

Contributors: Jason Kushner

### Findings Report for CTK Replication Project

In our replication of this study, our script was used to filter raw data from 1000 tweets down to 549 after eliminating duplicates and selecting for features which more accurately determine whether or not a tweet contains a Pronominal Relative Clause. Of these 549 tweets, 49 were filtered into the literary category, and 493 were filtered into the non-literary category. 150 sample tweets were then selected for annotation from the 493 non-literary tweets. These were tagged for relative clause type, head, and role. All tweets in our sample followed the pattern of “you + who” to form their relative clauses. None featured other pronouns before “who”. Annotation totals are as follows:

<b>RC-Type</b>	Appositive	Restrictive
Number	113	37

<b>RC-Head</b>	“you”	Other
Number	150	0

<b>Role</b>	Subject	Object	Other
Number	53	9	88

In comparison to the CTK data, our data differs pretty drastically in terms of the role played by the PRC and its head, but mirrors results regarding clause type. CTK data shows that PRC’s most commonly took the role of subject, followed by object, then other. Our data disagrees with this and finds other to be the most common role, followed by subject, then object. Our data also differs in clause heads, as CTK data showed that “you” was the least common head for both literary and non-literary PRC’s but it was our

most common and only head. CTK data showed a very low number of restrictive clause types, which fits with our data. Overall, our results tend to disagree with CTK data, but are similar in clause type.