Statistical Machine Translation between Myanmar Written Text and Myanmar SignWriting (SW)

Title Defence

Supervisor: Dr. Nandar Win Min
Co-supervisor: Dr. Ye Kyaw Thu, Daw Swe Zin Moe
Presented by: Ma Hnin Wai Wai Hlaing (ME-IST 23)
NLP Research Group

(20 July 2017)

Outlines

- Abstract
- Objectives
- Introduction
- Contribution
- Myanmar Sign Language
- SignWriting (SW)
- Example Transcription of Myanmar SignWriting
- Statistical Machine Translation (SMT)
- Neural Machine Translation (NMT)
- System Flow
- Conclusion
- References

Abstract

- In the field of machine translation, significant progress has been made by using statistical methods.
- The proposed system suggests a statistical machine translation system between Myanmar Written Text and Myanmar SignWriting.
- It takes Myanmar Written Text as input and the output is represented in from of Myanmar SignWriting.
- There is no Myanmar Written Text and Myanmar SignWriting parallel data yet, and thus we need to prepare it.
- It is based on and Statistical Machine Translation (SMT) and Neural Machine Translation (NMT).
- The proposed system will solve difficulties for deaf people to learn the basic concept of daily life, in communication and education.

Keywords: *Myanmar SignWriting, Statistical Machine Translation, Neural Machine Translation, Natural Language Processing.*

Objectives

- To learn Machine Translation between Myanmar Written Text and Myanmar SignWriting
- To develop Myanmar Written Text and Myanmar SignWriting parallel corpus
- To measure Machine Translation performance using Statistical Machine Translation (SMT) and Neural Machine Translation (NMT)
- To analyze detail error from translation experimental results
- To have easy translation between deaf people and hearing people

Introduction

- Myanmar Sign Language (MSL) is a natural language, which is used as a primary means of communication for about 2.3 million deaf people in Myanmar.
- There are limited resources and information written in Sign Language for Myanmar Deaf community.
- Deaf cannot read or write as normal people, so they have problems in communicating and knowledge sharing with normal people.
- To solve this problem, Sign Language translation systems are required.
- This proposed system will focus on translation between Myanmar Written Text and Myanmar SignWriting.

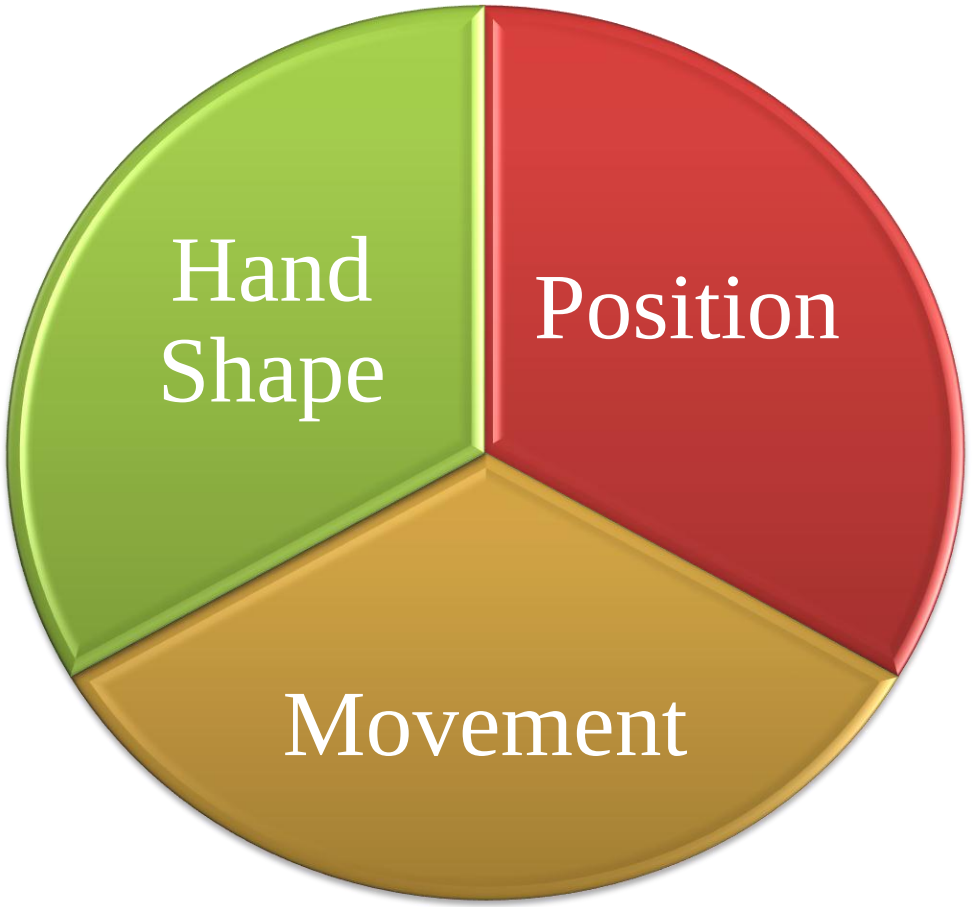
Contribution

- The proposed system is to investigate machine translation performance between Myanmar Written Text and Myanmar SignWriting using Statistical Machine Translation (SMT) and Neural Machine Translation (NMT) approaches.
- It intends to make detail error analysis (alignment error, reordering error, translation error, etc.) from the results of translation experiments.
- It contributes the first study of the Statistical Machine Translation and Neural Machine Translation approaches between Myanmar Written Text and Myanmar SignWriting.

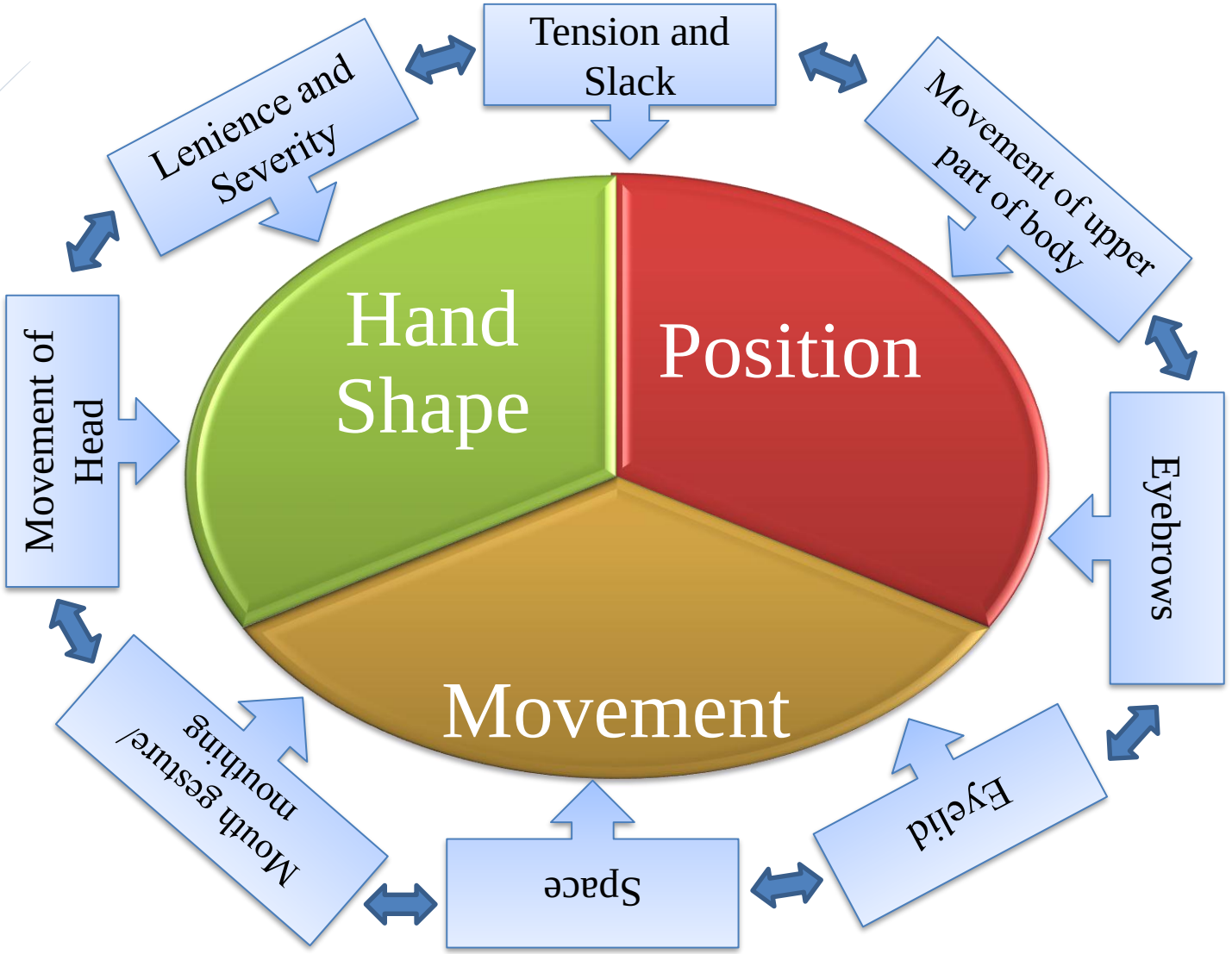
Myanmar Sign Language

- Each country has its own, native Sign Language according to their culture.
- Myanmar also has its own Sign Language.
- SL is the native language of the deaf community and they can express their needs and the formation of concepts by combining hand shapes, orientation and movement of the hands, arms or body, and facial expressions.
- SL consists of **Manual Features (MFs) and Non-Manual Features (NMFs).**

Structure of Manual Features (MFs)



Structure of Non-Manual Features (MFs)



SignWriting (SW)

- SignWriting (SW) is a writing system that is a sequence of symbols for deaf Sign Language.
- SW represents two perspective: **singer's perspective** and **observer's perspective.**
- SignWriting is based on how you see your own hands when you sign—the **signer's perspective**
- SW is written horizontally **(left to right)** and the **right hand is dominant**.
- SW symbols can be rotated in 8 directions and placed anywhere in the writing area.
- International Sign Writing Alphabet (ISWA) defines **7 categories, 30 groups** of symbols to form **639 base symbols** and **35,023 final symbols**.

International SignWriting Alphabet (ISWA) describes 7 categories:

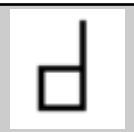
Category	Name	SignWriting Symbols	Description
1	Hand		• Hand shape differ number of finger and hand orientation.
2	Movements		• Movement is parallel to the front wall and parallel to the floor.
3	Dynamic and Timing		• Dynamic symbols are used to give the “feeling” or “tempo” to movement. • Timing symbols are used to show alternating or simultaneous movement.
4	Head and Faces		• Starting with the head and then from the top of the face and moving down.
5	Body		• Movement, shoulders, hips and the limbs are used in Sign Language as part of grammar.
6	Detailed Location		• Detailed location symbols is useful for sorting large dictionaries, refining animation.
7	Punctuation		• Punctuation symbols are used when writing complete sentences or documents in SignWriting.

Figure	Indication
	The **palm of the hand** is written in a **white or hollow symbol.**
	The side of the hand is written in a symbol that is **half-white and half-dark.**
	The back of the hand is written in **a black or fill-in a symbol.**

Table 1. Example of SignWriting HAND-FLAT Handshape parallel with Wall Plane (Front view)

Figure	Indication
	The **palm of the hand** is written in a **white or hollow symbol with small gap.**
	The side of the hand is written in a symbol that is **half-white and half-dark with small gap.**
	The back of the hand is written in **a black or fill-in a symbol with small gap.**

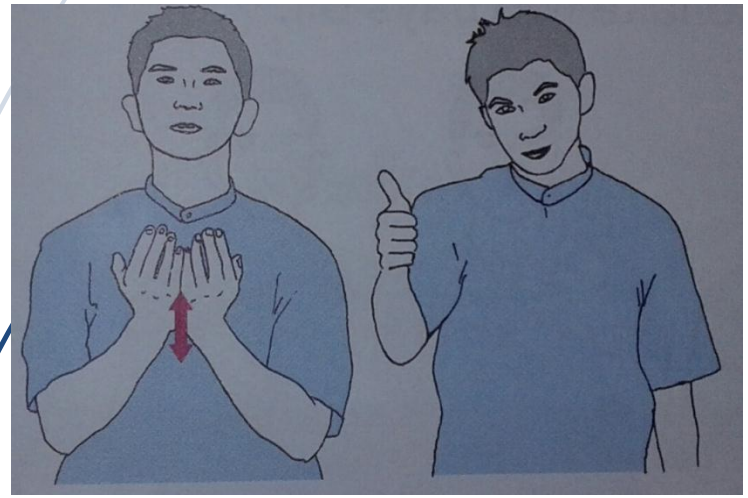
Table 2. Example of SignWriting HAND-FLAT Handshape parallel with Floor Plane (Top view)

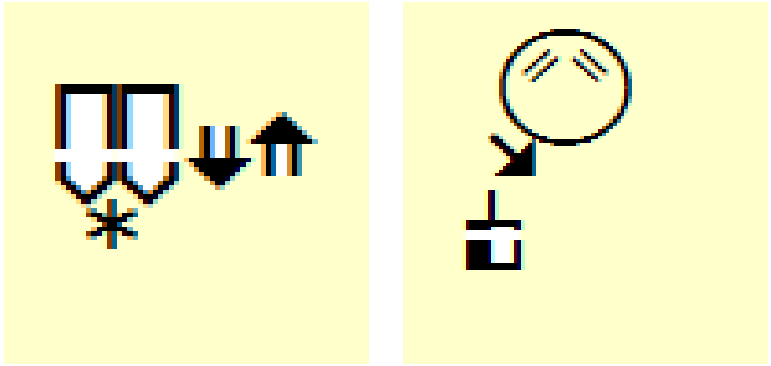
Transcription of Myanmar SignWriting

Example 1:

Myanmar Text : စာ လိုက်နိုင် သလား။

Sign Language: စာဖတ် ရလား။





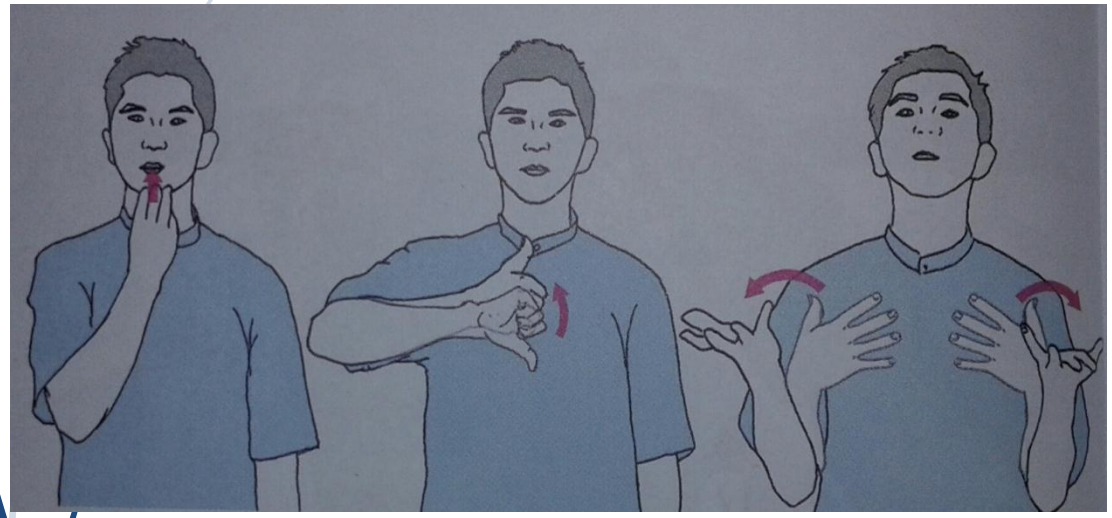
Sign Language **SignWriting**

Transcription of Myanmar SignWriting

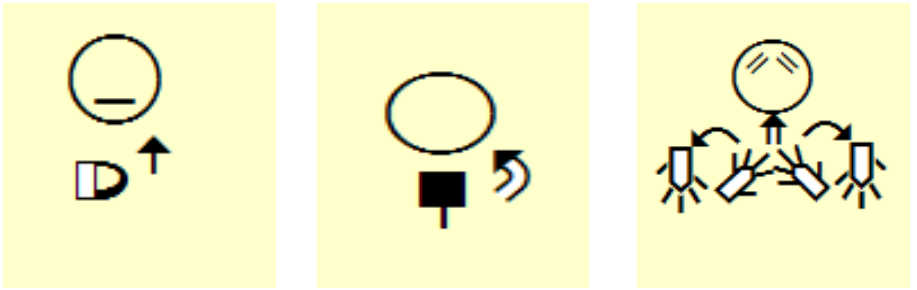
Example 2:

Myanmar Text : ဘာ အစားအစာ ကြိုက်လဲ ။

Sign Language: စား ကြိုက် ဘာလဲ ။



Sign Language



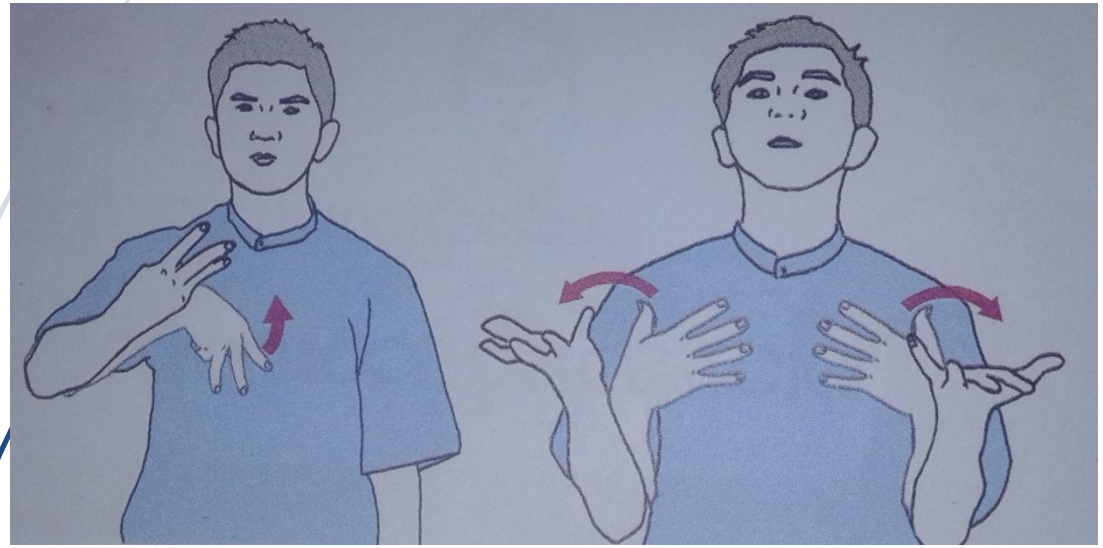
SignWriting

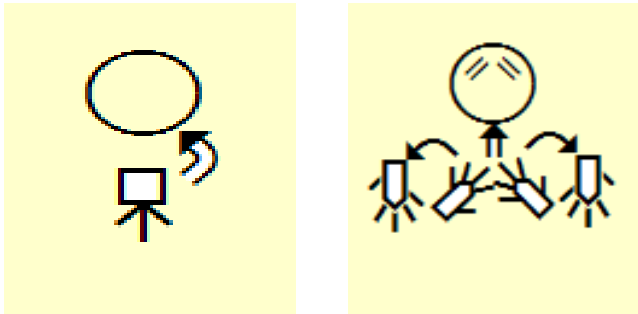
Transcription of Myanmar SignWriting

Example 3:

Myanmar Text : မင်းရဲ့ ဂါဝနာက ဘာလဲ။

Sign Language: ဂါဝနာ ဘာလဲ။





Sign Language **SignWriting**

Statistical Machine Translation (SMT)

- A machine translation paradigm where translations are generated on the basis of statistical models whose parameters are derived from the analysis of bilingual text copra.
- Currently, there are many kinds of statistical machine translation approaches.
- The proposed system will use three SMT approaches
  - Phrase-Based Statistical Machine Translation (PBSMT),
  - Hierarchical Phrase-Based Statistical Machine Translation (HPBSMT) and
  - Operation Sequence Model (OSM).

Neural Machine Translation (NMT)

- An approach to machine translation that uses a large neural network
- It departs from phrase-based statistical translation approaches.
- Google and Microsoft translation services now use NMT.
- Google uses Google Neural Machine Translation (GNMT).
- NMT models use deep learning and representation learning.
- All parts of the neural translation model are trained jointly (end-to-end) to maximize the translation performance.

System Flow

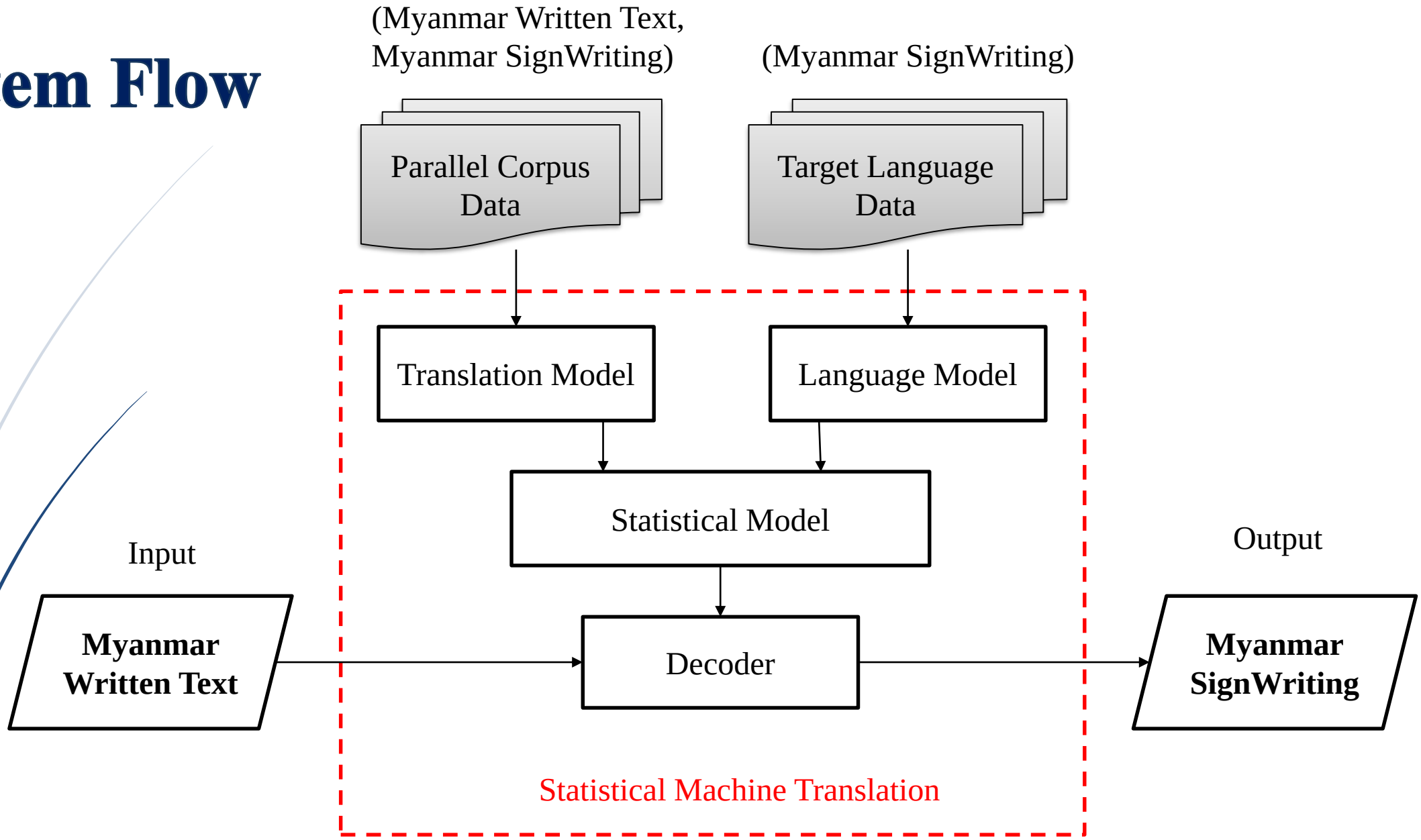


Figure : Flow diagram of the proposed system

Conclusion

- The proposed system focus on the first study of Statistical Machine Translation between Myanmar Written Text and Myanmar SignWriting.
- Building our own parallel corpus between Myanmar Written Text and Myanmar SignWriting will be able to use for later researches.
- This can make ease of communication between deaf people and non-deaf people by using Myanmar Written text to Myanmar SignWriting translation system.

References

- Cayley Guimaraes, Jeferson F. Guardezi, Luis Eduardo Oliveira, Sueli Fernandes ., “Deaf Cultures and Sign Language Writing System – a Database for a New Approach to Writing System Recognition Technology”
- Sara Morrissey and Andy Way., “Joining Hands: Developing a Sign Language Machine Translation System With and For the Deaf Community”
- Ameera M.Almasoud and Hend S. Al-Khalifa., “A Proposed Semantic Machine Translation System for translating Arabic text to Arabic sign language”
- Win Pa Pa, Ye Kyaw Thu, Andrew Finch, Eiichiro Sumita,. “A Study of Statistical Machine Translation Methods for Under Resourced Languages”
- https://en.wikipedia.org/wiki/Statistical_machine_translation
- https://en.wikipedia.org/wiki/Neural_machine_translation
- *The population and housing census of Myanmar,2014*

Thank You.