

Is the message control relevant? Is there any information in this sentence that indicates the author is being controlled by someone or something, or in control of someone or something?	Yes	Control relevant Is the author in control or being controlled by others or circumstances?	In control Internal control <ul style="list-style-type: none"> How much is the author in control over other people? (0-5) How much is the author in control of their situation? (0-5) Is the author solely in control? (0-2) 	<ul style="list-style-type: none"> “Hard but gotta do it” “Is sitting by the fire getting ready to start a new book” “I am satisfied with all I got done”
			Not in control External control <ul style="list-style-type: none"> How much is the author being controlled by other people? (0-5) How much is the author controlled by the situation? (0-5) Is the author controlled by just 1 or more people? 	<ul style="list-style-type: none"> “Why does Facebook keep spamming me with ads?” “Hate that stupid swap test they use to test for strep throat these days.” “I am now without my pocket book thanks to some wonderful person (or people) last night.” “Makes you wonder how life moves on”
	No	Not control relevant		<ul style="list-style-type: none"> “From a technical standpoint, how does Facebook store their data?” “He is 17 now” “The damn building is 4 blocks away” “Happy Halloween!” “Coffee is so good”

Figure 1: The annotation scheme for labeling sentences with control information. 4000 sentences were categorized as either control-irrelevant, internal or external control by three trained Master’s students.

	P	R	F1
Lexical features			
<i>n-gram</i>	.854	.918	.885
<i>n-gram,LIWC</i>	.857	.920	.888
<i>n-gram,LIWC,ANEW</i>	.857	.919	.887
<i>n-gram,LIWC,ANEW,w2v</i>	.857	.919	.887
<i>Pronouns</i>	.858	1.00	.923
<i>unigram</i>	.881	.865	.898
<i>bigram</i>	.854	.764	.968
<i>trigram</i>	.849	.744	.990
Syntax-based features			
<i>Pronouns-SVO</i>	.876	1.00	.934
<i>VerbCat-all (M)</i>	.878	.997	.934
<i>VerbCat-SVO (M)</i>	.877	1.00	.934
<i>VerbCat-1-2-3PP (M)</i>	.862	.988	.920
<i>VerbCat-1PP (M)</i>	.984	.910	.946
<i>VerbCat-all (Lev)</i>	.903	.997	.948
<i>VerbCat-SVO (Lev)</i>	.905	.996	.948
<i>VerbCat-1-2-3PP (Lev)</i>	.867	.989	.924
<i>VerbCat-1PP (Lev)</i>	.917	.990	.952
<i>VerbCat-M-P,Modals,Voice</i>	.737	.989	.845
<i>VerbCat-Lev-P,Modals,Voice</i>	.737	.994	.847
<i>POS-ngram</i>	.849	.913	.880
<i>POS-ngram,VerbCat-1PP (M)</i>	.852	.911	.880
<i>POS-ngram, VerbCat-1PP (M), modals</i>	.852	.911	.880
Best combinations of lexical and syntactic features			
<i>n-gram,LIWC,w2v,VerbCat-M-IP</i>	.859	.918	.888
<i>POS-ngram,ANEW</i>	.850	.914	.880
<i>POS-ngram,ANEW,w2v</i>	.850	.913	.880
<i>POS-ngram,ANEW,VerbCat-M-IP</i>	.851	.912	.880

	P	R	F1
Lexical features			
<i>n-gram</i>	.726	.682	.702
<i>n-gram,LIWC</i>	.731	.687	.700
<i>n-gram,LIWC,ANEW</i>	.728	.687	.706
<i>n-gram,LIWC,ANEW,w2v</i>	.728	.687	.706
<i>Pronouns</i>	.236	.204	.218
<i>unigram</i>	.700	.677	.727
<i>bigram</i>	.639	.614	.675
<i>trigram</i>	.617	.552	.754
Syntax-based features			
<i>Pronouns-SVO</i>	.573	.730	.639
<i>VerbCat-all (M)</i>	.532	.189	.280
<i>VerbCat-SVO (M)</i>	.562	.307	.395
<i>VerbCat-1-2-3PP (M)</i>	.569	.495	.526
<i>VerbCat-1PP (M)</i>	.576	.473	.513
<i>VerbCat-all (Lev)</i>	.597	.246	.347
<i>VerbCat-SVO (Lev)</i>	.568	.332	.415
<i>VerbCat-1-2-3PP (Lev)</i>	.524	.422	.455
<i>VerbCat-1PP (Lev)</i>	.638	.482	.549
<i>VerbCat-M-P,Modals,Voice</i>	.552	.804	.655
<i>VerbCat-Lev-P,Modals,Voice</i>	.583	.769	.662
<i>POS-ngram</i>	.635	.688	.659
<i>POS-ngram,VerbCat-1PP (M)</i>	.639	.691	.663
<i>POS-ngram, VerbCat-1PP (M), modals</i>	.637	.694	.664
Best combinations of lexical and syntactic features			
<i>n-gram,LIWC,w2v,VerbCat-M-IP</i>	.700	.737	.718
<i>POS-ngram,ANEW</i>	.635	.688	.659
<i>POS-ngram,ANEW,w2v</i>	.635	.688	.659
<i>POS-ngram,ANEW,VerbCat-M-IP</i>	.638	.694	.664

Table 1: Precision, recall and F1-measure for predicting control-relevance and internal/external control (10-fold CV) from lexical and syntactic features.

(a) Control Relevance

POS sequence	r	Example
VB (Verb base form)	.27	We will open the book
VBG (Verb gerund)	.27	I didn't no my job consist of raking leaves lmaooo
VBD (Verb past tense)	.18	Brianna got her second tooth yesterday and learner how to clap
PRP (Personal pronoun)	.27	We will open the book
IN PRP (Preposition Personal pronoun)	.22	He ate and drank some of his milk
TO DT PRP (To determiner Personal pronoun)	-.06	To all my Children
NN TO PDT (Noun to Predeterminer)	-.05	Commit to the LORD whatever you do
DT VBZ (Determiner Verb 3rd person singular)	-.05	This is me in the mornings lol
WP DT (Wh-pronoun Determiner)	-.04	What a wonderful way to start Thanksgiving
UH (Interjection)	-.04	Oh finance, I am totally not a fan
CC NNS RB (Conjunction Plural noun Adverb)	-.04	and 2 hours later

(b) Internal Control

POS sequence	r	Example
VBZ (Verb 3rd person singular)	.18	I would love to have one of these
WRB (Wh-Adverb)	.09	Good times are the best when they are shared with family
VBZ RB (Verb Adverb)	.10	Is very annoyed , why bring the kids back (...)??
RB JJ (Adverb Adjective)	.09	I have 2 very tired babies after spending around 3 hours
NN VBZ (Noun Verb)	.17	The first one is my favorite
NN PRP (Noun Personal-pron)	.10	back to DayCare for Ryan he's feeling much better
VBD NN IN (Verb Noun Prep)	-.07	(..) you can't do anything by yourself
TO VB DT (To Verb Determiner)	-.07	is hoping not to get the flu right now
JJ VBP VBG (Adjective Verb Verb-gerund)	-.07	(..) keep forgetting to upload the pictures
NN NN TO (Noun Noun To)	-.07	I just opened the strawberry container to discover (...)

Table 2: Linguistic features that capture POS Ngram sequences and their relationship with control (positive values are control relevance and external control). Features with positive coefficients are shaded blue and those with negative coefficients are shaded red. Example illustrates a sentence that contains the sequence. All the depicted associations are significant ($p < 0.05$ after Benjamini-Hochberg correction).