Simpler Voice: A Key Message & Visual Description Generator System for Illiteracy

Minh N. B. Nguyen, Samuel Thomas, Kush R. Varshney, Anne E. Gattiker, Sujatha Kashyap
IBM Research

IBM Research

LITERACY COALITION
OF CENTRAL TEXAS
BUILDING STRONGER COMMUNITIES

Introduction

The social issues of illiteracy

- Globally [1]
  - *750 million* illiterate adults.
  - UNESCO International literacy day: September 8th

- In the United States [2,3]
  - *32 million* adults are not able to read.
  - *50 percent* of adults cannot read a book written at an eighth-grade level.
  - *185,000* students are at the lowest ESL (English as second language) level, beginning literacy.

Proposed system SIMPLER VOICE

- AI for Social Good.
- Supporting adult basic education (ABE), and the English as a second language acquisition (SLA).
- System:
  - Decoding complex texts.
  - Simple key messages.
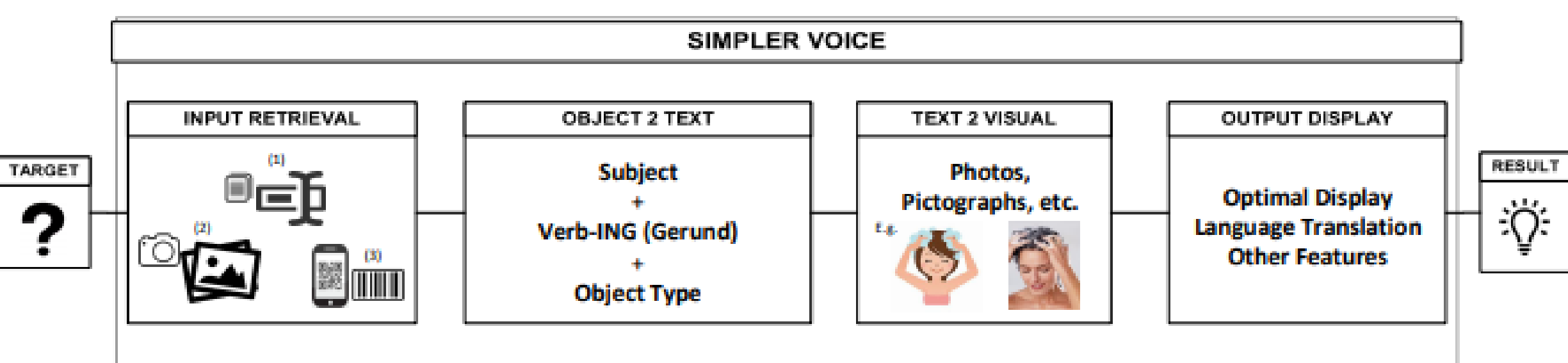 - Text-to-image, text-to-pictographs.
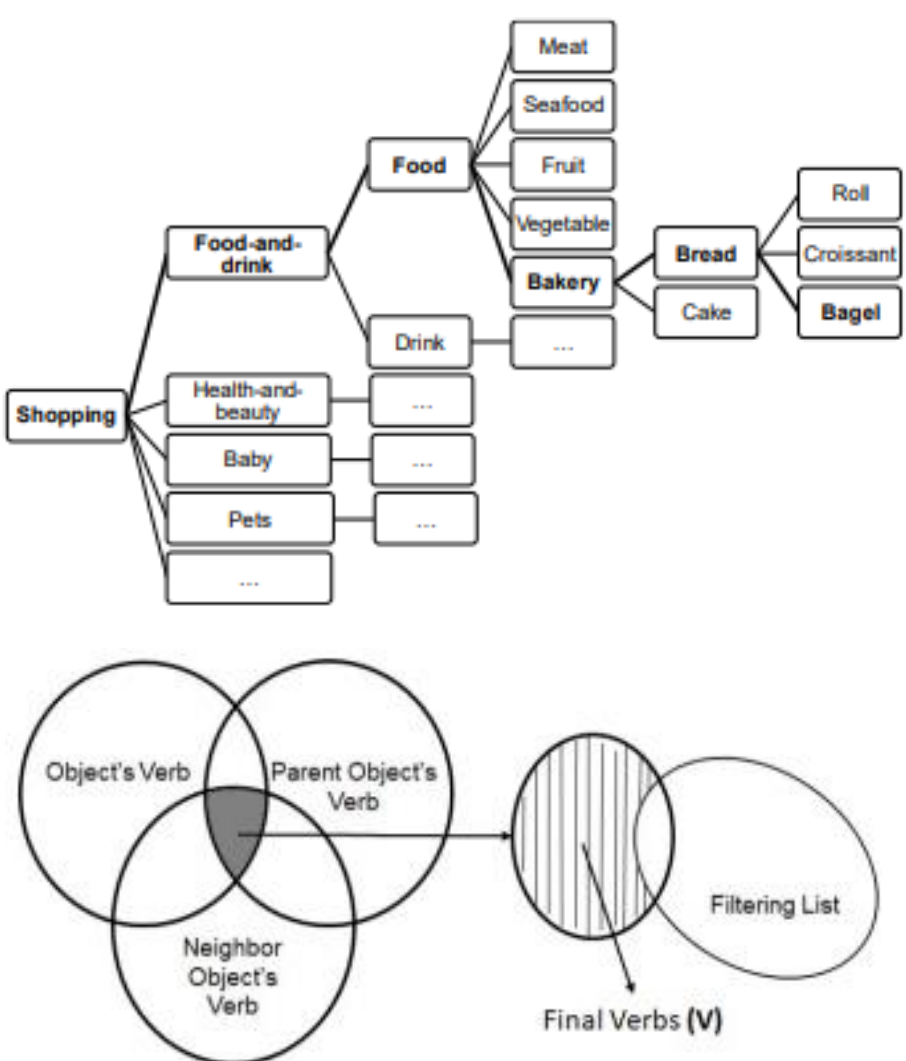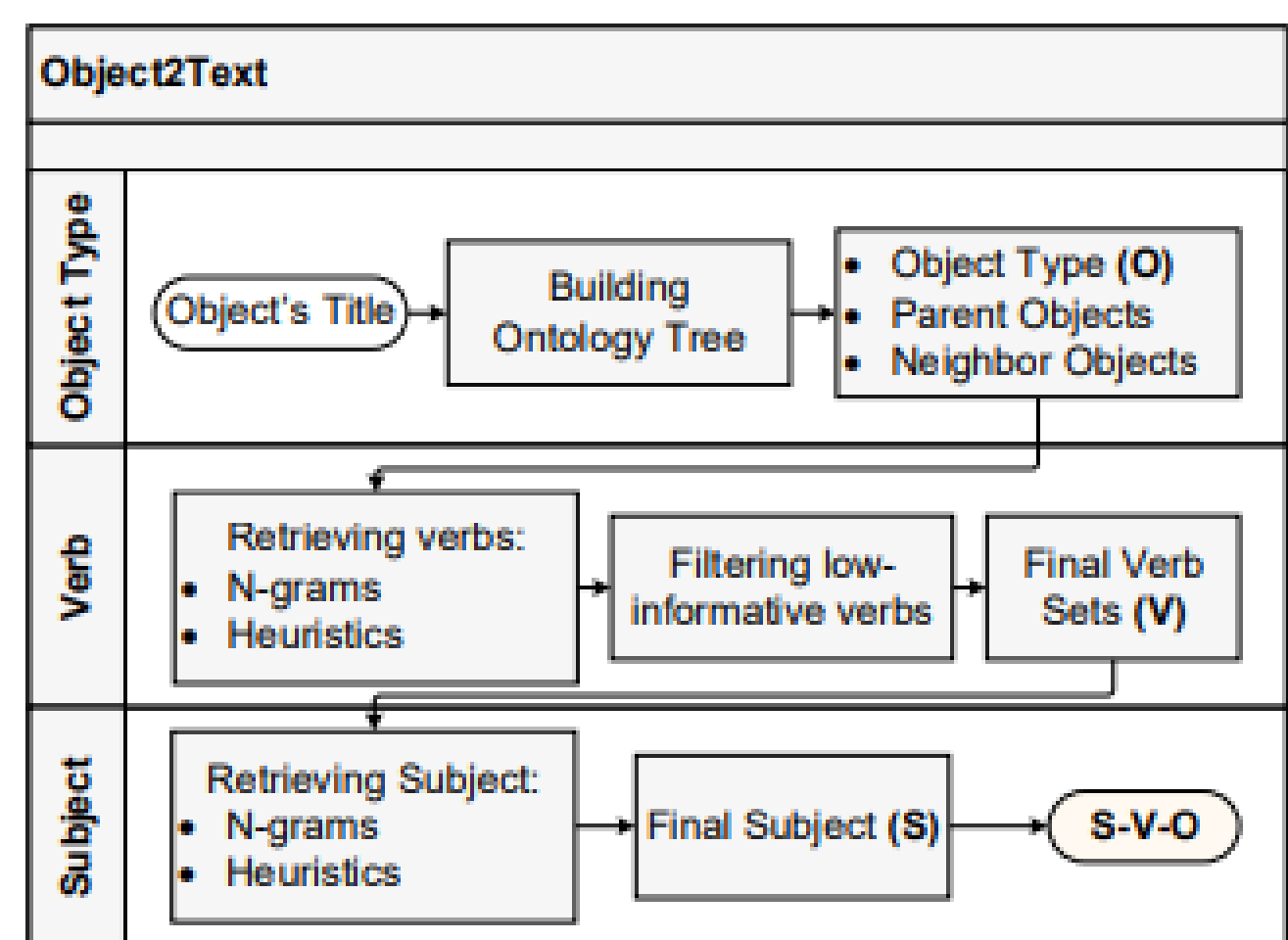
SOCIAL GOOD

System Design

Technical Challenges

- Semantic parsing / semantic analysis from unknown word to image queries.
- Word-sense-disambiguation mechanism.
- Linking words synsets to big datasets of descriptive images [5] .
- Ranking queries & evaluating metrics [4,5] .

System Components

- Input retrieval, Output display
- Object2text, Text2visual

- Result samples & System prototype:

Woman lighting candle

Evaluation

System Prototype

- Case study: grocery products shopping
- 114,522 products' description

Baselines & SIMPLER VOICE

- Controlled user study:
  - 15 foreign subjects does not speak/comprehend English.
  - 3 experts who were bilingual in English and the subjects' native language
  - A subset of random 20 U.S. products.

- Mean opinion score (MOS)

Approach	Approach Description	MOS Score Range	Mean MOS Score	Stdev
Baseline 1	1. Product title 2. Original product package	1 - 4.25	2.57	1.17
Baseline 2	1. Product title 2. Top product images using search engines: Google, and Bing	1 - 4.7	2.86	1.27
SimplerVoice	1. SimplerVoice generated text 2. SimplerVoice generated visual components: photorealistic images, pictographs	3.75 - 5	4.82	0.35

- Paired t-test of MOS scores of subjects' responses among products (baseline 1 & Simpler Voice): $t=-8.18224$, $p =1.19747e-07$

Feedback

- Participants in the Literacy Coalition of Central Texas's SPARK program

References

[1] UNESCO https://en.unesco.org/commemorations/literacyday (2018)
[2] C. Larry, et al. "The Impact of a Reading Intervention for Low-Literate Adult ESL Learners". National Center for Education Evaluation and Regional Assistance (2010).
[3] National Center for Education Statistics. "National Assessment of Adult Literacy" (2003).
[4] F. Schroff et al. "Harvesting Image Databases from the Web". IEEE transactions on pattern analysis and machine intelligence (2011)
[5] V. Vandeghinste et al. "Translating Text into Pictographs". Natural Language Engineering (2015)