

# FACTFILE: GCE DIGITAL TECHNOLOGY

## UNIT A2 1: INFORMATION SYSTEMS



### Natural Language and Voice Recognition

#### Learning Outcomes

##### Students should be able to:

- describe the technology required for natural language and voice recognition systems.
- evaluate the use of natural language and voice recognition systems.

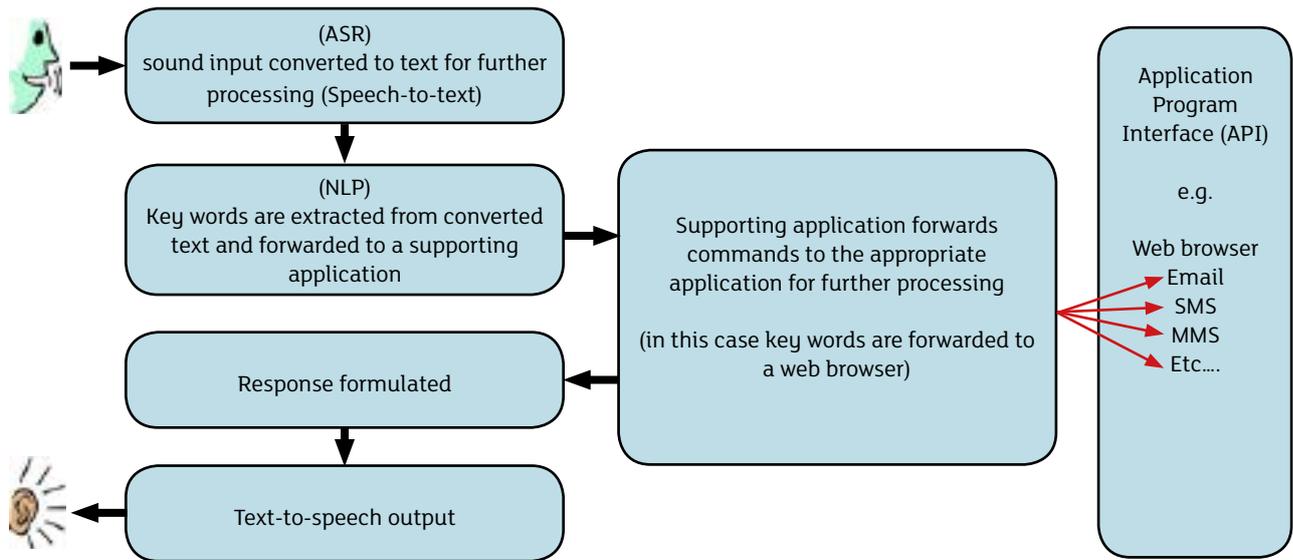
#### Content in Natural Language and Voice Recognition

- Natural Language and speech recognition systems.
- How natural language processing is used today.
- How does natural language processing work?
  - Part-of-Speech-Tagging.
  - Parse trees.
  - Semantics.
- How do voice recognition systems work?
- Evaluating the use of natural language and voice recognition systems.

#### Natural language and voice recognition systems

Natural language processing (NLP) is a branch of AI concerned with the ability of a computer to be able to understand human interaction in its natural spoken or written form. It can be defined as the ability of a machine to analyse, understand and generate human speech. The end goal of NLP is to make the interactions between humans and

computers as 'human like' as possible. Automated speech recognition (ASR) can form a part of the NLP process (although it does not always have to). Some applications (e.g. Siri) will use speech recognition to convert input sound to text before NLP takes place. The diagram shown next illustrates how ASR and NLP are used by Siri to support a user's request for data from the internet: -



### How natural language processing is used today

Some of the more common applications for natural language processing today include:

- **Spam filters** – many organisations such as Gmail use NLP as their first line of defence against spam. Spam filters use NLP to try and extract meaning from strings of text to help identify unwanted email and prevent it from entering clients in box in their email applications.
- **Answering questions** – search engines provide us with a wealth of information but rely on users being very specific with the key words used to support web based searches. Companies such as Google are focussing on the use of NLP to help with the processing of natural language questions posed by the user so that the meaning (key words) can be extracted and the appropriate answers provided (sometimes in natural language format).
- **Extracting information** - many organisations in the financial market are now using *algorithmic trading* as a means of managing investments. Financial investments are controlled primarily by technology which will evaluate news articles and extract relevant information to evaluate stock market patterns before determining if clients should buy, sell or hold onto stocks in their portfolios.
- **Summarising information** – information overload poses a problem for many digital users today, especially via social media

applications where we are constantly being bombarded with information and advertisements from other users. Social media providers such as Facebook use NLP to analyse information on users collected via social media to help determine their preferences and to help determine which articles and advertisements should be presented higher up in their news feed. Some of these companies have been in the news more recently as they have programmed their applications by default to access the user’s microphone on their mobile devices and they are using information collected this way to collect information for analysis.

### How does natural language processing work?

Computers can have trouble with natural language processing because they try to understand the meaning of individual words, rather than the whole sentence or phrase. Attempts at understanding whole sentences and phrases can be made more difficult due to the fact that many words have double meanings (e.g. to “pass” the milk / to “pass” an exam) so applications attempting to understand natural language input into the computer must also have an understanding of the context within which the text is being presented.

To support the understanding of the intricacies of human language those investigating natural language processing must have an understanding of the following terms: -

- **Morphology** – how words are formed and their relationship with other words (it considers for

example how context can change the meaning of a word).

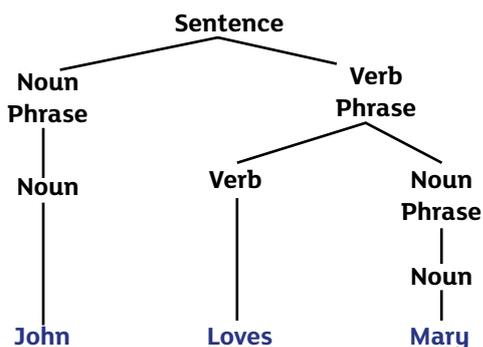
- Syntax – how words and sentence are put together.
- Semantics – the meaning of words and groups of words.
- Pragmatics – the context of spoken expressions.
- Phonology – the sound associated with spoken language (how the words and phrases sound when spoken).

### Part-of-Speech-Tagging (PoS)

The first step in natural language processing involves morphology (defining the functions of individual words; especially is there is ambiguity – e.g. the pass example from before). Modern applications will apply a self-learning algorithm which will tag words with multiple meanings. These applications will first determine the highest occurring meaning for the tagged word and then use this to try and understand the functions / meanings of the other words around it.

### Parse trees / diagrams

The next step in the process is to use knowledge derived from syntax to try to understand the structure of the sentence. The algorithm will repeatedly break the sentence down into noun and verb phrases to further aid the understanding of the sentence. The outcome of the parsing process is the production of a parse tree similar to the one shown below.



### Parse trees / diagrams

The third step in natural language processing considers the semantics of a sentence. If tagging and syntactical analysis determine that the word has the same meaning, then this stage will process the words that appear before and after the tagged word to help apply meaning to the sentence.

For example:

*Cut the bread on the board.*

*He is a member of the board.*

A human reading these two sentence will be able to determine the difference between a chopping board and a board of directors in each scenario but a natural language processing application must 'learn' that if the word 'board' is preceded or followed by the word 'bread' the reference is to a chopping board whilst if it is preceded or followed by 'member' it is more likely to refer to a board of directors.

Natural language processing is a complicated area but there is still room for improvement, especially in the area of **pragmatics**. Most sentences conform to a context that requires a general understanding of the human world and human emotions and this can be difficult to teach a computer especially with reference to things such as sarcasm for example.

### How do voice recognition systems work?

The term voice recognition (VR) refers to the combination of hardware and software systems which have the ability to decode a spoken command. VR is often used to operate devices or execute commands without the need to the use of peripherals such as keyboards, mice or tracker pads for example.

The first stage in voice recognition is the input and digitisation of spoken words into VR software. In order to achieve this a computer/device with a sound card is required, along with microphone or a headset. Some applications will use separate hardware devices to support voice recognition, while others such as smart phones have all the necessary hardware built into the device. Specialised software must also be installed on the device in order to support voice recognition.

### How computers recognise speech

Computers can take one of a number of approaches to speech recognition. These include: -

- Pattern matching where the words spoken by the user are recognised in their entirety. These types of systems are often used by business with automated switchboards. The user will be presented with questions with limited and simplistic responses (e.g. Yes/No). The computer will analyse the input from the user and try to match it with a list of potential sound patterns which represent each of the available answers.
- Pattern and feature analysis. Here the spoken input is recorded by a microphone and then

digitised using an ADC (Analogue to digital converter). This digital data is then analysed and compared to a stored dictionary which can then be used to identify what the user has said. More complex input can be analysed with this method of input and the user is not limited to the responses they are able to make.

- Statistical analysis. More complex systems can take a more statistical approach to the analysis of speech input. These systems can apply the rules of grammar to help predict words to support speech recognition; especially

in instances when the spoken word was not entirely clear.

- Artificial neural networks (ANN) are still being explored as a means to support voice recognition. Scientists are looking at how they can be trained through the use of examples to recognise spoken input. More recent studies are looking at the combination of ANN and statistical analysis to help improve the accuracy of voice recognition applications.

### Evaluating the use of natural language and voice recognition systems

	Advantages	Disadvantages
<b>Natural language</b>	<ul style="list-style-type: none"> <li>• No training is required in the use of the application</li> <li>• Increases accessibility of the application</li> <li>• If combined with speech recognition it can free users hands to complete other tasks</li> </ul>	<ul style="list-style-type: none"> <li>• Not all commands will be recognised / identified correctly, all the time</li> <li>• Applications are complex to create and are therefore expensive</li> </ul>
<b>Voice recognition</b>	<ul style="list-style-type: none"> <li>• No training is required in the use of the application</li> <li>• Accurate spelling of words for users with literacy problems</li> <li>• For those with limited mobility it is a more effective method of input</li> </ul>	<ul style="list-style-type: none"> <li>• Prone to interference</li> <li>• User may be required to speak more slowly and louder than normal</li> <li>• May only recognise a limited range of voices / must be trained specifically to recognise a user's voice</li> </ul>

### Questions

1. Explain the difference between natural language processing systems and voice recognition systems. [2]
2. Describe the following terms associated with natural language processing systems
  - a. Morphology [1]
  - b. Syntax [1]
  - c. Semantics [1]
3. Research the potential uses of NLP and make a list of potential uses. [3]
4. Explain how VR systems such as Siri make use of VR and NLP systems to process user requests. [6]
5. Explain the terms parsing and PoS in the context of NLP. [4]
6. Identify the main hardware elements associated with VR systems. [3]
7. Identify and describe 3 different approaches currently being taken to support voice recognition. [6]
8. List 3 items of hardware used to support voice recognition. [3]
9. Research the use of natural language processing and voice recognition. Produce a list of applications where each is used. Beside each application include a list of advantages and disadvantages of the use of natural language processing and voice recognition for that application.

## Bibliography

BCS Academy Glossary Working Party, 2013, BCS Glossary of Computing and ICT, 13th Edition, Swindon, BCS Learning and Development Ltd

## Links

- <http://blog.neospeech.com/what-is-natural-language-processing/>
- <https://www.lifewire.com/applications-of-natural-language-processing-technology-2495544>
- <https://www.1and1.co.uk/digitalguide/online-marketing/online-sales/how-does-natural-language-processing-work/>
- <https://www.computerhope.com/jargon/v/voicereco.htm>
- <http://www.explainthatstuff.com/voicerecognition.html>

