

Lingua. Mobila



A mobile advertising platform, delivering synthesized-voice multilingual phrasal translations to GSM users

GSM users submit free-form phrasal translation requests. These are processed by a multilingual translation server and speech synthesis modules. The waveforms are enveloped in MMS messages. Requests are parsed for geo-location and content; appropriate visual advertising is included in the MMS. This is a mobile 'talking' multilingual phrasal translation gadget, which can double as an advertising medium globally.

by Yiannis C Hatzopoulos
Electrical Engineer
3 Ipsilantou Str. SINDOS GR 574 00
GREECE
Tel +30-2310-798432
y.hatzopoulos@mycosmos.gr

Prototype description
Dec. 2005 Edition – Draft v1.2

Abstract:

LinguaMobila is a Java SIM-card (& J2ME) based advertising platform, which offers multilingual translation services to roaming cell-phone users.

Mobile users submit their translation request in a multitude of available source languages. The translation requests are parsed by a SYSTRAN™ machine translation server; the resulting translated content is processed by multilingual speech synthesis modules.

The produced sound waveforms are enveloped in MMS messages, which are sent back to the user's cell-phone. The translation request is parsed for content and user location: a corresponding visual advertising message is wrapped in GIF format, embedded in the response MMS. This mobile application is a 'talking' multilingual phrasal translation server, which doubles as a value oriented advertising medium, aimed at roaming cell-phone users.

So you've just arrived in Beijing for the 2008 Olympics, got into a taxi cab and need to go the wrestling venue fast. Unfortunately the taxi driver doesn't speak very good English, nor can you Chinese! You look in your phrase book for related phrases; after several pages of browsing, you point under his nose something like: 'I want to go there' - which reads 我想要去那里我想要去 in simplified Chinese, but how do you say 'wrestling venue' in Chinese? It is not in the book! You probably end up drawing a wrestler on a piece of paper to make the man understand...but you are not Picasso either!

This is a typical scenario of what LinguaMobila is focusing at. International travel had always involved difficulties in communicating with the locals; especially in world regions where there is limited knowledge and use of international languages.

Business travelers and tourists often resort to buying phrase books and translation PDA's. Phrase books have been used for many decades. They contain lists of commonly used phrases, literals, common names, in the source and target languages involved. You look up what you want – or at least what best matches your request, and either point the written translation to the foreign listener to read, or try to pronounce it yourself; which almost always sounds not only ridiculous but can also cause a severe tongue twisting!

Phrase books rarely cover more than 1500 common phrases, for the simple reason that they must be compact, light and cheap.

PDAs are the latest fad among international travelers. They function more or less like electronic phrase-books; they contain greater vocabulary collections and some even offer trainable speech recognition engines. You can say something in the source language, hope that your miraculous contraption understands it, and voices over it correctly into the target language. The US army has even issued similar devices to its troops in Iraq, containing common interrogation phrases from English into the Iraqi

language. Still, these devices are quite expensive, offer a limited phrasal content and language coverage sets.

There hasn't been much consideration for budget travelers that don't have a phrase book when they need to communicate in a foreign language; and don't want to spend big money on the latest PDA translation technology.

Wouldn't it be wonderful to get their ubiquitous mobile phones to do the translation work for them in any country they go to? Mass usage means a mass-market and that leads to advertising potential.

It would be a very interesting to call upon advertising revenue to subsidize the service cost. MasterCard and Visa, CocaCola, McDonalds, VirginAtlantic and Nestle, P&G and KFC might be interested to sponsor a good quality mobile phone translation engine. As long as it is compact and can work over the majority of mobile phones; and provided users can download it the moment they land in an international airport. This is all about utilitarian advertising; offer something useful in return for people's attention, instead of brainwashing their heads! It might make perfect sense.

The answer to this wish-list has been the development of the LinguaMobila prototype.

It is a thin-client application (around 3-4 Kbytes), which can offer translation services to mobile users, either roaming or in the native network. We have developed 2 versions, one running on Java SIM cards (for older phones) and one in J2ME, for newer ones.

The application can either be downloaded off an OTA server, or be pre-installed on a SIM card as part of a promotional value pack for international travelers.

There are additional features that complement this business model: Since LinguaMobila is operated over a wireless cellular network, the operator can identify the approximate user's location. Combine this knowledge with the actual translation request and you can reach a captive audience, ready to buy a service (e.g. people seeking food / retail outlets, money exchanges, banks, services, ATMs, travel, accommodation etc). You just have to 'push' the right advertisement content!

This feature could be attractive to companies with global commercial presence (soft-drink manufacturers, credit card companies, fast-food chains, hotel chains, money exchange bureaus, telecom operators offering callback schemes etc.) The whole application is multilingual by design – and can be accessible through any roaming GSM network in the world.

What is the most effective form of advertising? There one simple answer: Multimedia, and thus the chosen delivery vehicle of this service has been MMS.

The response MMS generated by LinguaMobila contains the requested translation playback in Voice content (actually in AMR format), and a video or static images which are controlled by the application's sponsor. MMS can be stored and played back as many times as ones wishes on his cellphone.

In our application, a challenging problem has been the way the mobile user submits his free-form phrase translation request. We have implemented 3 ways of communicating one's request to the system.

1. SMS-alike text edit (with T9 dictionary support)
2. Text-tree selection and editing.
3. IVR supporting T9 dictionaries (using key clicks + voice recognition – currently in English only using MultiTel's speech recognition engine)

The first choice is quite simple. The user selects the language combination he is interested in (e.g English to Greek), writes his phrase similar to writing an SMS (e.g I want to buy a train ticket to Paris). Presses the 'translate' button on his phone and an MMS is received by his Cellphone a few seconds later, which when played beams out the phrase 'Je voudrais acheter un billet du train pour Paris', audible by anyone within a 2 meter range.

The MMS generated contains sponsorship info usually as a banner or short video.

To speed up things a little, we have implemented a tree selection form which the mobile user can use to formulate the skeleton of his text request, which can be fine-tuned by further editing.

Category->want Buy /want Sell/ want Go/...

Objects->Tickets/ food/ stamps/ money/ clothes/...

Using a couple of menu selections and with a few clicks you have 'I want to buy a Ticket'. You just have to add by typing on the cellphone's keyboard 'to Paris' and the request phrase is ready to be submitted for translation. The selection content lists are maintained in the Server section of the LinguaMobila application and are downloaded dynamically on the air.

The most challenging and value added communication medium is our server based, T9 IVR dictionaries, a technology that is being developed by MultiTel – a Belgian research institute. This server based technology, uses key-clicks combined with voice recognition engines to propose words to the mobile user, speeding up the writing process on the Cellphone. The integration to LinguaMobila is in alpha testing.

Usage instructions to the system are also provided on demand through an automated multilingual voice portal or in text form. In the pages that follow, we will get a bit

more technical and describe the internal system architecture, which allows us to deliver the LinguaMobila services.

System architecture

The LinguaMobila is a mobile Client / Server system. On the client side resides a Java applet inside the SIM card (also a J2ME midlet version is available) –a very thin client actually –which has been appropriately named LinguaMobilaClient. This is the application's cell-phone based front end.

We will initiate the description on the prototype LinguaMobila SIM card Java application, developed using the Axalto Views package and USIMERA SIM card.

This applet offers basic menu selections, such as connecting to a LinguaMobila Voice portal, source and target language selection, it receives text input from the user, and submits the translation request and user location to the LinguaMobila OTA server.

In the demo applet that prototypes this submission, we supported four languages (English / French / German / Greek), which more-or-less reflect the language needs inside our local Greek market. The system's design is multilingual and can actually support any language combination which is backed by a commercially available server translation engine and supporting speech synthesis modules (Chinese included!)

A straight-forward menu selection provides the capability of connecting to a LinguaMobila voice portal through a GSM Setup Call proactive command. This telephone number can offer toll-free instructions on the usage of the system, and serves as an IVR front-end to voice recognition services that get linked to the system.

In a typical use scenario, the user is prompted to select the language translation set he is interested in (e.g. English into Greek). The user is requested to enter his source text using a Get Input proactive command on the SIM card. T9 text edit dictionaries are used to speedup the data-entry process on the cell-phone. The user menu interface experience on LinguaMobila is quite similar to creating a normal SMS.

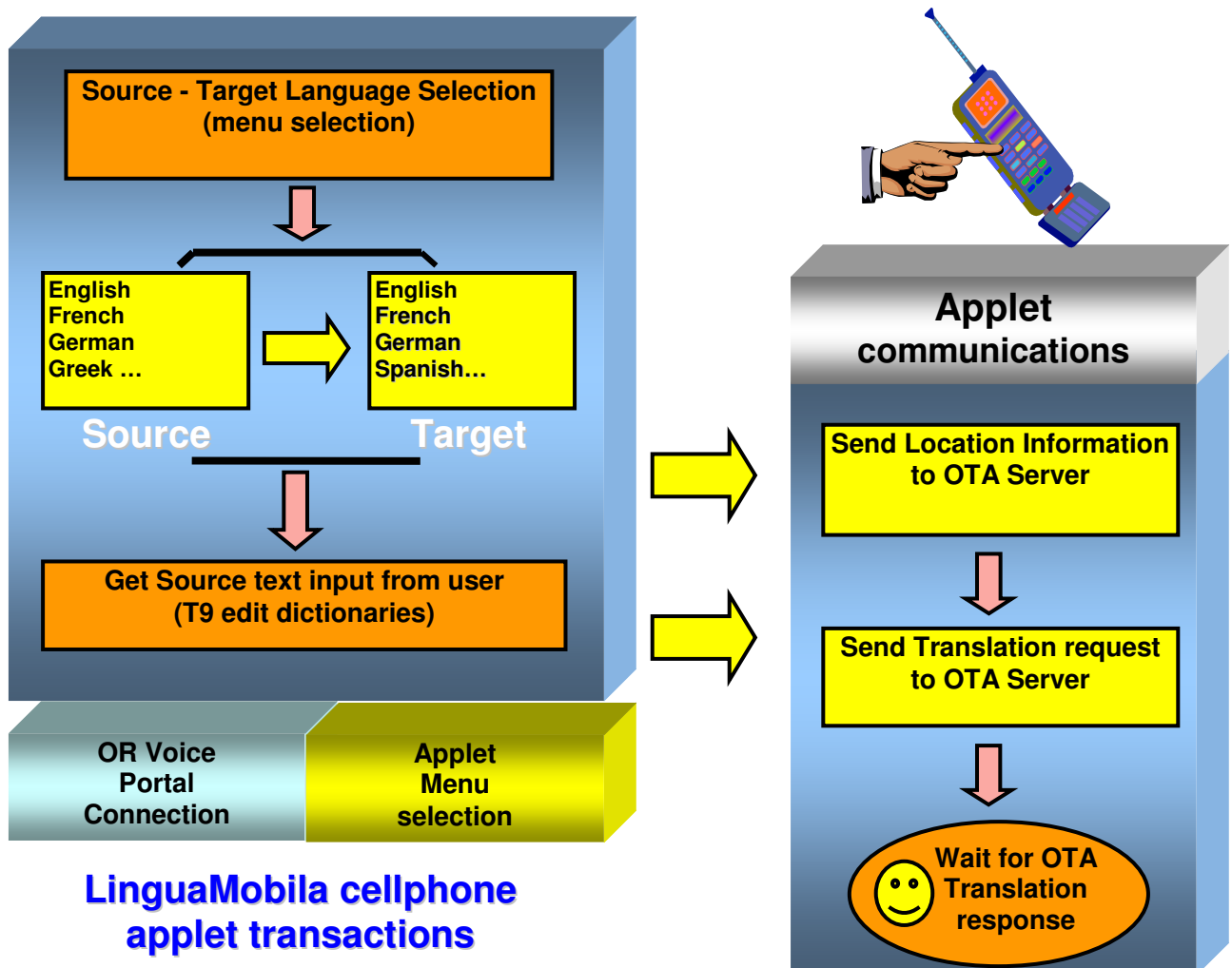
When all the above selections are made, the applet forwards a SMS to the LinguaMobila OTA server marked with a special LOC tag (locator tag), wrapping the cellphone's geographic location data. The OTA server registers this information on a location transaction database. It will use this data at a later moment to match location and content with advertising data broadcasts.

Next, the applet sends a second SMS to the OTA server, registering the user's language set selections and the actual translation source text. Finally, the applet enters a stand-by mode, waiting for the OTA server's translation response.

It is obvious that 99% of the work involving the LinguaMobila application is server based. Broadly speaking, an OTA server is a tool for receiving SMS messages from the cell-phone and sending response SMS messages back. We have used this communication channel to receive service requests from the mobile phone, connect to web-service gateways and establish socket connections / spawn requests to other PC applications in our network.

With this information we build the response content and send the necessary service status or information to the cell-phone, either with SMS or MMS.

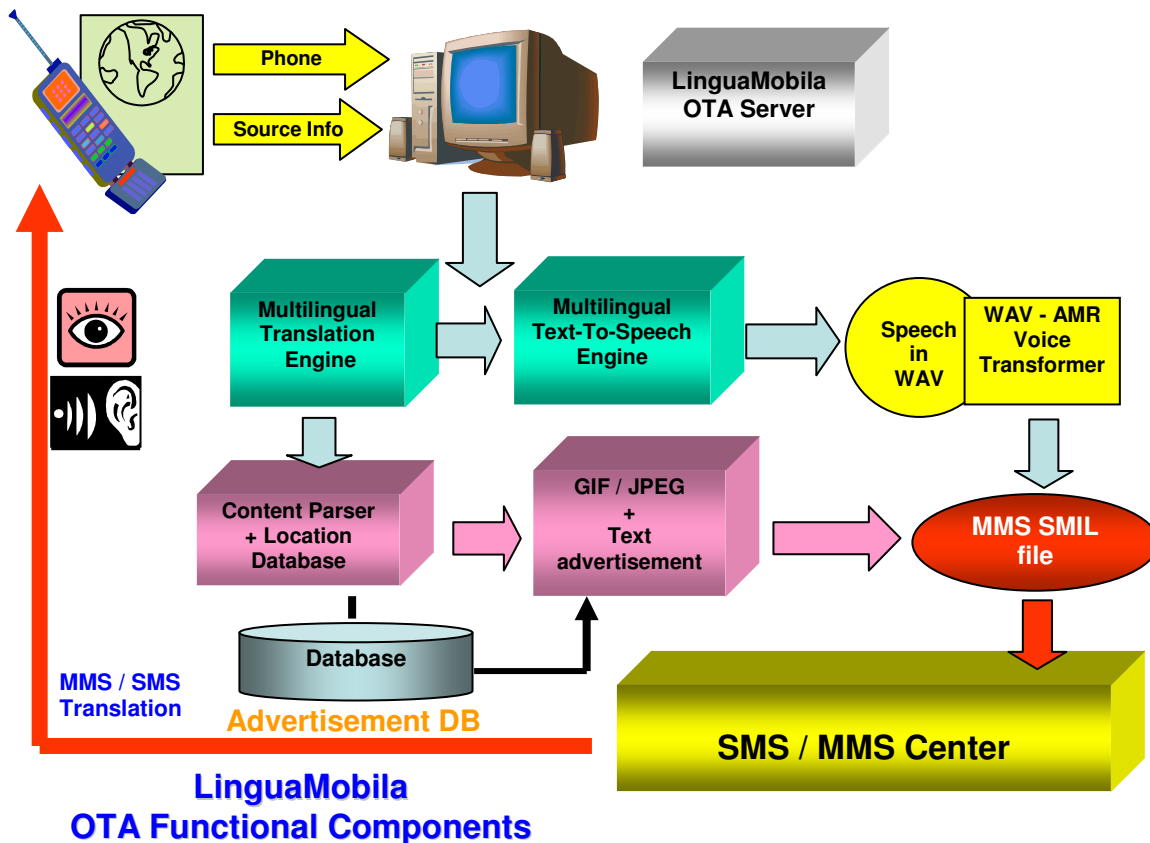
When the translation request reaches the OTA server, a connection to a SYSTRAN™ backed translation server - is performed to translate the source content, through a web portal interface.



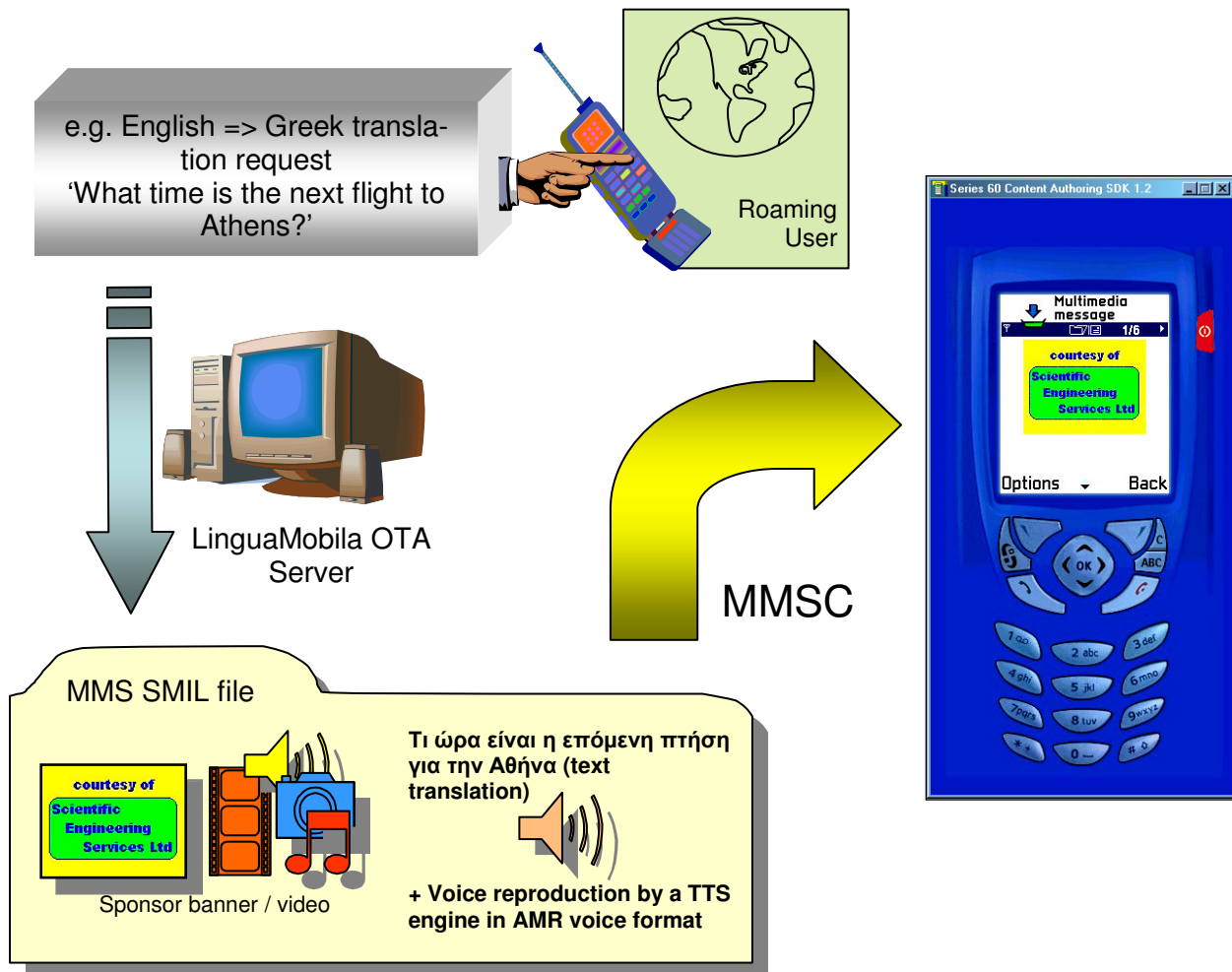
Systran returns the translation, embedded in html code, out of which we extract the useful translation information by an appropriate tag parser and filter. When the OTA server obtains the requested translation, it needs to get it read-out by a speech synthesis engine. In our prototype project, we have used the Loquendo TTS engine,

which works in 16 target languages. The Loquendo TTS engine produces WAV files containing the spoken-out translated text.

Wav files are too bulky to envelope directly in a MMS. We had to use a converter from WAV into the AMR format, which is optimized to transfer voice quality sound at an excellent compression to mobile-phones; AMR is the format by which voice recordings can be included in MMS.



By that moment we have all the basic elements we need to respond to the user's translation request: The source and target text, the user's location, his cell-phone number (we receive this as part of the SMS exchange between server – cell-phone), and the final translation in audio AMR format. We only need to append the proper advertisement in this mix to make the content deliverable to the user.



The response MMS built, contains the appropriate translation audio, plus sponsoring graphic and video/ text information.

Forming a MMS necessitates the buildup of a SMIL file, actually an XML file, which designates how the MMS message constituents (audio, voice, graphics, video and text) will be projected on the mobile screen, in terms of duration, transition, format, sound-video synchronization etc.



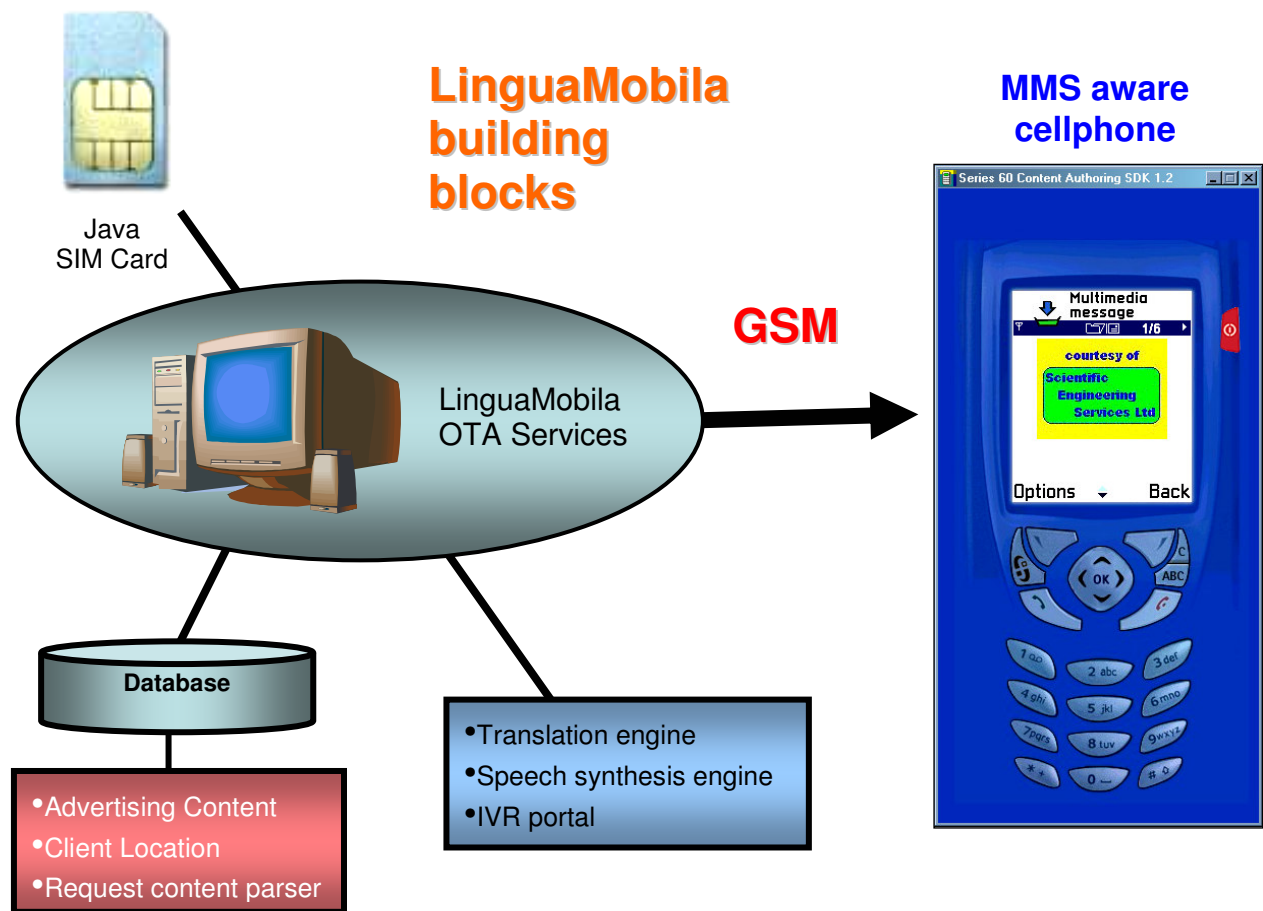
(LinguaMobila J2ME midlet prototype)

The LinguaMobila server builds an appropriate SMIL file and uses a MMS center (MMSC) to forward the compiled MMS to the user's cellphone, through a web interface. We should note that the selection of the sponsoring / advertisement content is done through access to the LinguaMobila client location database, a message content parser, and an advertisement database; all of which are matched against each other during the buildup of the MMS SMIL file.

A 10sec voice translation, attached with a static graphic banner and a few lines of text can be wrapped in around 10kbytes of MMS message, which is fast to download to a cell-phone and cheap to broadcast.

The J2ME version of this LinguaMobila application is more powerful than the SIM-Card based. J2ME offers a better graphical user interface, and direct interaction of the Cell-phone with a web server.

In other words the transaction request from the cellphone to the OTA server is done through http POST operations and not with background SMS exchanges. The web server part of the J2ME application has been written in Java and runs under the Jakarta Tomcat web-server.



The functional concept is similar but the data entry process is more powerful. The mobile user can submit his request through a text interface, speed things up combining semi formed phrases, or use a vocalized T9 server based function. The end result is always an MMS that contains the vocalized translated request for auditioning.

Business aspects

LinguaMobila is a modular application base. All necessary components are connected dynamically over the internet. The translation server can offer services in the States, while the speech generators can be working in Italy or France. The MMSC may be located in Germany.

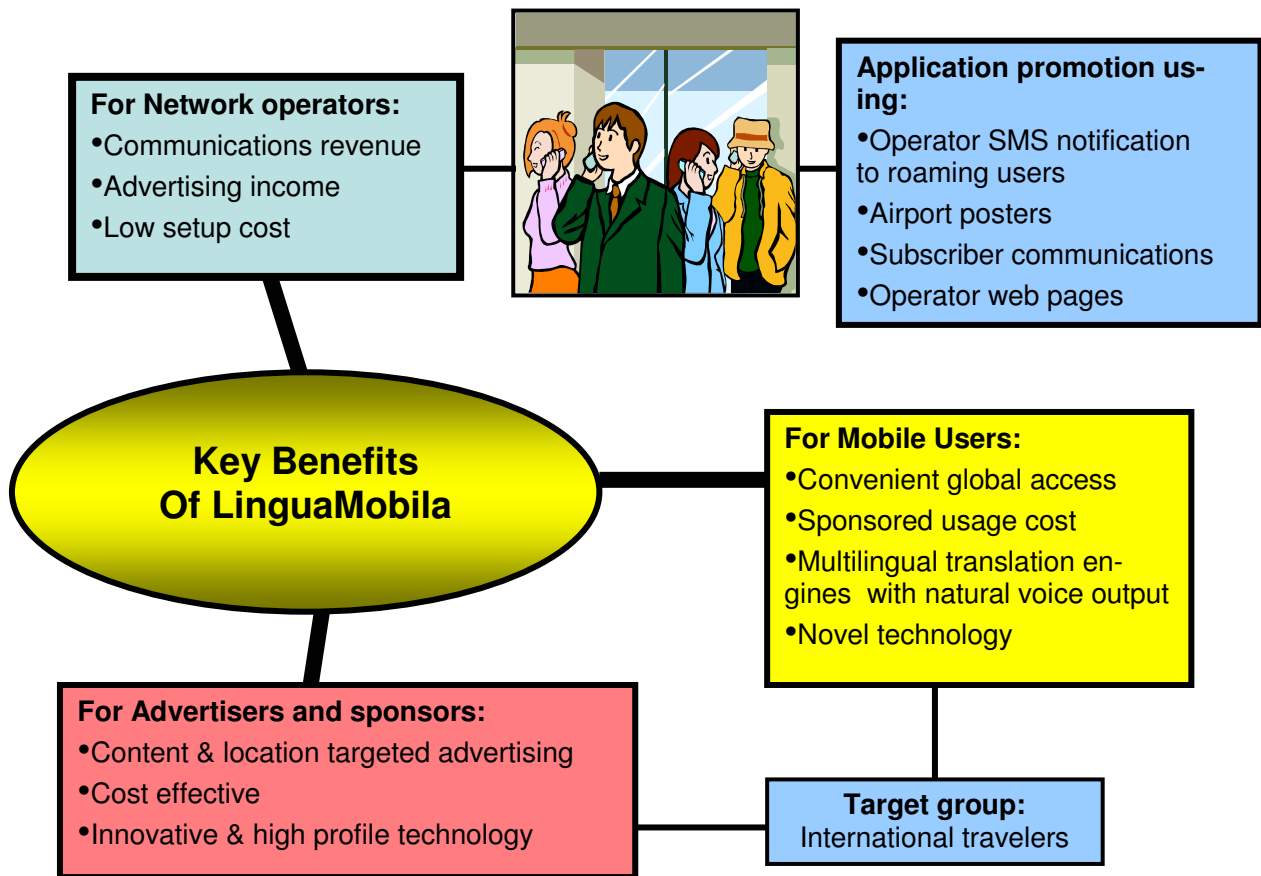
One can connect many language modules to the system, integrating them over a web service transparently. Most of the components involved can be either licensed or charged on a pay-as-you-use principle.



(MMS Message as received by the LinguaMobila MMSC)

Systran provides an industrial strength multilingual translation web portal which is charged around 12,500 USD/year for 2,500,000 translations. Less is necessary to license 'thinner' server solutions.

Top quality multilingual translation engines are also available from vendors such as IBM and AmiKai. Commercial multilingual Text-To-Speech server platforms are available from many firms (such as Loquendo and Acapela) at a fraction of this cost. Open source TTS server engines such as MBROLA are also available.



MMSC and OTA server software should already be part of any wireless network provider that would choose to host such a mobile translation service. In other words, the application's footprint is quite small and can be deployed with ease.

Promotion of this application might simply involve an automated SMS broadcast by the roaming network, delivered to the roaming user's cell-phone on his first connection; this would most probably be sufficient to make the user aware of the services offered.

Poster exposure at international airports, tourist shopping malls, subscriber mail / catalogue communications or web presence could also be utilized; all of which are routinely practiced by mobile network operators in their respective domestic market.

LinguaMobila could be a useful service in mobile networks that host roaming tourists and business travelers; and could also prove to be a useful and innovative advertising vehicle for commercial actors involved in international travel services.

The network operators can benefit both from communication traffic and advertising fees. The users can benefit by the subsidized access cost, which is introduced by the generated advertising potential.

The application works in approximately 30 language combination sets, which cover all the major language groups of the world; thus can be deployed globally and can function from a centralized server.

Credit must go to people at Axalto, Sun and Samsung for providing me with the tools and infrastructure necessary to materialize my ideas. I hope you enjoyed reading about the LinguaMobila project just as much as I enjoyed developing it!



Yiannis C Hatzopoulos . _

- The SIM card version of LinguaMobila has been developed using VIEWS by Axalto; a subsidiary of SchlumbergerSema. The prototyping has been done on USIMERA SIM Cards by the same manufacturer.





VIEWS Professional package

Includes :

A PC/SC card reader.

The VIEWS Professional software.

2 Simera cards and a plug-in adaptor.

From the Axalto web site...

- VIEWS Professional is a complete environment for developing end-to-end interoperable Java applications. It enables to develop efficient and fully tested Java applets, ready for deployment on Axalto Simera, USIMERA, Airflex and other Javacompliant SIMs
- VIEWS streamlines the development and testing of new interoperable mobile services, end-to-end. Cutting time to handset - not just time to prototype. At its heart is a unique simulation environment, spanning the SIM, USIM or R-UIM card, mobile phone and OTA server, which allows new applets to be tested on a range of smart cards from different manufacturers

Automatic code generation

A friendly GUI interface means creative people and Java developers new to SIM ToolKit can design their own services, fast and efficiently. VIEWS Professional automatically generates the code to create a fully interoperable Java applet, dramatically shortening design to prototype timescales.

Full simulation and debugging

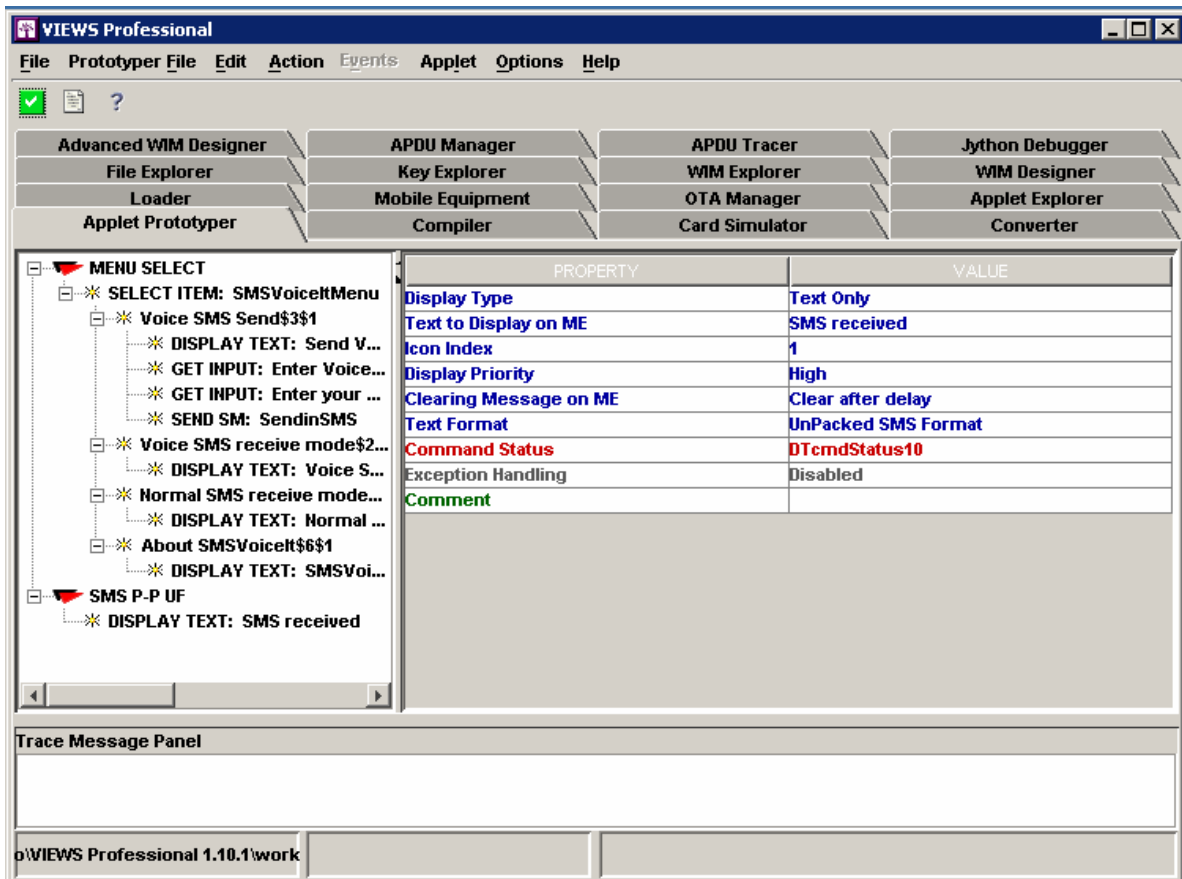
Using VIEWS Professional's simulation environment, and the library of server applications, the new service is loaded, tested and - if required - modified. When the applet? working in the simulation environment, on the required range of Java SIMs, it's ready for deployment.

In-house control

By bringing Java application in-house, VIEWS Professional allows sensitive plans for new services to be kept in-house, and communication between creative and technical personnel streamlined. And because VIEWS Professional is designed specifically for the open platform world, operators have the power and flexibility to test new interoperable services on any and all of the Java-compliant SIM cards subscribers are using in the field, regardless of manufacturer.

Mastering and managing the mobile phone smart card

It provides a powerful, performant and highly versatile environment for viewing, updating, testing, and debugging the contents of any mobile phone smart card, from any manufacturer - that's all SIMs and USIMs, and roaming cards including our own Simera Airflex and Simera GAIT.



The J2ME version has been compiled with Sun Wireless Toolkit, NetBeans 4.1 and Jakarta Tomcat web server. The viewer for the MMS emulated content has been Nokia's Series 60 SDK , Nokia's Mobile Internet Toolkit and Nokia MMS composer.