

S1-1. Antonyms – Candidates

The general processes of generating Lexicon antonyms are described as shown in Figure 1:

- 1) Generate antonym candidates from source of LEX|SD|PD|CC|SN
- 2) Manually tag the candidates by linguists
- 3) Validate and fix tagged file
- 4) Update tagged file to annual release tagged file.
- 5) Generate antonym and negation detection cue words

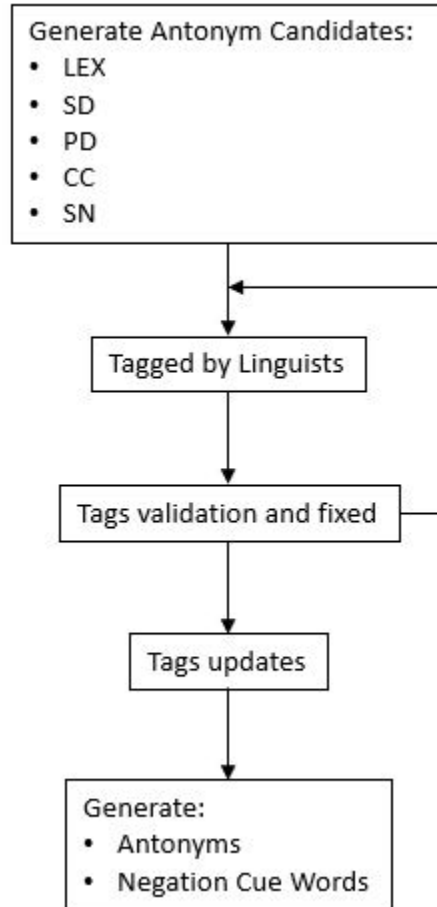


Figure 1. The general processes of generating Lexicon antonyms.

The processes in steps 2 to 5 and file format of the above five steps are similar for antonyms derived from the various sources (LEX|SD|PD|CC|SN) even when the antonym candidate generation is different. The associated programs are located under the directory of 0.Antonym. This document describes the format of antonym candidates from Step 1.

1. Antonym Candidates – Format and Examples

The output file of antonym candidates has 10 fields. Each line represents an antonym pair (aPair) candidate and includes all associated information. Table-1 and Table-2 show examples of antonym candidates from the Lexicon (LEX) and Train and Test Set (TT).

Ant1	EUI1	Ant2	EUI2	POS	Canon	Type	Negation	Domain	Source
like	E0419447	unlike	E0063275	Prep	Y	AB	BN2	quality	LEX
with	E0065516	without	E0065526	Prep	Y	B	N2	existence	LEX
is	E0012152	isn't	E0012152	Aux			N2		LEX
could	E0014877	couldn't	E0014877	modal			N2		LEX
		little	E0037839	Adv			BN2		LEX
		neither	E0042155	Det			N2		LEX
		neither	E0042156	Pron			N2		LEX
		neither	E0783690	Conj			N2		LEX
		hardly	E0030831	Adv			BN2		LEX
		neither	E0783690	Conj			N2		LEX
			LEX

Table 1 Examples of antonym candidates from Lexicon (lexical records with negative tags)

Please note that blank cells in the table are automatically filled up with [XX-TBD], such as [ANTONYM_TBD], [EUI_TBD], [POS_TBD], [CANON_TBD], [TYPE_TBD], [NEG_TBD] and [DOMAIN_TBD].

Ant1	EUI1	Ant2	EUI2	POS	Canon	Type	Negation	Domain	Source
accept	E0006699	refuse	E0052488	verb					CC
bye	E0785958	hi	E0785959	noun					SN
start	E0057681	stop	E0058047	noun					CC
start	E0057680	stop	E0058048	verb					CC
with	E0065517	without	E0065527	conj					CC
always	E0008403	never	E0042565	adv					LEX

able	E0006510	unable	E0062907	adj					PD
possible	E0049058	impossible	E0033808	adj					PD
careful	E0015340	careless	E0015344	adj					SD
worthy	E0065647	worthless	E0065643	adj					SD
...

Table 2 Examples of antonym candidates from training and test set

Please note that the blank cells are automatically filled up with [XX-TBD], such as [EUI_TBD], [POS_TBD], [CANON_TBD], [TYPE_TBD], [NEG_TBD] and [DOMAIN_TBD].