Language and Culture

[Author Name(s), First M. Last, Omit Titles and Degrees]

[Institutional Affiliation(s)]

Author Note

[Include any grant/funding information and a complete correspondence address.]

Language and Culture

Language is full of ambiguities. There are many different words with different meanings and pronunciations. Technology has provided humans with many gifts that prove to be helpful and useful in their daily lives. The devices known as speech recognizers are becoming common with every passing day. Speech recognition is a process in which spoken words are converted into digital data using human voice as the input. The devices such as IVRs, Siri, and Alexa interpret and recognize the human speech. These devices are helping human beings in many fields of life such as in education, business, call centers, etc. Applications for the disabled are also created where the disabled people can just give a command to the device and control their systems through their voice. They can perform their routine tasks such as writing, operating a computer, reading a book, turning a page, or controlling their home appliances like telephones, intercoms, radios, or curtains. Speech recognizers are currently changing many of our ways such as we shop, we search, we interact with our devices, and even with each other through these. The devices have influenced our lives to a larger extent but unfortunately, these devices are unable to interpret the language components such as morphemes, phonemes, syntax, semantics, pragmatics and discourse. Humans understand speech in a context that they are familiar with, including the components of language.

Speech recognition is not linked to the pronunciation and dictionary meaning only, rather the syntax should also be understood, which the speech recognition devices are unable to grasp. They rely only on the speech commands. Syntax in a speech that provides the rules which shows how sentences are built up using the smaller parts or units of language. The syntax rules are important for the human speech as it helps them understand the exact meaning of the words or sentences spoken. However, the devices are not able to understand this syntax in depth. Identical words tend to create difficulties for the devices in understanding language patterns. If a speech recognizer is asked, “There’s never been an error in this task, has there?” However, the recognizers would not be able to judge if the spoken word is “their” or “there.” The implicit grammatical knowledge that is used to understand the meaning of the words, phrases, and sentences with reference to the grammatical rules is not known to the computer technology. The spontaneous speech effects such as pauses, cross-talks, repetitions, and hesitations are very difficult to be judged or understood by the speech recognizing devices. These devices may also find it difficult to perceive phonemes used in the human language. For example, a device may perceive “I would like the coffee with sugar and dream,” which is actually “I would like the coffee with sugar and cream.” These devices also confuse the homophones that refer to the words with same sounds and different meanings. For example, if a person says’ “I read a book,” the speech recognition device may confuse the read with red that is a color.

Pragmatics in linguistics describe language use with relation to its context. The devices used for speech recognition have come so far in their achievements, still there needs to be a lot of improvement in their working. These devices can understand the language and commands, yet they are unable to understand the meaning in context. For example, if a woman says, “My daughter is a fox” she is probably referring to the qualities of a fox and relating them to those of her daughter. However, the speech recognition devices might not be able to get this context and infer this understanding that if the mother is calling her daughter a fox, may be the daughter is really a fox. Machines are not able to understand inferences in the use of language. Another example is calling an airline where the computer recognizes the human voice but cannot understand the meaning in context. This understanding is the key feature of human language.

Human speech also consists of some fillers such as “hmm,” “err,” “umm,” “ugh,” etc., and only human know the meanings and context of these fillers. A machine misunderstands these fillers and considers them incorrect. It is because our brains are incredible, we use all language components in our simple day to day conversation, while this may be extremely difficult for machines and technology.

There are also semantic errors in the speech recognition devices. Semantics refer to the meaning of any word or sentence in language. For example, if someone says, “let’s meet up Tuesday,” but the speech recognition device misinterprets it as, “let’s meet Tuesday.” Here the semantics of this sentence are not changed but the utterances are altered. Another example of the obstacles faced by speech recognizers in interpreting semantics is the use of words like cell and cell. If someone says, “Where is my cell?” the recognizer does not know whether he is talking about a cell which is a device used for communication or a cell which is referred to a room used to imprison someone. This shows that the technological devices are only understanding the words, rather than their meanings and context. For example, if a girl is asked, “Can you go for a movie tonight?” to which she replies, “I have to study,” means that she cannot join for the movie as she has to study at night. The voice-based devices would not be able to understand this semantic value.

Humans have their experiences in the world that they use in understanding of their language patterns. For example, when the word “mouse” is used by a person, we imagine images, associate symbols, and think of the characteristics related to cats based on our experiences of the world. It is very difficult for the machines to understand and associate this as well as humans can.

Language and human emotions are very broad and complex concepts for which one cannot totally rely on technological devices. Understanding of the human speech and language requires a lot of intelligence and knowledge. Today researchers, graphic artists, musicians, doctors, librarians, and magazine publishers, etc., are using speech recognition devices to digitalize their data and increase their sophisticated knowledge in this matter. Human beings are able to recognize speech in every way. The components of language can be understood by humans only. They are able to interpret and understand the tone, accent, attitude, and other clues that are displayed or spoken by the speaker. The intention, expectation, and meaning in a speech can also be judged or understood easily by another individual but the speech recognition devices are unable to understand these components and clues in a speech. They are trained for a certain understanding of the language patterns, which they are doing perfectly. Humans also have the ability to express and understand the body language of each other that is much helpful in the easy understanding of language. However, this understanding is also not possible for the automatic machines. Technology is progressing day by day and trying to make even those things possible that seemed impossible few years back. Machines are quickly gaining the ability to judge and understand what humans say. Still, there is a long way to go for the speech recognizers to work properly in the world of human beings.

# References

SHULEVITZ, J. (2018, Novenber). *Alexa, Should We Trust You?* Retrieved from: "The Atlantic https://www.theatlantic.com/magazine/archive/2018/11/alexa-how-will-you-change-us/570844/"