

UNDERGRADUATE THESIS

**GRAMMATICAL ERROR ANALYSIS ON ENGLISH SUBTITLES OF
NOMADLAND (2020) MOVIE PRODUCED BY SONIX AND MAESTRA
ONLINE AUTOMATIC SPEECH RECOGNITION SOFTWARE:
A COMPARATIVE STUDY**



By

IFTITA DAMAI AGITASARIA

NIM 121811233054

**ENGLISH DEPARTMENT
FACULTY OF HUMANITIES
UNIVERSITAS AIRLANGGA
SURABAYA**

2022

UNDERGRADUATE THESIS

**GRAMMATICAL ERROR ANALYSIS ON ENGLISH SUBTITLES OF
NOMADLAND (2020) MOVIE PRODUCED BY SONIX AND MAESTRA
ONLINE AUTOMATIC SPEECH RECOGNITION SOFTWARE:
A COMPARATIVE STUDY**



By

IFTITA DAMAI AGITASARIA

NIM 121811233054

**ENGLISH DEPARTMENT
FACULTY OF HUMANITIES
UNIVERSITAS AIRLANGGA
SURABAYA**

2022

**GRAMMATICAL ERROR ANALYSIS ON ENGLISH SUBTITLES OF
NOMADLAND (2020) MOVIE PRODUCED BY SONIX AND MAESTRA
ONLINE AUTOMATIC SPEECH RECOGNITION SOFTWARE:
A COMPARATIVE STUDY**

UNDERGRADUATE THESIS

By

IFTITA DAMAI AGITASARIA

NIM 121811233054

**ENGLISH DEPARTMENT
FACULTY OF HUMANITIES
UNIVERSITAS AIRLANGGA
SURABAYA**

2022

**GRAMMATICAL ERROR ANALYSIS ON ENGLISH SUBTITLES OF
NOMADLAND (2020) MOVIE PRODUCED BY SONIX AND MAESTRA
ONLINE AUTOMATIC SPEECH RECOGNITION SOFTWARE:
A COMPARATIVE STUDY**

UNDERGRADUATE THESIS

**Submitted in Partial Fulfillment of the Requirements
for the Sarjana Degree of English Department
Faculty of Humanities Universitas Airlangga**

**By
IFTITA DAMAI AGITASARIA
NIM 121811233054**

**ENGLISH DEPARTMENT
FACULTY OF HUMANITIES
UNIVERSITAS AIRLANGGA
SURABAYA**

2022

DECLARATION

I declare that this undergraduate thesis submitted in partial fulfillment of the degree of Sarjana Humaniora (S.Hum) of the English Department Faculty of Humanities Universitas Airlangga is entirely my own work and has been solely the result of my own original research. Hereby, I confirm that:

1. This thesis is not copied from any other person's work (published or unpublished);
2. I do not use the services of any professional agencies to produce this thesis;
3. This work has not previously been submitted in whole or in part for a degree or any other qualification either at Universitas Airlangga or elsewhere;
4. Any texts, tables, figures, data, or other materials copied from other sources (including, but not limited to, books, papers, and the internet) have been clearly acknowledged and references in the text. These details are also confirmed by a fuller reference at the end of the work.

I am fully aware that breaching any or all of the above points will result in the withdrawal of the awarded degree and disciplinary action in accordance with the regulations at Universitas Airlangga.

Surebaya, 4 January 2022



IFTITA DAMAI AGITASARIA

121811233054

I dedicate this thesis to

my dearest father,
my dearest mother,
and my little brother,
who always love and support me;

also to everyone
who has been very kind and helpful
for the past 3.5 years.

Thank you ♥

APPROVAL SHEET

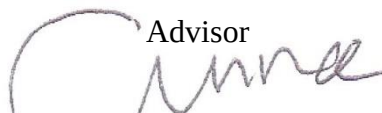
Title : Grammatical Error Analysis on English Subtitles of
Nomadland (2020) Movie produced by SONIX and
MAESTRA Online Automatic Speech Recognition
Software: A Comparative Study

Name : Iftita Damai Agitasaria

Student's Number : 121811233054

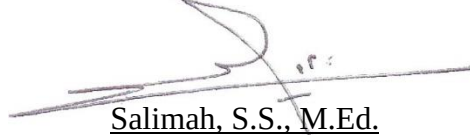
Department : English Department

Approved to be submitted on 6 January 2022 by:


Advisor
Dra. Anna Dewanti, Dipl. Tesl., M.Pd.
NIP. 196012011985112001

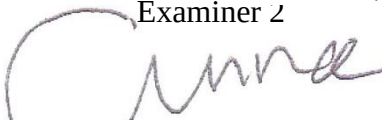
And has been examined and accepted by the Board of Examiners: Surabaya, 14
January 2022.

Chair/Examiner 1


Salimah, S.S., M.Ed.

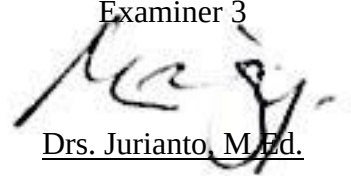
NIP. 197910062008122001

Examiner 2


Dra. Anna Dewanti, Dipl. Tesl., M.Pd.

NIP. 196012011985112001

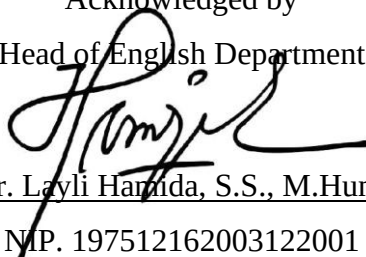
Examiner 3


Drs. Jurianto, M.Ed.

NIP. 196505191991031003

Acknowledged by

Head of English Department


Dr. Layli Hamida, S.S., M.Hum.
NIP. 197512162003122001

ACKNOWLEDGEMENTS

First and foremost, I would like to express all of my gratitude to Allah SWT for all of His blessings and guidance, so I can finish this undergraduate thesis entitled “Grammatical Error Analysis on English Subtitles of *Nomadland* (2020) Movie produced by SONIX and MAESTRA Online Automatic Speech Recognition Software: A Comparative Study.”

To begin with, I am deeply thankful to my thesis advisor, Mrs. Dra. Anna Dewanti, Dipl. TESL., M.Pd for always giving me guidance and consultation over a semester towards the completion of this undergraduate thesis. Thank you for all the valuable time in giving those meaningful feedback, comments, critiques, and advice to my work. I would also like to extend my gratitude to Mrs. Dra. Nur Wulan, M.A., Ph.D as my academic advisor, Mrs. Erlita Rusnaningtias, S.S., M.A., Mrs. Dra. Lilla Musyahda, M.Pd., and Mrs. Masitha Achmad Syukri, S.S., M.Hum as my Thesis Writing Design advisors, and all the lectures and staff of the English Department of Universitas Airlangga. Thank you for always being wonderfully cooperative. I would not have achieved this dream without your tremendous and immense help.

Sequentially, my deepest gratitude goes to my beloved parents, Bapak Agus Suprayitno and Ibu Dwi Puspitayaningati. Thank you so much for the love, support, and prayers. Thank you for always believing in me so that I can complete my study. To my little brother, Ilham Aditama Firdaus, thanks for always cheering me up. I love you. You are amazing just the way you are.

Moreover, I would like to thank Dea, Dhani, Rivi, Aulia, Nadhea, Amala, Nisya, Bima, Mbak Reza, Rico, Ari, April, Zhalsa, Ririn, Ainun, and all my fellows in English Department Universitas Airlangga batch 2018 for always supporting me during my ups and downs. Thanks for all the sweet moments, especially during my college years. To Fadli Alif and Ahnaf Hadwa, thank you for helping me in providing information about software and Artificial Intelligence (AI). Thank you for sharing fruitful knowledge so that I can complete this thesis. Subsequently, my special mention goes to Mas Diki Alifta for always being there whenever I get tired and lose motivation. Thank you for always understanding me during my fickle mood. Thank you for all the patience and kindness. May God bless you, always, Nurse!

Last but not least, I would like to thank me for having no days off and believing in me. Thank you for all the hard work, always being strong, and never giving up. Thank you for finishing this undergraduate thesis, so you can live your dream to graduate in 3.5 years. Thank you, Tita. You're so awesome!

The Writer

لَا حَوْلَ وَلَا قُوَّةَ إِلَّا بِاللَّهِ

“There is no power nor strength except with Allah.”

TABLE OF CONTENTS

Inside Cover Page	i
Inside Title Page.....	ii
Degree Requirement.....	iii
Declaration Page	iv
Dedication Page	v
Approval Sheet.....	vi
Acknowledgements	vii
Epigraph	ix
Table of Contents	x
List of Tables.....	xiii
List of Appendices	xvii
Abstrak	xviii
Abstract.....	xix
CHAPTER I INTRODUCTION	1
1.1 Background of the Study	1
1.2 Statement of the Problems	6
1.3 Objectives of the Study	7
1.4 Significance of the Study	7
1.5 Definition of Key Terms	8
CHAPTER II LITERATURE REVIEW.....	9
2.1 Theoretical Framework	9

2.1.1	Grammatical Error	9
2.1.1.1	Definition of Grammar.....	9
2.1.1.2	Definition of Error.....	10
2.1.1.3	Types of Grammatical Error	10
2.1.1.3.1	Errors Regarding Nouns and Noun Groups...	11
2.1.1.3.2	Errors Regarding Verbs and Verb Groups	13
2.1.1.3.3	Errors Regarding Prepositions.....	17
2.1.1.3.4	Errors Regarding Sentence Structures	20
2.1.2	Automatic Speech Recognition	24
2.1.2.1	Definition of Automatic Speech Recognition	24
2.1.2.2	How Automatic Speech Recognition Works	24
2.2	Review of Related Studies	25
CHAPTER III RESEARCH METHOD		30
3.1	Research Approach	30
3.2	Data Sources.....	31
3.3	Scope and Limitation	31
3.4	Technique of Data Collection	32
3.5	Technique of Data Analysis	33
CHAPTER IV RESULTS AND DISCUSSION.....		36
4.1	Results.....	36
4.1.1	Types of Grammatical Errors found.....	36

4.1.1.1	Types of Grammatical Errors found in the English Subtitle of <i>Nomadland</i> (2020) Movie produced by SONIX Online ASR Software	39
4.1.1.2	Types of Grammatical Errors found in the English Subtitle of <i>Nomadland</i> (2020) Movie produced by MAESTRA Online ASR Software.....	64
4.1.2	The Most Frequent Type of Grammatical Errors found.....	80
4.1.2.1	The Most Frequent Type of Grammatical Errors found in the English Subtitle of <i>Nomadland</i> (2020) Movie produced by SONIX Online ASR Software	80
4.1.2.2	The Most Frequent Type of Grammatical Errors found in the English Subtitle of <i>Nomadland</i> (2020) Movie produced by MAESTRA Online ASR Software.....	85
4.2	Discussion	89
4.2.1	Elaboration.....	89
4.2.2	Interpretation	91
CHAPTER V CONCLUSION		94
REFERENCES.....		96
APPENDICES		100

LIST OF TABLES

Table 4.1	Grammatical Errors found in the English subtitles of <i>Nomadland</i> (2020) movie produced by SONIX and MAESTRA online ASR software.....	39
Table 4.2	Miscellaneous Errors found in the English subtitle of <i>Nomadland</i> (2020) movie produced by SONIX online ASR software	62
Table 4.3	Miscellaneous Errors found in the English subtitle of <i>Nomadland</i> (2020) movie produced by MAESTRA online ASR software	78
Table 4.4	Frequency of Errors regarding Nouns and Noun Groups found in the English subtitle of <i>Nomadland</i> (2020) movie produced by SONIX online ASR software	81
Table 4.5	Frequency of Errors regarding Verbs and Verb Groups found in the English subtitle of <i>Nomadland</i> (2020) movie produced by SONIX online ASR software	82
Table 4.6	Frequency of Errors regarding Prepositions found in the English subtitle of <i>Nomadland</i> (2020) movie produced by SONIX online ASR software	82
Table 4.7	Frequency of Errors regarding Sentence Structures found in the English subtitle of <i>Nomadland</i> (2020) movie produced by SONIX online ASR software	83
Table 4.8	Frequency of Miscellaneous Errors found in the English subtitle of <i>Nomadland</i> (2020) movie produced by SONIX online ASR software	84

Table 4.9	Frequency of Transcription Errors found in the English subtitle of <i>Nomadland</i> (2020) movie produced by SONIX online ASR software	84
Table 4.10	Frequency of Errors regarding Nouns and Noun Groups found in the English subtitle of <i>Nomadland</i> (2020) movie produced by MAESTRA online ASR software	86
Table 4.11	Frequency of Errors regarding Verbs and Verb Groups found in the English subtitle of <i>Nomadland</i> (2020) movie produced by MAESTRA online ASR software	86
Table 4.12	Frequency of Errors regarding Prepositions found in the English subtitle of <i>Nomadland</i> (2020) movie produced by MAESTRA online ASR software	87
Table 4.13	Frequency of Errors regarding Sentence Structures found in the English subtitle of <i>Nomadland</i> (2020) movie produced by MAESTRA online ASR software	87
Table 4.14	Frequency of Miscellaneous Errors found in the English subtitle of <i>Nomadland</i> (2020) movie produced by MAESTRA online ASR software	88
Table 4.15	Frequency of Transcription Errors found in the English subtitle of <i>Nomadland</i> (2020) movie produced by MAESTRA online ASR software	89

Table 4.16	Errors regarding Nouns and Noun Groups found in the English subtitle of <i>Nomadland</i> (2020) movie produced by SONIX online ASR software	116
Table 4.17	Errors regarding Verbs and Verb Groups found in the English subtitle of <i>Nomadland</i> (2020) movie produced by SONIX online ASR software	120
Table 4.18	Errors regarding Prepositions found in the English subtitle of <i>Nomadland</i> (2020) movie produced by SONIX online ASR software	121
Table 4.19	Errors regarding Sentence Structures found in the English subtitle of <i>Nomadland</i> (2020) movie produced by SONIX online ASR software	123
Table 4.20	Miscellaneous Errors found in the English subtitle of <i>Nomadland</i> (2020) movie produced by SONIX online ASR software	133
Table 4.21	Transcription Errors found in the English subtitle of <i>Nomadland</i> (2020) movie produced by SONIX online ASR software	134
Table 4.22	Errors regarding Nouns and Noun Groups found in the English subtitle of <i>Nomadland</i> (2020) movie produced by MAESTRA online ASR software	149
Table 4.23	Errors regarding Verbs and Verb Groups found in the English subtitle of <i>Nomadland</i> (2020) movie produced by MAESTRA online ASR software	150

Table 4.24	Errors regarding Prepositions found in the English subtitle of <i>Nomadland</i> (2020) movie produced by MAESTRA online ASR software	151
Table 4.25	Errors regarding Sentence Structures found in the English subtitle of <i>Nomadland</i> (2020) movie produced by MAESTRA online ASR software	153
Table 4.26	Miscellaneous Errors found in the English subtitle of <i>Nomadland</i> (2020) movie produced by MAESTRA online ASR software.....	155
Table 4.27	Transcription Errors found in the English subtitle of <i>Nomadland</i> (2020) movie produced by MAESTRA online ASR software	157

LIST OF APPENDICES

Appendix 1	The English Subtitle of <i>Nomadland</i> (2020) Movie produced by SONIX Online Automatic Speech Recognition Software	100
Appendix 2	The Grammatical Errors Found in the English Subtitle of <i>Nomadland</i> (2020) Movie produced by SONIX Online Automatic Speech Recognition Software	116
Appendix 3	The English Subtitle of <i>Nomadland</i> (2020) Movie produced by MAESTRA Online Automatic Speech Recognition Software ..	135
Appendix 4	The Grammatical Errors Found in the English Subtitle of <i>Nomadland</i> (2020) Movie produced by MAESTRA Online Automatic Speech Recognition Software	148

ABSTRAK

Kesalahan tata bahasa merupakan fenomena yang sering dijumpai dalam karya-karya tulis yang tersebar di masyarakat, termasuk teks terjemahan berbahasa Inggris sebuah film yang diproduksi oleh perangkat lunak pengenalan ucapan otomatis yang telah dirancang untuk memiliki kecerdasan setara dengan otak manusia. Penelitian ini bertujuan untuk mengeksplorasi jenis-jenis kesalahan tata bahasa yang terdapat dalam teks terjemahan berbahasa Inggris film *Nomadland* (2020) yang diproduksi oleh perangkat lunak pengenalan ucapan otomatis bernama SONIX dan MAESTRA. Penelitian ini menggunakan metode kualitatif. Selain itu, teori *Taxonomies of Errors* oleh Ho (2005) juga digunakan sebagai acuan penulis dalam menganalisis data. Hasil penelitian menunjukkan bahwa enam jenis kesalahan tata bahasa ditemukan dalam teks terjemahan berbahasa Inggris film *Nomadland* yang diproduksi oleh kedua perangkat lunak. Empat jenis kesalahan tata bahasa—*errors regarding nouns and noun groups*, *errors regarding verbs and verb groups*, *errors regarding prepositions*, dan *errors regarding sentence structures*—ditemukan berdasarkan teori *Taxonomies of Errors* oleh Ho (2005) sedangkan dua jenis kesalahan tata bahasa lainnya—*miscellaneous errors* dan *transcription errors*—merupakan penemuan di lapangan. Jenis kesalahan tata bahasa yang paling dominan adalah *errors regarding sentence structures* dengan persentase 36.27% dalam teks terjemahan berbahasa Inggris film *Nomadland* yang diproduksi oleh perangkat lunak SONIX dan 35.04% dalam teks terjemahan berbahasa Inggris film *Nomadland* yang diproduksi oleh perangkat lunak MAESTRA. Selain itu, kategori kesalahan tata bahasa yang paling dominan adalah *incomplete or fragmented sentences* dengan persentase 13.16% dalam teks terjemahan berbahasa Inggris film *Nomadland* yang diproduksi oleh SONIX dan 26.75% dalam teks terjemahan berbahasa Inggris film *Nomadland* yang diproduksi oleh MAESTRA.

Kata Kunci: *Kesalahan Tata Bahasa, Perangkat Lunak Pengenalan Ucapan Otomatis MAESTRA, Perangkat Lunak Pengenalan Ucapan Otomatis SONIX, Teks Terjemahan Film Nomadland Berbahasa Inggris*

ABSTRACT

Grammatical error is a common phenomenon in written works which have been spread in society, including the English subtitle of a movie produced by an Automatic Speech Recognition (ASR) software that has been designed to have the same intelligence as humans. This study aims to explore the types and the most frequent type of grammatical errors found in the English subtitles of *Nomadland* (2020) movie produced by SONIX and MAESTRA online ASR software. This study applied a qualitative method and used Ho's (2005) theory about Taxonomies of Errors. The result shows that there are six types of grammatical errors found in the English subtitles of *Nomadland* (2020) movie produced by both software. Four of them—errors regarding nouns and noun groups, errors regarding verbs and verb groups, errors regarding prepositions, and errors regarding sentence structures—are found based on Ho's theory, whereas the other two types of errors—miscellaneous errors and transcription errors—are found as field discovery. The most frequent type of grammatical errors found in the English subtitle of *Nomadland* (2020) movie produced by SONIX is errors regarding sentence structures (36.27%) which is the same as the most frequent type of grammatical errors found in the subtitle produced by MAESTRA (35.04%). Moreover, the most frequent category of grammatical errors found in the English subtitle of *Nomadland* (2020) movie produced by SONIX is incomplete or fragmented sentences (13.16%) which is also the same as the most frequent category of grammatical errors found in the subtitle produced by MAESTRA (26.75%).

Keywords: *English Subtitles of Nomadland Movie, Grammatical Errors, MAESTRA Automatic Speech Recognition Software, SONIX Automatic Speech Recognition Software*

CHAPTER I

INTRODUCTION

1.1 Background of the Study

Grammar is a part of a language which cannot be separated from the language (Wilcox 2004). It is a set of rules which tells us how to recognize the expression of words (Newson et al. 2006). Hence, grammar is a resource for successful communication, not just a guideline for organizing and arranging words (Hyland 2002). However, there are so many people who still make errors in using grammar within their conversations or communications. Those errors, then, are called grammatical errors.

As stated by Garner (2012), grammatical error is a concept used in prescriptive grammar to denote a case of incorrect, unusual, or controversial use, such as a misplaced modifier or an unsuitable verb tense. Grammatical errors are usually found in scientific writings, students' writings, and translational texts (Abdulkareem 2013). In the scientific writings, grammatical errors are usually found in journal articles, theses, dissertations, papers, etc.; in the students' writings, they are mostly found in the writings of almost all non-native students (Abdulkareem 2013) and rarely found in native speakers' writing (Kraichoke 2017); and in the translational texts, they can occur in the writings that are translated from one language (L1) to another language (L2). The occurrence of those grammatical errors, especially in the scientific writings, students' writings, and translational texts, is usually due to a lack of knowledge and understanding of the