Representing sentences:

X: Hello   I   am   Asier (...) $\Rightarrow$ Input sequence

$\quad x^{<1>} \quad\quad x^{<2>} \quad x^{<3>} \quad x^{<4>} \quad (...) x^{<T>}$

- $T_x \Rightarrow$ length of input sequence
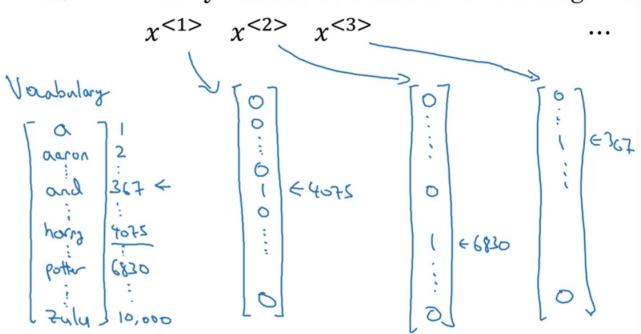
- $x^{(i)<t>} \Rightarrow$ the input at the position t of the ith training example.

- $T_x^{(i)} \Rightarrow$ length of input sequence of the ith training example

Representing words in a sentence:

One-hot representation. The goal is, given the mapping, to learn a sequence model to the target output Y.

For a word not present in the vocabulary, we define a token <UNK>.

RNNs:

$a^{<i>} \Rightarrow$ activation entering layer i. $a^{<0>}$ usually $\vec{0}$

$W_{ax} \Rightarrow$ W is multiplied by $x^{<i>}$

W is used to compute $a^{<i>}$

Deep RNNs:

$a^{[\ell]<t>}$: activation of layer ℓ, at time t.