

## Building Sentences with Web 2.0 and the Tatoeba Database

Paul Raine<sup>1</sup>

*J. F. Oberlin University*

### ABSTRACT

Sentence translation activities have seen a resurgence in popularity in recent years, with a number of hugely successful language learning websites adopting translation as their pedagogical basis. This paper introduces *Sentence Builder*, a web-based tool that utilizes the Tatoeba sentence database and Web 2.0 technologies to generate effective and engaging sentence translation tasks. Pedagogical justifications for this tool include the argument that translation itself should be regarded as the fifth macro skill; that translation is highly amenable to gamification; and that the process of contrastive analysis and translation offer numerous benefits to language learners. The paper concludes by arguing that CALL-powered translation-based activities could be effectively exploited by both English language teachers and learners.

### INTRODUCTION

This paper introduces the web-based sentence translation activity *Sentence Builder*<sup>2</sup>, and examines the pedagogical justification for this tool and others like it. *Sentence Builder* utilizes the Tatoeba sentence database, which will be briefly described in the first part of this paper. Following this, the *Sentence Builder* activity itself will be explained. The final section of the paper will evaluate the possible pedagogical justifications for this type of activity, and conclude that *Sentence Builder* is an effective and engaging tool for teachers and learners of English as a second or foreign language.

### TATOEBA SENTENCE DATABASE

The Tatoeba sentence database ([www.tatoeba.org](http://www.tatoeba.org)) is a collection of approximate 6.7 million example sentences in 330 languages. It is crowd-sourced and available under a creative commons license. It includes sentences relating to a variety of topics and representing a wide range of grammatical structures. A typical sentence is shown in *Figure 1* (below):

#### FIGURE 1

---

<sup>1</sup> Paul Raine (BA, LLB, MA) has taught EFL in Japan since 2006, and lectures at three universities in the Tokyo area. He is particularly interested in Computer Assisted Language Learning. Correspondence should be sent to: [paul.raine@gmail.com](mailto:paul.raine@gmail.com)

<sup>2</sup> [http://www.apps4efl.com/activities/sentence\\_builder](http://www.apps4efl.com/activities/sentence_builder)

### A typical sentence pair from the Tatoeba database



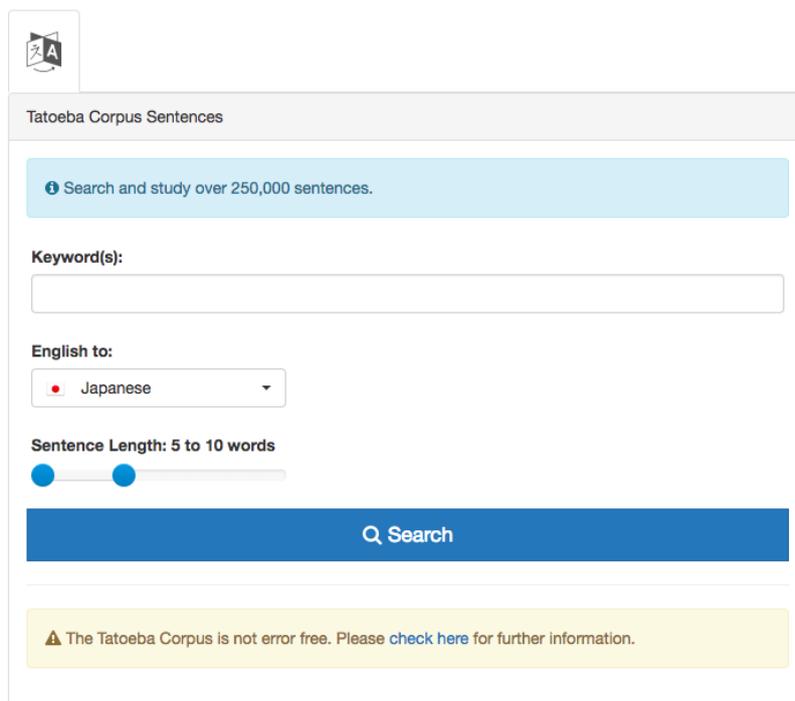
The entire Tatoeba database is available to download, and *Sentence Builder* incorporates a smaller subset of the database which includes only those sentences that are written in English and have Arabic, Dutch, French, German, Greek, Italian, Japanese, Korean, Chinese, Portuguese, Russian, Spanish, or Swedish translations.

## SENTENCE BUILDER

*Sentence Builder* is a web-based language learning activity based on a structuralist view of language. For example, it focuses on “breaking down and reconstructing the elements of language” (Raine, 2018, p.131).

*Sentence Builder* first requires users to input a keyword and select their native language. Users are also able to specify the approximate length of the sentences they wish to study (*Figure 2*, below).

**FIGURE 2**  
**Sentence Builder keyword input and language selection screen**



Tatoeba Corpus Sentences

Search and study over 250,000 sentences.

Keyword(s):

English to:

Japanese

Sentence Length: 5 to 10 words

Search

The Tatoeba Corpus is not error free. Please [check here](#) for further information.

A list of sentences containing the keyword input by the user will be returned after clicking “Search”. The user can then select the sentences they wish to study from these results, or choose a randomized selection by clicking on the “Random 10” button (Figure 3, below).

**FIGURE 3**  
Sentences returned for keyword “soccer” and native language “Japanese”

Found 95 sentences for 'soccer'

Study Select all Deselect all Random 10

#	English	Japanese
1	You will be able to play soccer.	貴方達は、明日サッカーの試合をする事ができます。
2	Our soccer team beat all the other teams in the town.	我々のサッカーチームは、町の全チームを負かした。
3	I was extremely disappointed to see our national soccer team suffer a historic loss.	わが国のサッカー代表チームが歴史的な大敗を喫しガックリした。
4	Our soccer team was covered in glory.	わがサッカーチームは栄光に輝いた。
5	Rugby, American football, and Australian rules football all come from soccer.	ラグビー、アメリカン・フットボール、そしてオーストラリア式フットボールはすべてサッカーに由来するものなのです。
6	In hockey and soccer, goals are counted as points.	ホッケーやサッカーでは点数はゴール数で数えられる。
7	No other sport in Brazil is so popular as soccer.	ブラジルではサッカーほど人気のあるスポーツはない。
8	Tom leads the soccer team.	トムがサッカーのチームを引率する。

After selecting the sentences that they wish to study, or choosing a randomized selection, the user has an initial chance to see each English sentence and its L1 translation. This information is

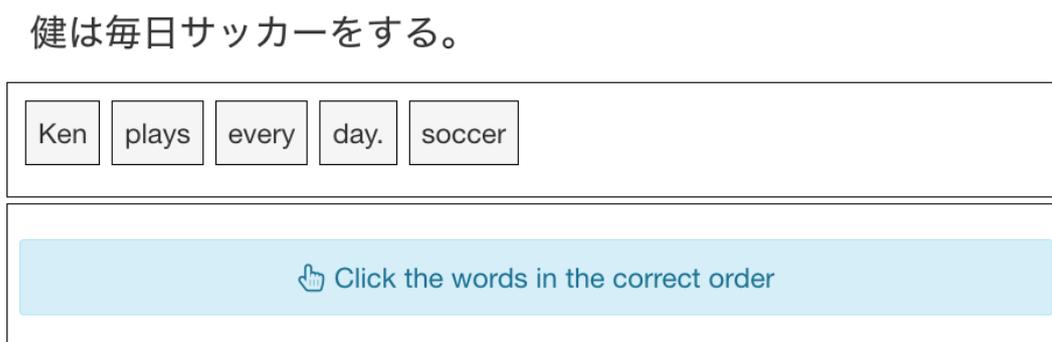
presented in the form of “digital flashcards” that can be “flipped” from one side to the other (*Figure 4*, below).

**FIGURE 4**  
**The sentences selected by the user are presented in the style of “digital flashcards”**



When the user is ready to begin the test, they can click on the “Start Test” button, and they are then presented with each sentence in a randomized order. The L1 translation of the sentence is shown without modification, but the words of the L2 version of the sentence are shown in a jumbled order. The aim of the activity is to click the English words in the correct order to form a sentence that matches the meaning of the L1 translation (*Figure 5*, below).

**FIGURE 5**  
**Reordering English words to match the meaning of an L1 translation**



The user is allowed to make three incorrect word selections before the question is marked “incorrect” and the correct L2 sentence is displayed. The activity concludes with a review of all the questions answered, and whether or not the user responded correctly. If the user is registered with the Apps 4 EFL web-based language learning platform (Raine, 2017), they will be awarded points

for each correct answer and their teacher will be able to track their progress over the course of a semester.

## **PEDAGOGICAL JUSTIFICATIONS**

Although most ELT experts (e.g., Larsen-Freeman & Anderson, 2011; Mukalel, 2005; Richards & Rodgers, 2016) do not recommend that Grammar Translation be adopted as the sole method of teaching or learning a foreign language, we should be cautious about dismissing the utility of translation activities altogether. Indeed, there are several persuasive arguments in favor of integrating translation-based activities into contemporary foreign language learning classrooms.

### **Translation is an essential macro skill**

In addition to listening, reading, writing, and speaking as the four well-established macro skills, it has been posited that translation should be regarded as the fifth macro skill (Campbell, 2002). The argument states that the forced separation of translation and language teaching is unnatural given the fact that two disciplines “share the same general aim of enhancing communication across language boundaries” (Campbell, 2002, p.59). Furthermore, macro functions of translation, such as “giving a rough idea in another language of the meaning of a document” (Campbell, 2002, p.59), seem highly likely to be encountered by language learners in both academic and business situations. This line of reasoning is supported by the Common European Framework of Reference (CEFR), which highlights “the necessity for language programmes to enhance the development of language users’ interlingual strategies and plurilingual competences” (Stathopoulou, 2016, p.761).

### **Translation is amenable to gamification**

Another argument in favor of translation activities in the context of CALL is that it is amenable to gamification, particularly “game-informed” learning, such as goal-oriented, rule-based, playful and voluntary activities (Reinhardt & Sykes, 2014). Translation can be scaffolded, and the words of the target translation can be provided in a random order, rather than requiring the user to generate the entire translation from scratch. Gamification techniques, such as sound or user interface effects can then be integrated into the answer process. Most importantly, instant feedback can be given, and points can be awarded, both of which have been shown to be effective in engaging and maintaining user attention and motivation (Garcia, 2013).

### **Benefits of contrastive analysis and translation**

Finally, the benefits of translation and contrastive analysis and translation should also be acknowledged. *Sentence Builder* could help highlight the differences and similarities between English and learners’ L1. It could also draw attention to the interaction between the two languages and highlight the possible interference of one’s L1 (Kaye, 2009). By doing so, it might also be possible to prevent the development of errors (Baker, 2003). There is also strong evidence that translation tasks are more effective than other “focus-on-form” activities in causing learners to acquire new vocabulary and improve their collocational knowledge (Laufer & Girsai, 2008).

## **CONCLUSION**

CALL-powered translation-based activities can be both engaging and effective. *Sentence Builder* aims to provide an easy and convenient way for students to understand and appreciate how certain ideas in their first language can be expressed in written English. It incorporates elements of game-informed learning in order to keep users engaged and motivated. The type of contrastive analysis and translation activities offered by *Sentence Builder* have been shown to be useful for improving vocabulary and collocational knowledge, as well as allowing students to develop essential interlingual strategies and competencies. The *Sentence Builder* activity and other similar tools are hoped to be developed further in the future.

## REFERENCES

- Baker, M. (2003). *Routledge Encyclopedia of Translation Studies*. Routledge.
- Campbell, S. (2002). *Translation in the Context of EFL-The Fifth Macroskill? TEFLIN Journal*, 13(1), 58-72.
- Garcia, I. (2013). *Learning a Language for Free While Translating the Web. Does Duolingo Work? International Journal of English Linguistics*, 3(1), 19-25.
- Kaye, P. (2009). In *Translation activities in the language classroom*. Retrieved from <https://www.teachingenglish.org.uk/article/translation-activities-language-classroom>
- Larsen-Freeman, D., & Anderson, M. (2011). *Techniques & principles in language teaching*. Oxford: Oxford University Press.
- Laufer, B. & Girsai, N. (2008). Form-focused Instruction in Second Language Vocabulary Learning: A Case for Contrastive Analysis and Translation. *Applied Linguistics*, 29(4), 694–716.
- Mukalel, J. C. (2005). *Approaches to English language teaching*. Discovery Publishing House.
- Raine, P. (2017). Web-based language learning with creative commons data. *Accents Asia*, 9(2), 41-55.
- Raine, P. (2018). *Developing Web-based English Learning Applications: Principles and Practice. CALL-EJ*, 19(2), 125-138
- Reinhardt, J., & Sykes, J. M. (2014). Special Issue Commentary: Digital Game Activity in L2 Teaching and Learning. *Language Learning & Technology*, 18(2), 2–8. Retrieved from <http://llt.msu.edu/issues/june2014/commentary.pdf>
- Richards, J. C., & Rodgers, T. S. (2016). *Approaches and methods in language teaching*. Cambridge: Cambridge University Press.
- Stathopoulou, M. (2016). *From 'languaging' to 'translanguaging': Reconsidering foreign language teaching and testing through a multilingual lens*. Retrieved from <http://ejournals.lib.auth.gr/thal/article/view/5267>