INGIT – Limited Domain Formulaic Translation from Hindi to ISL

Purushottam Kar
Madhusudan Reddy
Amitabha Mukerjee
Achla Raina

Indian Institute of Technology Kanpur
Introduction

• Objective
  – Create a scalable proof-of-concept system
  – Exhibit challenges in cross modal translation
  – Provide tangible solutions

• Scope
  – Railway Counter Domain
  – Handle Speech → Sign translation only
Past Research

• Speech $\rightarrow$ Sign
  – Form based approaches [Veale94, Zhao02, Speers02]
  – Semantic based approaches [Marshall03, Wray04]

• Sign Recognition (Sign $\rightarrow$ Speech)
  – Reasonable accuracy rates
  – Very small vocabularies handled
Sign Language

• Spatial Modality

• Extensive use of space
  – Iconic signs
  – Role play
  – Use of person and space deixis
  – Directional verbs
  – Non-manual markers
Indian Sign Language

THREE

AC

TICKET

NEG
The Proposed Model

Spoken Language
- Visual Auditory
- Schemas
  - Grounded
  - Abstract
  - Compositional
  - Constructional
  - Partially Compositional

Sign Language
- Visual Spatial
- Schemas
  - Abstract
  - Grounded
  - Compositional
  - Constructional
  - Partially Compositional

Conceptual Representation
- Linguistic Representation
  - Grounded
  - Abstract
  - Compositional
  - Constructional
  - Partially Compositional

Articulatory Perceptual System
- Aural-Auditory
- Sequential
- Sequential + Parallel
- Motor-Spatial
Framework Adopted

- **Constructions**
  - Form-meaning maps at
  - morphological, lexical, and syntactic levels

- **Construction Grammars** [Kay, 2002]
  - The fly is buzzing.
  - What is the fly doing in my soup?
  - . . . fly in the ointment.

- **Unification based approach**
Constituent-Level operations in cross-modal mapping

• Constituent Level
  – Complete
    • गाडी सात बजे जाएगी → \{train time seven go\}
      \emph{train seven O’clock go-fem-future}
  – Partial
    • Constituent Deletion
      राजधानी रात में चलती है → \{rAjdhAni night go\}
      \emph{rAjdhAni night in go-fem}
    • Constituent Insertion
      आप दस रुपये दीजिए → \{money ten give \{you → me\}\}
      \emph{you-hon ten rupees give-hon}
Construction-Level operations in cross-modal mapping

• Construction Level
  – Compositional
    • टिकट नहीं मिलेगा क्योंकि वेटिंग है
      \[\text{ticket neg get because waiting be} \rightarrow\]
      \[@n \{\text{ticket get neg}\} \{\text{Q-why}\} \{\text{waiting-list}\}\]
  – Non-Compositional
    • \(X \text{ में } Y \text{ वेटिंग है} \rightarrow \{x \text{ waiting-list y}\}\]
      \[X \text{ in } Y \text{ waiting be}\]
Polysemy and Anaphora

• Polysemous Expressions
  – attributive and existential senses of ‘है’ (hai, be)
  – alienable vs. inalienable possession as in मेरी किताब (meri kitab, my book) and मेरा भाई (merA bhai, my brother)
  – transactional and non-transactional senses of the verb ‘ले’ (le, take)

• Anaphoric expressions
  – May be resolved to saturate event semantics
  – May be replaced with deictic signs
    • Default deixis
    • वह गाडी कानपुर नहीं जाएगी → {train -dei @n{kAnpur go neg}}
      that train kAnpur neg go-future
Input Parser

- Accepts transcribed spoken expressions
- Analyses the input morphologically, identifying phrases (viz. modifiers, polarity items)
- Uses a simple dictionary look-up approach on the input language construct-icon
- Outputs a (possibly incomplete) semantic structure
INGIT - Architecture

• शताब्दी शाम को कानपुर नहीं जाती

\textit{shatAbdi evening in kAnpur neg goes}

• जा(jA, go)

\textit{((MOTION-VERB EV) (GO EV) (VERB-CLASS EV UNARY)
(ARGUMENT-1 EV OBJ) (TIME-FRAME EV X)
(ARGUMENT-1-PREREQ EV MOBILE))}

• \{SUB-NOMINATIVE MODIFIER-1 MODIFIER-2 NEGATION UNARY-VERB\}
INGIT - Architecture

• शताब्दी शाम को कानपुर नहीं जाती

• ((MOTION-VERB X-95) (GO X-95) (VERB-CLASS X-95 UNARY) (ARGUMENT-1 X-95 X-96) (TIME-FRAME X-95 PRESENT))
• (SHATABDI X-96) (DISCOURSE-ROLE X-96 EXTERNAL) (GENDER X-96 FEMININE) (MOBILITY X-96 MOBILE)
• (NEG X-73)
• (KANPUR X-61)
• (EVENING X-58) (TEMPORAL-MODIFIER X-30 X-58)
• (SATURATED X-41) (EVENT X-41 X-95) (MODIFIERS X-41 X-95 X-30 X-61) (NEGATOR X-41 X-95 X-73))
INGIT - Architecture

• Ellipsis Resolution Module
  – दस रुपये दीजिए → {money ten give {you → me}}

  ten rupees give-hon

  – Attempts to saturate the incomplete semantic structures given by input parser

  – Assumes simple constraints imposed by the limited domain

• Viz. constraints of animacy on participants in a transactional event
INGIT - Architecture

• ISL Generator
  – Very similar to input parser in operation
  – Uses ISL construct-icon to build up ISL sign-gloss
  – Constituent reordering
  – Negation and Q scope resolution

• \{\text{SUB \text{@n}\{\text{MODIFIER-1 MODIFIER-2 UNARY-VERB NEGATION}\}}\}

• \{\text{SHATABDI \text{@n}\{\text{EVENING KANPUR GO NEG}\}}\}
INGIT - Architecture

• Graphical Simulator
  – Convert the tagged ISL strings into HamNoSys notation
  – Graphics generation final step in the system
Summing Up

• Problem of cross-modal translation
• Semantically mediated procedure
• Adapt CG to specific objectives
• Develop working implementation
• Groundwork
  – Identify representational, conceptual issues in cross-modal linguistic processing
  – Provide implementable solutions for the same
Future Directions

• Describe ISL in terms of a framework allowing parallel processing
  – Define such a framework and develop formalisms for the same
• Develop robust graphical front end
• Explore cognitive processes underlying generation of spatial language