

Analysis of Voice Recognition System on Translator for Daily Use

Yudi Aryatama Fonggi^{1*}, Tio Oktavianus²

^{1,2}Computer Science Department, School of Computer Science,
Bina Nusantara University,
Jakarta, Indonesia 11480
fonggi.aryatama@binus.ac.id; tio.oktavianus@binus.ac.id

*Correspondence: fonggi.aryatama@binus.ac.id

Abstract - This research paper's aim is to analyse the usage and the implementation of voice recognition system in translator device by reviewing some of literatures in aspect of linguistic and voice recognition system. This research purpose is to understand the procedure of how the voice recognition system is implemented and what the impact of using the translator device for daily use is in a lot of field. This paper also include a review about what is the importance of multilinguistic skill, what is the problem that we usually have in multilinguistic communication usually in global environment, How the basic system that usually use in voice recognition system so the reader can learn by describing the procedure of voice recognition system, In what way the implementation of voice recognition in translator device by providing some example of the devices, and How is the development of the translation devices and the example of translator device usage in society

Keywords: Linguistic; Multilinguistic; Voice Recognition; Translator; Voice Recognition Translator; Speech Translator

I. INTRODUCTION

With the ever-advancing modern era, definitely a lot of new technologies was found and became the foundation for humankind to keep moving forward, to upgrade themselves. In order to achieve so, they will need a better understanding on those matters. Unfortunately, they will surely face with the language barrier and the huge amount of existing language, so people will be having a hard time exchanging information without the help to do so with precise accuracy. Therefore, it is needed to have a machine that helps with translating the incoming language into the understandable language, so the user can understand the

information better. The other thing that is worth noting is translator. The use of a translator has also been developed and has become one of the areas affected by technological developments which resulting in various forms of translator tools. With the various language used when communicating, a lot of problems arise especially when doing direct communication. The subjects who need the translation process only have a little time in their conversation. The translation system developed by using a computer input system that can be used by all types of circles and can eliminate the time problem that exists, namely using a voice recognition system so that the conversation process between languages can be carried out more efficiently

Linguistics is the scientific study of language. It involves an analysis of language form, language meaning, language in context as well as an analysis of the social, cultural, historical, and political factors that influence the language.

The reason why the studies about linguistics are important in the learning circles is because those languages spoken by each individual definitely affect the other individual on how to perceive the incoming information. Choice of words and how the words are spoken are the main reasons which affect how clear the messages delivered. It has been said that the more fluent a person can say the language, the more understanding can be given when talking said language which will boost the confidence of the person and encourage the person to talk more and to understand more of the language.

II. METHODS

This review has used various kinds of data collection and analysis from various articles related to language science, multilinguistic, voice recognition systems and translators. We used 30 articles for the research, but after a deeper screening we only used 20 articles that really fit and really useful for the topic of this research.

III. RESULT AND DISCUSSION

3.1. Multilinguistic

Since willingness to communicate (WTC) varies from person to person depending on the communicative situation, personality type, and even one's mood, a clearer insight into how much volatility in verbal behaviour is rooted in the individual variable will help in designing an effective educational model in an academic bilingualism setting. Willingness to communicate supposes that language learners are actively looking for speaking opportunities and really communicate, leaving the comfort zone of the classroom [1].

Therefore, it is needed a vessel which can help the user to understand more about other languages plus the willingness to learn the languages. With the mastery of several languages, of course the user will be able to understand the subject of discussion and can express and define what he or she knows about the subject better.

Despite the efforts of teaching staff and the introduction of innovations in the educational space of the universities, language, cultural and psychological barriers impede contacts with native speakers and other bilinguals and remain an unsolvable task. There is a need for purposeful actions on the part of all involved in the organization of the learning setting to analyse the existing practices of creating appropriate conditions for the effective multilingual environment [1].

In the process of learning from other language, it is important to note that it is not just about learning a new language. But more importantly, it is to understand the meaning of the chosen word or letter and how to use it to make an understandable sentence for the known languages.

The educative aspect of multilingual education is not limited to the formation of any one value in one lesson. It is a gradual process with a focus on integrated lessons and the use of non-standard forms and methods of work with students. Otherwise, this process will be ineffective and educational goals will become a pro-forma [2].

In other aspects, being able to speak a lot of language can help in the psychotherapy section. People with mental problems tend to speak more open-mindedly in their home language. With that in mind, it is important to be able to speak a lot of language in order to help those people [3].

3.2. Voice Recognition

Voice recognition or Speech Recognition is a computer analytic system that purposes is to receive a

human voice as an input mechanism to be used as a data for the computer processing and this input can be detect or differentiate from each human voice, voice recognition can be applied in a lot of field as Forensic, Telephony, Access control, voice assistant, etc[4].

From the start of its development voice recognition has been developed in mayor development especially from 1971 that was started by Allen Newel In 1976, the computational power available was only adequate to perform speech recognition on highly constrained tasks with low branching factors (perplexity). Today, we are able to handle nearly unlimited vocabularies with much larger branching factors, and for 40 years we expect speech recognition to become a bridge between human and machine to enhance a conversation among peoples regardless of barriers of location and languages [5].

As an input mechanism voice recognition have a lower impact compared to others input mechanism. The most popular method is by typing method which is more commonly used and usually has a lot more interest to be used for data input purposes, thus causing a significance lower user interest and development for voice recognition. But from the research voice recognition system can be used to help disabilities user or some user that has a significance lower typing speed and give more satisfaction result from thus slow-typing user [6].

The method of a computer to recognize a voice have different approach in each system, as discussed in this article the genetic algorithm is used to recognize a word which is spoken from 25 different user and resulted in 100% of the words [7], Fourier method can be used by passing some input data the computer system to recognize a different voice signal in a numeric pattern and can be used by the computer system to learn and recognize a human voice [8], another method is used in this research which uses magnitude analysis to compute a pattern of signal and can be used as a basic method for a voice recognition system and can be developed with future system or technology to different application [9].

As for nowadays implementation in application we can easily find and use it with a simple effort like for a simple example is used by using python programming language to build a simple Voice recognition application that can be used for computer user, and make a voice recognition system more easily build and developed from each passing year [10].

3.3. Voice Recognition for Translator

From the beginning, the translation process was developed, the methods carried out were divided into 3 types in the aspect of their involvement in computerization, namely machine translation which is a direct translation by computers, computer assistance translation where the user must have knowledge in the initial language and the second language where computers only help the translation process only, and the third is using human labour by translating it manually such as using a dictionary, or writing directly. These methods have their respective obstacles and the main obstacle that is most often found is customary problems

which result in translation cannot be done raw [11]. In the process of technological development, these constraints have begun to be analysed and solutions have been found that can help the translation process become smoother, namely by using Multilingual Neural Machine Translation which can remember and detect the input entered and adjust it based on input from various languages so that over time it can detect and translate Language to another language without any strangeness [12].

The development of Speech Translation and the use of speech translation can still be developed in various aspects and by adding innovations and new methods that can increase speech translation capabilities.

As discussed in this article regarding error detection in the translation process using 2 main schematic divisions of the translation process, namely ASR and MT translation [13]. Examples of other developments can be seen in the form of the test of implementing the new method in this article which aims to translate paralinguistic information using the Semi-Markov hidden Linear Regression method and it is proven that this method is effective for Modelling emphasized speech, making it possible to utilize all acoustic features rather than individual ones [14]. And from end Article End-to-end Automatic Speech translation of audiobooks which discusses the development of the translation model, namely end to end speech to text translation without using source language transcription during learning or during the coding process, and produces a conclusion where in cascading two neural models for ASR and MT gives the best result, end-to-end method that incorporate the source language transcript come close in performance [15].

Another example of the development or variation of the speech translator model is the use of multilingual corpora and machine learning algorithms in the translation process, where current translation requires development in translating long sentences commonly found in general speech, and this problem can be solved by using a long sentence division model that is difficult to translate. into sentences that are easier to translate [16]. This article provides an example of the use of speech to speech translation in the health sector using the S-Minds (Speaking Multilingual Interactive Natural Dialog System) system which is expected in the future to help nurses and medical personnel in their work process, especially in communicating with patients [17].

There are some example application of voice recognition in translation device and their system

In the research about online voice recognition translator the algorithm is describe as. The initial use of a voice recognition translator device is that the user must provide initial input to provide initial voice information, then the user is asked to confirm the information that has been obtained by the device, the computer then translates the information that has been obtained, then matches it with the second user according to the input. Which is done by the second user, then the computer will reconcile the information that has been obtained from the first user and then output the appropriate translation results. In this case,

the input information given for the first time greatly affects the accuracy of the translation process carried out, because it will be used as the basis for further translation results [18].

A portable voice translator in device that easy to carry around for the algorithm, using a voice receiving device as a voice recorder to detect sound as a voice signal then the processor embedded in the device will process the sound and match it with various kinds of voice segments that have been previously inputted to adjust what language is inputted by user, translate the language according to the required voice segment and the output device will output the result of the translation [19].

3.4. The Impact of Translator in Daily Society

The usefulness of modern technologies can lead to different impact on the user depending on how they use the technologies, even the translators

There are some examples about the impact about the use of translator in daily lives of the users:

As the results of the qualitative content analysis of the study suggest, some participating teachers believe that when students' use of mobile devices are more related to classroom activities, the instructors are more likely to encourage the use of mobile devices in the classrooms [1]. This confirms about the importance of using mobile devices on classes if used correctly like helping the user to understand more about a subject with the help of machine translation. Else, it will definitely affect negatively like distracting the user from learning if wrongly used in the class.

Summing up the entire research work done, we can now conclude that this approach of developing such a software would lead to a more advanced and user-friendly system for interaction and communication [20]. From the review of the article above, it was proven that by using a language translator, it will definitely bring positive results. Such as, giving more opportunity for the user to interact more and communicate easier with the help of the device.

IV. CONCLUSION

From the data collected above, it can be concluded that the use of voice recognition system translator devices is commonly used, has a positive impact, and can still be developed in various aspects, the use of voice recognition in multilingual communication really helps the translation process between different languages, especially communication during the same period both online and offline using a portable device. The basic algorithm used in the voice recognition translator is to use a voice recorder as a receiving device for the user's voice input then it will be matched with the language that has been inputted in the system then the voice signal will be translated according to the desired language in the database system or by a second user as the first user partner.

If the translator is used correctly it will bring many positive benefits. For example, the use of a translator can increase the user's confidence to socialize more with other people. Starting from making users understand language better to being able to speak fluently with other people with an understanding of multiple languages. The result of this course of action is that the user will tend to socialize more with language that is easy to understand. But still, that doesn't mean the use of this technology will always run smoothly. If used incorrectly, the translator will become just an ordinary machine or can even be used to do bad things like prank or teaching bad words.

REFERENCES

- [1] L. Studies, A. International, and A. Centre, "The Application of Mobile Devices in the Translation Classroom," *Adv. Lang. Lit. Stud.*, vol. 7, no. 6, 2016, doi: 10.7575/aiac.all.v.7n.6p.237.
- [2] Y. Kharlamova, "Educative Aspect of Multilingual Education," *KnE Soc. Sci.*, vol. 2020, pp. 92–98, 2020, doi: 10.18502/kss.v4i13.7701.
- [3] J. Dewaele, B. Costa, and L. Rolland, "INTERACTIONS AND MEDIATION BETWEEN MULTILINGUAL CLIENTS AND THEIR Sally Rachel Cook | University of London," no. November, pp. 1–4, 2020.
- [4] S. K. Arts, "A Voice Recognition Controlled Web Browser 1," vol. 5, no. 3, pp. 218–220, 2020.
- [5] X. Huang, J. Baker, and R. Reddy, "A historical perspective of speech recognition," *Commun. ACM*, vol. 57, no. 1, pp. 94–103, 2014, doi: 10.1145/2500887.
- [6] M. Q. Khairuzzaman, "No Title血清及尿液特定蛋白检测在糖尿病肾病早期诊断中的意义," vol. 4, no. 1, pp. 64–75, 2016.
- [7] H. S. I. Harba and E. S. I. Harba, "Voice Recognition with Genetic Algorithms," *Int. J. Mod. Trends Eng. Res.*, vol. 2, no. 12, pp. 144–155, 2018.
- [8] G. C. Hansen, K. H. Falkenbach, and I. Yaghmai, "Voice recognition system.," *Radiology*, vol. 169, no. 2, p. 580, 1988, doi: 10.1148/radiology.169.2.3175016.
- [9] C. G. Krishnan, Y. H. Robinson, and N. Chilamkurti, "Machine Learning Techniques for Speech Recognition using the Magnitude," *J. Multimed. Inf. Syst.*, vol. 7, no. 1, pp. 33–40, 2020, doi: 10.33851/jmis.2020.7.1.33.
- [10] R. Shanmugam, R. K. Patjoshi, S. R. Jena, and V. Gaur, "Desktop Assistant Based on Voice Recognition and Face Detection," no. July, 2020.
- [11] Othman Saleh Mahdy, "A Review of Literature of Computer-Assisted Translation," *Lang. India*, vol. 18, no. 9, pp. 340–359, 2018.
- [12] R. Dabre, C. Chu, and A. Kunchukuttan, "A Survey of Multilingual Neural Machine Translation," *ACM Comput. Surv.*, vol. 53, no. 5, 2020, doi: 10.1145/3406095.
- [13] N. T. Le, B. Lecouteux, and L. Besacier, "Disentangling ASR and MT Errors in Speech Translation," *arXiv*, 2017.
- [14] Q. T. Do, T. Toda, G. Neubig, S. Sakti, and S. Nakamura, "Preserving word-level emphasis in speech-to-speech translation," *IEEE/ACM Trans. Audio Speech Lang. Process.*, vol. 25, no. 3, pp. 544–556, 2017, doi: 10.1109/TASLP.2016.2643280.
- [15] A. Berard, L. Besacier, A. C. Kocabiyikoglu, and O. Pietquin, "End-to-End Automatic Speech Translation of Audiobooks," *ICASSP, IEEE Int. Conf. Acoust. Speech Signal Process. - Proc.*, vol. 2018-April, no. February, pp. 6224–6228, 2018, doi: 10.1109/ICASSP.2018.8461690.
- [16] S. Nakamura *et al.*, "The ATR multilingual speech-to-speech translation system," *IEEE Trans. Audio, Speech Lang. Process.*, vol. 14, no. 2, pp. 365–375, 2006, doi: 10.1109/TSA.2005.860774.
- [17] F. Ehsani, J. Kimzey, D. Master, K. Sudre, and H. Park, "Speech to speech translation for medical triage in Korean," no. June, pp. 13–19, 2006, doi: 10.3115/1706257.1706260.
- [18] C. Search and P. P. Data, "h hao wakati mwilini kauno auto," 2018.
- [19] M. F. O. R. Voice, "(12) United States Patent," 2019.
- [20] M.- Volume, "International Journal of Engineering Research & Management Technology," vol. 0181, no. December, p. 333001, 2018.