APPENDIX RESULTS

We examined the differences between Human and ChatGPT stories in each retelling for all categories: nouns, verbs, adjectives, adverbs, pronouns, and prepositions. At the beginning, humans are using more words from all categories. In the first retelling, human used more nouns ($M\_{human}$ = 16.40, $M\_{GPT}$ = 13.30, \textit{U =} 5402, \textit{p =} .009), verbs ($M\_{human}$ = 16.19, $M\_{GPT}$ = 10.13, \textit{U =} 2730, \textit{p <} .001), adjectives ($M\_{human}$ = 4.49, $M\_{GPT}$ = 3.42, \textit{U =} 5205, \textit{p =} .003), adverbs ($M\_{human}$ = 5.17, $M\_{GPT}$ = 2.16, \textit{U =} 2775.5, \textit{p <} .001), pronouns ($M\_{human}$ = 9.53, $M\_{GPT}$ = 4.59, \textit{U =} 2940, \textit{p <} .001), and prepositions ($M\_{human}$ = 6.49, $M\_{GPT}$ = 5.18, \textit{U =} 5375.5, \textit{p =} .008, see Figure 6). So, humans were benefiting from words more.

When, the stories are retold second time, there are no differences between Human and ChatGPT in the number of nouns (\textit{p =} .859), adjectives (\textit{p =} .512), and prepositions (\textit{p =} .788). However, humans still used significantly higher number of verbs ($M\_{human}$ = 11.64, $M\_{GPT}$ = 8.39, \textit{U =} 4472, \textit{p <} .001), adverbs ($M\_{human}$ = 3.65, $M\_{GPT}$ = 1.58, \textit{U =} 3393.5, \textit{p <} .001), and pronouns ($M\_{human}$ = 6.84, $M\_{GPT}$ = 3.69, \textit{U =} 3776.5, \textit{p <} .001).

Although humans seems to preserved their advantage on several categories, in the third retelling their use of words in these grammatical categories decreased. So that, ChatGPT used more nouns ($M\_{human}$ = 9.49, $M\_{GPT}$ = 10.66, \textit{U =} 8512.5, \textit{p <} .001), adjectives ($M\_{human}$ = 2.34, $M\_{GPT}$ = 2.75, \textit{U =} 7995.5, \textit{p =} .012), while humans still used higher number of verbs ($M\_{human}$ = 9.49, $M\_{GPT}$ = 7.84, \textit{U =} 5542, \textit{p =} .020), adverbs ($M\_{human}$ = 2.97, $M\_{GPT}$ = 1.39, \textit{U =} 4124, \textit{p <} .001), and pronouns ($M\_{human}$ = 5.49, $M\_{GPT}$ = 3.43, \textit{U =} 3776.5, \textit{p <} .001).

To get a better understanding, we also compared the ratio of these words from their grammatical categories to total number of words in a given sentence. These ratios displayed a clearer picture, so that ChatGPT used more nouns (first iteration: $M\_{human}$ = .21, $M\_{GPT}$ = .24, \textit{z =} -2.56, \textit{p =} .01, second iteration: $M\_{human}$ = .22, $M\_{GPT}$ = .25, \textit{z =} -2.78, \textit{p =} .005, third iteration: $M\_{human}$ = .20, $M\_{GPT}$ = .25, \textit{z =} -6.15, \textit{p <} .001), adjectives (no difference in first iteration, but in second iteration: $M\_{human}$ = .05, $M\_{GPT}$ = .07, \textit{z =} -2.97, \textit{p =} .003, and third iteration: $M\_{human}$ = .05, $M\_{GPT}$ = .06, \textit{z =} -3.45 \textit{p <} .001), and prepositions in first and second iteration ($M\_{human}$ = .08 $M\_{GPT}$ = .09, \textit{z =} -3.17, \textit{p =} .002, second iteration: $M\_{human}$ = .08, $M\_{GPT}$ = .10, \textit{z =} -2.91, \textit{p =} .004, respectively.) On the other hand, humans use more verbs (first iteration: $M\_{human}$ = .19, $M\_{GPT}$ = .18, \textit{z =} 2.42, \textit{p =} .02, second iteration: $M\_{human}$ = .19, $M\_{GPT}$ = .17, \textit{z =} 2.90, \textit{p =} .004, third iteration: $M\_{human}$ = .19, $M\_{GPT}$ = .17, \textit{z =} 3.33, \textit{p <} .001), adjectives (no difference in the first iteration, but in second: $M\_{human}$ = .05, $M\_{GPT}$ = .07, \textit{z =} -2.97 , \textit{p =} .003, and third: $M\_{human}$ = .05, $M\_{GPT}$ = .06, \textit{z =} -3.45, \textit{p <} .001), and pronouns (first iteration: $M\_{human}$ = .11 $M\_{GPT}$ = .08, \textit{z =} 5.17, \textit{p <} .001, second iteration: $M\_{human}$ = .11, $M\_{GPT}$ = .08, \textit{z =} 5.53, \textit{p <} .001, third iteration: $M\_{human}$ = .11, $M\_{GPT}$ = .08, \textit{z =} 5.65, \textit{p <} .001).

We observed that the usage of adverbs and verbs were high in all retellings of human compared to ChatGPT. While on average 6\% of a sentence of is adverbs and 19\& is verbs in humans' stories when we compare the number of adverbs to total words, it is only 3\% for adverbs, and 17\& for verbs in ChatGPT’s sentences. Controlling for the role of verbs to see whether adverb use is only depending on the verb use \textit{the ratio of adverbs to verbs}, showed that adverb use is significantly higher in humans in all iterations compared to ChatGPT (first iteration: $M\_{human}$ = .32, $M\_{GPT}$ = .21, \textit{z =} 4.54, \textit{p <} .001, second iteration: $M\_{human}$ = .32, $M\_{GPT}$ = .19, \textit{z =} 5.19, \textit{p <} .001, third iteration: $M\_{human}$ = .31, $M\_{GPT}$ = .17, \textit{z =} 5.17, \textit{p <} .001). Thus, humans use approximately 1 adverb for 3 verbs through the iterations whereas ChatGPT uses 1 adverb per 5 verbs and this ratio decreases with iterations.

When we examined the most commonly used adverbs, we observed both agents used negations (\textit{"not"} and \textit{"n't"}) in their stories, however, there is a huge difference between them regarding the amount they used. Humans used negations more than ChatGPT in all retelllings (first retelling: $M\_{human}$ = .72, $M\_{GPT}$ = .27, \textit{U =} 4692, \textit{p <} .001, second retelling: $M\_{human}$ = .56, $M\_{GPT}$ = .21, \textit{U =} 5168, \textit{p <} .001, third retelling: $M\_{human}$ = .45, $M\_{GPT}$ = .16, \textit{U=} 5147, \textit{p <} .001). Overall, humans use 201 negations and ChatGPT uses 73 negations, excluding the original story.

We specifically focused on how the number of negations changes through iterations. From original story to first iteration, the change in negation was significant ($M\_{human}$ = -.39, $SD\_{human}$ = 1.67, $M\_{GPT}$ = -.82, $SD\_{GPT}$ = .10, \textit{t(152)=} -4.90, \textit{p <} .001). However, changes in negations for other iterations were not significant (to-second: ($M\_{human}$ = -.43, $SD\_{human}$ = .34, $M\_{GPT}$ = -.37, $SD\_{GPT}$ = .29, \textit{t(152)=} .42 \textit{p =} .68; to-third: ($M\_{human}$ = -.37, $SD\_{human}$ = .84, $M\_{GPT}$ = -.32, $SD\_{GPT}$ = .31, \textit{t(152)=} .42 \textit{p =} .72).

When we separately examined the change in happy and sad stories, we observed that humans use more negations, and preserve the negations better in sad stories ($M\_{human}$ = -.34, $SD\_{human}$ = .60, $M\_{GPT}$ = -.83, $SD\_{GPT}$ = .08, \textit{t(152)=} -3.88, \textit{p <} .001), and in happy stories ($M\_{human}$ = -.48, $SD\_{human}$ = .24, $M\_{GPT}$ = -.80, $SD\_{GPT}$ = .13, \textit{t(152)=} -3.01, \textit{p <} .001). Contrary to our expectations, there is no difference between happy and sad stories.