Simalungun language phonotactics: the study of phonology

A R Purba\(^1\) and J Siahaan\(^1\)

\(^1\)Department of Batak Literature, Faculty of Cultural Sciences, University of North Sumatra, Jalan Universitas Number 19 Padang Bulan Medan

E-mail: asriiatyr@gmail.com

Abstract. Phonotactics research is a procedure for the discovery and determination of the sequence and phonemes in a language based on distribution (beginning, middle and end of a word). Each language has a special characteristic in phonotactics. Therefore, phonotactics research needs to be done. The purpose of this study is to describe a series of vocal, diphthong, a sequence of consonants, and syllables. The method used was descriptive qualitative methods by collecting data through observation, interview and study library by using the phonotactics theory initiated by O’Grady, et al (1989) and Pulgram (1970). Based on the analysis of phonotactics in Simalungun language, vowel sequence was found to occupy the position of beginning, middle and end of a Word. Diphthongs only occupied the position of the end of a word, whereas the consonant sequence was only found in the middle of word. The results of the analysis showed that the vowel sequence consists of sixteen (16), namely: vowel sequence /ai/, /au/, /ao/, /ia/, /iu/, /io/, /ua/, /ue/, /uo/, /ui/, /ea/, /eu/, /eo/, /oa/ and /oi/. Difftong consists of three (3), they are /ou/, /ei/ and /ui/. There are ten (10) consonant sequences, namely: a). The sequence starts with a consonant /h/ as : /-hb-/ , /-hk-/ , /-hp-/ , and /-ht-/, b). The sequence starts with a consonant /k/ as : /-kb-/ , /-kd-/ , /-kj-/ , /-kp-/ , and /-kt-/, c). The sequence starts with a consonant /l/ as : /-lb-/ , /-ld-/ , /-lh-/ , /-lg-/ , /-lj-/ , /-ln-/ , /-lp-/ , and /-lt-/, d). The sequence starts with a consonant /m/ as : /-mb-/ , /-mm-/ , /-ml-/ , and /-mp-/ , e). The sequence starts with a consonant /n/ as : /-nd-/ , /-nj-/ , /-ns-/ , and /-nt-/, f). The sequence starts with a consonant /ŋ/ as : /-ng-/ , /-nj-/ , /-ŋk-/ , /-ŋl-/ , /-ŋm-/ , /-ŋr-/ , /-ŋs-/ , and /-ŋt-/ , g). The sequence starts with a consonant /p/ as : /-pd-/ , /-ph-/ , /-pr-/ , /-ps-/ , and /-pt-/ , h). The sequence starts with a consonant /r/ as : /-rb-/ , /-rd-/ , /-rq-/ , /-rk-/ , /-rl-/ , /-rp-/ , and /-rt-/ , i). The sequence starts with a consonant /s/ as : /-sd-/ , /-sq-/ , /-sk-/ , /-sp-/ and /-st-/ , j). The sequence starts with a consonant /t/ as : /-tg-/ , /-tp-/ , and /-tr-/. The syllables in the vowel sequence consist of V, VK, KV, and KVK. The syllable in diphthongs consists of KVV at the end of the word and the consonant sequence consists of the KK which occupies the middle position of the word.

1. Introduction

Humans communicate through language with oral and written. Oral communication involves tools of human speech. Otherwise written communication does not involve tools of human talk. The tool of human speech forms the sounds of the language used in oral communication learned in the science of phonology.

Language is a system of Heraldry arbiter sound, which is produced by the instrument, uttered by humans, and used by communities to communicate, work together, and identify themselves. The language is created by rules and patterns that should not be violated so as not to disrupt at the time of communication takes place. Rules, precept, and patterns formed include the sound, formatting, and the...
form of the sentence. Therefore, communication can go well between the recipient and the sender, if both master the Lingo language.

Each language has its provisions relating to linguistic rules, including the rules of phonemes, syllables. The rules govern which rows are present in the language and which cannot be learned in phonotactics. Phonotactic is the possible sequence of phonemes in a language or description for the phoneme sequence. Kridalaksana (1982: 58) [1] gives three definitions for the term phonotactic, namely: Phoneme sequences that are possible in a language, Descriptions of the sequence, Grammatical stratification, Regulatory system in phonetic strata. Hawkins (1984: 61) [2] says that the study of sound combinations is generally known as phonotactics.

In language activities, the arrangement of phonotactic rules can support the determination of syllabication (syllabic) systems in a language. The syllable pattern is the part that needs to be complemented in a phonotactic study of a language. With the phonotactic knowledge of a language, language researchers can determine the pattern of certain syllables.

The phonotactic research of the Simalungun language is interesting to examine because the spoken word is often incompatible with the prevailing rules. It is due to the lack of knowledge of the speaker about the science of how the sounds should be pronounced.

2. Phonotactic theory
Pulgram [3] said a combination of two or more consonants occurring in different syllables though adjacent is called a sequence, while consonant sequences belonging to the same syllable are called consonant clusters.

O'Grady et al. [4] stated that phonotactics is a set of constraints on how the pattern of sequences of sounds (segments) is formed, forming part of the phonological abilities and knowledge possessed by the speakers of that language. Syllable phonetically consists of two parts, namely:
1. The consonants preceding the vowel is called the Onset (O)
2. Rima (R). Rima consists of 2 parts, namely:
   a. The nucleus or ‘Peak’ Hyman [5].
   b. The consonant following the vowel is called coda

\[ \Sigma = \text{suku kata} \]

\[ \text{Onset (O)} \]
\[ \text{Rhyme (R)} \]
\[ \text{Nucleus (N)} \]
\[ \text{Coda (C)} \]

Figure 1. Syllable phonetically.

Describing the bookkeeping in two or more syllables in Simalungun Language, four steps opinions of O'Grady, et al [4] were followed: The first step: because the nucleus of the syllable is mandatory on a constituent syllables, then the core of the syllables is determined on each tribe that is typically vowel, and above each symbol of the nucleus (N) placed Rima (R), and above the respective Rima (R) placed the sigma symbol ( ) for dividing the tribe said. The second step: the longest consonant sequence to the left of each nucleus (N) that does not violate the phonotactic barrier of a language is called the onset (O) of the syllable. The third step: it means that every remaining consonant on the right of each nucleus (N) forms a coda (C). This coda is combined with the syllable core to the left of rima (R). The syllable ending with coda (C) is called closed syllable. The fourth step: the second syllable is connected to form words (Wd).
Figure 2. The second syllable in connected to from words. Clockwise, number (1) to (4) illustrate the transformation of syllable in creating word, as example: ekstrim (Extreme – in English).

3. Research methodology
The method used in this research was a descriptive qualitative method which aims to describe or explain events or occurrence objectively based on the data found in the field.

4. Discussion
Phonotactics is a procedure of finding and determining the sequence of phonemes in a language based on the distribution (beginning, middle, and end of a word). Thus, phonotactics is a science that regulates the order of phonemes that form a word producing acceptable sounds in a language. These rules are not the same in all languages, but each is different.

4.1. Vowel sequence
Vowel sequence is two or more lined vowels that are located on a different syllable. In Simalungun language, the sequence of the vowels was found in the beginning, middle and end of a word. Nevertheless, some vowel sequences can only occupy two positions and one position only.

The vowel sequence in the Simalungun language is / ai /, / au /, / ae /, / ao /, / ia /, / iu /, / io /, / ua /, / ue /, / uo /, / ui /, / ea /, / eu /, / eo /, / oa / and / oi /. The sequence of the vowel is not pronounced simultaneously but pronounced separately, which means that each contains different syllable. This statement is in accordance with the opinion of Alwi, et al [6] which said that the vowel sequence is the same breath or almost the same as the two vowels that are included in different syllables and also as mentioned by Aminoedin et al. (1984: 140)[7] said that what is meant by a series vowels are two or more vowels that line up, but each is the peak of speech loudness.

| Vowel Sequence | Position | Beginning | Middle | End |
|----------------|----------|-----------|--------|-----|
| ai             | Words    | lift      | sait   | dai |
|                | Meaning  |           | canine tooth | taste |
| au             | auge     | buffalo yoke | haut | bau |
|                | [aʷuga]  |           | [haʷut] | [baʷu] |
| ae             | aer      | mix       | paet   | sae |
|                | [aer]    |           | [paet] | [sae] |
|                |          |           | bitter | finish |
Table 1. Vowel sequence.

| Vowel Sequence | Beginning | Middle | End |
|----------------|-----------|--------|-----|
| ao             | saor [sa^or] | mix [ta^o] | lake |
| ia             | [i^a] he/she/it | piar [pi^ar] | ambia [ambi^a] |
| iu             | [i^u] piuk [pi^uk] | shrink [-] |
| io             | iok [io?] pity | dior [di^or] | pio [pi^o] |
| ua             | uak [u^ak] poke | buat [bu^at] | tua [tu^a] |
| ue             | - - - - | tuok [tu^ok] | bue [bu^e] |
| uo             | [u^wos] blow | [u^wos] gambling game |
| ui             | [u^i] the name of a type of bird | kuit [ku^i] | dab [-] |
| ea             | eak [eak] yes | dear [dear] | lea [lea] |
| eu             | - - - - | beuk [beu?] | maheu [maheu] |
| eo             | eor [efor] scattering | beot [beot] | beo [beo] |
| oa             | - - - - | boan [boan] bring [-] |
| oi             | - - - - | bois [bois] out | boi [boi] get |

4.2. Diphthong

Diphthongs are two vowels or more rows contained in the same syllable. In Simalungun language, there are three diphthongs, they are / ou /, / ei /, and / ui /. Diphthongs are vowel cluster that belongs in a single syllable and is pronounced simultaneously. Furthermore, Wolfram and Johnson [8] said that diphthong is a combination of two vowels which one acts as the core (vowels) and the one acts as the sound of the slide. The diphthong position is at the end of the base word.

Table 2. Diphthong.

| Diphthong | Words | Meaning |
|-----------|-------|---------|
| ou        | logou [lo^ou] | 'wind' |
| ei        | hio [hi^ou] | 'Custom fabric' |
| ui        | matei [mat^e] | 'died' |
|           | mabei [ma^be] | 'mature' |
|           | surui [sur^u] | 'comb' |
|           | halahu [la^hu] | 'Eagle' |

4.3. The Consonant sequence

In Simalungun language, the sequence of consonants was found in middle of basic words. All sequence of consonants were found in the basic word between two syllables. It is in accordance with the opinion of Pulgram [3] stating that the sequence of consonants is divided into adjoining syllables called a sequence, and Lauder, (1996: 148)[9] also states consonant series are consonants that are in different syllables.
Table 3. The consonant sequence.

| Consonant Sequence | Position | Middle | Words | Meaning |
|--------------------|----------|--------|-------|---------|
| h hb               |          |        | behbeh [behbeh] | down to the bottom |
| hk                 |          |        | kahkah [kahkah] | hit |
| hp                 |          |        | pahpah [pahpah] | food from the sliced rice |
| ht                 |          |        | tehteh [tehteh] | saliva |
| k kb               |          |        | sabokbok [sabokbok] | contemporary |
| kd                 |          |        | dakdak [dakdak] | porridge |
| kj                 |          |        | jukjuk [jukjuk] | poke |
| kl                 |          |        | laklak [laklak] | tree bark |
| km                 |          |        | hatimukmuk [hatimu?mu?] | caterpillars |
| kp                 |          |        | pokpok [pokpok] | hit |
| kt                 |          |        | tuktuk [tuktuk] | tap |
| l lb               |          |        | holbung [holbuŋ] | valley |
| ld                 |          |        | aldom [aldom] | plug |
| lg                 |          |        | bolgang [bolgaŋ] | boiled |
| lj                 |          |        | juljul [juljul] | push |
| lm                 |          |        | jolma [jolma] | human |
| ln                 |          |        | julfang [julnaŋ] | decline |
| lj                 |          |        | nglingil [nilnil] | laugh |
| lp                 |          |        | tolpus [tolpus] | leaking |
| lt                 |          |        | boltok [boltok] | stomach |
| n nd               |          |        | gondang [gondaŋ] | drum |
| nj                 |          |        | ganjo [ganjoʷ] | crab |
| ns                 |          |        | pansing [passiŋ] | holy |
| nt                 |          |        | bintang [bittan] | star |
| ng                 |          |        | anggi [anggi] | brother |
| nj                 |          |        | jongjonj [jonjonj] | stand up |
| nk                 |          |        | angkat [ankat] | lift |
| nl                 |          |        | langlang [lanlan] | hindered |
| nm                 |          |        | mangmang [mangmanj] | spell |
| nr                 |          |        | rangrang [ranranj] | rarely |
| ns                 |          |        | angsur [ansur] | repay |
| nt                 |          |        | tingting [tiŋtiŋ] | announcement |
| p pd               |          |        | dapdap [dapdap] | waru |
| ph                 |          |        | hophop [hophop] | protect |
| pr                 |          |        | raprap [raprap] | trip it up |
| ps                 |          |        | sapsap [sapsap] | spend it |
| pt                 |          |        | tiptip [tiptip] | flatten |
| r rb               |          |        | arbab [arbab] | musical instrument |
| rd                 |          |        | ardan [ardan] | stairs |
| rg                 |          |        | margar [marga] | family name |
| rj                 |          |        | horja [horja] | work |
| rk                 |          |        | porkis [porkis] | ant |
| rl                 |          |        | mirlak [mirlak] | shine |
| rm                 |          |        | normar [narnar] | live coals |
| rp                 |          |        | shirput [sirput] | mimosa |
| rt                 |          |        | arata [arta] | treasure |
| rŋ                 |          |        | borgin [borgin] | night |
| s sd               |          |        | dosdos [dosdos] | same |
Table 3. The consonant sequence.

| Consonant Sequence | Position | Words | Meaning |
|--------------------|----------|-------|---------|
| sg gusgus [gusgus] | Middle   | scrub |         |
| sk kuskus [kuskus] |          | crust |         |
| sp pispis [pispis] |          | sprinkle |     |
| st pistar [pistar] |          | smart |         |
| t tg gatgat [gatgat] |          | chopped |   |
| tp pitpit [pitpit] |          | eyes closed | |
| tr rutrut [rutrut] |          | pull |         |
| m mb tambar [tambar] |          | drug |         |
| ml lamlam [lamlam] |          | gentle |        |
| mm summah [summah] |          | kiss |         |
| mp hampung [hampun] |          | light |         |

Table 4. Syllables in a sequence of vowel.

| Syllables | Vowel Sequence | Words | Meaning |
|-----------|----------------|-------|---------|
| V         | ia [iʰa]       | he/she/it | |
| V-VK      | ea [eʰak]      | agree |         |
| KV-V      | ao [aʰo]       | lake |         |
| KV-VK     | ua [suʰan]     | planting | |
| KVK-KV-V  | ua [sintuʰa]   | elders |         |

b. Syllables in a diphthong

Table 5. Syllables in diphthong.

| Syllables | Diphthongs Sequence | Words | Meaning |
|-----------|---------------------|-------|---------|
| KV-KVV    | ei [mateʰ]          | die |         |
| KV-KVV    | ou [logoʰ]          | wind |         |
| KV-KVV    | ui [apuʰ]           | fire |         |

c. The syllable in consonant sequence

The syllable in consonant sequence of Simalungun language consists of KK which occupied the middle position of words and was found in between the two syllables. Example:
Table 6. The syllable in consonant sequence.

| Consonants Sequence | Position | Meaning               |
|---------------------|----------|-----------------------|
|                     | Middle   |                       |
| hb                  | behbeh   | down to the bottom    |
| kb                  | sabokbok | contemporary          |
| lb                  | holbung  | swamp                 |
| nd                  | gondang  | drum                  |
| ng                  | anggi    | brother               |
| rb                  | arbab    | musical instrument    |
| sd                  | dosdos   | same                  |
| tg                  | gatgat   | Chopped               |
| mb                  | tambar   | drug                  |

5. Conclusion and suggestion

5.1. Conclusion

Based on the results of data analysis, the phonotactic structure of Simalungun language in terms of vowel sequence, diphthongs, consonant sequence, and syllables can be concluded as follows:

1. There are sixteen (16) vowel sequences of the Simalungun language, namely / ai /, / au /, / ae /, / ao /, / ia /, / iu /, / io /, / ua /, / ui /, / ea /, / eo /, / eu /, / oa /, / oi /. The vowel sequences that occupy the initial, middle and end positions are the vowel series of / ai /, / au /, / ae /, / ia /, / io /, / ua /, / ea /, and / eo /, that occupy the beginning and middle position are / uo /, / ui / and / oa /. The vowel sequence that occupies the middle position is / iu /, / ui / while the vowel sequence that occupies the middle position and the end of the word is / ao /, / eu / and which vowel / ua / only occupies the end of the word.

2. There are three (3) diphthongs found in Simalungun language, they are / ei /, / ou / and / ui /. The diphthong is only found at the end of the word.

3. There are eleven (11) consonant sequences in the Simalungun language, namely:
   a. Consonant sequences starting with / h / are all in the middle position of the consonant sequence / h / as in: / -hb- /, / -hk- /, / -hp- / and / -ht- /.
   b. Consonant sequences starting with / k / are all in the middle position, consonant sequence / k / is as in: / -kb- /, / -kd- /, / -kj- /, / -km- /, / -kl- /, / -kt- /, and / -kp- /.
   c. Consonant sequences beginning with / l / are all in the middle position, the consonant sequence / l / is as in: / -lb- /, / -lh- /, / -lm- /, / -lg- /, / -lj- /, / -ln- /, / -lp- /, and / -lt- /.
   d. Consonant sequences starting with / m / are all in the middle position of the consonant sequence / m / as in / -mb- /, / -ml- /, / -mm- /, and / -mp- /.
   e. Consonant sequences beginning with / n / are all in the middle position of the consonant sequence / n / as in: / -nd- /, / -nj- /, / -ns- /, and / -nt- /.
   f. Consonant sequences starting with / ŋ / are all in the middle position of the consonant sequence / ŋ / as in: / -ng- /, / -nj- /, / -nk- /, / -nj- /, / -nm- /, / -nr- /, / -ns- /, and / -nt- /.
   g. Consonant sequences starting with / p / are all in the middle position, consonant sequence / p / as in / -pd- /, / -ph- /, / -pr- /, / -ps- /, and / -pt- /.
   h. Consonant sequences starting with / r / are all in the middle position, consonant sequence / r / as in: / -rb- /, / -rd- /, / -rg- /, / -rk- /, / -rp- /, / -pr- /, and / -pt- /.
   i. Consonant sequences starting with / s / are all in the middle position, consonant sequence / s / as in: / -sd- /, / -sg- /, / -sk- /, / -sp- /, and / -st- /.
   j. Consonant sequences starting with / t / are all in the middle position, consonant sequence / t / as in: / -tg- /, / -tp- /, and / -tr- /.

4. Consonant clusters in Simalungun language were not found.

5. Syllables
a. The syllables in the vowel sequence consist of v, vk, kv and kvk.
b. The syllable diphthong consists of kvv, found at the end of word.
c. The syllable in consonant sequence consists kk, found in the middle of basic word.

5.2. Suggestions

1. The results of this research are expected to be input in the preparation of Simalungun language textbooks as a lesson in local content.
2. Researchers who are interested in examining matters relating to linguistic are recommended to follow-up for further research to strengthen the rule of phonotactics.

6. References

[1] Kridalaksana, H. 1982. Linguistic dictionary (in Indonesia). Jakarta: Gramedia.
[2] Hawkins, P. 1984. *Introduction Phonology*. Great Britain: Hutchinso & Co, Ltd
[3] Pulgram, E. 1970. *Syllable, Words, Nexus, Cursus*. The Netherlands: Mouton & Co.
[4] O’Grady, William; Micael Dobrovolsky, and Mark Aronoff. 1989. *Contemporary Linguistics: An Introduction*. New York: St. Martini’s Press.
[5] Hyman, L.M. 1975. *Phonology: Theory and Practice*. New York: Holt, Rinehart and Winston.
[6] Alwi, Hasan. Soejono, D.; Hans, L, dan Anton, M.M. 1998. Structure of standard Indonesian language (in Indonesia). Jakarta: Balai Pustaka.
[7] Aminoedin, dkk. 1984. Phonology of Indonesian language : a perspective study (in Indonesia). Jakarta: Pusat Pembangunan dan Pengembangan Bahasa
[8] Wolfram, W. and Johnson., R. 1982. *Phonological Analysis: Focus on American English*. Washington, DC: Centerfor Applied Linguistics.
[9] Lauder, Multamia R.M.T. 1996. “Indonesian language phonem vocabulary: Analysis of frequency and phonotactic (in Indonesian). Bahasa Nasional Kita dari Sumpah Pemuda ke Pesta Emas Kemerdekaan. Bandung: ITB.
[10] Verhaar, J.W.M. 2010. Principles of basic linguistic (in Indonesia). Gadjahmada University Press.