



1. Intro and background

This paper shows how gender is represented in an iconic game - Genshin Impact, and how this question could be addressed by computational linguistic methods, specially, the corpus-assisted discourse study.

Why interest in this topic?

Being one of the prominent forms of entertainment, video games have become an important part of our daily life. And their influence on people's behaviour or mental status have been widely discussed and also raised concerns (e.g. the blur of real and virtual worlds/persons, the tendency of violent act among young and unemployed men players, the misrepresentation of gender). If media such as games indeed has a corruptive influence on people (the Cultivation Theory proposed by Gerbner et al., 2006), then the long-term effects of video games on players should be part of that influence and thus worth exploring. But we could not do this without rich datasets.

What make this paper worth presenting are:

- Game corpora are relatively rare. And this paper has created a corpus containing over 800,000 words of in-game text.
- There is a research gap in the gender representation in video games, that is most studies analyse this issue in content level, lexical level is rare. Most studies use one-shot games, leaving long term exposure untouched, and Genshin Impact on the contrary offers long-lasting, daily gameplay and introduces new regions, quests, characters with each patch since its release in 2020. This study tried to fill this gap. This attempt also highlights the usefulness of computational linguistics in gender issues.

2. Method

Corpus-assisted discourse study (CADS), which includes the quantitative **Keyword Analysis** (the most frequent 100 words), and the quantitative **Collocation Analysis** is combined with the qualitative **Concordance Analysis** to explore the most frequent verbs, adjectives, and nouns associated with the he and she pronouns. More specifically, what are these words and what gender representations have they implied, such as social roles, powers, status, and such.

This paper applies no machine learning method or algorithm.

3. Data

Two corpora:

- New Model Corpus (NMC) as a reference corpus. It is a general English corpus.
- Specialised Genshin-Impact corpus with 800,231 words from in-game quest dialogues. All of the released quests from the game's first patch (1.0) until the end of 2.6 were collected from a fan website. 1011 daily commissions and 84 main story, 98 character, 41 event and 1065 world quests (side quests) that contain dialogues were added to the database. Junk words like dialogue markers were found and deleted by using Excel applications.

Two tools are used too:

- **Sketch Engine**: with simple clicks the corpus analysis could be done, such as keyword analysis, word lists, collocations, and concordance lines.

- AntConc, similar to Sketch Engine.

4. Results

Top 50 frequent words:

- Even though gendered words were not frequent, the pronoun *he* appeared on the keywords lists while the pronoun *she* did not make into the list.

Verb collocations of *He*

- Males were represented most with bodily/physical verbs while mental/cognitive collocates were few.
- Concordance analysis of verbs like *hides*, *recover*, *die* and *disappeared* found that male characters were not actors of physically violent acts but receivers of them, unlike in other games where males were the source of actions like murder and rape.
- When verbs that denote negative representations were examined, it was found that drinking was associated with males who felt guilty over something or had a reckless personality

Adjective collocations of *He*

- The representation of male characters affected by external forces which was presented in the previous section can also be seen in the results of the collocation analysis. Male characters were either victims of some events or they were in bad psychological states.

Verb collocations of *She*

- Both bodily/physical and mental/cognitive verbs used almost similar percentages while mentioning female characters. There were no traces of violent acts that denote masculine powers. This is a relatively new finding compared to previous studies where women were nearly always represented with mental/cognitive verbs, suggesting that unlike other games where women are associated with mental actions, women have a much more solid presence in Genshin.
- When verbs that denote negative representations were examined, verbs described women as *playing* and *acting*, implying that some female characters are not reliable and sneaky.
- Another verb is *need*, frequent occurrence of this verb in the corpus has a negative connotation that women are subjects rather than actors who need help and exist with the presence of other characters.

Adjective collocations of *She*

- Females were frequently represented with their psychological conditions. They would get angry quickly and do things that were not appropriate, or they would be happy out of simple things.
- Women were described with either unstable psychological states or happy moods more than men.

Noun collocations of *She* and *He*

- When the social roles that appear with both genders were examined, usages such as son, dad and daughter, and mother, which indicate familial ties in both genders, had emerged. However, this usage in men has a meaning that includes patriarchy and

status. In women, this use expresses a domestic bond as stated in previous gender studies which reveals the gender role distinctions.

- Various professions and occupations came to the fore when looking at other social roles that define women. Although the unreal women with supernatural powers and characters mentioned in previous studies are also present in Genshin, there are also realistic characters such as chiefs, administrators, merchants, and doctors that are famous for their professions and achievements.

5. Conclusions

The paper's conclusion:

In general, although women in Genshin Impact are linguistically overrepresented, emotionally prominent, and defined by some negative traits and traditional views, there are also positive changes in their representation compared to previous studies where they were victims of male violence. These positive changes can also be seen in depictions of social roles and concrete actions. While other games depict women with unrealistic body features, women in Genshin came to the fore with their supernatural powers and achievements in their professions. Additionally, although men being victims of violence are related to the struggle for survival that can be seen in gender roles studies, when compared to previous games, the dominant, abusive and violent male figure was not found in Genshin Impact.

Conclusion on the presentation:

- This paper's research methods are not advanced or sophisticated, namely this is not a very fancy article, I chose this mainly because of the efforts the researcher made to compensate the lack of in-game text corpus, and also because this article not only "provide valuable data for future work on the long-term effects of video games on player", as the author said, on the other hand, it might also inspire game developers to reflect and evaluate their dialogue system, their textual contents and such, to see if indeed they have avoided implicit gender misrepresentation for both men and women characters.
- As large language models continue to advance, it is only a matter of time before the game industry involves them in game development or gaming communities, which means there will be larger and larger volumes of in-game text, making this kind of evaluation more urgent. This paper provides some criteria for such evaluation. We might foresee the emergence of such evaluation mechanisms in the future, which personally I think is a very meaningful application of CL in DH.

Further Reading:

Heritage, F. (2020). Applying corpus linguistics to videogame data: Exploring the representation of gender in videogames at a lexical level. *Game Studies*, 20(3). http://gamestudies.org/2003/articles/heritage_frazer.

Heritage, F. (2021a). *Language, Gender and Videogames: Using Corpora to Analyse the Representation of Gender in Fantasy Videogames* (1st ed.). Palgrave Macmillan. <https://doi.org/10.1007/978-3-030-74398-7>.

Heritage, F. (2021b). *Maidens and Monsters: A corpus assisted critical discourse analysis of the representation of gender in The Witcher videogame series* (Doctoral Dissertation, Lancaster University).). <https://doi.org/10.17635/lancaster/thesis/1333>.