

# TRANS-VERBAL-ATOR

Group 10

October 14, 2003

Professor Man \_\_\_\_\_

Neesha Blake \_\_\_\_\_

Roy Guzman \_\_\_\_\_

Pearl Odu \_\_\_\_\_

Cecilia Ogude \_\_\_\_\_

Joyce Reyes \_\_\_\_\_

I pledge my honor that I have abided by Stevens Honor Code.

## **Table of Content**

I. Abstract.....	2
II. Project Proposal Plan.....	3
II.1 Introduction.....	3
II.2 Design Requirements.....	4
II.3 Design Approaches.....	6
II.4 Financial Budget.....	11
II.5 Project Schedule.....	12
III. Conclusion.....	13
IV. References.....	14
V. Appendix.....	15

# **I. Abstract**

## II.1 Introduction

Communication is and has always been an important aspect of life. With the growing technological advances and the growing need for these technologies, communication has still been able to withstand the hands of time. We are now more than ever able to communicate with people over incredible distances and in instantaneous periods of time. Considering how one can communicate as far and as fast as possible, why not make it possible to communicate to whomever possible. That is, make it possible for an English speaking American to communicate with a Spanish speaking person from his or her respective country overcoming the boundaries of space, time and now, language. The Trans-verbal-ator will do exactly this.

The Trans-verbal-ator; a sophisticated compilation of different software put together to create a device that makes it possible for two people of different tongues to communicate with each other both visually and audibly while comfortably speaking in their respective languages.

Group 10's tasks is to create this device utilizing the theoretical skills we have learnt from the past years as students of Stevens Institute of Technology and New York University as well as the practical skills we gathered from experience in our college careers. The group has to accomplish this task ensuring that we are within a prescribed budget and an allotted time frame.

In this proposal we hope to deliver a basic idea of the technical actions we suggest are relevant and intend to implement in the fabrication of our device as well as the time and money it would take for such a device to be fabricated.

## II.2 Design Requirements for the Trans-verbal-ator

Our project, the trans-verbal-ator, will consist of many components that will come together to create an application that will eliminate language barriers over any network. Like every application there will be some software aspects as well as the hardware components. First and foremost, our project will need to convert anything the users say into text that the software can readily translate. The software must be able to recognize the word(s) being said by the user and convert it to text, which is easily translated into the text version of the other language. The group has seen a product on the market that can already accomplish this. It is called dragon naturally speaking 7.0. Fortunately, there are two versions for English and for Spanish. This package brings a headset along with the hardware that gives this product its voice recognition capabilities and the program that shows the text output of the voice input. This microphone would obviously have to be a very sophisticated one since voice recognition depends a lot on accuracy.

Secondly, our Software will also have to have a text to audio component. For this part no hardware will be necessary, since every computer already has this capability of producing sound. It will all just require some programming or using any software out in the market that can already do this. The group came across a product called textAloud. This product actually converts text according to pronunciation of the text. Our application should then be able to convert this text into the other language. This brings us to the actual translating part of the project. There are readily available programs that can translate from English to Spanish and vice versa, but they are not 100% accurate. Grammar can become an issue since every language follows different grammar rules.

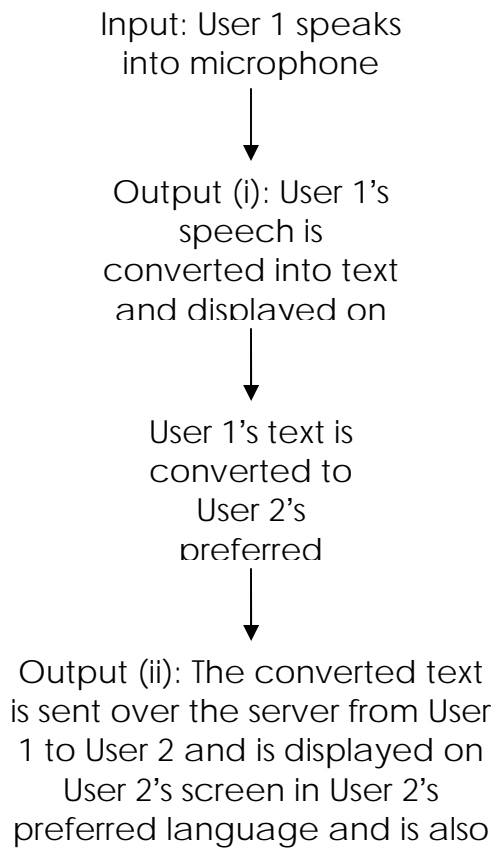
This can be corrected by grammar and spelling checking programs. These two applications can come together to become a very good means of communicating through a network even though there is a language barrier.

The next step would be to create a user interface such as a chat room or an instant messaging server to connect two or more users. Of course, each user would have to have the software installed to have audio and text capabilities of the program. One can see an example of what this interface may look like and how the output and inputs of this program may work from figure A1 in the appendix. The chat server would be created using MySQL / Java. For now these programming tools seem to be ideal for our project. Many of the group members are also familiar with these programming tools, which gives us a base to begin working up from. A java interface will be used to administer an SQL database. Queries and other commands will be used to retrieve the information that the user requests. This provides the group with many options as a Java applet can consist of many features and appearances that can appeal to the users.

Lastly, a server will be needed where all the information will be stored and retrieved from. For now this will consist of a computer where our users will be represented by 2 computers connected by a network. The only requirements are that both computers have the software installed and that they are somehow connected to each other through a network. The Tans-verbal-ator has endless room for upgrades as other components can be added. For example, a video feature can be implemented to give the program a more personal feel.

### **II.3 Approaching the design of the Trans-verbal-ator**

As has been explained in previous sections of this proposal, in order for the Trans-verbal-ator to perform its prescribed task it will have to have the ability to convert speech-to-text, then convert that text to the appropriate language representation, and then send that through a client-server connection to another computer and then perform a final text-to-speech conversion. The diagram below describes the procedures the Trans-verbal-ator should take for one cycle of operation.



The same process described above is repeated for User 2 with User 2's voice intake as the input and User 1 now on the receiving end of the interaction.

In order for the events shown above to occur properly and efficiently certain specifications have to be met by the Trans-verbal-ator. It would have to perform accurate speech-to-text and text-to-speech conversions, it would have to be able to translate accurately from one language to another without losing most of the "gist" of the message and it would have to be able to send this information promptly across a server to another user.

Accurate Speech-To-Text and Text-To-Speech conversion:

Speech to text has been an on-going endeavor for engineering and computer scientists and till today has not been perfected. For accurate speech-to-text conversion the minute changes of decibels in one's voice from one person to another would have to be accounted for as well as the numerous languages one can choose to speak in. Different people with their respective accents and respective tones all trying to say the same word all have the possibility of saying that one word differently. This is the voice recognition area of speech-to-text which seems to be the largest hurdle to leap over. After voice-recognition is performed what remains is matching the "voice-recognized" word to its equivalent in text format. Another tedious process considering there are over 100,000 words that need to be identified, mapped and displayed and depending on which language you choose to use this number increases. Considering all these obstacles in the way of one performing precise speech-to-text recognition and the amount of time it has taken a lot of engineers and scientists to come up with what we have out now in the market, we decided to use an already existing speech-to-text converting software and then try from there to figure out ways to manipulate it for our own device. In this way we are able to provide with the Trans-verbal-ator a device that has the most accurate as-of-the-moment voice recognition capabilities. As this goes for speech-to-text the same goes for text-to-speech. The numerous amounts of words and the time it would take to recognize each word and its corresponding pronunciation in its corresponding language as well as the engineering time that would be necessary would be far longer than the time we have available for the fabrication of the device thus, like before we have decided to use a speech-to-text software which like the former are available in different languages.

In terms of languages for the moment we are concentrating on the trans-verbal-ator being able to work between two languages; English and Spanish implying that we would have need for two speech-to-text software applications and two text-to-speech software applications (each one in English and Spanish).

Accurate language translation:

Considering that we wish for users of two different languages to communicate, it is very important that they understand each other. That is to say that during an interaction with the device, we wish that important information from one user is not lost before it reaches the other user. In order to achieve this there has to be close to accurate language translation between the two languages. For this purpose because we are working with devices that perform rather precise recognition of Speech-to-text thereby implying a rather large amount of words available for translation we decided to use a translation software which in turn would have a word bank close to or as much as that of the voice recognition device. Using this method we decrease the time it would take for us to gather, store and match all these words as well as possibly the inaccurateness and errors of us trying to perform the grammatical aspect of the conversion ourselves. The software would also have to be able to allow for minor manipulations since we intend to implement it in the Java code. The entire process of the language translation part of the trans-verbal-ator's operation would have to occur at the same time as that of sending the information over the server. This would mean it would have to occur almost instantaneously and in the background of the trans-verbal-ator's operation. Giving us

another reason for the necessity of a good software application as well as it's ability to be manipulated to work in different environments.

Client-Server chat capability:

We wish for the Trans-verbal-ator to be able to allow for two users to communicate with each other through a server. We intend to setup a server that would allow for clients (the Users) to connect themselves to and in turn allow for the clients to connect with each other. There are different kinds of programming methods that are available for us to use in order to achieve the goal set above. We intend to use Java programming with the SQL API (Application Programming Interface). The basic idea behind programming Java with the SQL API for chat rooms is that it enables one to start a connection and then allows a user to store their information using databases on the server and then sends this information to the user who requests it or who is open to a request for information. Using this method we should be able to manipulate the information received from the user before it is sent to another user (since we intend to translate whatever User 1 said in one language into User 2's language).

In terms of manipulating the different aspects of the Trans-verbal-ator; Speech-to-text, text-to-speech and language translation so that they can work with the client-server and allow two users to send and receive information we intend to let each process occur step by step. That is to say, we wish that once the information is converted into text from speech it can then be converted into the respective language, then that information can be sent to the database which is where we use the text-to-speech software and at the same

time send the information over to the other user so they can view what the other user wrote but now in the language they prefer.

The bulk of the project is concentrated in creating algorithms and code so that we can successfully link all aspects of the Trans-verbal-ator concisely and efficiently so in the end the Trans-verbal-ator, as described in the introduction will perform to the best of its ability its prescribed tasks.

## II.4 Financial Budget

We have provided a table showing our estimated costs for the 28 weeks that we are spending on the project with 5 team members. On average, each team member spends 4 hours each week on the project. The table shows four estimated costs for direct labor, materials and parts, software, test equipments, documentation costs, miscellaneous expenses like phone charges to vendors, and traveling expenses. These estimated costs will eventually be compared with the costs incurred as refined in the final design report and finally evaluated in the final report in the second semester.

Components	Description	Quantity	Price	Total
<b>Direct Labor</b>				
	Estimated 560 hours	5	\$19/hr.	\$10,640
<b>Materials/Parts</b>				
	Compaq Armada E500	2	\$1,200	\$2,400
	Audio Technical Unidirectional Voice/Entrustment Microphones	2	\$19.99	\$38.98
	Cyber Acoustic 3-Piece Flat-Panel Speaker System	2	\$19.99	\$38.98
<b>Software</b>				
	Java Tools for Extreme Programming	1	\$27.99	\$27.99
	Dragon NaturallySpeaking 7 Preferred Upgrade	2	\$199.99	\$399.98
	SQL Database	1	\$0.00	\$0.00
	Text Aloud	2	\$49.95	\$99.90
	Abode Photoshop	1	\$649.00	\$649.00
<b>Test Equipment</b>				
	Lab supplied by school		0	0
<b>Documentation Costs</b>				
	Printing Documents	100	\$0.06/page	\$6.00
	Printing Advertisements	50	\$1.00/page	\$50.00
<b>Miscellaneous Expenses</b>				
	Phone charges to vendors		\$20	\$20.00
<b>Traveling Costs</b>				
	Traveling to and from Hoboken and NYC		\$15/wk	\$420
<b>Total</b>				\$14,791



## **II.5 Project Schedule**

### **III. Conclusion**

Finally, after extensive research on the different components that make up the Trans-verbal-ator and serious considerations on how to manipulate these components so they work together to perform the actions required by the Trans-verbal-ator, we believe that our ideas will be realized and our goals achieved.

If the assumptions, ideas and thoughts we have outlined here in this proposal from speech recognition to text displaying to language translation and finally to audible sound hold true, the Trans-verbal-ator will follow all these ideal steps utilizing JAVA, SQL, Networking, Voice recognition, text-to-speech and all other programming tools necessary so as to create the desired result of the Trans-verbal-ator.

This project is not only feasible but will also help to revolutionize communication in the social level, cooperate level as well as the government sector. The Trans-verbal-ator will surely be a wanted commodity.

## IV. References

1. Circuit City <<http://www.circuitcity.com/init.jsp>>
2. Dragon NaturallySpeaking 7 <http://www.scansoft.com/naturallyspeaking/>>
3. Electrical and Computer Engineering Senior Design Website <http://koala.ece.stevens-tech.edu/sd/>
4. Free Translation <http://ets.freetranslation.com>
5. Google Search <http://www.google.com/>
6. Bloch, Joshua. (2001) *Effective Java Programming Language Guide*  
[Addison-Wesley]
7. MySQL <http://www.mysql.com>
8. Brown, Steven. (1998) *Visual Basic 6 Complete* [Sybex]
9. Textaloud <http://www.textaloud.com>

## V. Appendix

### A.1

