INDONESIAN FORMANTS IN ACOUSTIC ANALYSIS

Ichwan Suyudi & Debyo Saptono
Gunadarma University
INDONESIA

ABSTRACT

This research has a purpose to analyze the characteristic of speech sound or usually called as formants analysis or vocal tract. The contour of vowel sound was analyzed just if its fundamental frequency showed the up-down pattern clearly. For that importance, so the fundamental frequencies (f0, f1, and f2) were counted separately on certain points. As a comparison, it was also counted the value of f0, f1 and f2 in words. The used method was Praat software and VisArtico application. The result of the research shows that vowel sound of /a/, shows f1 and f2 point that is different when the utterance data was done by two different recording processes. This difference is assumed as a result of allophonic that influenced the production of that vocal sound.

Keywords: Vocal Tract, Acoustic, Vocal Sound, Bahasa Indonesia.

BACKGROUND OF THE STUDY

Vowel system of Bahasa Indonesia is relatively simple. This system just has six monophtongs. There is no nasalization vowel in Bahasa Indonesia (Zanten, 1989) and there are no differences between vowel phonemes. If it was compared to English vowel that had 12 monophtong vocals since English phonology system differentiate the long vowel and short vowel. Bahasa Indonesia vowel system, commonly divided into six phonemes, they are /i, ə, a, o, a/ and one central vocal (Lapoliwa, 1983). Vowels of Bahasa Indonesia can be pronounced in different ways, especially depended on its emergence as opened syllable or closed syllable. Generally, it was agreed that /i, ə, o/ had a chance to have allophonic variation, either /ə/ isn't. Numerous linguists put /u/ in that group (Halim, 1974; Soedjarwo, 1983). According to Poedjosudarmo (1982) the appearance of two allophones for /o/ was not determined by context in Bahasa Indonesia so those two sounds become two different vowel phonemes, they were /o/ and / ə/.

There is no any agreement among the linguists about what context of the appearance of low allophone and high allophone is. According to Stokhof (1980), the closed syllable does not contain high allophones, except on a loanword. Ekowardono (1983) emphasizes the facts that there are numerous exceptions about this principle, especially related to /i/ and /u/. Other linguists set measures to the appearance of lower allophone on the last closed syllable, also on the syllable (penultimate) ended by consonant non nasal, but those shapes do not have any differences in Bahasa Indonesia vocabulary (Soedjarwo, 1987). However, in some words, the higher allophones of /i/ and /u/ on the loan word emerge on the last closed syllable. The harmony of allophones were stated by many linguists, which was about the low allophone of a phoneme that appeared in both closed syllable and opened syllable penultimate that was directly followed by the closed syllable, that contained the low allophone of the same vocal phoneme. Zanten (1989) states that he agrees with all other linguists about variation of allophonic for /e/ and /o/ and seem that position of /i/ and /u/ are weaker with some exceptions. Allophonic variations for /a/ is stated by some people, but no one who states about / ə/. The experimental phonetic methods could be used so that became more explicit and maybe more objective. By comparing the measurements of physical quality of utterances,

other differences of utterances' realization could be revealed. Like what had been done by Azhar (2001). In his research, Azhar explained that the sound was produced by two processes, i. e. generation and filtering. Generation process was the first time the utterance was produced by the vibration of vocal cord and vocal fold inside the larynx for producing periodic sound. Periodic sound had the characteristic of being constant and next would be filtered by vocal tract or articulator that contained of tongue, teeth, lips, palate and others so the sound became an output of vowel and or consonant that formed meaningful words and could be analyzed next for understanding the utterance or could be said as voice recognizion. Azhar stated that principally the speech sound was divided into some components, such as pitch, formants and spectrogram that could be used for identifying someone's characteristic for the importance of voice recognition. For the importance of speech sound analysis, there were some components to be analyzed, they were:

1. Pitch or could be called as fundamental frequency (basic) with f0 notation. Each people had habitual pitch that was so influenced by physiological aspects of human larynx. On the condition of normal conversation, the habitual pitch level was between 50 to 250 Hz for men and 120 to 500 Hz for women. The changing of f0 constantly could give linguistic information, for example to differentiate between intonation and emotion. Pitch analyze could be used for doing voice recognition to someone's voice through statistic analyzing to the value of minimum pitch, maximum pitch and mean pitch.

2. The utterance could be divided through the pitch, loudness, and vocal quality (Ladefoged and Johnson, 2011: 7). Wedhawati, et al. (2001: 24) explained that the quality of vocal cord was defined by four factors, such as high-low of the tongue, a part of the tongue that moved around, positional relation of the active articulator with passive articulator, and lips' shape. While Bickford and Rick (2006:32) stated that vocal quality was influenced by the differences of two tongue positions and mouth, for both in front-backs, up-down. The vocal differentiation was explained by Ladefoged and Johnson (2011: 22-23) could be seen by pitch and additional tone that related hardly to the differentiation of front vocal and back vocal. High-low of the vocal pitch was decided by tongue's position, the high vocal pitch was when a tongue's position was low and on the contrary, the low vocal pitch was when the tongue's position was high.

3. Forman was resonance' frequencies of the filter, which was vocal tract (articulator) continued and filtered the vocal output, consonant or words. Cohn in Aronoff and Janie (2003), explained that the utterance generally had characteristics in tongue or jaw's height (height, mid, low) and tongue's part that could move (front, mid, back). Besides, the utterance appeared because of the utterance does not too close so the airflow did not have any obstacle (Ladefoged and Johnson, 2011). Ladefoged and Johnson (2011) explained that the utterance was better explained by the explanation of acoustic structure than by the explanation of movement's influences as articulatoric.

   According to the background of Bahasa Indonesia vowel experimentally, this research had a purpose to show the result of formants analysis used Praat analysis and compared it by vocal tract animation. Through the counting of f1 and f2 on Praat equipment and vocal tract animation, it was expected to get points that were precision enough about the position and tongue's movement for producing the vocal so it could be seen the error point of the phonetic.

RESEARCH'S METHOD

The research's method that was used was acoustic study (van Zanten and van Heuven, 1983). The production process was done by recording to the informant of native and non-native Bahasa Indonesia speakers. These informants produce vowel and consonant by 2 experiments, i.e. producing the vowel sound isolated and vocal sound and consonant based on (certain) word (s). The recording process used Sony recorder with ideal and natural space of recording. The informants were adult male and are able to produce vowel and consonant clearly. The data were analyzed using Praat software (for formants analysis),Visartico application and macro media flash to make the vocal tract that had been designed before.

ANALYSIS AND DISCUSSION

The counting result of f0, f1, and f2 was executed in two steps. Step number one was using Praat software. The next step was inserting the speech sound in to Visartico application to get the two comparisons of f0, f1, and f2. The first result of Praat software analysis is like follow.

1. The pitch curve of an example word 'pergi' ('go') compared to the sound countour curve of vowel /ə/ isolated.

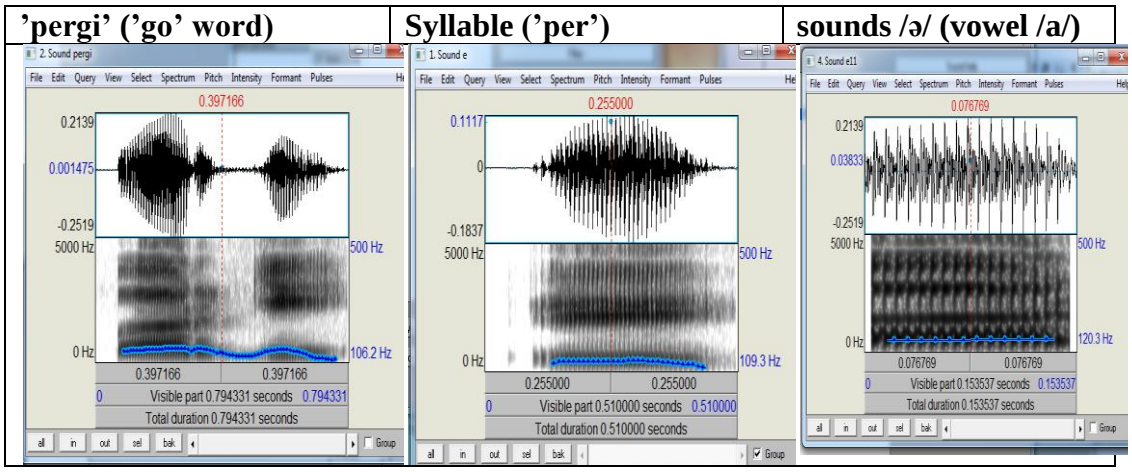

Figure 1: Segmentation process of vowel /ə/

Figure 1 shows the analysis result of accoustic vowel /ə/ or ê (pepet) that was produced with duration of 0.51 second, with f0 is 106.43 hz. The word 'pergi' (go) was segmented and taken its 'per-'syllable for further segmentation to get /ə/ sound. The next step was vowel segmentation of /ə/ was counted to get the mean f0, minimum f0, and maximum f0. Further step was forman analysis by counting f1 and f2 to figure out the characteristic of speech sound in the vowel isolated (gotten from single recording), and vowel produced in context (in word). Through the counting of mean pitch analysis (min f0), minimum f0, and maximum f0 of each vowel sound were presented in following table.

Tabel 1: Comparison of Mean Pitch, Minimum Pitch, Maximum Pitch of vowel (hz)

Vowel			
	mean Pitch	**min pitch**	**max pitch**
/i/(in a word)	125.52	116.06	127.92
/i/ isolated	115.49	85.67	128.42
/a/(in a word)	121.84	120.07	123.28
/a/ isolated	108.31	92.18	120.3
/u/ (in a word)	102.89	87.26	123.97
/u/ isolated	114.35	85.95	124.64
/ə/ /(in a word)	118.84	116.68	121.42
/ə/ isolated	188.69	111.76	498.7
/o/(in a word)	127.73	127.33	128
/o/ isolated	113.41	79.33	135.6

Through the comparison between the isolated vowel sounds and in words vowels sound obtained mean values f0 for each vowel. Characteristics f0 vowel /a/ and /ə/ show the results of a significant difference compared to other vowels. The calculation of f1 and f2 values can be seen in the table 2 below:

Tabel 2: The Comparation of f1 and f2 Value in Vowels Sounds

Vowel	**f1**	**f2**
/i/(in word)	367.32	2482.93
/i/ isolated	426.09	2538.35
/a/(in word)	861.61	1460.681
/a/ isolated	120.3	1334.58
/u/ (in word)	373.21	782.85
/u/ isolated	389.04	74.611
/ə/ /(in word)	635.42	1411.69
/ə/ isolated	484.24	1353.01
/o/(in word)	511.34	1002.57
/o/ isolated	543.6	825.68

Based on the table above, we can see a significant difference in the production of isolated vowels and in words vowels sound. The values obtained in the table above is almost like the table mean f0 value calculation, the value of f1 and f2 from the Indonesian native informant with Javanese language background. Through his research, Zanten measured f0, f1 and f2 in isolation way and put them in the word. The mean values of f1 and f2 are in the following table

Table 3: The comparison of the isolated vowels sound value and in words vowels sound value

Vowel Sound	Mean f1 Isolated	Mean F2 Isolated	Mean f1 Word	Mean F2 Word
/i/	295	2378	316	1988

/ə/	467	1464	492	1130
/a/	868	1389	680	1142
/o/	495	1010	586	970
/u/	299	1005	342	996

Table 3 shows the Characteristic of the isolated vowel sounds and in words vowels sound in Van Zanten's research. Van Zanten concerned his research on the production of vocal tract on Indonesian people from various cultures.

Table 4. isolated f1 and f2
Between Van Zanten and researcher's calculation

Vowel Sound	Mean f1 (VZ)	Mean F2 (VZ)	Mean f1 (P)	Mean F2 (P)
	Isolated	Isolated	Isolated	Isolated
/i/	295	2378	426.09	2538.35
/ə/	467	1464	484.24	1353.01
/a/	868	1389	120.3	1334.58
/o/	495	1010	543.6	825.68
/u/	299	1005	389.04	74.611

Table 5: The value f1 and f2 in the word
Between Van Zanten and researcher's calculation

Vowel Sound	VZ	VZ	P	P
	Mean f1	Mean F2	Mean f1	Mean F2
	Word	Word	Word	Word
/i/	316	1988	367.32	2482.93
/ə/	492	1130	635.42	1411.69
/a/	680	1142	861.61	1460.681
/o/	586	970	511.34	1002.57
/u/	342	996	373.21	782.85

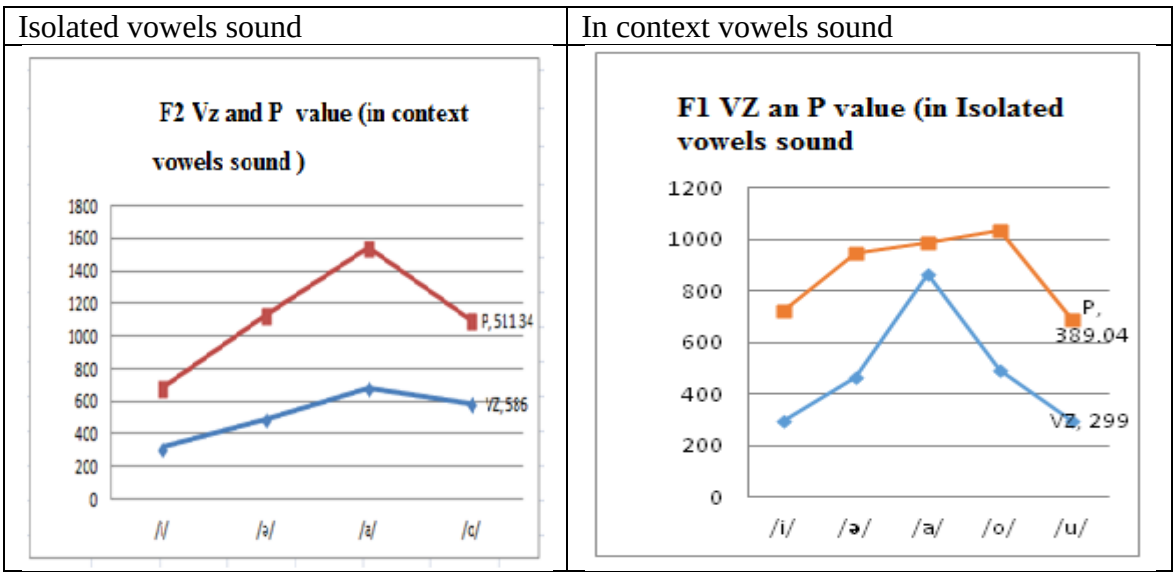

Figure 2: The Comparison value of f1 in isolated vowels sounds and in context vowels sound

Tables 4, 5, and figure 2 explains the tendency of almost equal value between the calculations performed by Van Zanten and calculations have been carried out by investigators. Vowel /a/ tends to provide different value for f1 and f2 between the production of isolated vowels sounds and in words vowels sound. Van Zanten (1989) had examined the characteristics of vowels and consonants which are produced by speakers from the Javanese, Sundanese and Batak Toba.

DISCUSSION

The fundamental frequency is also known as f0 which is coherent in form of formant transition f1, f2, and so forth. The dominant frequency components that characterize phonemes associate with the resonance frequency component of the system, it is further defined as a vowel formants. The spoken voices, in particular are vowels, usually have 3 pieces formants and often referred to as the first, second, and third formant. It is starting with the lowest frequency component. All three are always written as f1, f2 and f3. Formant 4 and 5 are needed to get formant parameter values to be more detail. Since if we only have sound signal less than 3 pieces, then certainly the analysis of these data will fail (Zanten, 1989). Ladefoged (1975: 173) explains that the lowest formant frequencies, f1, in reverse figures an altitude vocals; the second formant, F2, (or just the difference between the frequencies f1 and f2) in line with the level to back in the traditional vowel diagram and sometimes until f3. This is decisive in differentiating vowels from one another. The third center frequencies are measured by using narrow band spectrograms (Kay Sonograph, 6061), filtered with bandwidths of 50 hz wave which are created in the mid vowel. On 1 mm a scale or equal with frequency axis 82 hz.

Phoneticians experimental stated that vocal quality is determined by the center frequencies of the formants, means the resonant voice channel. The lowest formant frequency is F1, it is the inverse ratio if it is associated with a vowel high based on traditional vowel diagra. The second formant, F2 is illustrated in inverse diagram as backwardness vowels compared with those vowel. Therefore the lowest vowel, namely F1, F2, or F3 is a determinant in differentiating vowels from one another. The third center frequencies of formant analyzed and determined from spectrograms through Linear Predictive Coding (LPC) and by analyzing the time spectrum. Here is the production of vowel /a/ use VisArtico application. Through visArtico production, vowel sounds are clearer and easier to imagine. The tongue movement also supports the clarity of position of the tongue.

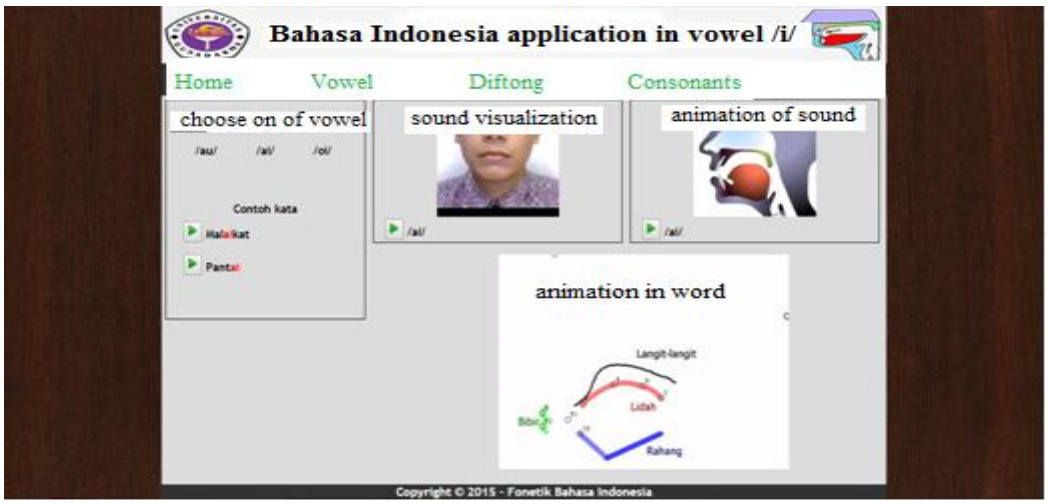

Figure 3: Illustration of the production of the vowel /i/ by using VisArtico application

Is actually rather difficult to find the overlap between the realizations of in word spoken vowels, despite all kinds of realization each vowel has been calculated. The distribution of /e/, /o/ and /u/ which are spoken isolation and the greater overlapping for isolated vowels are much clear, especially in vowel /a/, and some appear in the distribution of another vowel, as in /e/. Through the application of Vis Artico, researcher tries to present a form of computation f1 and f2 in the form of animation. By using the same word example, the word 'pergi'(go), the researcher obtained propensity score that similar to the calculation results of the Praat application. In acoustic vowel, it is clearly seen the formants center frequencies are lower, near overton group amplified by the resonance characteristics of the measured vocal tract.

CONCLUSION

Through the analysis of production, has performed the measurement values of f0, f1 and f2 on Indonesian native informants. Vowels measured were 6 vowels, that is, /a/, /i/, /u/, /ə/ and /o/. Vowel sounds were measured in isolation and in words which are then compared. The comparison showed the same propensity score. Researcher also compared the measurements that have been conducted by researcher himself and Van Zanten measurement results with the same variable. Visually, the measurement of f1 and f2 were not possible been done, but by using VisArtico and Praat, the calculation of vocal tract has been successfully carried out and have obtained a precise value. However, it is still needed to do a different calculation for the isolated vowels or in words vowels. Phonologically vocal quality is specified by the levels of openness mouth. If all variables are constant, the more wide-open mouth and their articulatory movement, the vowel will be much longer. To measure the length of isolated vowels, for more detailed analysis, it can use oscillograms tool.

REFERENCES

Boersma, P., & Weenink, D. (2009). Praat: Doing phonetics by computer. (version 5.1.05).

Clark, J., and Yallop, C. (1995). *Phonetics and Phonology.* Oxford: Blackwell. Ch 7.1-7.13.

Denes, P.B., and Pinson, E.N. (1973). *The Speech Chain.* Murray Hill, NJ: Bell Telephone. Ch 3, 4.

Johnson, K. (1997). *Acoustic and Auditory Phonetics*. Oxford: Blackwell. Chs. 1, 2, (3), 4, 5.

Hayward, K. (2000). *Experimental Phonetics.* London: Longman. Chs. 2-4.

Borden, J. & Harris, K.S. (1980) *Speech Science Primer.* Baltimore: Williams & Wilkins. Ch 3; ch 4 pp 89-130.

Fry, D.B. (1979). *The Physics of Speech.* Cambridge: CUP. Ch 1-9.

M. N. Al-Azhar. (2001). Audio Forensic: Theory And Analysis. Pusat Laboratorium Forensik Polri Bidang Fisika dan KKomputer Forensik.

Pickett, J.M. (1999). *The Acoustics of Speech Communication: Fuandmentals, Speech Perception Theory, and Technology*. Needham Heights, MA: Allyn & Bacon. Ch. 2-4

Or Pickett, J.M. (1980) *The Sounds of Speech Communication.* Baltimore: University Park Press. Ch 1-4.

Rosen, S. and Howell, P. (1991). *Signals and Systems for Speech and Hearing.* Academic Press. Clearly written and relatively nontechnical, but much more detail than the above books. Try it if you enjoy more scientific approaches. Relevant sections are spread over several chapters—find them in Table of Contents & Index.

Liberman, A. M. (1954). *The Role of Consonant-Vowel Transitions in the Perception of the Stop and Nasal Consonants*. Washington, DC: American Psychological Association.

Palmer, A., & Shamma, S. (2004). Physiological Representations of Speech. In S. Greenberg, W. A. Ainsworth, A. N. Popper, R. R. Fay (Eds). *Speech Processing in the Auditory System: Springer Handbook of Auditory Research* (Vol, 18, pp, xiv, 476). New York: Springer.

Poedjosoedarmo, Soepomo. 1982. Javanese influence on Indonesian. Canberra: Department of Linguistic, Research School of Pasific Studies, The Australian National University (Pasific Linguistics, Series D, no. 38; Materials in Languages of Indonesia, 7).

Zanten, van. 1989. Vokal-vokal Bahasa Indonesia. Jakarta: Balai Pustaka.