Creating Interactive Macaronic Interfaces for Language Learning

Adithya Renduchintala and **Rebecca Knowles** and **Philipp Koehn** and **Jason Eisner**
Department of Computer Science
Johns Hopkins University
`{adi.r,rknowles,phi,eisner}@jhu.edu`

Abstract

We present a prototype of a novel technology for second language instruction. Our learn-by-reading approach lets a human learner acquire new words and constructions by encountering them in context. To facilitate reading comprehension, our technology presents mixed native language (L1) and second language (L2) sentences to a learner and allows them to interact with the sentences to make the sentences easier (more L1-like) or harder (more L2-like) to read. Eventually, our system should continuously track a learner's knowledge and learning style by modeling their interactions, including performance on a pop quiz feature. This will allow our system to generate personalized mixed-language texts for learners.

1 Introduction

Growing interest in self-directed language learning methods like Duolingo (von Ahn, 2013), along with recent advances in machine translation and the widespread ease of access to a variety of texts in a large number of languages, has given rise to a number of web-based tools related to language learning (ranging from dictionary apps to more interactive tools like Alpheios (Nelson, 2007) or Lingua.ly (2013)). Most of these either focus on vocabulary learning or require hand-curated lesson plans. We present a prototype of a system for learning to read in a foreign language, which presents learners with text consisting of a mix of their native language (L1) and the language they are interested in learning (L2). We refer to sentences containing a mix of L1 and L2 text as **macaronic**[1] sentences. Along the continuum from fully L1 to fully L2 text are sentences with any combination of L1 and L2 vocabulary, syntax, and (potentially) morphology.

Proponents of language acquisition through extensive reading, such as Krashen (1989), argue that much of language acquisition takes place through incidental learning—when a learner is exposed to novel vocabulary or structures and must find a way to understand them in order to comprehend the text. The trouble is that learning by reading already requires considerable L2 fluency. To bootstrap, we propose making L2 sentences more accessible to early learners by shifting these sentences along the macaronic spectrum towards L1, stopping at the "zone of proximal development" (Vygotskiǐ, 2012) where the learner is able to comprehend the text but only by stretching their L2 capacity. We aim in the future to customize macaronic sentences to each individual learner.

A reasonable concern is whether exposure to macaronic language might actually harm acquisition of correct L2 (even though our interface uses color and font to mark the L1 "intrusions" into the L2 sentence). As some reassurance, our approach is analogous to the well-established paradigm of inventive spelling (or "invented spelling"),[2] in which early writers are encouraged to write in their native language without concern for correct spelling, in part so they can more fully and happily engage with the writing challenge of composing longer and more authentic texts (Clarke, 1988). We also observe that simultaneous dual language acquisition—from multilingual and code-switched language—is common for young children in many countries, who employ code-switching in a socially appropriate way and as "a resource ... to fill gaps in their developing

---

[1] The term "macaronic" traditionally refers to a mashup of languages, often intended to be humorous. We use this term, rather than "code-switching," since code-switching requires the speaker/writer to be fluent in both languages. Code-switching is governed by syntactic and pragmatic considerations, rather than by pedagogical or humorous ones.

[2] Spelling, like L2, is a type of linguistic knowledge that is acquired after L1 fluency and largely through incidental learning (Krashen, 1993).

languages" (Genesee, 2009). Still, it remains an open question whether older students can successfully unlearn initial habits and move toward an increasingly complete and correct L2 model.

We envision our technology being used *alongside* traditional classroom L2 instruction—the same instructional mix that leads parents to accept inventive spelling (Gentry, 2000). Traditional grammar-based instruction and assessment, which use "toy" sentences in pure L2, should provide further scaffolding for our users to acquire language by reading more advanced (but macaronic) text.

We provide details of the current user interface and discuss how content for our system can be automatically generated using existing statistical machine translation (SMT) methods, enabling learners or teachers to choose their own texts to read. Our prototype is currently running on `http://www.clsp.jhu.edu:3030/` with sample content. Our interface lets the user navigate through the spectrum from L2 to L1, going beyond the single-word or single-phrase translations offered by other online tools such as Swych (2015), or dictionary-like browser plugins.

Finally, we discuss plans to extend this prototype and to integrate it with a continuously adapting user model. To this end, our companion paper (Renduchintala et al., 2016) develops an initial model of macaronic sentence comprehension by novice L2 learners, using data collected from human subjects via Amazon's Mechanical Turk service. In another paper (Knowles et al., 2016), we carry out a controlled study of comprehension of individual L2 words in isolation and in L1 context.

2 Macaronic Interface

For the purposes of this demo we assume a native English speaker (L1=English) who is learning German (L2=German). However, our existing interface can accommodate any pair of languages whose writing systems share directionality.[3] The primary goal of the interface is to empower a learner to translate and reorder parts of a confusing foreign language sentence. These translations and reorderings serve to make the German sentence more English-like. The interface also permits reverse transformations, letting the curious learner "peek ahead" at how specific English words and constructions would surface in German.

---

[3] We also assume that the text is segmented into words.

Der Preis ist nach Georg Büchner benannt .

(a) Initial sentence state.

Der Preis ist nach Georg Büchner benannt .

(b) Mouse hovered under `Preis`.

Der *prize* ist nach Georg Büchner benannt .

(c) `Preis` translated to `prize`.

Der *prize* ist nach Georg Büchner benannt .

(d) Mouse hovered above `prize`. Clicking above will revert the sentence back to the initial state 1a.

Der *prize* ist nach Georg Büchner *named* .

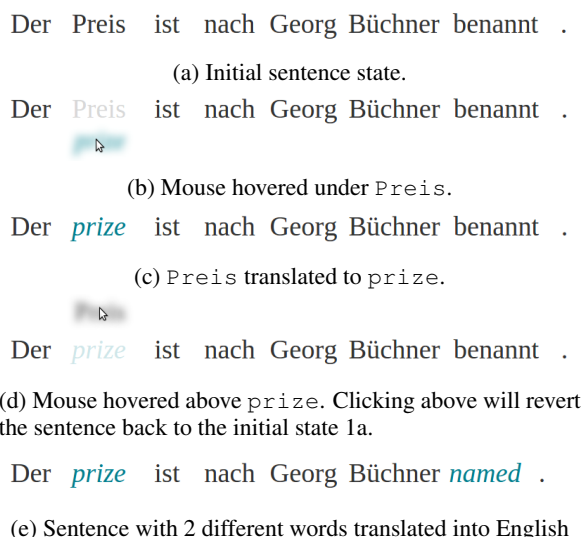(e) Sentence with 2 different words translated into English

Figure 1: Actions that translate words.

Using these fundamental interactions as building blocks, we create an interactive framework for a language learner to explore this continuum of "English-like" to "foreign-like" sentences. By repeated interaction with new content and exposure to recurring vocabulary items and linguistic patterns, we believe a learner can pick up vocabulary and other linguistic rules of the foreign language.

2.1 Translation

The basic interface idea is that a line of macaronic text is equipped with hidden interlinear annotations. Notionally, English translations lurk below the macaronic text, and German ones above.

The *Translation* interaction allows the learner to change the text in the macaronic sentence from one language to another. Consider a macaronic sentence that is completely in the foreign state (i.e.,, entirely in German), as shown in Fig. 1a. Hovering on or under a German word shows a *preview* of a translation (Fig. 1b). Clicking on the preview will cause the translation to "rise up" and replace the German word (Fig. 1c).

To translate in the reverse direction, the user can hover and click above an English word (Fig. 1d).

Since the same mechanism applies to all the words in the sentence, a learner can manipulate translations for each word independently. For example, Fig. 1e shows two words in English.

The version of our prototype displayed in Figure 1 blurs the preview tokens when a learner is hovering above or below a word. This blurred preview acts as a visual indication of a potential change to the sentence state (if clicked) but it also
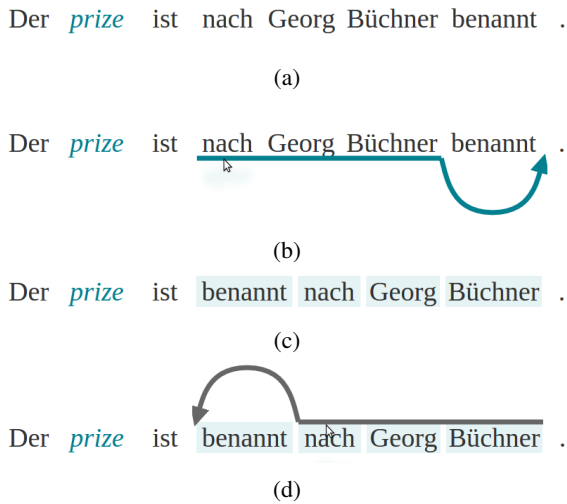
Der *prize* ist nach Georg Büchner benannt .

(a)

Der *prize* ist nach Georg Büchner benannt .

(b)

Der *prize* ist benannt nach Georg Büchner .

(c)

Der *prize* ist benannt nach Georg Büchner .

(d)

Figure 2: Actions that reorder phrases.

Figure 3: State diagram of learner interaction (edges) and system's response(vertices). Edges can be traversed by clicking (c), hovering above (a), hovering below (b) or the enter (e) key. Unmarked edges indicate an automatic transition.

gives the learner a chance to think about what the translation might be, based on visual clues such as length and shape of the blurred text.

2.2 Reordering

When the learner hovers slightly *below* the words `nach Georg Büchner` a *Reordering* arrow is displayed (as shown in Figure 2). The arrow is an indicator of reordering. In this example, the German past participle `benannt` appears at the end of the sentence (the conjugated form of the verb is `ist benannt`, or `is named`); this is the grammatically correct location for the participle in German, while the English form should appear earlier in the equivalent English sentence.

Similar to the translation actions, reordering actions also have a directional attribute. Figure 2b shows a German-to-English direction arrow. When the learner clicks the arrow, the interface rearranges all the words involved in the reordering. The new word positions are shown in 2c. Once again, the user can undo: hovering just *above* `nach Georg Büchner` now shows a gray arrow, which if clicked returns the phrase to its German word order (shown in 2d).

German phrases that are not in original German order are highlighted as a warning (Figure 2c).

2.3 "Pop Quiz" Feature

So far, we have described the system's standard responses to a learner's actions. We now add occasional "pop quizzes." When a learner hovers below a German word (s_0 in Figure 3) and clicks the blurry English text, the system can either reveal the translation of the German word (state s_2) as de-
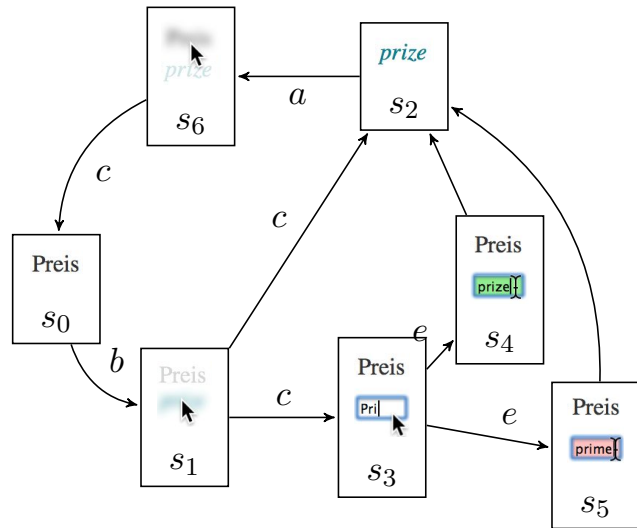
scribed in section 2.1 or quiz the learner (state s_3). We implement the quiz by presenting a text input box to the learner: here the learner is expected to type what they believe the German word means. Once a guess is typed, the system indicates if the guess is correct (s_4) or incorrect(s_5) by flashing green or red highlights in the text box. The box then disappears (after 700ms) and the system automatically proceeds to the reveal state s_2. As this imposes a high cognitive load and increases the interaction complexity (typing vs. clicking), we intend to use the pop quiz infrequently.

The pop quiz serves two vital functions. First, it further incentivizes the user to retain learned vocabulary. Second, it allows the system to update its model of the user's current L2 lexicon, macaronic comprehension, and learning style; this is work in progress (see section 4.2).

2.4 Interaction Consistency

Again, we regard the macaronic sentence as a kind of interlinear text, written between two mostly invisible sentences: German above and English below. In general, hovering above the macaronic sentence will reveal German words or word orders, which fall down into the macaronic sentence upon clicking. Hovering below will reveal English translations, which rise up upon clicking.

The words in the macaronic sentence are colored according to their language. We want the user to become accustomed to reading German, so the German words are in plain black text by de-

Action	Direction	Trigger	Preview	Preview Color	Confirm	Result
Translation	E-to-G	Hover above English	Blurry German translation above	Gray Blur	Click on Blurry Text	translation replaces English word(s)
	G-to-E	Hover under German token	Blurry English translation below	Blue Blur	Click on Blurry Text	translation replaces German word(s)
Reordering	E-to-G	Hover above token	Arrow above reordering tokens	Gray Arrow	Click on Arrow	tokens reorder
	G-to-E	Hover under token	Arrow below reordering tokens	Blue Arrow	Click on Arrow	tokens reorder

Table 1: Summary of learner triggered interactions in the Macaronic Interface.

fault, while the English words use a marked color and font (italic blue). Reordering arrows also follow the same color scheme: arrows that will make the macaronic sentence more "German-like" are gray, while arrows that make the sentence more "English-like" are blue. The summary of interactions is shown in Table 1.

3 Constructing Macaronic Translations

In this section, we describe the details of the underlying data structures needed to allow all the interactions mentioned in the previous section. A key requirement in the design of the data structure was to support orthogonal actions in each sentence. Making all translation and reordering actions independent of one another creates a large space of macaronic states for a learner to explore.

At present, the input to our macaronic interface is bitext with word-to-word alignments provided by a phrase-based SMT system (or, if desired, by hand). We employ Moses (Koehn et al., 2007) to translate German sentences and generate phrase alignments. News articles written in simple German from `nachrichtenleicht.de` (Deutschlandfunk, 2016) were translated after training the SMT system on the WMT15 German-English corpus (Bojar et al., 2015).

We convert the word alignments into "minimal alignments" that are either one-to-one, one-to-many or many-to-one.[4] This step ensures consistent reversibility of actions and prevents large phrases from being translated with a single click.[5] The resulting bipartite graph can be regarded as

Figure 4: The dotted lines show word-to-word alignments between the German sentence $f_0, f_1, \ldots, f_7$ and its English translation $e_0, e_1, \ldots, e_6$. The figure highlights 3 of the 7 units: u_2, u_3, u_4.
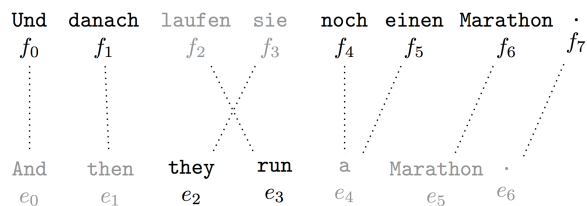
Figure 5: A possible state of the sentence, which renders a subset of the tokens (shown in black). The rendering order (section 3.2) is not shown but is also part of the state. The string displayed in this case is "Und danach *they run* noch einen Marathon." (assuming no reordering).

a collection of connected components, or *units* (Fig. 4).[6]

3.1 Translation Mechanism

In a given state of the macaronic sentence, each unit is displayed in either English or German. A translation action toggles the display language of the unit, leaving it in place. For example, in Figure 5, where the macaronic sentence is currently displaying $f_4 f_5 = $ `noch einen`, a translation action will replace this with $e_4 = $ `a`.

3.2 Reordering Mechanism

A reordering action changes the unit order of the current macaronic sentence. The out-

---

[4]For each many-to-many alignment returned by the SMT system, we remove alignment edges (lowest probability first) until the alignment is no longer many-to-many. Then we greedily add edges from unaligned tokens (highest probability first), subject to not creating many-to-many alignments and subject to minimizing the number of crossing edges, until all tokens are aligned.

[5]Preliminary experiments showed that allowing large phrases to translate with one click resulted in abrupt jumps in the visualization, which users found hard to follow.
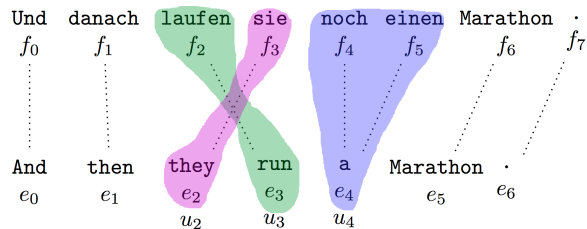
[6]In the sections below, we gloss over cases where a unit is discontiguous (in one language). Such units are handled specially (we omit details for reasons of space). If a unit would fall outside the bounds of what our special handling can handle, we fuse it with another unit.

put string "Und danach *they run* noch einen Marathon." is obtained from Figure 5 only if unit u_2 (as labeled in Figure 4) is rendered (in its current language) to the left of unit u_3, which we write as $u_2 < u_3$. In this case, it is possible for the user to change the order of these units, because $u_3 < u_2$ in German. Table 2 shows the 8 possible combinations of ordering and translation choices for this pair of units.

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

Table 2: Generating reordered strings using units.

The space of possible orderings for a sentence pair is defined by a bracketing ITG tree (Wu, 1997), which transforms the German ordering of the units into the English ordering by a collection of nested binary swaps of subsequences.[7] The ordering state of the macaronic sentence is given by the subset of these swaps that have been performed. A reordering action toggles one of the swaps in this collection.

Since we have a parser for German (Rafferty and Manning, 2008), we take care to select an ITG tree that is "compatible" with the German sentence's dependency structure, in the following sense: if the ITG tree combines two spans A and B, then there are not dependencies from words in A to words in B *and* vice-versa.

4 Discussion and Future Work

4.1 Machine Translation Challenges

When the English version of the sentence is produced by an MT system, it may suffer from MT errors and/or poor alignments.

Even with correct MT, a given syntactic construction may be handled inconsistently on different occasions, depending on the particular words involved (as these affect what phrasal alignment is found and how we convert it to a minimal alignment). Syntax-based MT could be used to design a more consistent interface that is also more closely tied to classroom L2 lessons.

---

[7]Occasionally no such ITG tree exists, in which case we fuse units as needed until one does.

Cross-linguistic divergences in the expression of information (Dorr, 1994) could be confusing. For example, when moving through macaronic space from Kaffee gefällt Menschen (coffee pleases humans) to its translation *humans like coffee*, it may not be clear to the learner that the reordering is triggered by the fact that *like* is not a literal translation of gefällt. One way to improve this might be to have the system pass smoothly through a range of intermediate translations from word-by-word glosses to idiomatic phrasal translations, rather than always directly translating idioms. We might also see benefit in guiding our gradual translations with cognates (for example, rather than translate directly from the German Möhre to the English *carrot*, we might offer the cognate Karotte as an intermediate step).

We also plan to transition through words that are macaronic at the sub-word level. For example, hovering over the unfamiliar German word gesprochen might decompose it into ge-sprochen; then clicking on one of those morphemes might yield ge-*talk* or sprech-*ed* before reaching *talked*. This could guide learners towards an understanding of German tense marking and stem changes.

4.2 User Adaptation and Evaluation

We would prefer to show the learner a macaronic sentence that provides just enough clues for the learner to be able to comprehend it, while still pushing them to figure out new vocabulary or new structures. Thus, we plan to situate this interface in a framework that continuously adapts as the user progresses. As the user learns new vocabulary, the system will automatically present them with more challenging sentences (containing less L1). In (Renduchintala et al., 2016) we show that we can predict a novice learner's guesses of L2 word meanings in macaronic sentences using a few simple features. We will subsequently track the user's learning by observing their mouse actions and "pop quiz" responses (section 2).

While we have had users interact with our system in order to collect data about novice learners' guesses, we are working toward an evaluation where our system is used to supplement classroom instruction for real foreign-language students.

5 Conclusion

In this work we present a prototype of an interactive interface for learning to read in a foreign language. We expose the learner to L2 vocabulary and constructions in contexts that are comprehensible because they have been *partially* translated into the learner's native language, using statistical MT. Using MT affords flexibility: learners or instructors can choose which texts to read, and learners or the system can control which parts of a sentence are translated.

We are working towards integrating models of learner understanding (Renduchintala et al., 2016; Knowles et al., 2016) to produce personalized macaronic texts that give each learner just the right amount of challenge and support. In the long term, we would like to extend the approach to allow users also to *produce* macaronic language, drawing on techniques from grammatical error correction or computer-aided translation to help them gradually remove L1 features from their writing (or speech) and make it more L2-like.

Acknowledgments

Supplemental Material

- A video demonstration can be found here: `https://youtu.be/d5lxyeHIDWI`
- A live sample version is here: `http://www.clsp.jhu.edu:3030/signin`

References

Ondřej Bojar, Rajen Chatterjee, Christian Federmann, Barry Haddow, Matthias Huck, Chris Hokamp, Philipp Koehn, Varvara Logacheva, Christof Monz, Matteo Negri, Matt Post, Carolina Scarton, Lucia Specia, and Marco Turchi. 2015. Findings of the 2015 Workshop on Statistical Machine Translation. In *Proceedings of the Tenth Workshop on Statistical Machine Translation*, pages 1–46.

Linda K. Clarke. 1988. Invented versus traditional spelling in first graders' writings: Effects on learning to spell and read. *Research in the Teaching of English*, pages 281–309, October.

Deutschlandfunk. 2016. nachrichtenleicht. `http://www.nachrichtenleicht.de/`. Accessed: 2015-09-30.

Bonnie J. Dorr. 1994. Machine translation divergences: A formal description and proposed solution. *Computational Linguistics*, 20(4):597–633, December.

Fred H. Genesee. 2009. Early childhood bilingualism: Perils and possibilities. *Journal of Applied Research on Learning*, 2(Article 2):1–21, April.

J. Richard Gentry. 2000. A retrospective on invented spelling and a look forward. *The Reading Teacher*, 54(3):318–332, November.

Rebecca Knowles, Adithya Renduchintala, Philipp Koehn, and Jason Eisner. 2016. Analyzing learner understanding of novel L2 vocabulary. To appear.

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: Interactive Poster and Demonstration Sessions*, pages 177–180.

Stephen Krashen. 1989. We acquire vocabulary and spelling by reading: Additional evidence for the input hypothesis. *The Modern Language Journal*, 73(4):440–464.

S. Krashen. 1993. How well do people spell? *Reading Improvement*, 30(1).

Lingua.ly. 2013. Lingua.ly. `https://lingua.ly/`. Accessed: 2016-04-04.

Mark Nelson. 2007. The Alpheios project. `http://alpheios.net/`. Accessed: 2016-04-05.

Anna N Rafferty and Christopher D Manning. 2008. Parsing three German treebanks: Lexicalized and unlexicalized baselines. In *Proceedings of the Workshop on Parsing German*, pages 40–46. Association for Computational Linguistics.

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

Swych. 2015. Swych. `http://swych.it/`. Accessed: 2016-04-05.

Luis von Ahn. 2013. Duolingo: Learn a language for free while helping to translate the web. In *Proceedings of the 2013 International Conference on Intelligent User Interfaces*, pages 1–2.

Lev Vygotskiĭ. 2012. *Thought and Language (Revised and Expanded Edition)*. MIT press.

Dekai Wu. 1997. Stochastic inversion transduction grammars and bilingual parsing of parallel corpora. *Computational Linguistics*, 23(3):377–404.