

## PRONUNCIATION MISTAKE WHICH LEADS TO INCORRECT TRANSCRIPTION BY AUTOMATIC CAPTION GENERATOR

Arif Haryadi<sup>1)</sup>, Rahmadsyah Rangkuti<sup>2)</sup>

<sup>1)</sup> Student of Universitas Sumatera Utara  
Email: imortabeingx@gmail.com

<sup>2)</sup> Lecturer of Universitas Sumatera Utara  
Email: syahkuti@gmail.com

### Abstract

*This study focuses on the analysis of pronunciation mistakes in the speech which was delivered by the Indonesian President during the 2014 APEC CEO Summit in Beijing, China. This study will provide description on the pronunciation mistakes, the word classes of the words with pronunciation mistakes and its place of articulation. From the result of the analysis, it can be seen that in the Indonesian President's speech, there were 108 pronunciation mistakes. The main mistakes were made in the part of speech word class of noun 69 times and the place of articulation where the mistakes occurred mostly is in the alveolar area 54 times.*

**Keywords:** *Caption Generator, Mistake, Pronunciation, Transcript,*

### 1. INTRODUCTION

The phenomenon where incorrect English word pronunciation in society is not a mistake for Indonesians, ideally, the spelling system should accurately reflect the pronunciation, which is reflected in many languages. Each English sound is represented by multiple letters or a sequence of letters. Also, each letter in English can represent multiple sounds or no sound at all. There is an inconsistency. English is emphasized and uncompromising, and that English spelling is only related to spoken language.

English indicates inconsistent pronunciation of words because one letter of the English alphabet can be pronounced differently in different words. Lagefaged (2006, p. 35) stated that most of the sound symbols are the same letters used in spelling the English words, but there are a few differences. One letter of the alphabet cannot contain one phonetic symbol. In most cases, phonetic symbols were displayed in pronunciation textbooks, but the letters of the alphabet in the phonetic symbols of words were not displayed. In English, there are twelve vowels. They are /I:/, /I/, /e/, /æ/, /ə/, /ɜ:/, /ʌ/, /ɑ:/, /ɒ/, /ɔ:/, /u:/, /ʊ/. There are two kinds of vowels in English. They are long vowels and short vowels. The long vowels like /I:/, /ɜ:/, /ɑ:/, /ɔ:/, /u:/. The short vowels like /I/, /e/, /ə/, /ʌ/, /ʊ/. Compare with Indonesian vowels, they are /A/, /I/, /U/, /E/, /O/. English has 24 consonants. They are [p], [b], [t], [d], [k], [g], [f], [v], [θ], [ð], [s], [z], [ʃ], [ʒ], [h], [tʃ], [dʒ], [m], [n], [ŋ], [l], [r], [w].

English is a foreign language in Indonesia, so there are certainly differences in pronunciation and grammar rules. The difference of structure between Indonesian language and English could cause the speakers to make mistakes. Brown (1980: 165) error is a noticeable deviation from the adult grammar of native speaker, reflecting the interlanguage competence of the learners. As the learner of the foreign language, non-native speakers find difficulties in speaking foreign language which generally produces either grammatical errors or pronunciation errors.

English is the international languages used everywhere around the world. Halliday (1994: 3). "Language is a human system of communication that uses arbitrary signals, such as voice sounds, gestures, or written symbols. Language is also an expression of ideas by voice; sounds, expression of thought, articulated by organs of throat and mouth. By this definition, when a person communicates to an international grade audience, the speaker must use English as the default language. During Communication as long as the speaker is able to speak with the minimum ability requirements, so the speakers are able convey their ideas to the audience.

Launched in 1996, the APEC Leadership Summit provides an opportunity for business leaders in the Asia Pacific region to discuss the challenges facing the region with APEC economic leaders, politicians, academia and other leaders. The summit, held in APEC countries just before the annual APEC Leadership Meeting, showcases APEC's efforts to promote dialogue and collaboration between nations. This research aims to analyze the pronunciation mistakes made by a non-native English speaker conveying his/her thought to an international audience.

## **2. METHOD**

### **Research Design**

The objective of this study is to determine the comparison and analysis for the speech which was delivered by the Indonesian President during the 2014 APEC CEO Summit in Beijing, China. It will describe the differences between the Indonesian's President predetermined speech transcript and the transcript which is produced by an automatic caption generator in Youtube. The study aimed to collect data in the form of words and sentences rather than numbers. The researcher will conduct a qualitative approach by using an observation method. The researchers used non-participatory observational methods to gather data on pronunciation mistakes in Joko Widodo's speech at the 2014 APEC CEO Summit in Beijing, China. This study is an interpretative study using an inductive way of thinking which steps from social phenomenon, feelings and daily facts (Emzir, 2010,p.2). This study is naturalistic, whereas it takes place during the APEC CEO Summit as the factual background to see process in which the speech was delivered.

### **Respondent of the study**

The focus area of this study will be President of Indonesia. The President's name is Joko Widodo but he is more known and recognized by the Indonesian

citizens as Jokowi. The President will be observed during his speech session in the APEC CEO Summit in Beijing, China.

#### **Data Procedure and Analysis**

The researcher attempts to present the existing phenomena or natural setting on the Indonesian President’s speech during the 2014 APEC CEO Summit in Beijing, China. The researcher will acquire the video of the speech from the beginning until the end and also the original transcript of the speech which was to be delivered during the summit. This study identifies the pronunciation mistakes of Joko Widodo's speech at the 2014 APEC CEO Summit in Beijing, China. The researcher identifies the mistake and compared it to the original script that had to be said in the speech. Minor mistakes, which does not change the meaning or classes of words was ignored because the meaning are still the same. Only mistakes in which it causes change in meaning or transforming into another word will be analyzed. The study results were finally analyzed by the researchers.

### **3. FINDINGS AND DISCUSSION**

This research analyzed the pronunciation mistake that occurs in the Indonesian President’s speech during the 2014 APEC CEO Summit in Beijing, China. In total there were 71 data’s (sentences) to be analyzed but after the analysis, only 48 of the 71 sentences contain pronunciation mistake. The data (sentences) in the table below has been selected to only show the data to reduce space usage and only show necessary details which is appropriate for the research analysis as displayed bellow:

Table 1: Words pronounced with mistakes along with its IPA form and words from the original transcripts which were supposedly uttered along with its IPA form and its word class.

<b>No. [Data.]</b>	<b>Auto-Generated Captions From the Speech</b>	<b>IPA format of Auto-Generated Captions</b>	<b>Original Transcript</b>	<b>Part of speech</b>
1. [5.]	of	/ʌv/	shows (/ʃoʊz/)	Verb
2. [6.]	piston K	/'pɪstən/ /keɪ/	distance (/'dɪstəns/) UK (/ju-keɪ/)	Noun Noun
3. [7.]	island island	/'aɪlənd/ /'aɪlənd/	islands (/'aɪləndz/). islands (/'aɪləndz/).	Noun Noun
4. [8.]	believed	/brɪ'livd/	billion (/'bɪljən/)	Noun
5. [9.]	suits	/suts/	It's huge. (/ɪts hjuːdʒ./)	Noun, Adjective
6. [10.]	above your	/ə'bʌv jʊər/	our fuel (/'aʊər 'fjuəl/)	Pronoun, Noun
7. [12.]	above the family	/ə'bʌv ðə/ /'fæməli/	our fuel (/'aʊər 'fjuəl/) farm (/farm/)	Pronoun, Noun Noun
8. [13.]	beat the dam them	/bɪt ðə dæm/ /ðəm/	build dams(/bɪld dæmz/) dams (/dæmz/)	Verb, Noun Noun

9. [14.]	get good get refrigerator	/gɛt/ /gʊd/ /gɛt/ /rə'frɪdʒə'reɪtər/	give (/gɪv/) boat (/boʊt/) give (/gɪv/) refrigerators (/rɪ'frɪdʒə'reɪtəz/).	Verb Noun Verb Noun
10. [16.]	sub zero channel	/sʌb 'zɪroʊ/ /'ʃænəl/	fuel (/'fjuəl /) give (/gɪv/)	Noun Verb
11. [18.]	head	/hɛd/	health (/hɛlθ/).	Noun
12. [20.]	bid support tip support	/bɪd/ /sə'pɔrt/ /tɪp sə'pɔrt/	build (/bɪld/) seaports (/si,pɔrts /) deep seaports (/dɪp 'si,pɔrts/).	Verb Noun Adjective, Noun
13. [21.]	Island support tip support	/'aɪlənd/ /sə'pɔrt/ /tɪp sə'pɔrt/	islands (/'aɪləndz seaports (/'si,pɔrts/) deep seaports (/dɪp 'si,pɔrts/).	Noun Noun Adjective, Noun
14. [22.]	support support	/sə'pɔrt/ /sə'pɔrt/	seaports (/'si,pɔrts/) seaports (/'si,pɔrts/).	Noun Noun
15. [23.]	pot	/pɑt/	Port (/pɔrt/)	Noun
16. [24.]	these tears	/ðɪz/ /tɜrz/	TEUs (/ti-i-juz/) TEUs (/ti-i-juz /)	Noun Noun
17. [25.]	pot	/pɑt/	Port (/pɔrt/)	Noun
18. [27.]	be	/bi/	build (/bɪld/)	Verb
19. [28.]	pit ready ready	/pɪt/ /'rɛdi/ /'rɛdi/	build (/bɪld/) railway (/'reɪl,wɛɪ/) railway (/'reɪl,wɛɪ/)	Verb Noun Noun
20. [29.]	pit	/pɪt/	build (/bɪld/)	Verb
21. [32.]	pay	/peɪ/	build (/bɪld/)	Verb
22. [33.]	pit	/pɪt/	build (/bɪld/)	Verb
23. [36.]	beat Sita	/bɪt Sɪtə/	build sea toll (/bɪld si toʊl/).	Verb, Noun
24. [37.]	acetyl-coa	/acetyl-coa/	sea toll? Sea toll (/si toʊl? si toʊl/)	Noun
25. [38.]	bed	/bɛd/	build (/bɪld/)	Verb
26. [39.]	vessel my professor  invert this it all	/'vɛsəl/ /maɪ prə'fɛsər/  /'ɪnvɜrt ðɪs ɪt əl/	vessels (/'vɛsəlz /) mother vessels (/'mʌðər 'vɛsəlz /) enter the sea toll (/'ɛntər ðə si toʊl/).	Noun Noun  Verb, Determiner, Noun
27. [41.]	when saxman sick for saxman	/wɛn 'sæksmən/ /sɪk/ /fɔr 'sæksmən/	sack cement (/sæk sə'mɛnt/), 6 (/sɪks/) per sack cement (/pɜr sæk sə'mɛnt/).	Noun  Noun Preposition, Noun
28. [42.]	for sex mean	/fɔr sɛks mɪn/	per sack cement (/pɜr sæk sə'mɛnt/).	Preposition, Noun

29. [43.]	time	/taɪm/	times(/taɪms/).	Noun
30. [44.]	versatile all of the island	/ə 'vɜrsətəl/ /əl əv ði ' aɪlənd /	our sea toll (/si tɔʊl/) our islands (/'aʊər 'aɪləndz/)	Pronoun,Noun Pronoun, Noun
31. [46.]	plane	/pleɪn/	plants (/plænts/).	Noun
32. [47.]	run beat so bit of a tune	/rʌn/ /bɪt/ /soʊ bɪt əv ə tʌn/	around (/ə'raʊnd/) build (/bɪld/) to build our (/tu bɪld 'aʊər/) zones (/zəʊnz/).	Prepositio Verb Verb, Pronoun Noun
33. [48.]	plane	/pleɪn/	plants (/plænts/).	Noun
34. [49.]	plan zone	/plæn/ /zəʊn/	plants (/plænts/) zones (/zəʊnz/).	Noun Noun
35. [50.]	infests are all most	/ɪn 'fests ər/ /əl moʊst/	investors (/ɪn 'vestərz/), most (/moʊst/)	Noun Adjective
36. [51.]	clearing	/'klɪrɪŋ/	clear (/klɪr/)	Adjective
37. [53.]	project uttering had family	/'prɒdʒekt/ /'ʌtərɪŋ/ /hæd/ /'fæməli/	a project (/ 'prɒdʒekt/), Outer (/ 'aʊtər/) have (/hæv/) families (/ 'fæməliz/)	Determiner, Noun Noun Verb Noun
38. [55.]	the grown for	/ðə grəʊn fɔr/	them (/ðem/) to (/tu/)	Pronoun
39. [56.]	For time	/fɔr taɪm/	Four times (/fɔr taɪmz/).	Noun, Noun
40. [59.]	for time	/fɔr taɪm/	Four times (/fɔr taɪmz/).	Noun, Noun
41. [60.]	for time	/fɔr taɪm/	Four times (/fɔr taɪmz/)	Noun, Noun
42. [61.]	clear	/klɪr/	cleared (/klɪrd/).	Verb
43. [62.]	power rod	/'paʊər rɒd/	toll road (/tɔʊl rɒd/)	Noun, Noun
44. [63.]	This now without	/ðɪs naʊ wɪ 'θaʊt/	Now we talk (/naʊ wɪ tɔk/)	Adverb, Pronoun, Noun
45. [64.]	bands top surface... one sub surface	/bændz tɒp 'sɜrf əs/... /wʌn sʌb 'sɜrfəs/	national one-stop service (/'næʃənəl wʌn- stɒp 'sɜrvəs/)	Adjective, Adjective, Noun
46. [65.]	DS	/di-ɛs/	days (/deɪz/)	Noun
47. [66.]	inducing	/ɪn 'dusɪŋ/	Indonesian (/ɪndəʊ 'nɪzən/)	Adjective
48. [67.]	a Tunisia	/ə ,tu 'nɪzə/	Indonesia (/ɪndəʊ 'nɪzə/).	Noun

During the speech, there were 71 sentences which consist of 886 words that were uttered by the speaker (The Indonesian President). The 886 words consist of; 269 Nouns, 117 Pronouns, 140 Verbs, 37 Adverbs, 41 Adjectives, 28 Conjunctions, 59 Determiners, 93 Prepositions and one Interjection while the other 90 words in the original transcript indicates numbers and symbols. Out of the 71 sentences, 48 sentences contained some pronunciation mistakes which lead to the change of word form and meaning. In the 48 sentences, the pronunciation mistakes occurred 108 times which includes; 69 Nouns, 7 Pronouns, 18 Verbs, 1 Adverb, 8 Adjectives, 2 Determiners and 3 Prepositions.

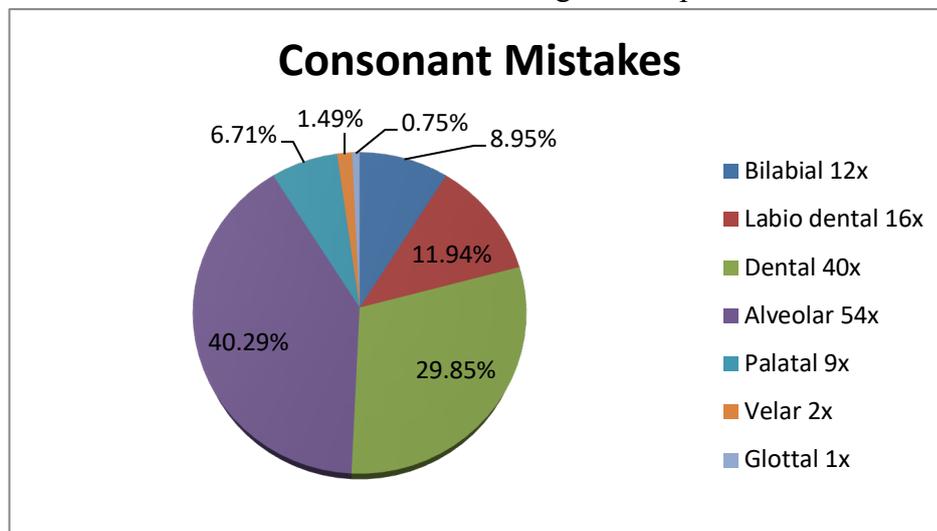
By using the IPA consonants for general american english (the standard international language) like in the table below:

**IPA Consonants for General American English**  
as used by the Cambridge Online Dictionary

		Place of Articulation							
		Front						Back	
(numbers refer to positions on articulation diagram)		Bilabial (both lips) 8 - 1	Labio-dental (lip + teeth) 8 - 2	Dental (teeth) 11 - 2 11 - 2 + 10	Alveolar (behind teeth) 11 - 3, 4	Post-Alveolar (near a-ridge) 11 - 4, 5	Palatal (hard palate) 12 - 6	Velar (soft palate) 13 - 7	Glottal (throat) 16
Airflow	<b>Stop</b> (air is stopped then released)	p • b			t • d			k • g	ʔ <sup>1</sup> •
	<b>Nasal</b> (air passes through nose)	• m			• n			• ŋ	
	<b>Fricative</b> (air is constricted)		f • v	θ • ð	s • z	ʃ • ʒ			h •
	<b>Affricate</b> (stop + fricative)					tʃ • dʒ			
	<b>Approximant</b> (vowel acts as syllable boundary)				• r <sup>2</sup>		• j	• w <sup>3</sup>	
	<b>Lateral Approximant</b> (vowel acts as syllable boundary)				• l				

The researcher has further analyzed the pronunciation mistake by using the IPA consonant for general american english as the standard to identify in which place of articulation does the mistake occur at. After comparing the original transcript IPA form with the auto generated caption IPA form (only the consonants). By using the original transcript IPA form as the base for comparison, the analysis result as following the pie chart:

Pie Chart 1: Consonant mistakes according to their place of articulation



Matthews (1997) divides the place of articulation as the following:

Bilabial consonants occur when airflow is blocked out of the mouth by bringing the lips together. From the speech analysis, mistakes done in the bilabial area occurred 12 times, mainly when producing the consonant [p], [b].

9. [14.] **Good** (/gʊd/) - **boat** (/boʊt/) || 19. [28.] **Pit** (/pɪt/) - **build** (/bɪld/)

Labio-Dental consonants occur when airflow is blocked by curling the lower lip back and raising it to touch the upper row of teeth. From the speech analysis, mistakes done in the labio-dental area occurred 16 times, mainly when producing the consonant [f], [v].

1. [5.] **Of** (/ʌv/) **shows** (/ʃoʊz/) || 9. [14.] **Get** (/ɡet/) **Give** (/ɡɪv/)

Dental consonants occur when airflow is blocked by placing the tongue against the upper teeth. From the speech analysis, mistakes done in the Dental area occurred 40 times, mainly when producing the consonant [t], [d].

31. [46.] **Plane** (/pleɪn/) **plants** (/plænts/). || 2. [6.] **Piston** (/ˈpɪstən/) **distance** (/ˈdɪstəns/)

Alveolar is where the teeth meet the gums. Alveolar consonants are created when the tongue is raised to the alveolar ridge to block or constrict airflow. From the speech analysis, mistakes done in the alveolar area occurred 54 times, mainly when producing the consonant [s], [z].

13. [21.] **Island** (/ˈaɪlənd/) **islands** (/ˈaɪləndz/). || 29. [43.] **Time** (/taɪm/) **times** (/taɪms/).

Palatal is the roof of the mouth is the hard palate. Palatal consonants are produced when the tongue rises to certain height and constrict airflow. From the speech analysis, mistakes done in the palatal area occurred 9 times, mainly when producing the consonant [j].

4. [8.] **believed** (/bɪˈlivd/) **billion** (/ˈbɪljən/)

Velar is behind the hard palate is the velum or soft palate. Unlike the hard palate in front of it, this consists of soft tissue. Velar Consonants are produced when the back of the tongue rises to the velum to block or restrict airflow. From the speech analysis, mistakes done in the velar area occurred 2 times, mainly when producing the consonant [k]

44. [63.] **now without** (/naʊ wɪˈθaʊt/) **Now we talk** (/naʊ wɪ tɔk/)

Glottal is the two vocal folds (i.e. vocal cords) located in the larynx. From the speech analysis, mistakes done in the glottal area occurred 1 time, mainly when producing the consonant [h].

5. [9.] **Suits** (/sʊts/) **it's huge**. (/ɪts hʒudʒ/)

#### 4. CONCLUSION

From the result of the analysis, it can be seen that in the Indonesian President's speech during the 2014 APEC CEO Summit in Beijing made 108 pronunciation mistakes. The main mistakes were made in the part of speech word class of noun 69 times and the place of articulation where the mistakes occurred mostly is in the alveolar area 54 times. As a non-native English speaker the difficulty in mastering the pronunciation can be considered difficult because English indicates inconsistent pronunciation of words where one letter of the

English alphabet can be pronounced differently in different words. The difference of structure between Indonesian language and English could cause the speakers to make mistakes.

## 5. REFERENCE

- APEC CEO Summit - <https://www.apec.org/About-Us/About-APEC/Business-Resources/APEC-CEO-Summit> - Accessed on 23rd December 2021
- Brown, H, D. (1980). "*PRINCIPLES of LANGUAGE LEARNING AND TEACHING: 5th Edittion*". San Fransisco: Pearson Education. ISBN: 0-13-199128-0
- Emzir. (2010). Metodologi penelitian kualitatif: Analisis data. Jakarta: PT. Raja Grafindo Persada.
- Halliday, M.A.K. (1994). "*An Introduction to Functional Grammar*". London: Edward Arnold.
- IPA symbols for Vowels and Consonant. <https://www.oxfordlearnersdictionaries.com/about/practical-english-usage/phonetic-alphabet>
- Joko Widodo, President of Indonesia, at the APEC CEO Summit [https://www.youtube.com/watch?v=Lo2jx\\_IFAoU&t=36s](https://www.youtube.com/watch?v=Lo2jx_IFAoU&t=36s) - Accessed on 23rd December 2021
- Joko Widodo, President of Indonesia, at the APEC CEO Summit, English (auto-generated) [https://downsub.com/?url=https%3A%2F%2Fwww.youtube.com%2Fwatch%3Fv%3DLo2jx\\_IFAoU%26t%3D36s](https://downsub.com/?url=https%3A%2F%2Fwww.youtube.com%2Fwatch%3Fv%3DLo2jx_IFAoU%26t%3D36s) - Accessed on 23rd December 2021
- Ladefoged, P. (2006). "*A course of phonetics (5th ed.)*". Boston: Thomson Wadsworth.
- Matthews, P. H. (1997). "*The Concise Oxford Dictionary of Linguistics*". Oxford: Oxford University Press.
- Remarks by Indonesian President Joko Widodo at the APEC CEO Summit on November 10, 2014, in Beijing, China. <https://setkab.go.id/remarks-by-indonesian-president-joko-jokowi-widodo-at-the-apec-ceo-summit-on-november-10-2014-in-beijing-china/> - Accessed on 23rd December 2021
- Roach, P. (2009). "*English Phonetics and Phonology: A Practical Course, 4th Ed*". Cambridge: Cambridge University Press, ISBN 978-0-521-78613-3