ISSN 2229-5518

# Determining the possible Noise Input from the amplitude and intensity of each sound wave in Speech Signal

Nelson O Madu, Obikwelu R Okonkwo

**Abstract:**

In this paper, an efficient method of modelling possible noise from amplitude and Intensity of each sound wave in speech signal is proposed. Here the focus is on speech signals which manifests as a sound wave form. The principle of determining sound wave intensity using the wave amplitude was applied. The paper elucidates how beamforming technique was practical in the evaluation of the Amplitude of the signals and the various noise inputs given the distance variations. The obtained result shows that noise input increases as the angular distance increased which is as a result of amplitude increase

**Keywords:** *amplitude, intensity, beamforming, sound wave, microphones, signals*

## 1. INTRODUCTION

In a speech automated system, microphone is installed far away from the user; as such there are severe problems associated with it such as poor sound quality and acoustic feedback from the far-end side. In other words, Speech transmitter and receiver are located at remote places separated by certain distance. In hand-held telephony the microphone is close to the speaker which suffers from environmental noise and interfering sounds corrupting the received speech signal. These conflicts additionally impair received speech signal which is due to the reverberations of the voice from the walls or ceilings.

The recorded speech signals in speech automated systems are often corrupted by acoustic background noise. Generally, background noise is broadband and non-stationary. The signal to noise ratio of microphone is low. Therefore, the speech quality and intelligibility reduce.

## 2. REVIEW OF RELATED LITERATURE

In recent years, tremendous and remarkable progress has been made in the domain of speech and noise separation. The pioneering work of Wiener et al. gives an optimum approach [1,5,6] for deriving a filter that tends to suppress the noise while leaving the desired signal relatively unchanged. The design of these filters requires that the signal and the noise be stationary, and the statistics of both signals be known a priori. In practice these conditions are hardly ever met.

[4] speech-in-noise research showed that speech signals combine several acoustic properties that contribute to compensating for signal distortions and noisy interferences. In [3], it was revealed that Paul Lueg was the first to realize the possibility of attenuating background noise by superimposing a phase flipped wave for the concept of active noise cancellation. Also, this approach requires spatial selectivity. To achieve this, [2], noted that the beamformer needs to distinguish between the components from different directions of an incoming signal in order to suppress hugely and efficiently the parts of the signal that are not coming from the target source(s). Consequently, it becomes necessary to analyse the incoming signal and differentiate between different angles of incidence. Achieving this requires a beamforming technique. In practical beamforming, when the microphone array picks up a signal coming from an angle other than 0 or 180, every consecutive microphone will experience an increased delay. This is because the signal entering from an angle needs to travel an additional distance, Δx, to the next microphone in the array. Δx is proportional to the distance between the microphones and the angle of incidence, as has been expressed in figure 2.1

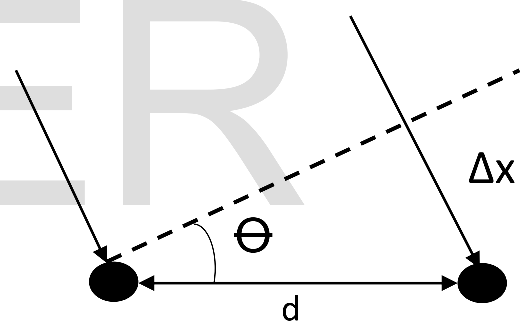


Figure 2.1 - The signal travels an additional distance Δu to the next microphone [2]

Given that the microphones are fixed, the distance between them is fixed too. It is therefore possible to relate the additional distance to the angle of incidence of the signal. However, since the speed of sound is also fixed, [2] noted that the time delay can be deduced from the direction of the signal by

$$t = \frac{d}{C}\sin(\theta): \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \textit{(2.1)}$$

Where t is the time delay in [s] and d, are the distance between the microphones in [m], c the speed of sound in [m/s] and θ the angle between the microphones in [rad].

For all linear systems, the net response caused by two or more stimuli is the sum of the responses that would have been caused by each stimulus individually. This is known as the superposition principle. Therefore, if input $A$ produces response $X$ and input $B$ produces response $Y$ then input $(A + B)$ produces response $(X + Y)$.

A function $F(x)$ that satisfies the superposition principle is called a linear function. Superposition can be defined by two simpler properties; additivity and homogeneity; however, our interest lies on the additivity.

$$F(x_1 + x_2) = F(x_1) + F(x_2) \qquad (2.2)$$

Waves are usually described by variations in some parameter via space and time. The value of this parameter is called the amplitude of the wave, and the wave itself is a function specifying the amplitude at each point. In many cases such as classic wave equation, the equation describing the wave is linear. Once this is real, the superposition principle can be applied. As expected, the undesired sine has amplitude larger than the amplitude of the desired sine.

## 3. METHODOLOGY

In this work the focus is on speech signals and the noise input which manifest as a sound wave form. In signal processing, delays are applied to control the signals coming from all microphones to hit a given point, called focal point. The signals received by the elements (raw signal data) are delayed such that they sum up signals coming from the same directions (correlated signals) constructively and destructively sum those signals coming from different directions (uncorrelated signals).

The Figure 3.1 below shows a vector superposition of a large amplitude N of simple harmonic vibration of equal amplitude A and equal successive phase difference δ.

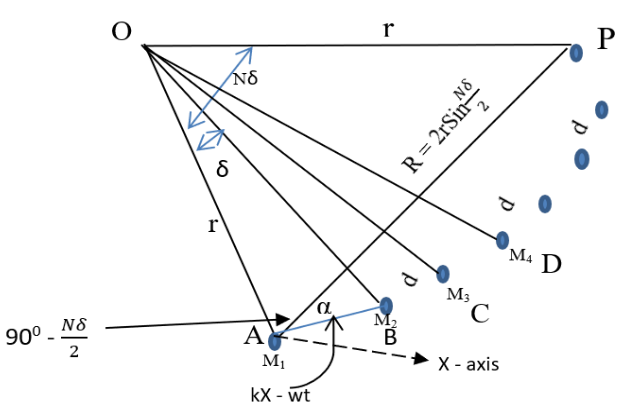


*Figure 3.1 - Vector Superposition*

The amplitude of all the resultant (ie Chord)

$$R = 2\,r\,\text{Sin}\,\frac{N\delta}{2} = A\,\frac{Sin\frac{N\delta}{2}}{Sin\frac{\delta}{2}} \qquad (3.1)$$

Consider the triangle AOP. Splitting the triangle AOP into two right angles, we have;

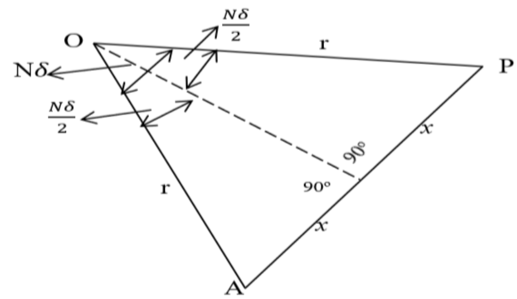


*Figure 3.2 – Splitting the Triangles for Microphone positions*

Therefore, for one of the Right-Angled triangles,

$$\text{Sin}\frac{N\delta}{2} = \frac{x}{r} \qquad (3.2)$$

Where $x$ is Opposite, and r is the Hypotenuse.

$$X = r\,\text{Sin}\frac{N\delta}{2} \qquad (3.3)$$

Then, for both right angled triangles,

$$x + x = r\,\text{Sin}\frac{N\delta}{2} + r\,\text{Sin}\frac{N\delta}{2} = 2r\,\text{Sin}\frac{N\delta}{2} \qquad (3.4)$$

therefore, $R = 2r\,\text{Sin}\frac{N\delta}{2}$ ……………………………(3.5)

Its phase with respect to the first contribution is given by $\alpha$;

To get the $\alpha$ which is angle between the microphones ($M_1$); we subtract the angle OAP from angle OAB. However, first let's calculate the two angles. From the previous, OÂP is gotten, by subtracting $\frac{N\delta}{2}$ from $90^o$;

$$O\hat{A}P = 90^o - \frac{N\delta}{2} \qquad (3.6)$$

For angle OÂB ;

ISSN 2229-5518

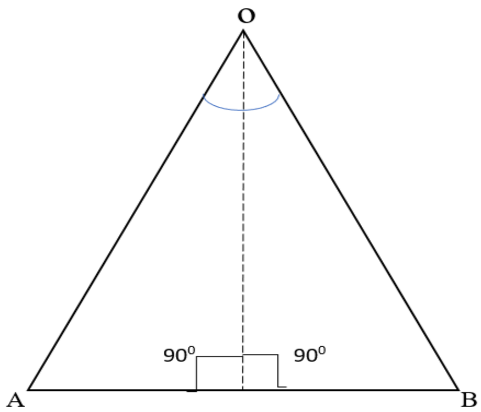


*Figure 3.3(a) - Delayed distance*

$O\hat{A}B = 90^{o} - \frac{\delta}{2}$ ……………………………………….*(3.7)*

Therefore, $\theta = O\hat{A}B - O\hat{A}P$

$= 90^{o} - \frac{\delta}{2} - (90^{o} - \frac{N\delta}{2})$……………. *(3.8)*

$= 90^{o} - \frac{\delta}{2} - 90^{o} + \frac{N\delta}{2} = \frac{N\delta}{2} - \frac{\delta}{2} = (N-1)\frac{\delta}{2}$…..*(3.9)*

Therefore $\alpha = (N-1)\frac{\delta}{2}$ ……………………..……*(3.10)*

From triangle OAP, following arrangement of microphones symmetry on y-axis we use the Cosine rule

$R = r^2 + r^2 - 2r^2 \text{ Cos } N\delta$ ………………….*(3.11)*

$= 2r^2(1 - \text{Cos}\delta)$ ………………….*(3.12)*

$R^2 = \frac{4r^2(1 - \text{Cos}N\delta)}{2}$ …………………….*(3.13)*

$R = \frac{2r\sqrt{1 - CosN\delta}}{2}$ ……………………..*(3.14)*

$R = 2r\text{Sin}\frac{N\delta}{2}$ …………………………..…*(3.15)*

From triangle OAB, it can be seen that delay distance from the speech source to the microphone is $dSin\,\theta$ where $\theta$ is the angle substituted by the microphones within an area

Speed of Sound (c) $= \frac{distance(\text{dSin}\frac{(N-1)\delta}{2})}{time(t)}$ …...*(3.16)*

Therefore, $t = \frac{\text{dSin}\frac{(N-1)\delta}{2}}{c}$ ………………..…..*(3.17)*

Where $c$ is the Speed of Speech = 343m/s

Also, from triangle OAB using same Cosine rule;

$A = 2r\text{Sin}\frac{\delta}{2}$ …………………………...…*(3.18)*

**3.1 Modelling the Intensity and Amplitude in beamforming System**

Consider a speech from a source to different microphones with a reference point.

Studying the graph on Figure 3.4(b) which is a graph showing different microphone at different distances, we would understand that at a reference point, the phase difference can be used to acquire the amplitude as in Figure 3.4(a).

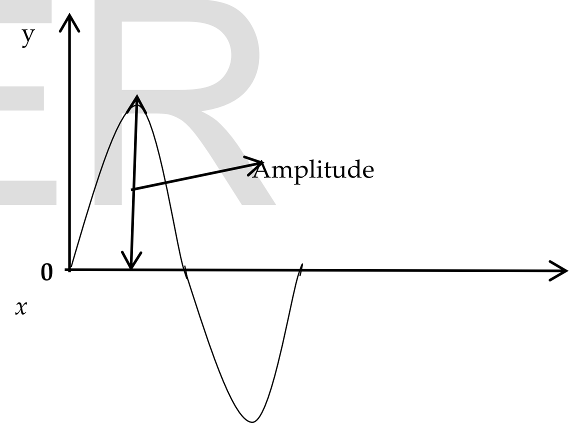


***Figure 3.4 (a) - Amplitude of a speech signal***

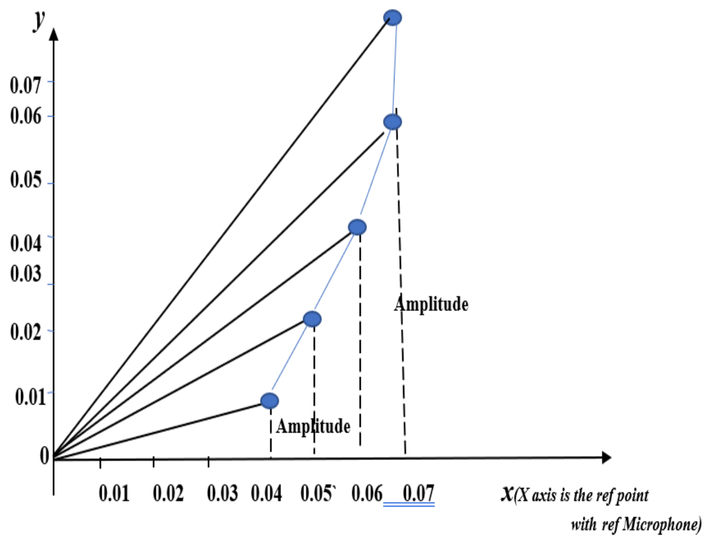


*Figure 3.4(b) - Graph of the amplitudes*

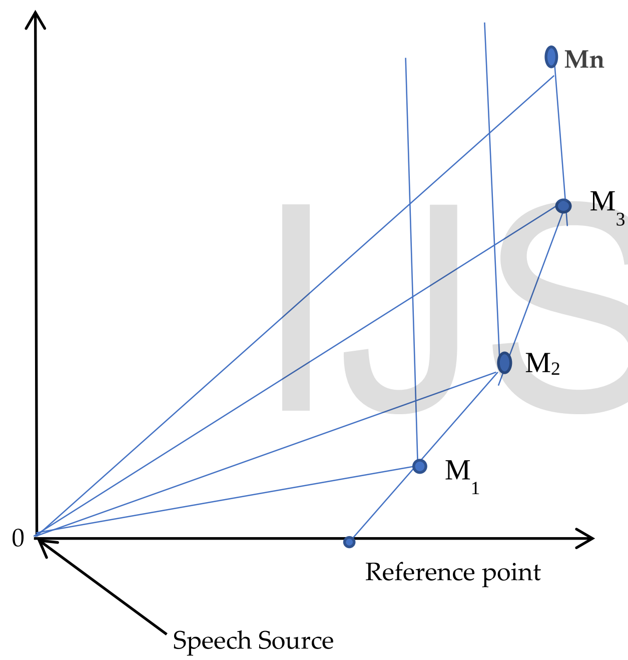


*Figure 3.5 – Array of microphones*

By defining, the average amount of energy passing through a unit area per unit of time in a specified direction is called the intensity of the wave. Thus, as the amplitude of the sound wave increases, the intensity of the sound increases. The intensity of a sound wave is related to its amplitude squared.

$I = A^2$ ……………………………..*(3.19)*

From Figure 3.6 above, there is an array of microphone – $M_1$, $M_2$, $M_3$, $M_4$ --- Mn with a reference point on the $X$ –axis. Taking a distance of 0.015m between all microphones, we calculate the amplitude and intensity of each sound wave to estimate the increase in amplitude and possible noise input.

From, equation (3.20), above ; $A = 2r\text{Sin}\frac{\delta}{2}$

For microphone $M_1$, with a radian distance of 0.06m and phase difference of 10;

$A = 2 \times 0.06 \times \text{Sin}\left(\frac{10}{2}\right) = 0.010458\omega$

$\text{Intensity} = A^2 = (0.010458)^2 = 0.0001094\ \frac{10}{m^2}$

Converting to decibel

$\text{Dbel} = 10\text{Log}_{10}\left(\frac{I}{I_c}\right) = 10\text{Log}10\left(\frac{0.0001094}{10^{-12}}\right)$

$= 80.4\text{dbel}.$

Now for Microphone $M_2$, it is important to note the increase in radia distance of the speech source. This distance must first be calculated.

Recall from equation (3.12) ; $\alpha = (N\text{-}1)\frac{\delta}{2}$;

To get increase in distance,

$X = d\text{Sin}\alpha = d\text{Sin}(N\text{-}1)\frac{\delta}{2}$

Distance between both Microphones is 0.015

Going by our initial estimation, distance increase,

$X = d\text{Sin}(N\text{-}1)\frac{\delta}{2}$

$= 0.015\text{Sin}\ (2\text{-}1)\frac{10}{2}$

$= 0.001307$

Therefore, complete distance from the speech source to microphone, $M_2$ is

$0.06 + 0.001307 = 0.061307$

then, from $A = 2r\text{Sin}\frac{\delta}{2}$ we would obtain;

$= 2 \times 0.0061307 \times \text{Sin}\frac{10}{2} = 0.01069$

$I = A^2 = (0.01069)^2 = 0.0001142\ \frac{10}{m^2}$

Converting to decibel

$\text{Dbel} = 10\text{Log}_{10}\left(\frac{I}{I_0}\right) = 10\text{Log}_{10}\left(\frac{0.0001142}{10^{-12}}\right)$

***= 80.58dbel***

For $M_3$ distance increase $x = d\text{Sin}(N\text{-}1)\frac{\delta}{2}$

$= 0.03\text{Sin}\ (3\text{-}1)\ \frac{10}{2}$

` $= 0.005209$

$M_3 = 0.06 + 0.005209 = 0.065209$

$A = 2r\text{Sin}\frac{\delta}{2}$

$= 2 \times 0.065209 \times \text{Sin}\frac{10}{2} = 0.01137$

$I = A^2 = (0.01137)^2 = 0.0001292769\ \frac{10}{m^2}$

Dbel $= 10\text{Log}10\,(\frac{I}{I_o}) = 10\text{Log}10\,(\frac{0.0001292769}{10^{-12}}) =$ ***81.115dbel***

This shows an increase of 0.54dbel noise addition over the distance increase.

## 4. CONCLUSION

Overall, it is possible to estimate the level of noise in a mix of sound via the determination if the amplitude of the sounds. The obtained result is indicative of the fact that noise input increases as the angular distance increased (amplitude) increase

## REFERENCES

[1] Brady N. M. Laska, MiodragBolic, and Rafik A. Goubran, "Practical Filter Enhancement of Speech Spectral Amplitude" , IEEE Transactions on audio, speech and language processing, Vol 18, No. 8, November 2010

[2] Ganesh T., Sekhar Rao R.V.Ch., Prasad P.H.V; Ashok kumar N, Vasundhra P. (2016) ; *'Noise Reduction Using Beamforming Algorithms'* - International Journal of Engineering Science and Computing, April.

[3] Harris, Bill (2007). "How Noise-canceling Headphones Work." How stuff works. 16 July http://electronics.howstuffworks.com/noise-canceling-headphone.htm

[4] Meyer J, Dentel L ; Meunier F (2013), Speech Recognition in Natural Background Noise Published: November 19, {https://doi.org/10.1371/journal.pone.0079279)

[5] Khalifa, Z. H Ahmad, A-H A. Hashim And T. S. Gunawan "SMaTalk: Standard Malay Text to Speech Talk System", Vol 2 no 5.

[6] Welker, D.P., Greenberg, J E., Desloge, J.G., and M,. Z. P. "Microphone-array hearing aids with binaural output-part II: A Two microphone adaptive system," IEEE Trans. Speech Audio Processing, 1997



**Authors:**

**Obikwelu R Okonkwo**
Department of Computer Science; Nnamdi Azikiwe University, Awka Nigeria.
*ro.okonkwo@unizik.edu.ng*

**Nelson O Madu**
Department of Computer Science; Nnamdi Azikiwe University, Awka, Nigeria.
*noge_m2@yahoo.com*