Successful linguistic communication relies on comprehenders making predictions about upcoming words. Here we ask: 1) Is such prediction an all-or-none phenomenon (in which a word is either pre-activated or not) or a probabilistic phenomenon (in which a word’s pre-activation is proportional to its predicted probability)? and 2) Does the need to suppress competing alternatives increase the difficulty of prediction and/or lexical processing when a small number of candidates are likely? In an Event-Related Potential (ERP) experiment, we manipulate whether constraining sentence contexts have only one or more than one likely continuation, and whether participants see the most expected word, the second-most expected word (in cases with more than one likely continuation), or an unexpected word. Consistent with a probabilistic prediction account, we find a graded N400 effect in which N400 amplitudes increase with decreased word probability, replicating DeLong, Urbach, & Kutas (2005). We additionally find a late anterior positivity in response to unexpected words in constraining sentence contexts, replicating Federmeier et al. (2007). In future work, we plan to investigate whether this positivity varies gradiently with sentence constraint, as predicted by the probabilistic prediction account. Comparing responses to the most expected word in contexts with more than one likely continuation versus contexts with only one likely continuation, we saw no differences in N400 responses, providing no evidence for an effect of competition between alternatives on the difficulty of lexical processing.