



# **Amharic Poem Wax and Gold Meanings Identification System**

**CoSc 6405 - Natural Language Processing**

**Surafiel H., Feven T. and Alemayehu E.**

**February 27, 2023**

# Introduction

- ❖ Poetry is a type of literature that conveys a thought, describes a scene or tells a story in a concentrated, lyrical arrangement of words.
- ❖ Qene (ቅኔ) is an Ethiopian style of speech expressed in poetic form, where one says one thing while implying a different meaning at the same time.
- ❖ Its origins are supposed to date back to the 14<sup>th</sup> century.
- ❖ The Amharic poem (ቅኔ) articulates three elements:
  1. The wax (ሰፆ),
  2. The gold (ወርቅ), and
  3. The double-meaning word (ህብረቃል), which is also known as the "harmonizer" or the bridge

# Introduction

- ❖ A great deal of Amharic and Geez Qene depends on one word having, or being interpreted, in more than one way.
- ❖ The Wax and Gold (ሰፆ እና ወርቅ) method plays with double-layered meaning.
- ❖ The apparent meaning, on the surface, is known as “wax” (ሰፆ).
- ❖ The underlying truth, hidden meaning, or knowledge residing in the poem is known as “gold” (ወርቅ).
- ❖ Knowledge of the Amharic poetry (ቅኒ) tradition is necessary to identify the gold (ወርቅ) meaning in Amharic poem.

# Introduction

- ❖ In Amharic poem hermeneutics, knowledge would function as a tool to determine the wax and gold meanings corresponding to a double-meaning word (ህብረ ቃል).
- ❖ Amharic lacks a symbol for accent as stress indicator on words.
- ❖ Example

ያልታወቀ ቅመም ከወጥ ገብቷል አሉ፤  
በሽታ ጨረሰው ተጋባዥን ሁሉ።

- ✓ The double meaning word: በሽታ (be-sheeta)
- ✓ **wax**: unknown spice was giving off a strong aroma to all the invited guests.
- ✓ **gold**: all the invited guests were consumed by sickness due to some unknown spice which had entered the dish.

# Introduction

❖ Another technique of Qene (ቅኔ) is where two or more words may be joined to become one, or a word may be split into two or more parts, to change the meaning.

❖ Example:

አንጥረኛው ብዙ በቴዎድሮስ ቤት፣  
ባለአልቦ አደረጉት ይህን ሁሉ ሴት።

✓ ህብረቃል (The double-meaning word): ባለአልቦ

✓ ስም: አጼ ቴዎድሮስ ሴቱን ሁሉ አልቦ የተሰኘው ጌጥ ባለቤት አደረጉት

✓ ወርቅ: አጼ ቴዎድሮስ የሴቱን ሁሉ ባል በመግደላቸው ሴቱ ሁሉ ባል የሌለው ሆነ  
(King Tewodros killed the husbands of many women)

❖ The Gold's (ወርቅ) meaning is deduced from the hidden knowledge in the two independent words “ባል” and “አልቦ”.

# Problem Statement

- ❖ In the Amharic poem, exploring the gold, which represents the actual hidden meaning of the poem, requires identifying the double-meaning word of the poem.
- ❖ Due to the complexity of Amharic Qene poetry, fully and computationally modeling its symbolic representation or meaning had not yet been accomplished.
- ❖ A computational model based on non-poetic Amharic literature can account for at least some aspects of the semantics in the Qene.
- ❖ As a result, we propose a knowledge-based Amharic poem interpreter that finds words with dual meanings and outputs both the symbolic meaning (wax) and the hidden true meaning (gold).

# General Objective

- ❖ To design and build a system that can be used to show users the wax (ሰፆ) and gold (ወርቅ) meanings of a given Amharic poem by identifying the double-meaning holder word (ህብረ ቃል) using word embedding and a knowledge base.

# Specific Objective

- ❖ To explore and understand the methods of Amharic poem interpretation from literature reviews and related works,
- ❖ To collect Amharic poems from different resources and prepare a corpus.
- ❖ To collect words from the corpus and other sources like the Amharic dictionary to develop the Amharic WordNet,
- ❖ To design the architecture and algorithm for wax and gold Amharic poem interpretation,
- ❖ To build the automated double-meaning word identifier from the Amharic poem with its literal meaning (wax) as well as the hidden meaning (gold), and
- ❖ To develop, test, and evaluate the prototype of the system



# Scope

- ❖ Only takes into account the semantic level of the Amharic poem's context
- ❖ Provide an Amharic WordNet with a small number of senses and be specialized for the purpose of prototype testing
- ❖ Use only one of several techniques to obtain the wax and gold meanings, namely Word Sense Disambiguation (WSD), to solve the meaning of double-meaning words (ህብረቃል)

# Limitations

- ❖ Will not address Amharic poem that employs techniques of joining words or splitting words.
- ❖ It doesn't work for double-meaning words (ህብረቃሎች) that don't exist in the corresponding Amharic WordNet.
- ❖ In our case, when two or more dual-meaning words appear in the same Amharic poem, which is a rare occurrence, it is impossible to pinpoint the precise term dedicated to the poem
- ❖ Example:

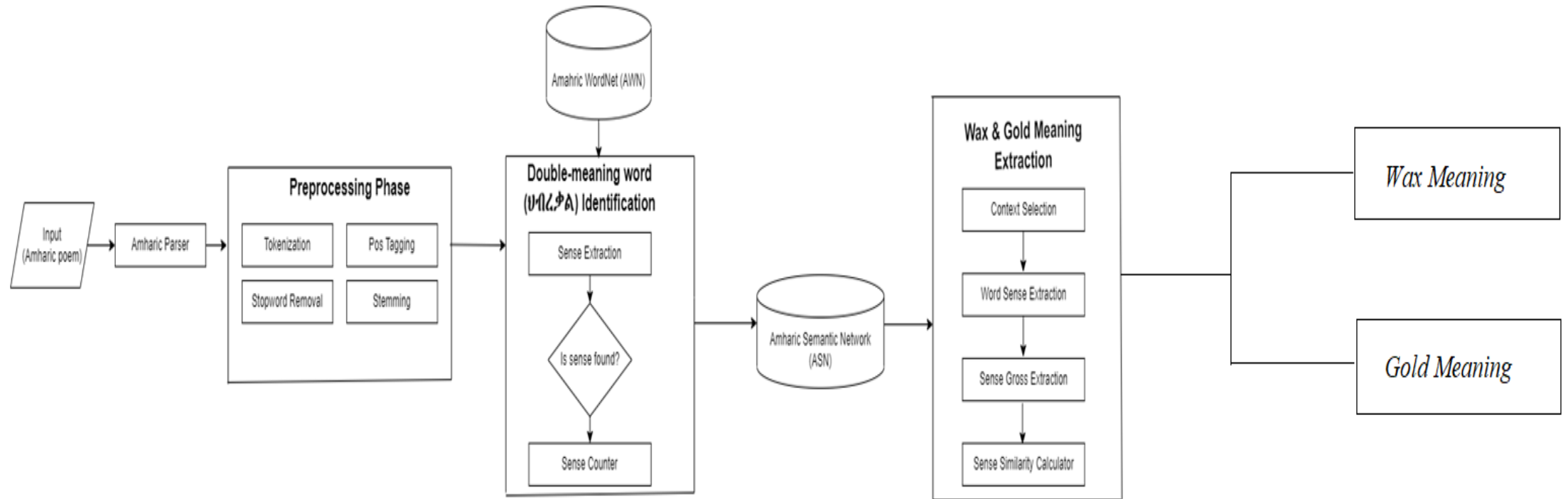
አዲስ አበባ ላይ ቡና ቢወደድ፣  
ከጅማ ሰው ካለ ይመጣል በግድ።

- ❖ There are two double-meaning words (ህብረቃሎች) in the above poem: “አበባ” and “ከጅማ”, but the precise double-meaning word in the context is “ከጅማ”

# Method

- ❖ The data that is used to develop the Amharic WordNet (AWN) and the Amharic poetry corpus is collected from various sources, such as elementary and high school Amharic textbooks, the Internet, and journals
- ❖ The technique of Word Sense Disambiguation (WSD) is used to detect the double-meaning word (ህብረቃል)

# Proposed System Architecture



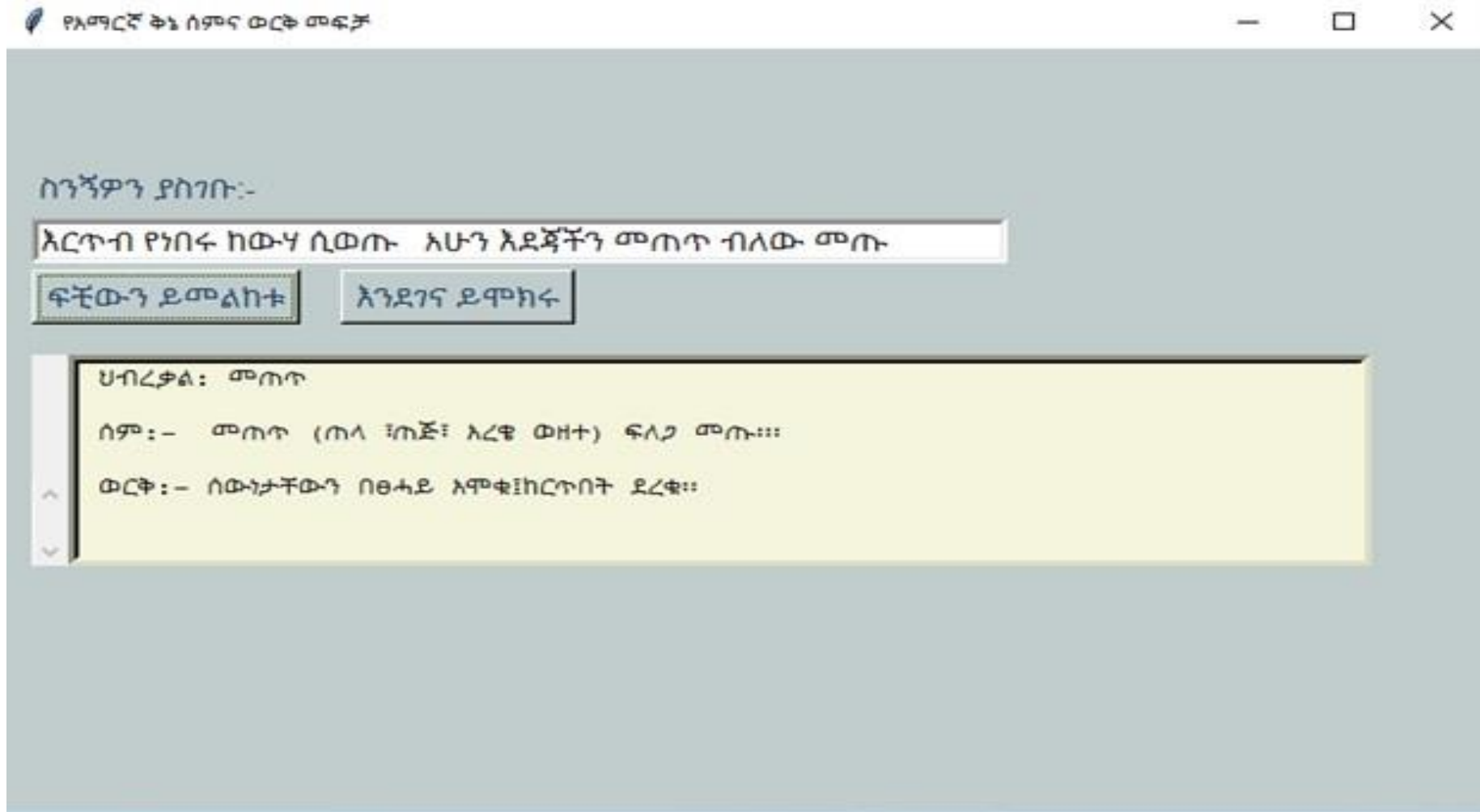
*Adopted and customized, from a research work entitled "Amharic Word Sense Disambiguation Using WordNet" by Seid Hassen and Yaregal Assabie in 2015.*

# Implementation

- ❖ We have created a Python prototype to demonstrate the usability of the proposed Amharic poem wax and gold meaning identification system, which uses Amharic WordNet and Amharic Semantic Network as knowledge bases



# Implementation



# Evaluation

- ❖ Precision and recall have been employed as performance measures
- ❖ The precision is used to measure how much of the poem entered by a user is processed correctly
- ❖ The recall is used to measure how many correct meanings are displayed, given the total number of poems fed to the system

Precision is calculated as follows:

$$\frac{\text{Number of correctly identified Hibre kal and its double meaning}}{\text{Number of total results displayed by the system}}$$

Recall is calculated as follows:

$$\frac{\text{Number of Poem (Qine) with correct interpretation}}{\text{Number of total poems}}$$

# System Test Result

❖ The results obtained during the testing procedure are listed below, with a total of 70 tested Amharic poems

Test No.	Total No. of poems fed to the system	Correctly interpreted poems	Incorrectly interpreted poems	Precision	Recall
1	20	17	3	85%	20%
2	20	19	1	95%	27%
3	30	24	6	80%	34%



# Conclusion and Future Works

- ❖ Based on our work on this project, even though the process is very complex, we understood that using Amharic WordNet (AWN) and the Amharic Semantic Network (ASN), it's possible to develop a system that identifies the double-meaning word (ህብረቃል) and the wax and gold meanings of Amharic poems using the "Word Sense Disambiguation" technique.
- ❖ The limitations of this project can be improved by providing:
  - ✓ An integrated Amharic morphological analyzer,
  - ✓ An Amharic spelling and grammar checker, and
  - ✓ An algorithmic solution for detecting the exact double-meaning word (ህብረቃል) with its wax and gold meanings when two or more appear in a given Amharic poem.

**Thank you!**