naturallyspeaking/

[5] DiMicco, J.M., Changing Small Group Interaction through Visual Reflections of Social Behavior, Doctoral Thesis, Massachusetts Institute of Technology, 2005, pp 31-44

[6] Fang Chen, Benjamin Close, Peter Eades, Julien Epps, Peter Hutterer, Serge Lichman, Masa Takatsuka, Bruce Thomas, Mike Wu1 ViCAT: Visualization and Interaction on a Collaborative Access Table, IEEE, Workshop on Horizontal Interactive Systems, Australia, 2006

[7] Han, J. Y., Low-cost multi-touch sensing through frustrated total internal reflection. ACM Symposium on Unser Interface Software and Technology, Seattle, WA, 2005, UIST '05. ACM Press, New York, 115-118.

[8] Karahalios, K., Bergstrom, T., "Visualizing Audio in Group Table Conversation", IEEE, Workshop on Horizontal Interactive Systems, 2006

[9] Lindley S., Durrant A., Kirk D., Taylor A, "Collocated Social Practices Surrounding Photos" CHI 2008

[10] LissaJous Curve. http://en.wikipedia.org/wiki/Lissajous_curve

[11] Microsoft Surface. http://www.microsoft.com/surface/

[12] Morris, M. R., Huang, A., Paepcke, A., and Winograd T. 2006. "Cooperative gestures: multi-user gestural interactions for co-located groupware." CHI '06. ACM Press, New York, NY 1201-1210

[13] Natural Language Toolkit, software for natural language processing, http://nltk.sourceforge.net/index.php/Main_Page

[14] Rekimoto, J., Saitoh, M., Augmented Surfaces: A Spatially Continuous Work Space for Hybrid Computing Environments, CHI, 1999

[15] Schiphorst, T., Motamedi, N., and Jaffe, N, "Applying an Aesthetic Framework of Touch for Table-Top Interactions", IEEE, Workshop on Horizontal Interactive Systems, 2007, p. 74

[16] Shahram Izadi, Harry Brignull, Tom Rodden, "Dynamo: A public interactive surface supporting the cooperative sharing and exchange of media" UIST03, pp159-168, 2003.

[17] Scott, S. Grant, K, Regan, M, "System Guidelines for Co-located, Collaborative Work on a Tabletop Display", Presented in the Co-Located Tabletop Collaboration: Technologies and Directions CSCW2002 Workshop

[18] Smith, D., Graham, T.C.N, Holman, D., Borchers, J., "Low-Cost Malleable Surfaces with Multi-Touch Pressure Sensitivity", IEEE, Workshop on Horizontal Interactive Systems, 2007, pp. 206-7

[19] TouchLib, a library for creating multi-touch interaction surfaces, http://www.nuigroup.com/touchlib/

[20] Visual Thesaurus, ThinkMap Inc. http://www.visualthesaurus.com

[21] Wigdor, D.; Shen, C.; Forlines, C.; Balakrishnan, R., MERL Research, DiamondTouch Table "Table-Centric Interactive Spaces for Real-Time Collaboration", AVI 2006

[22] WordNet, a lexical database for the English language, Cognitive Science Library, Princeton University, http://wordnet.princeton.edu/

[23] Xiaoyuan Tu and Demetri Terzopoulos. "Artificial Fishes: Physics, Locomotion, Perception, Behavior", ACM Computer Graphics, Proceedings of SIGGRAPH'94, July 1994.