

959- Voice, a new interface to GIS

by Ir Charles Kemper & Benoit Ballieu - Voice-Insight SA/NV

1. Introduction

Information Technology has been evolving rapidly in the last 25 years.

Milestones like the introduction of virtual memory operating systems, data base technology, object oriented applications and operating systems on one hand and the fabulous microprocessor evolution on the other hand have shaped the IT scenery of today, without forgetting the internet and the telecommunication evolution.

In a period where the Personal Computer is becoming more and more portable "growing", or should we say "de-growing" into palm devices and into multimedia capabilities, in a period where the borders between a Television, a PC and a Telephone are vanishing, there is one emerging technology to watch: the combination of Voice Technology and Natural Language techniques.

The drivers from a technology point of view of the demand for voice enabled products are:

- Handheld wireless devices such as
 - cell phones,
 - PDAs (with WiFi and or Blue Tooth),
 - wireless microphones
 - Portable data repositories like MP3 players etc...
- Natural language and Relational Database Systems Technology
- Telecommunications and Internet/Intranet telephony (Convergence of Voice, Data and Video)
- Emerging Interactive TV

This evolution is creating the perfect setting for the convergence of Voice, Data and Video.

Pioneers of the Voice Recognition and Synthesis like Philips, Dragon, L&H, SpeechWorks(regrouped now under ScanSoft) and Nuance have been establishing the basic ground for applications to emerge, while Operating Systems companies like Microsoft and infrastructure companies like IBM and Panasonic are standardizing the incorporation of this technology into the supporting layers of the IT/Multimedia sets of tomorrow.

The Microsoft SAPI (Speech Application Programmable Interface) is the de facto standard for embedded solutions while SALT and VoiceXML are other emerging standards in the internet server space.

To really understand the impact of voice technology each of us should imagine for a while that he is sitting in front of a device accessing his favourite application, the application he has been building in the last years, in which he has invest in the last year time and money. He should imagine himself manipulating his favourite application simply using his voice in a natural language....

Imagine, now that he could do this remotely, just using his voice and a telephone or a portable device using voice over IP.

Imagine now that he could browse from one application to the other, from GIS to digital imaging, from digital imaging to document management etc..., simply using his voice.

While doing so, browsing with his voice, input, processing, display and network devices are redistributed in the architecture of tomorrow....

The application drivers in a GIS environment are :

- The need for up to date data with a more rapid turnover requiring more efficient field data acquisition
- The need for direct mobile access to given data in specific mission critical domains
- The need of an natural easy to use interface to generalize the use of GIS to all literate as well as illiterate PC users

To imagine the benefit for GIS users this paper will use practical examples.

2. Possibilities of Voice Technology for GIS

Potential users, developers and integrators have to see the voice channel as an additional way to pilot their application, in some cases easier and more natural according the circumstances. By essence, it is adding voice interaction to existing graphical and traditional user interfaces to offer the user more efficiency, and mobility. Speech is the most natural and quickest tool man possesses to interact with their environment.

Essentially, there are 5 functionalities Voice Technology allows to do :

- Navigate in a GIS environment,
- Query any information contained in a database and
- Update data in the field.
- Initiate some commands like digitizing, digital image capture,
- Manipulate objects like copy from one layer to another layer etc...

2.1) Navigate in a GIS environment



Entering a location by voice, into the PC or handheld device is so within everybody's reach, even while driving. Next the system would offer a selection of similar names in other cities from which the user can choose the appropriate one, displaying then the map he is looking for. The input starts from the information available: This could be an address composed by a street name, a city and a country but also a building name, a district, an area, a department or a section of a site or city, a shelf number in a depot or store, explicit geographical coordinates; for short: any information stored in the database. This all can be used to find the destination.

On the obtained map or floor plan, the user can use his voice for display and navigation commands like zoom in/out, pan or fit map or selection/item. To illustrate this the commands of the new Voice-Assistant for ArcPad version 6.02 for instance are allowing to rotate the map x degrees to the left or to the right. Now, he can also activate and disable the automatic rotation mode should he be using a GPS for Geolocation. More complex queries to find specific objects are also possible.

2.2) Data query and information request

A GIS voice interface offers the possibility to ask for all attached attributes, values of a selected object (GIS or not) in a direct way by passing the usual tree menu structure required by traditional GUI. The answer is in function of the used device and requirements vocal and or graphical. Voice application could for instance allow the seamless jump from one database or database application to another.

An hyperlink to a pdf file, picture, video, soundtrack, URL or website (embedded, on a local network or on the internet by GPRS mode) can be activated by voice delivering immediately the available and required information. With the emerging quick wireless connections, the information must not be embedded on the device, but can be retrieved from remote servers at any given time. Even searches in large databases can be done remotely where more CPU-power is available like for example in security management or police applications. Using a mobile phone, the needed information can be returned in a SMS or MMS, depending the size of the information and available network.



■ Facility & Building Management



"I'm looking for an office,

- On floor level,
- With 2 desks,
- With more than 2 windows,
- With Internet connection,

List all offices"



"1 office was found.
Room 124 of the Copernicus building"



"How many electricity sockets are there in this room?"



"6 sockets on the walls and 4 sockets on the ground, in the middle of the room"

■ Real Estate

Applied to Real Estate, this would create a voice interface that helps to find a house, an apartment or an office, simply by giving the conditions about the real estate asset you are looking for and this 24 hours by 7. Based on the attribute values (surface, city, number of bedrooms, presence of a garden, swimming pool or garage, etc...) the user can vocally select an asset on the result list and ask to display graphical information like pictures, floor lay-outs and maps, but also all relevant information concerning the contract sale or rent price and so on.



"I'm looking for an office,

- In Brussels,
- Near the European Communities,
- With a surface between 200 and 300 square meters

List all assets."



"3 assets were found

- Avenue Cortenberg 112,
- place du Luxembourg 1-2,
- rue du Trône 26B "



"Select the first asset"



"Give me the surface"



"234 square meters"



"Give me the rent price"



"75 € per month and square meter"

■ People and Occupancy Management

A vocal interface on applications in this domain, will deliver the freedom to consult the people database by phone or through a PDA. The user just list the characteristics of the desired person, the same way as the police proceeds for wanted criminals or missing persons. Once selected, he can ask for all relevant information about this person to make the best choice in a fraction of the time it would normally take.



- Younger than 35,
- With 7 years active experience
- With a driving licence
- Specialized in electronics

List all engineers



"23 engineers were found"



"Select first one"



"I'm ready"



"Show CV"



Orin Hagg

1215 South Allen Street # 2
State College, PA 16801
(717) 337-7000
ohagg@psu.edu
www.vql.com/orin-hagg

327 Thomas Building
Department of Statistics
The Pennsylvania State University
University Park, PA 16802
(814) 863-1710 (office)
(814) 863-1519 (fax)

Education

1990-present Ph.D. student, Department of Statistics, The Pennsylvania State University, University Park, PA
• GPA: 3.94/4.0
• Dissertation Topic: "Two Stage Multiple Imputation," advisor: Joseph Schallin, Ph.D.

1987 B.S., Department of Statistics and Department of Economics, University of Idaho, Idaho Falls, ID
• GPA: 3.94/4.0 - Statistics
• GPA: 3.24/4.0 - Economics

Experience

2000-present Research Assistant, Methodology Center, Penn State University
• Develop scoring data techniques, Supervisor: Joseph Schallin, Ph.D.

1999-2000 Graduate Coordinator, Statistical Consulting Center, Penn State University
• Conducted on research projects with graduate students and faculty
• Held key on consulting hours for graduate students and faculty

1995-2000 Graduate Student of Statistics, General Clinical Research Center (GCRC), Penn State University
• Conducted on research projects with graduate students and faculty



**VQL for
Facilities Management**

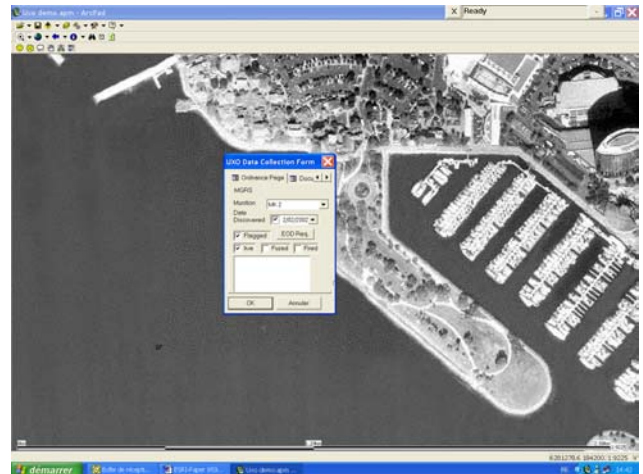
Younger than 35,
With 7 years active experience
With a driving licence
Specialized in electronics

List all Engineers

2.3) Data entry and information updating

Voice is probably best to be used in GIS for field data collection and update. This functionality allows the user to open forms for verification of the stored information, comparing it with the real world and goes to the edit mode when the reality and the database are presenting discrepancies that need to be addressed accordingly. Updating the information on the server can be done by regular synchronization in the evening after a days work or instantaneously by a wireless connection so that content on the server has the highest possible reality or accuracy level. In addition to normal text entries, digital pictures with geographical references can be taken when needed and added to the central database, all this by voice.

Voice-Insight's Voice Assistant for ESRI ArcPad™ is an example of Voice interface for GIS : it allows for instance to voice-activate editable user forms by writing simple visual basic scripts. In following example we illustrate the data collection form for green management.

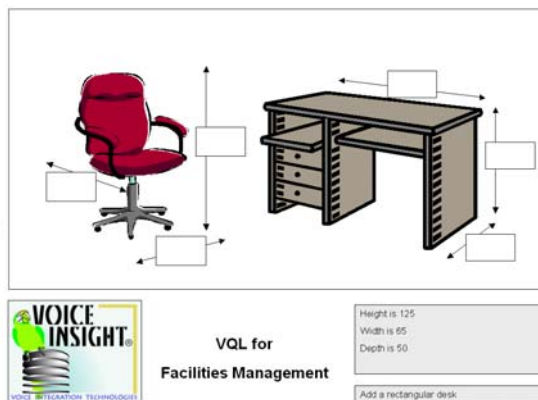


■ Green management

- "Open the data collection form"
- "I'm ready"
- "Tree is birch"
- "Height is between 5 and 7 meters"
- "State is very good"
- "Update information"
- "Data was successfully updated"



■ Inventories, Storage or Furniture Management, technical objects for utilities



- "Add a chair to the list"
- "What kind of chair ?"
- "A rolling office chair with armrests"
- "Give the dimensions please"
- "Height is 1meter 25"
- "Width is 65 centimeters"
- "Depth is 50 centimeters"
- "Something else ?"
- "Yes, a rectangular desk"

4) Initiate a command : Voice digitizing

More complex commands can be initiated like request to create a symbol at the current GPS position with eventually a digital picture been taken at the same time and geo-referenced to the voice digitized position. Equally when reviewing existing maps in the field selected objects could be move from layer to the other...

5) Manipulate and Manage Objects on the map

Objects on the map can be manipulated and meta data updated to become in line with the real modified circumstances.

Equally, using voice technology technology, it is easy to navigate on a site map of factories, offices, buildings, networks or any type of object , simply by using the voice.

In the same way, facility & object management can be done using voice : floor drawing plans for these managed sites , highlight some specific zones or areas as open spaces, closed room areas, green areas, functional zones, depots, false ceilings, free zones, pipes, ...

The user can ask then for attributes of objects present in the maps: buildings, floors, offices, rooms but also their building components like walls, windows, doors, stairs, ... For each of them, he can ask for available information and details as long as those are stored in the database and he eventually can enter information like : building year, floor level, surface type, floor covering (status: need for cleaning or repair/replacement), location and condition (expiration date) of the fire extinguishers, vertical or horizontal surface, volume, space allocation, furniture and equipment list, ...

3. Practical example

**Imagination is the limit,
Return on Investment the final driver;
the example given here below is courtesy (data and script) of the Belgian NGI:**

The configuration is based on ESRI ArcMap, Voice-Insight VQL and Siemens Tablet PC.
The goal is to allow field map verification and modification for the 1:50.000 NGI maps.

There are 2 methods to create a map on the scale 1:50.000

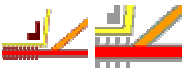
- From air pictures
- From more detailed maps



Symbolisation



Simplification



Moving and scaling some objects

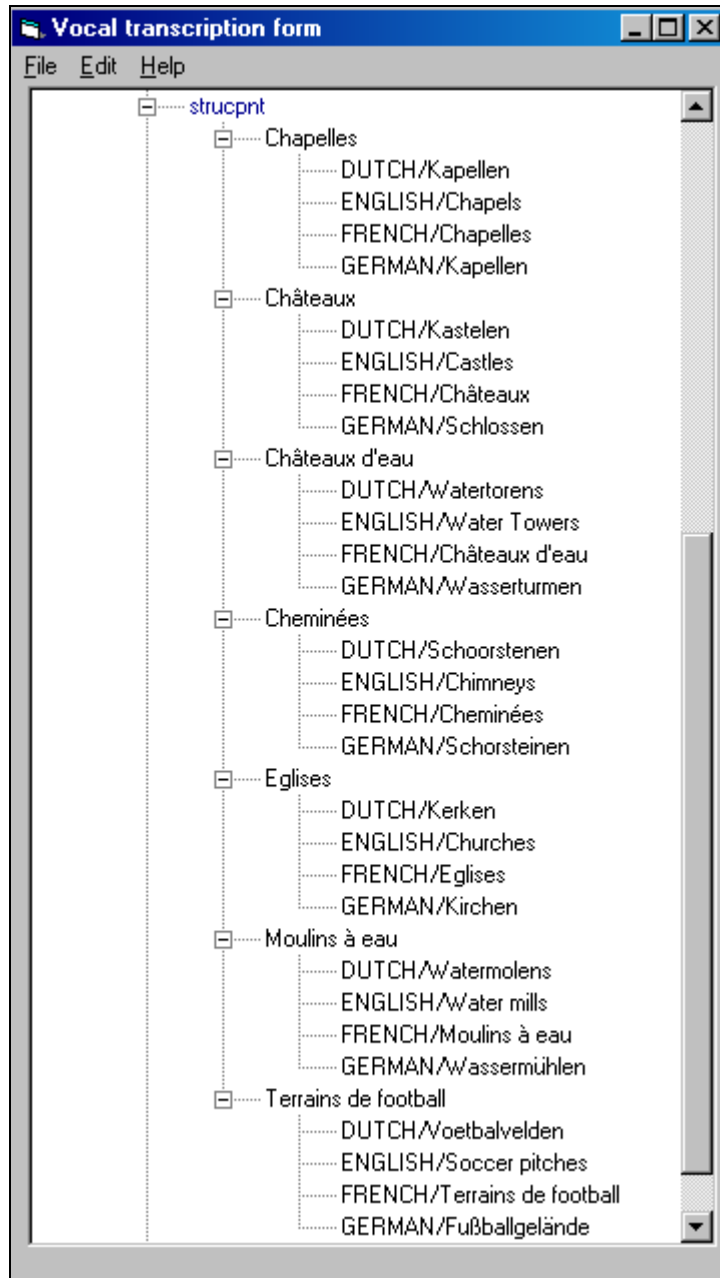
In both cases, once the map is created, field update is needed to :
maintain the map up-to-date
maintain the quality standards
This field update is done by using a car equipped with ArcMap on a Tablet PC.
The voice activation of ArcMap is briefly described here after :

3.1 MULTILINGUAL VOCAL DISPLAY CONTROL:

Belgium being a multilingual country, the interface should work in 3 to 4 languages : Dutch, French, German and for completeness English.

Display / Hide all points / curves / areas

Display / Hide the structural points

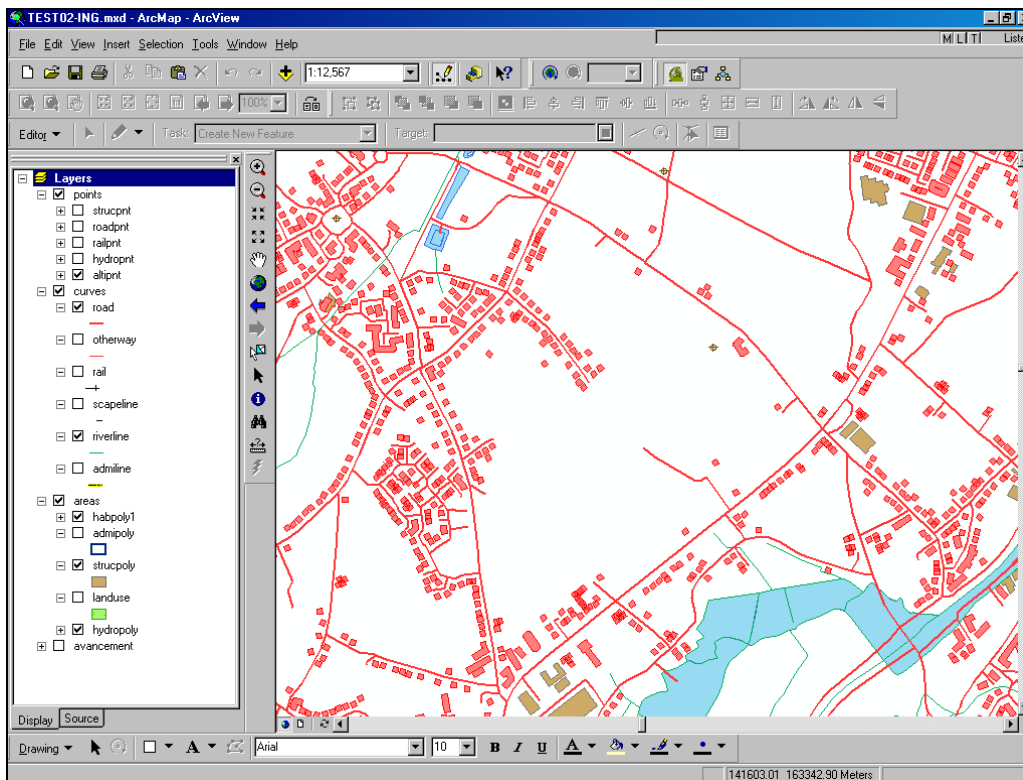
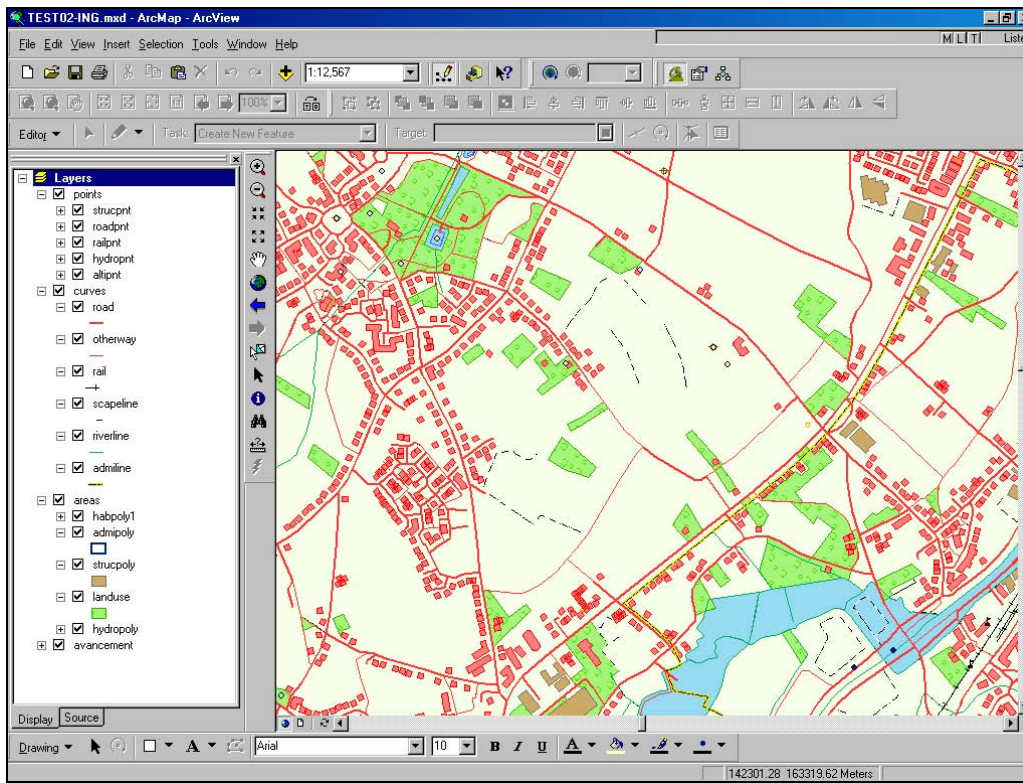


Display / Hide the chapels

Display / Hide the castles

Display / Hide the water towers

...



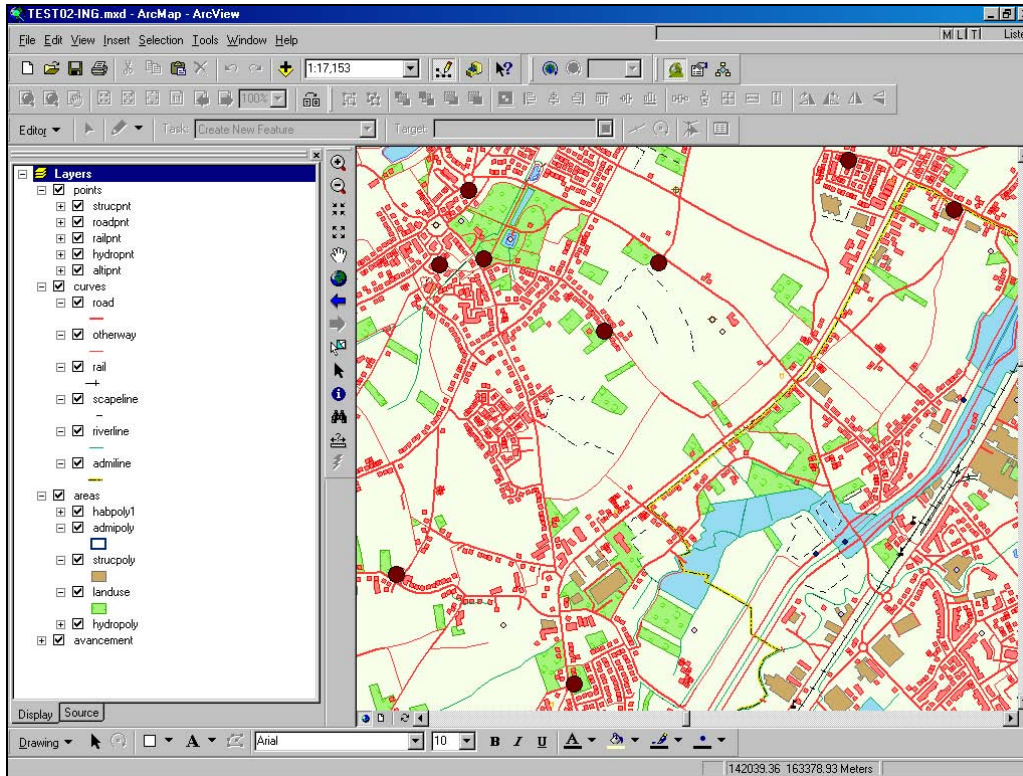
3.2 VOCAL ITEM SELECTION

Zoom-Select and Highlight the item you want vocally to execute another command just after it.

Select all the navigable rivers
Select all schools
...

With a GPS, the user can also say :

- Within x miles (meters), select all milestones
- Clear all selections



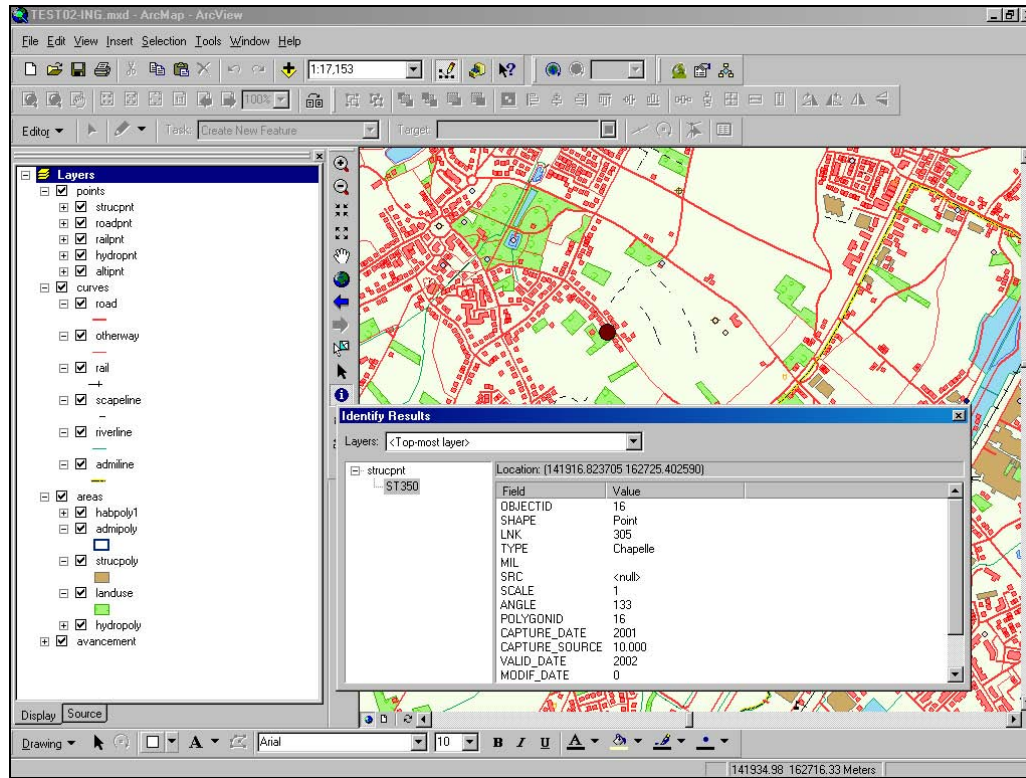
3.3 VOCAL SCRIPT EXECUTION

For instance copy an item to another layer because the affection has changed.

- Copy this factory as apartment building
- Copy this chapel as church

An new item is created with the new attributes

The old item is modified so that history could be followed up.



4. Conclusions

Voice technology allows to give a new mobile dimension and efficiency to GIS and FM applications, in every specific market segment or vertical niche.

The drivers are :

1. Mobility – delivering hands free modus
2. Speed
3. Natural ease of use by all

One will notice that in the ESRI GIS world almost all products are voice enable capable: from ArcInfo to ArcPad.

Author information :

Ir Charles Kemper, CEO of VOICE-INSIGHT SA/NV
EEBIC Building
40, avenue J. Wybranlaan
B-1070 Brussels (Belgium)

TEL : +32 2 529 5 812

FAX : +32 2 529 5 983

Email : charles.kemper@voice-insight.com