**SilenSwear: An Automatic Swearword Detection**

**and Muting Video Player using Limited Vocabulary**

**Continuous Speech Recognition**

A Thesis

Presented to the Faculty of the

College of Computer and Information Sciences

Polytechnic University of the Philippines

In Fulfilment

Of the Requirements for the Degree

Bachelor of Science in Computer Science

Computer Science Research

Medrano, Jamie Paul R.

Sefe, Marlon M.

October 2015

Table of Contents

[List of Figures iii](#_Toc430610761)

[Chapter 1: The Problem and Its Background 1](#_Toc430610762)

[1.1 Introduction 1](#_Toc430610763)

[1.2 Background of the Study 1](#_Toc430610764)

[1.3 Conceptual Framework 2](#_Toc430610765)

[1.3.1 Conceptual Framework of the System 2](#_Toc430610766)

[1.3.2 Conceptual Framework of the Study 3](#_Toc430610767)

[1.4 Statement of the Problem 3](#_Toc430610768)

[1.5 Scope and Limitation 3](#_Toc430610769)

[1.6.1 Scope and Limitation of the System 4](#_Toc430610770)

[1.6.2 Scope and Limitation of the Study 4](#_Toc430610771)

[1.6 Significance of the Study 4](#_Toc430610772)

[1.7 Definition of Terms 4](#_Toc430610773)

[Chapter 2: Review of Related Literature and Studies 6](#_Toc430610774)

[2.1 Review of Related Literature 6](#_Toc430610775)

[2.2 Review of Related Studies 6](#_Toc430610776)

[Chapter 3: Research Methodology 8](#_Toc430610777)

[3.1 Research Method Used 8](#_Toc430610778)

[3.3 System Architecture 8](#_Toc430610779)

[3.4 Statistical Treatment 8](#_Toc430610780)

[REFERENCES 11](#_Toc430610781)

# List of Figures

[Figure 1.1 Conceptual Framework of the System 3](C:\\Users\\Medrano\\Documents\\THESIS TOPICS\\Chapter 1-3.docx" \l "_Toc430610782)

[Figure 2 Conceptual Framework of the Study 3](file:///C:\Users\Medrano\Documents\THESIS%20TOPICS\Chapter%201-3.docx#_Toc430610783)

[Figure 3 System Architecture 8](file:///C:\Users\Medrano\Documents\THESIS%20TOPICS\Chapter%201-3.docx#_Toc430610784)

Chapter 1: The Problem and Its Background

## 1.1 Introduction

Foul language or swear words in Hollywood films are very common. Its use rose dramatically in the recent years which made parents worry about the movies their children watch. Exposure to profane language can trigger violence in children, therefore it is dangerous. Because of this problem an automatic swear word filtering is needed to ensure that children won’t be exposed to bad words.

## 1.2 Background of the Study

The rate of usage of foul language or swear words in Hollywood films is increasing yearly. Nowadays, it is common to hear dozens of swear words in just a single movie. [[Warner 2014](#War14)] According to the researchers of the website vidangel.com, the average swear words spoken in movies rose from 5 words per movie in 1945 to 67 words per movie in 2013 [[VidAngel Inc. 2013](#Vid13)]. Profanity in films promotes offensiveness, verbal abuse, hatred, disrespect, and lack of empathy toward other people [[Howard 2013](#How13)].

Exposure to obscenity and profanity in media is correlated with aggression and violence. A study by researchers at Brigham Young University in 2011 revealed that the more children were exposed to profanity, the more likely they were to use swear words themselves, and the more they use profanity, the more likely they become aggressive toward others [[Park 2011](#Par11)]. This is a disturbing truth especially for parents who doesn’t want their children to be exposed to profanity and become violent.

In the Philippines, foreign films holds 80 percent of the domestic market, mostly are Hollywood films, and only 20 percent are local films, which means that Filipinos watch more Hollywood-made movies than local-made ones. While there are TV parental guidelines that offer ratings of television shows and movies based on language, sex and violent content, these are often not followed and many children even has access to movies with contents of profanity [[Agence France-Presse 2012](#Age12)].

There are various ways to filter profanity from films. You can remove swear words by using a video editor and manually mute swear words before children can watch it [[Watts 2012](#Wat12)], but this is a tedious and long work. Another solution is using sites that offer movie streaming with content filtering for profanity and foul language [[Warner 2014](#War14)], but these services are usually expensive and considering the average internet speed in the Philippines, this is considered to be of little help. Another is by using a device called TV Guardian that continuously monitors for foul language by reading closed-captions data, just slightly ahead of the spoken words [[TVGuardian 2015](#The15)], but without closed captioning it is not possible to filter out the swear words, and also this is unavailable in the Philippines. These solutions are unavailable up to some extent and costly.

Automatic speech recognition is the process by which a computer maps an acoustic speech signal to text. A speaker independent system is developed to operate for any speaker of a particular type [[cmu.edu](#Wha15)]. To recognize speech, the waveform must be taken, split on utterances by silences then try to recognize what's being said in each utterance. The model of speech is called Hidden Markov Model or HMM, a model that has been proven to be really practical for speech decoding [[CMUSphinx 2015](#CMU15)].

By using automatic speech recognition, a system can be created to automatically detect and mute swearwords. This is important to allow filtering profanity in movies. It is specifically useful to parents who want to protect and lessen the swearword exposure of their children from films.

## 1.3 Conceptual Framework

### 1.3.1 Conceptual Framework of the System

Figure 1.1 shows the conceptual framework of the System where the swearword audio, the movie and extracted audio are inputs. The processes are swearword feature extraction, swearword detection, and swearword muting. The outputs are swearword feature collection and muted swearwords from movie.

Figure 1.1 Conceptual Framework of the System

**Input**

*Training:*

* Swearword audio

*Simulation:*

* Movie
* Extracted audio

**Process**

*Training:*

* Swearword feature extraction

*Simulation:*

* Swearword detection
* Swearword muting

**Output**

*Training:*

* Swearword feature collection

*Simulation:*

* Automatically muted swearwords in movie

### 1.3.2 Conceptual Framework of the Study

Figure 1.2 shows the conceptual framework of the study wherein the developed system and the experiment paper are inputs. The development, implementation, experimentation and analysis produces the output of the study which is the accuracy rate and reliability of the system in detecting and muting swear words.

Figure 2 Conceptual Framework of the Study

**Input**

* Developed system
* Implementation

**Process**

* Development
* Implementation
* Experimentation
* Analysis

**Output**

* Accuracy rate and reliability of the system in detecting and muting swear words.

## 1.4 Statement of the Problem

The study aims to automatically detect and mute foul words using Automatic Speech Recognition with Hidden Markov Model and plans to answer the following questions:

1. What is the accuracy of the system in detecting and muting foul words from movies?
2. What is the reliability of the system in detecting and muting foul words from movies?

## 1.5 Scope and Limitation

### 1.6.1 Scope and Limitation of the System

The system will be deployed as a video player. The system will be dependent of the tools such as Sphinx4 and SphinxTrain. The system that will be developed will be dependent on the Latest Java Virtual Machine and Runtime Environment.

The algorithms to be used is Hidden Markov Model for the training and decoding and of the swearwords in movie audios.

### 1.6.2 Scope and Limitation of the Study

The time frame for the development of the system is an estimated time of 3 months. The study only includes the swearwords in a Hollywood movie.

## 1.6 Significance of the Study

SilenSwear focuses on detecting and muting swearwords on a movie. The system will benefit the following:

**Children –** these are people below 18 that are restricted in watching and exposure to movies with high level of profanity in language.

**Parents –** this will benefit parents in preventing exposure of their children to profanity.

**General Audience –** these are people with different preferences and opinions about swearing in movies.

**Future Researchers –** These are the people who will conduct future research to improve the existing studies about speech recognition.

## 1.7 Definition of Terms

**Bad Words -** words that are said in order to make someone or something look or feel bad.

**Continuous Speech Recognition -** speech recognition of continuous stream of words.

**Foul Words or Foul Language -** unpleasant words to hear.

**Hidden Marcov Models -** s a statistical Markov model in which the system being modeled is assumed to be a Markov process with unobserved states. A HMM can be presented as the simplest dynamic Bayesian network.

**Hollywood Films -** Movies Made in Hollywood.

**Limited Vocabulary Speech Recognition -** limited translated text of spoken words.

**Profanity or profane Words -** Offensive words or phrases.

**Speech Recognition-** the translation of spoken words into text.

**Swear Words -** an Expression of anger and an offensive word.

# Chapter 2: Review of Related Literature and Studies

## 2.1 Review of Related Literature

Sangeeta Danhay Chaubal said that watching movies has known to have both positive as well as negative impact on a child's mind as well on his overall development. The author also said that movies with social message or pure fun can be worth watching but movies with too much of violence and mean messages are of no good to children [[Chaubal](#Cha)].

Ronald E. Riggio Ph.D. said that it is clear that swearing can take form of verbal abuse and harassment, and it is this sort of swearing that can be potentially harmful- a substitute for physical aggression. According to the researcher hearing swear words in a movie or TV show or on playground harm children not because of swearing itself but the factors associated with swearing [[Riggio 2012](#Rig12)].

"Because many films transmit ideas through emotion rather than intellect, they can neutralize the instinct to suppress feelings and trigger emotional release,” said Birgit Wolz, a psychologist focusing on movies as therapy, and author of E-motion Picture Magic [[Braff 2011](#Bra11)].

According to Dr. Diane Howard, family sitting down to watch a movie is shot with paintballs every time a swear word is uttered on screen. Every word has impact [[Howard 2013](#How13)].

A Child developmental capacity for abstraction is a crucial factor in appreciating his or her capacity to understand a movie. A related but separate factor is the child's capacity for moral reasoning. Jean Piaget discovered that children younger that eight or nine had trouble judging people on their intentions. Rather, when read hypothetical stories, younger children would judge people on the final outcome of their actions [[Minnow 2014](#Min14)].

## 2.2 Review of Related Studies

Automatic speech recognition is the process by which a computer maps an acoustic speech signal to text. A speaker independent system is developed to operate for any speaker of a particular type [[cmu.edu](#Wha15)]. To recognize speech, the waveform must be taken, split on utterances by silences then try to recognize what's being said in each utterance. The model of speech is called Hidden Markov Model or HMM, a model that has been proven to be really practical for speech decoding [[CMUSphinx 2015](#CMU15)].

Automatic continuous speech recognition (CSR) has many potential applications including command and control, dictation, transcription of recorded speech, searching audio documents and interactive spoken dialogues. The core of all speech recognition systems consists of a set of statistical models representing the various sounds of the language to be recognized. Since speech has temporal structure and can be encoded as a sequence of spectral vectors spanning the audio frequency range, the hidden Markov model (HMM) provides a natural framework for constructing such models [[Eagon](#Eag)].

# Chapter 3: Research Methodology

## 3.1 Research Method Used

The proponents will be using descriptive research as its research methodology to test the accuracy and reliability of the SilenSwear.

## 3.3 System Architecture

Figure 3 System Architecture

Video Player

Movie

Swear words audio

Feature extractor

Speech decoder

Audio extractor

Swearwords

Swear word detection

.dat File

Movie with muted swear words

The system accepts swear words audio for feature extraction and Hidden Markov Model training. The audio of the Movie extracted and its features are decoded for swear words detection. It will now produce a .dat file for the video player. The video player requires both the movie and the .dat file because the .dat file contains the time when the sound will be muted.

## 3.4 Statistical Treatment

**Percentage**

The percentage computes the percentage of detection per movie.

P = (Total No. of Swear words identified / Total No. of Swear words) \* 100

**Accuracy Rate**

This is used to determine the accuracy rate of the system by computing the true positive, false negative, true negative, and false negative with the use of the observation sheet for accuracy.

AR = ( tp + tn ) / ( tp + tn + fp + fn )  
where:  
tp – true positive (total number of correctly identified swear words)  
tn – true negative  
fp – false positive (total number of incorrectly identified Swear words)  
fn – false negative (total number of unidentified Swear words)

**True Negative Rate**

This is used to get the correct absence of result

tn = tp / ( tp + fp )

**Recall**

The portion of relevant characters that are not recognized by the system yet present in the database.

**F-score**

The weighted average mean of recall and precision. The F-Score evaluates the system performance.

**Error rate**

The number of incorrectly identified faces over the correctly identified faces.

# **REFERENCES**

x

|  |  |
| --- | --- |
| [Agence France-Presse 2012] | AGENCE FRANCE-PRESSE. 2012. Philippine film industry decline. from *Inquirer.net*: http://entertainment.inquirer.net/30455/philippine-film-industry-in-decline |
| [Braff 2011] | Braff, Daniel. Movies may cause special effects on the body (June 2011). |
| [Chaubal] | Chaubal, Sangeeta Danhay. How watching movies impact children. |
| [cmu.edu] | CMU.EDU. What is speech recognition? from *Carnegie Mellon University*: http://www.speech.cs.cmu.edu/comp.speech/Section6/Q6.1.html |
| [CMUSphinx 2015] | CMUSPHINX. 2015. Basic Concepts of Speech. from *CMUSphinx*: http://cmusphinx.sourceforge.net/wiki/tutorialconcepts |
| [Daily Mail] | DAILY MAIL. Too much bad language on TV. from *Daily Mail Online*: http://www.dailymail.co.uk/tvshowbiz/article-188713/Too-bad-language-TV.html |
| [Eagon] | Eagon, L. E. Baum and J. A. An inequality with applications to statistical. *Bulletin of American Mathematical Society*, 73, 360-363. |
| [Howard 2013] | Howard, Diane. 2013. A Media-Wise Report on Foul Language. from *Movie Guide*: http://www.movieguide.org/news-articles/media-wise-report-foul-language.html |
| [Minnow 2014] | Minnow, Nell. *The Movie Mom's Guide to Family Movies 2nd Edition*. 2014. |
| [Park 2011] | Park, Alice. 2011. Children Who Hear Swear Words on TV Are More Aggressive. from *Time*: http://healthland.time.com/2011/10/17/children-who-hear-swear-words-on-tv-are-more-aggressive/ |
| [Riggio 2012] | Riggio, Ronald E. Will Swearing Harm your Child? (May 2012). |
| [TVGuardian 2015] | TVGUARDIAN. 2015. The TVGuardian Mutes Foul Language From TV Shows and Movies. from *TVGuardian*: http://www.tvguardian.com/ |
| [VidAngel Inc. 2013] | VIDANGEL INC. 2013. Hollywood vs. YouTube. from *VidAngel*: https://www.vidangel.com/hollywood-vs-youtube/swearing/ |
| [Warner 2014] | Warner, Russ. 2014. Too Much Profanity in Movies? There’s a Great New Solution. from *Net Nanny*: https://www.netnanny.com/blog/too-much-profanity-in-movies-theres-a-great-new-solution/ |
| [Watts 2012] | Watts, Gregory. 2012. Remove The Curse Words. from *Videomaker*: http://www.videomaker.com/community/forums/topic/remove-the-curse-words |

x