

Erratum to: A Statistical Approach to Machine Translation

Peter F. Brown
Stephen A. Della Pietra
Fredrick Jelinek
Robert L. Mercer

John Cocke
Vincent J. Della Pietra
John D. Lafferty
Paul S. Roossin

In Section 6 of "A statistical approach to machine translation" (*Computational Linguistics* 16(2), 79-85), we reported the results of two experiments in which we estimated parameters of a statistical model of translation from English to French.

In the first experiment, the English and French vocabularies each consisted of 9,000 common words, and the model parameters were estimated from 40,000 pairs of sentences 25 words or less in length. Words outside the 9,000-word vocabularies in these sentences were mapped to special *unknown* words.

In the second experiment, the vocabularies were limited to 1,000 common English words and 1,700 common French words, and the model parameters were estimated from 117,000 pairs of sentences 10 words or less in length that were completely covered by the respective vocabularies.

In Figures 4, 5, and 6 of the paper, we erroneously presented parameter estimates from the 1,000-word experiment, while claiming in the text that they were from the 9,000-word experiment. The parameter estimates for these two experiments differ considerably because of the restriction of the training corpus in the 1,000-word experiment to short, covered sentences. For example, the probability that *hear* is translated as *bravo*

English: <i>the</i>			
French	Probability	Fertility	Probability
le	.443	1	.856
la	.207	0	.140
les	.184		
l'	.097		
ce	.018		
il	.012		

Figure 4
Probabilities for *the*.

English: <i>not</i>			
French	Probability	Fertility	Probability
ne	.482	2	.728
pas	.455	0	.153
non	.029	1	.114
rein	.012		

Figure 5
Probabilities for *not*.

English: <i>hear</i>			
French	Probability	Fertility	Probability
bravo	.808	1	.527
entendre	.079	0	.472
entendu	.026		
entends	.024		
entendons	.013		

Figure 6
Probabilities for *hear*.

is .992 in the 1,000-word experiment (see Figure 6 of the paper)¹, while it is only .808 in the 9,000-word experiment (see Figure 6 above). This difference is due to the fact that the sentence pair (*Hear, hear!* | *Bravo!*) is extremely common in our data and is completely covered by the 1,000-word and 1,700-word vocabularies.

Figures 4, 5, and 6 contain the parameter estimates from the 9,000-word experiment. Only probabilities greater than or equal to .01 are reported.

¹ We thank Ken Church for pointing out that this estimate is not consistent with the frequency with which *hear* translates to *bravo* in other data from the same source.