Frequency Effects on Letter Transpositions Including Computer Cascade Modeling

Ruth M. Jones (RUTHJONE@CATS.UCSC.EDU)
Alan H. Kawamoto (AHK@CATS.UCSC.EDU)
Program in Cognitive Psychology at University of California at Santa Cruz
Social Sciences II, Santa Cruz, CA 95064 USA

In reading, errors are sometimes created due to letter transposition; e.g., reading FROM when seeing the word FORM. According to Estes (1975), these transposition letter errors were systematically related to the familiarity (e.g., frequency) ratio of the bigrams involved, while Grainger and Jacobs (1993) reported that the best predictor of partial-word priming effects was the positional letter frequency. An experiment with its corresponding simulation was conducted to test for the location of the transposition letter errors.

In the experiment, an investigation of the frequency ratios of letters, bigrams, and words using a forced-choice task with the four-letter monosyllabic word pairs varying only by the transposition of their middle two letters (e.g., FROM and FORM) was conducted. A stepwise multiple regression procedure was used for analysis based on the data from 72 participants. This research showed that the positional single letter frequency ratio accounted for 0.45 of the variance with bigram frequency and word frequency ratios adding only an additional 0.06 in the forced-choice task of word recognition. In addition within the word pairs of the four filler groups which varied in their letter position change, there was a significant difference between the groups with a monotonic increase in average errors across the letter positions from left to right (i.e., first as BELT and MELT with 3.67, second as FIND and FOND with 4.65, third as CROW and CREW with 5.32, and fourth as KIND and KING with 8.07).

This was followed by carrying out a computer simulation. A feedforward network with one layer of 52 hidden units was used. The input for each four-letter word had 312 units, i.e. 12 units per letter of the alphabet with the first, second, third, and fourth word letter positions respectively in overlapping units 1-5, 4-8, 6-10, and 9-12 (see Table 1). The output had 104 units, i.e. 26 units for the letters of the alphabet in first, second, third, and fourth letter position. After training the network using the generalized delta rule on the 1653 four-letter words based on their relative frequencies from Kucera and Francis (1967) with weights updated after each word trial, the words from the experiment were processed by cascading them through the feedforward network. To specifically illustrate the results of the experiment, three word pairs (SNUG/SUNG, FROM/FORM, and SLAT/SALT) will be discussed.

From the forced-choice experiment, the positional single letter frequency ratio was the best predictor over word and/or bigram frequency ratios. For the word, SNUG (with nine errors in the experiment), having lower word, bigram, and positional letter frequencies than its counterpart, SUNG (with two errors), the most strongly activated letters throughout the cascade process were those of SUNG and reflected the results of the experiment. For FROM (with nine errors) having higher word and bigram frequencies than FORM (with one error), yet lower positional letter frequencies, the second letter position in the cascade process initially had a stronger 'O' than 'R' and then were reversed much later. The model indicated that a quicker response to FROM would be the error, FORM, which was shown in the experiment. For SLAT (with nine errors) having higher bigram frequency but lower word and positional letter frequencies except for the third letter position than SALT (with one error), the cascade process noted little difference in strength of 'A' or 'L' in either position between words, but a noted difference between 'A' and 'L' at a specific position. For the SLAT/SALT pair, the word and positional letter frequencies supported each other in correctly choosing SALT more often than SLAT; while for the FROM/FORM pair, the positional letter frequency supported correctly choosing FORM more often than FROM.

The modeling of the experiment with the feedforward cascade model has supported the finding that specific letter position frequencies have a greater effect in word recognition for reading than either word or bigram frequency.

Table 1: Letter Positional Representation for Input

Word Length	Letter Position	Units On 1 2 3 4 5 6 7 8 9 0 1 2	Active Unit Total
4	1	x x x x x	
	2	x x x x x	
	3	x x x x x	
	4	x x x x x	20

References

Estes, W. K. (1975). The locus of inferential and perceptual processes in letter identification. *Journal of Experimental Psychology: General, 104,*(2), 122-145.

Grainger, J., & Jacobs, A. M. (1993). Masked partial-word priming in visual word recognition: Effects of positional letter frequency. *Journal of Experimental Psychology: Human Perception and Performance, 19, 951-964.*