Solution

(a) H, I, S, T, and U are the only letters whose tile count either increases or decreases by more than one.

(b) It explains H, S, T, and U.

H, S, and T are all more common in this list even than their *overall* token frequency, which is already higher than their Scrabble frequency.

U does not appear in any of the most common words, so its underrepresentation in overall token frequency is also accounted for.

The list does not explain the letter I, however. I is less common in overall token frequency than its Scrabble frequency would suggest, but it is *more* common in the most frequent twenty words.

(c) The frequencies of the words in the list explain H and T even better, because they both occur in *the*. It continues to explain U, which continues to be absent.

The data no longer explain S as well, because S does not occur in the most frequent of these words.

I, however, is now better explained. It is underrepresented in the top seven words (those more frequent than one percent, comprising 21% of all tokens) among vowels. E occurs only once, but in the most common (by far) word. A and O both occur twice in the rest of the top five. On the other hand, I only occurs in *in*. Therefore these new data *do* help explain why I might have about the same number of Scrabble tiles as A and O, though it has a smaller corpus frequency.