Creating Macaronic Interface for Language Learning

Adithya Renduchintala, Rebecca Knowles, Philipp Koehn, Jason Eisner

CENTER FOR LANGUAGE AND SPEECH PROCESSING

JOHNS HOPKINS UNIVERSITY

RESEARCH GOAL

- Investigate the feasibility of using *Macaronic Text* to assist language learning.
- Macaronic: mixture of native language (L1) with second language (L2).
- Present a learner with appropriately leveled macaronic text to allow them to read and learn new vocabulary and simple linguistic structures.
- Examples:

The	*Police*	verhaftete	*the*	*criminals.*
Die	Polizei	verhaftete	*the*	*criminals.*
Die	Polizei	verhaftete	die	Straftäter.

TASK & REQUIREMENTS

- Automatically generate macaronic text from monolingual (L2) content.
- Generate full spectrum of macaronic states.
- Provide a learner with the ability to explore and interact with the macaronic content.

MACARONIC INTERFACE INTERACTION

- A Learner can interact with and explore the text via two main actions: Translation and Reordering.

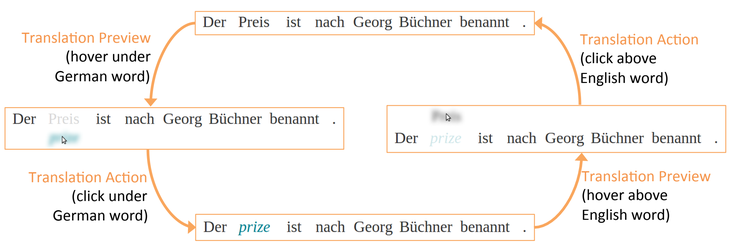

Fig 1: Translation action and translation preview interaction flow.

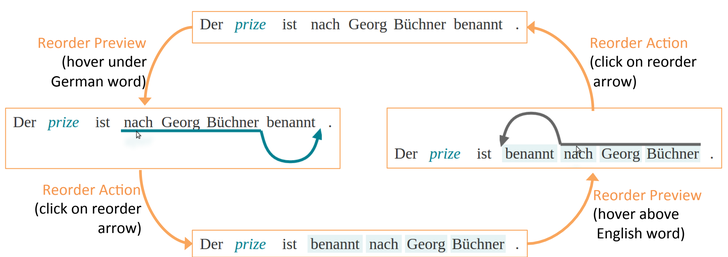

Fig 2: Reordering action and reordering preview interaction flow.

MACARONIC TEXT GENERATION

- Moses (Koehn et al., 2007) was used to translate L2 content to L1 with associated word alignments.
- Alignments were converted to *Minimal Alignments* i.e. 1-to-1, 1-to-many, or many-to-1, with no null alignments.
- Minimal Alignments ensure consistent reversibility of actions.
- Alignments form small connected components called *Units*.

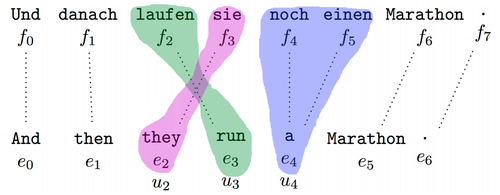

Fig 3: Source sentence (L2) with translation (L1) and minimal alignments constituting 7 units (only units 2,3 and 4 are highlighted for clarity of image).

- A unit forms a bipartite graph; the language of the words (which form nodes) defines two disjoint sets.
- Macaronic sentences can then be produced by selecting the display language for each unit.

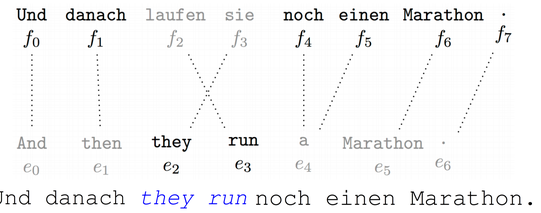

Fig 4: Selecting English as the display language for units u2 and u3 results in the following displayed macaronic sentence.

- Reordering is handled by changing the unit order of the macaronic sentence.
- Possible orderings for a sentence pair are defined by a bracketing ITG tree (Wu, 1997)
- Translation and Reordering act independently, which allows for a large space of macaronic states to be displayed.

String Rendered	Unit Ordering
...*they run*...	$\{u_2\} < \{u_3\}$
...*they* laufen...	
...sie *run*...	
...sie laufen...	
...*run they*...	$\{u_2\} > \{u_3\}$
...*run* sie...	
...laufen *they*...	
...laufen sie...	

Table 1: Generating reordered strings using units along with possible translations for each unit ordering.

ADDITIONAL FEATURES

- *Slider Feature:* An alternate version of the UI allows a learner to control a slider, which changes the "macaronicity" of the displayed content in real-time.
- *Pop Quiz Feature:* Occasionally, when a learner requests a translation action from German (L2) to English (L1), the system responds with a "Pop Quiz." The learner is prompted to guess the translation. Once a guess is entered, the system scores the guess and gives the learner feedback.

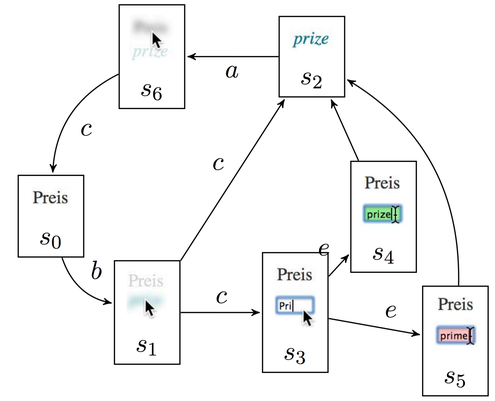

Fig 5: State diagram of learner interaction (edges) and the system's response (vertices). Edges can be traversed by clicking (c), hovering above (a), hovering below (b) or pressing the enter (e) key. Unmarked edges indicate an automatic transition.

- This feature allows the system to update its model of the user's comprehension of macaronic sentences.

FUTURE WORK

- Reduce translation errors that result in poor quality macaronic sentence.
- Address issues with cross-linguistic divergence(Dorr, 1994).
- Provide ability for sub-word exploration and more interaction refinements.
- Integrate macaronic interface with continuously updated model of learner's comprehension(Renduchintala et al. 2016).

REFERENCES

Bonnie J. Dorr. 1994. Machine translation divergences: A formal description and proposed solution. *Computational Linguistics*.

Philipp Koehn, Hieu Hoang, Alexandra Birch, Chris Callison-Burch, Marcello Federico, Nicola Bertoldi, Brooke Cowan, Wade Shen, Christine Moran, Richard Zens, et al. 2007. Moses: Open source toolkit for statistical machine translation. In *Proceedings of ACL*

Adithya Renduchintala, Rebecca Knowles, Philipp Koehn, and Jason Eisner. 2016. User modeling in language learning with macaronic texts. In *Proceedings of ACL*.

Dekai Wu. 1997. Stochastic inversion transduction grammars and bilingual parsing of parallel corpora. *Computational Linguistics*