
The AMODEUS Project

ESPRIT Basic Research Action 7040

VoicePaint:
a Voice and Mouse-Controlled Drawing Program

Salber D., Coutaz J. & Nigay L.
Laboratoire de Génie Informatique, University of Grenoble

23rd February 1993

Amodeus Project Document: System Modelling/WP9

AMODEUS Partners:

MRC Applied Psychology Unit, Cambridge, UK (APU)
Depts of Computer Science & Psychology, University of York, UK. (YORK)
Laboratoire de Génie Informatique, University of Grenoble, France. (LGI)
Department of Psychology, University of Copenhagen, Denmark. (CUP)
Dept. of Computer & Information Science Linköping University, S. (IDA)
Dept. of Mathematics, University of the Aegean Greece (UoA)
Centre for Cognitive Informatics, Roskilde, Denmark (CCI)
Rank Xerox EuroPARC, Cambridge, UK. (RXEP)
CNR CNUCE, Pisa Italy (CNR, CNUCE)

VoicePaint:

A Voice and Mouse-controlled drawing program

**Daniel Salber
Joëlle Coutaz
Laurence Nigay**

February 1993

Abstract

This document presents VoicePaint along the lines defined for the description of common exemplars: overview, reference material available, hardware and software platforms, usage, and future plans. VoicePaint supports the simultaneous use of voice commands to modify brush attributes and color while drawing with the mouse. It illustrates the synergistic use of multiple modalities for input (mouse and voice). It has been implemented on the Macintosh using VoiceNavigator, a discrete, speaker dependent speech recognition system.

Table of contents

1. Overview of VoicePaint.....	1
2. Reference material available.....	2
3. VoicePaint and the criteria for exemplar selection	2
4. Hardware and Software platforms.....	4
5. What we have done with VoicePaint and the future	4
6. References.....	4

1. Overview of VoicePaint

VoicePaint is a colour, pixel-oriented drawing program (*à la* MacPaint). It handles speech, keyboard and mouse inputs. VoicePaint is a traditional Macintosh application with all of the standard interface elements, as defined in the Apple User Interface Guidelines [Apple 86]. Because of this, mouse and keyboard can be used to manipulate windows, menus, dialogues. But because it is also a voice-enabled program, commands may also be issued by voice. For example it would possible to create a new document or print the current one using voice commands. VoicePaint also allows to use voice commands to modify brush attributes while drawing; this feature relies on the synergistic fusion of input data.

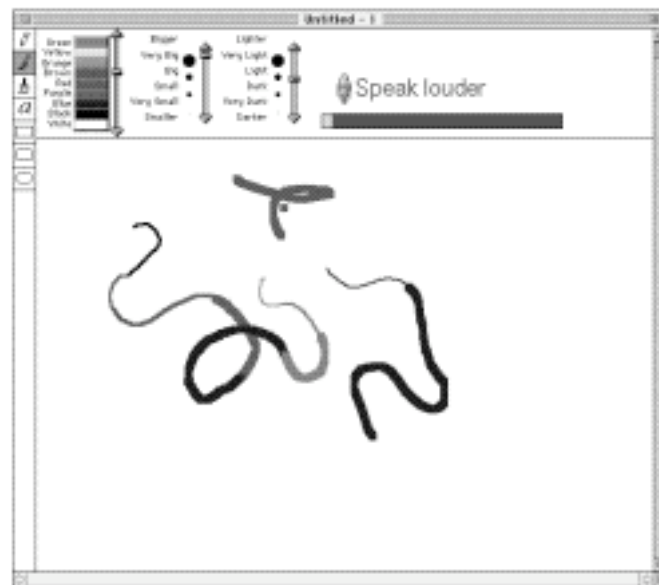


Figure 1. A snapshot from VoicePaint.

As shown in Figure 1, the user interface of VoicePaint is composed of the following elements:

- on the left hand side of the window, a palette of tools allows the user to choose among geometric shapes and brush types;
- at the top, a palette shows the set of available colors;
- then two sliders are used for setting the size of the brush and the luminance of the paint (the luminance can be modified in black and white mode only). The sliders can be moved with the mouse by direct manipulation or moved verbally. For example, the user may say “green” to set up the current color to green. The cursor of the slider moves to position “green” and the paint turns to green;

- in the upper right part of the window, an icon and a bargraph respectively represent the microphone and the intensity of the voice input signal. If the sound level is too low, the system displays a warning: “speak louder”. If the user talks loud enough, then the recognized token is displayed instead;
- the bottom part of the window is dedicated to the drawing area. Pictures can be drawn using the usual graphical Macintosh paradigm. In addition, the user can modify the size of the brush, or the color and luminance of the paint while drawing free hand pictures. In particular, while the mouse button is pressed, the mouse can be moved and simple sentences such as “green”, “bigger”, “very small”, “blue” can be uttered to modify the graphics context until the mouse button is released.

According to the framework described in [Coutaz 93], VoicePaint is a multimodal interactive system for input only. At the command level, it supports fusion of input data issued simultaneously from two distinct channels (microphone on the one hand, mouse on the other hand). It also performs fusion of data produced through a single channel such as speech or mouse only. Thus VoicePaint supports both the synergistic and exclusive use of modalities. It also supports interleaving at the task level: multiple windows such as the one shown in Figure 1 can be opened by voice or through the File menu of the menubar. Then it becomes possible for the user to elaborate multiple documents in an interleaved way.

2. Reference material available

VoicePaint is fully described in A. Gourdol’s Master Thesis [Gourdol 91]. This document is written in French.

A three minute PAL video is also available. It shows an example of multimodal synergistic inputs (voice & mouse) using VoicePaint.

The VoicePaint software itself and its source code are available. It runs on any Macintosh and requires VoiceNavigator for the recognition of voice input.

3. VoicePaint and the criteria for exemplar selection

1) Real world validity

The real world validity of VoicePaint is poor. It has been designed and developed as a toy example to study the software dimension of multimodal interaction and to demonstrate the feasibility and the interest of multimodal systems.

VoicePaint has been used for demonstration purpose only.

2) Design process as well as product

Does not apply.

3) Influence over design

Does not apply.

4) Available design knowledge

VoicePaint has been developed by the Grenoble team.

5) Domains of interest

VoicePaint is a synergistic multimodal system. It provides a way to draw free hand pictures that cannot be performed easily in the real world: as you draw, you can change the color, the thickness and the luminance. Parallelism at the physical level through distinct communication channels makes this possible with VoicePaint.

6) Support development of integration

VoicePaint was a first attempt to perform data fusion between modalities. It has emphasized the problem of the common representation of data from different modalities.

VoicePaint supports :

- multiple resource types,
- multiple channels,
- multiple interaction styles.

7) Familiarity

The domain of VoicePaint is drawing (pixel-oriented). It is a familiar domain.

8) Communicable

The reference material available is described in paragraph 2 of this document.

A running version of VoicePaint is available in Grenoble.

9) Coverage of models

To be discussed. From the system point of view, however, VoicePaint illustrates a number of interesting phenomena: parallelism at the interface, fusion at low level, fission at high level, multithreading at the task level.

10) Commonality of analysis

To be discussed.

Table 1 presents the features of VoicePaint with regard to the exemplar criteria.

1. Real-World validity	LOW
2. Design process as well as product	?
3. Influence over design	?
4. Available design knowledge	YES
5. Domains of interest	YES
6. Support development of integration	?
7. Familiarity	YES
8. Communicable	YES
9. Coverage of models	SYSTEM+?
10. Commonality of analysis	?

Table 1. VoicePaint and the criteria for the exemplar selection.

4. Hardware and Software platforms

VoicePaint runs on any Macintosh with system software version 6.0 or higher. VoicePaint has been developed using Think Pascal (object-oriented Pascal), and HIK, a custom application framework developed by A. Gourdol [Gourdol 89].

The voice services interface and voice acquisition hardware is VoiceNavigator by Articulate Systems [Articulate 90]. VoiceNavigator provides discrete, speaker dependent voice recognition. The speech recognition associates recognized utterances with tokens that are then handled by the application. The recognizer needs to be trained for every user and command.

5. What we have done with VoicePaint and the future

VoicePaint has been used to study the feasibility and the interest of multimodal systems. It has shown that synergistic multimodal systems are of great interest since they allow the user to perform commands that are not otherwise possible; with VoicePaint, the user can change the color or the size of the brush *while drawing*. This is not possible with mouse-only drawing programs such as MacPaint. VoicePaint was also a successful implementation of a multimodal system using PAC agents.

We do not plan to develop VoicePaint any further.

6. References

[Apple 86]

Apple Computer, Inc. Human Interface Guidelines: The Apple Desktop Interface, Addison Wesley, New-York, 1986.

[Articulate 90]

Articulate Systems Inc., The Voice Navigator Developer Toolkit, 1990.

[Coutaz 93]

J. Coutaz, L. Nigay, D. Salber: The MSM framework, A Software Design Space for Multi-Sensori-Motor Systems, Amodeus 7040 Working paper, SystemModelling/WP4, Jan. 1993.

[Gourdol 89]

Human Interface Kernel: Developer's Guide, 1989.

[Gourdol 91]

Architecture des Interfaces Homme-Machine Multimodales, Master's Thesis, Université Joseph Fourier, 1991.

[Gourdol 92a]

A. Gourdol, L.Nigay, D.Salber and J.Coutaz: Two Case Studies of Software Architecture for Multimodal Interactive Systems: Voice-Paint and a Voice-enabled Graphical Notebook, IFIP WG2.7 Proceedings, 1992.

[Gourdol 92b]

A. Gourdol, L.Nigay, D.Salber and J.Coutaz: Multimodal Systems: Aspects of Events Fusion and a Taxonomy, IFIP Conference '92 Proceedings.