Lexical	We use the lemmas, POS tags, and phrase lengths of p_1 and p_2, the substrings shared by p_1 and p_2, and the Levenstein, Jaccard, and Hamming distances between p_1 and p_2.
Distributional	Given a dependency context vectors for p_1 and p_2, we compute the number of shared contexts, and the Jaccard, Cosine, Lin1998, Weeds2004, Clarke2009, and Szpektor2008 similarities between the vectors.
Paraphrase	We include 33 paraphrase features distributed with PPDB, which include the paraphrase probabilities as computed in Bannard and Callison-Burch (2005). We refer the reader to Ganitkevitch and Callison-Burch (2014) for a complete description of all of the features included with PPDB.
Translation	We include the number of foreign language "pivots" (translations) shared by p_1 and p_2 for each of 24 languages used in the construction of PPDB, as a fraction of the total number of translations observed for each of p_1 and p_2.
Path	We include a sparse vector of all lexico-syntactic patterns (paths through a dependency parse) which are observed between p_1 and p_2 in the Annotated Gigaword corpus (Napoles et al., 2012).
WordNet	We include binary features indicating whether WordNet classifies p_1 and p_2 according to any of the following relations: synonym, hypernym, hyponym, antonym, holonym, meronym, cause, entailment, derivationally-related, similar-to, also-see, or attribute.