Machine Translation and Gender biases in video game localisation: a corpus-based analysis

María Isabel Rivas Ginel
Universités de Bourgogne Franche-Comté et Valladolid
Sarah Theroine
Université de Bourgogne Franche-Comté
The video game industry

- 175.8 billion dollars worldwide in 2021
- "COVID-19-related lockdown orders led to more time spent at home than ever before." (Newzoo, 2021).
- "The year 2020 was an unprecedented time for the games market, shattering revenues, user, and growth records." (Newzoo, 2021).
- "up +5.3% year on year from 2020, showcasing that 2020’s gaming boom has led to a lasting increase in players, with room for further growth." (Newzoo, 2021)
Internationalisation (I18N), the process to create a product that is “translation friendly”.

**Languages**
- FIGS
- E-FIGS
- BRIC
- CJK

**Level of Localisation**
- Box and docs
- Partial localisation
- Full localisation
- Enhanced localisation

**Release Model**
- Simultaneous shipment
- Versus
- Post-gold
The video game industry

- Strings that need to be localized are sent in Excel files that sometimes lack internal coherence.
- Unnatural split between text and image that leaves many translators working blind.
- Bernal Merino (2013:119) describes this as a “a double-blind process (no audiovisual context, no text linearity) that is bound to produce more linguistic and culturalisation mistakes than they otherwise would”.
- “Enter %place to obtain %object.”

The video game industry

Only 30% of the localisers have access to the game.

Only 11.6% participants stated using MT regularly

More than 70% of video game developers chose Unity as their main video game development engine.

LSCs: In which technologies do you expect to invest in 2020?

Translation departments: In which technologies do you expect to invest in 2020?
The video game industry

Auto-Translation

Any term or language can be automatically translated by clicking a single button or by linking the language source with the Google Spreadsheet’s translation features.

Using automatic translations you can support any language from the very start and correct it later on.
Google, DeepL & SmartCat

- GNMT uses a hybrid model architecture replacing the previous RNN-based GNMT model with a transformer encoder and an RNN decoder for predictive typing.
- Google uses a Neural Machine Translation model to web crawl instead of the previous phrase-based machine translation model (2016).
- Google’s training phase uses parallel texts.

- DeepL uses Deep Learning and Neural Network and draws linguistic data from Linguee.
- DeepL uses Deep Learning to constantly train paired languages.
- DeepL uses an API to provide restricted access to a glossary of terms which increases the accuracy.

- SmartCat uses Yandex Free by default, it uses statistical machine translation.
- Yandex's language training is built around three models: a translation model, a language model and a decoder.
- The decoder's function can be assimilated to a phrased-based model or a Seq2Seq model.
Main reasons

- The results of the surveys
- Inclusivity in the video game industry
- Gender bias in Machine Translation
METHODOLOGY
The games

Los Rostros del Bosque
The Faces of the Forest

The Devil's Womp

Deltarune
Methodology

- **Blue:** Female character translated as male
- **Green:** Feminine object translated as masculine
- **Violet:** Non-binary character
- **Red:** Male character translated as female
- **Orange:** Masculine object translated as feminine

<table>
<thead>
<tr>
<th>Source (EN)</th>
<th>GOOGLE</th>
<th>DEEPL</th>
<th>SMARTCAT</th>
</tr>
</thead>
<tbody>
<tr>
<td>ARE WE CONNECTED?</td>
<td>SOMMES NOUS CONNECTÉS?</td>
<td>SOMMES-NOUS CONNECTÉS?</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>EXCELLENT.</th>
<th>EXCELLENT.</th>
<th>EXCELLENT.</th>
</tr>
</thead>
<tbody>
<tr>
<td>NOW.</td>
<td>A PRESENT.</td>
<td>MAINTENANT.</td>
</tr>
<tr>
<td>WE MAY BEGIN.</td>
<td>NOUS POUVONS COMMENCER.</td>
<td>NOUS POUVONS COMMENCER.</td>
</tr>
<tr>
<td>FIRST.</td>
<td>PREMIER.</td>
<td>D’ABORD.</td>
</tr>
<tr>
<td>YOU MUST CREATE A VESSEL.</td>
<td>VOUS DEVEZ CRÉER UN NAVIRE.</td>
<td>VOUS DEVEZ CRÉER UN VAISSEAU.</td>
</tr>
<tr>
<td>SELECT THE HEAD THAT YOU PREFER.</td>
<td>SELECTIONNEZ LA TÊTE QUE VOUS PREFERENCEZ.</td>
<td>CHOOSEZ LA TÊTE QUE VOUS PREFERENCEZ.</td>
</tr>
<tr>
<td>SELECT THE TORSO THAT YOU PREFER.</td>
<td>SELECTIONNEZ LE TORSE QUE VOUS PREFERENCEZ.</td>
<td>SELECTIONNEZ LE TORSE QUE VOUS PREFERENCEZ.</td>
</tr>
<tr>
<td>SELECT THE LEGS THAT YOU PREFER.</td>
<td>SELECTIONNEZ LES JAMBES QUE VOUS PREFERENCEZ.</td>
<td>SELECTIONNEZ LES JAMBES QUE VOUS PREFERENCEZ.</td>
</tr>
<tr>
<td>THIS IS YOUR BODY.</td>
<td>CECI EST VOTRE CORPS.</td>
<td>C’EST VOTRE CORPS.</td>
</tr>
<tr>
<td>DO YOU ACCEPT IT?</td>
<td>L’ACCEPTEZ-VOUS?</td>
<td>AVEZ-VOUS L’ACCEPTER?</td>
</tr>
<tr>
<td>EXCELLENT.</td>
<td>EXCELLENT.</td>
<td>EXCELLENT.</td>
</tr>
<tr>
<td>NOW.</td>
<td>A PRESENT.</td>
<td>MAINTENANT.</td>
</tr>
<tr>
<td>YOU HAVE CREATED A WONDERFUL FORM.</td>
<td>VOUS AVEZ CRÉE UNE FORMULE MERVEILLEUSE.</td>
<td></td>
</tr>
<tr>
<td>WHAT IS ITS FAVORITE FOOD?</td>
<td>QUELLE EST SA NOURRITURE?</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Nom</th>
<th>Description</th>
<th>Genre</th>
<th>Image</th>
</tr>
</thead>
<tbody>
<tr>
<td>Abbig</td>
<td>She is the teacher of Kits and other students</td>
<td>Female</td>
<td></td>
</tr>
<tr>
<td>A Man</td>
<td>The most hidden NPC the party can encounter in Forest.</td>
<td>Male</td>
<td>none</td>
</tr>
<tr>
<td>Agnes Creamer</td>
<td>His hidden in the flower shop “Flower King” in hometown, and is Kits and Abby’s father</td>
<td>Male</td>
<td>none</td>
</tr>
<tr>
<td>Aziz</td>
<td>He is mentioned by numerous characters in hometown, but is away at university and does not appear in the game.</td>
<td>Male</td>
<td>none</td>
</tr>
<tr>
<td>Azzy</td>
<td>One mention</td>
<td>Male</td>
<td>none</td>
</tr>
<tr>
<td>Barry</td>
<td>He sits in front row of Abbig’s class, with Termini to his right, Mouille Holiday’s side, and an empty seat behind him.</td>
<td>Male</td>
<td></td>
</tr>
<tr>
<td>Bior</td>
<td>An enemy that appears in the Forest.</td>
<td>Male</td>
<td>none</td>
</tr>
<tr>
<td>Blue Blocks</td>
<td>The Blue Blocks are Darkness the party can encounter in Forest.</td>
<td>Male</td>
<td>none</td>
</tr>
<tr>
<td>Blue Tracks</td>
<td>An employee of ICE’s PIZZA along with Buffins. The</td>
<td>Male</td>
<td>none</td>
</tr>
</tbody>
</table>
Methodology

- Blue: Female character translated as male
- Green: Feminine object translated as masculine
- Violet: Non-binary character
- Red: Male character translated as female
- Orange: Masculine object translated as feminine

Annotated and reviewed three times with comments for reference and/or discussion.

Once the non-problematic segments were removed, final check by the native speaker.

Creation of a single document per colour for better visualisation.
Methodology

- **Blue**: Female character translated as male
- **Green**: Feminine object translated as masculine
- **Violet**: Non-binary character
- **Red**: Male character translated as female
- **Orange**: Masculine object translated as feminine
RESULTS: colour code
## Distribution of types of errors per tool

<table>
<thead>
<tr>
<th>TOOL</th>
<th>ALL ERRORS</th>
<th>BLUE</th>
<th>GREEN</th>
<th>VIOLET</th>
<th>RED</th>
<th>ORANGE</th>
</tr>
</thead>
<tbody>
<tr>
<td>GOOGLE</td>
<td>333</td>
<td>156</td>
<td>95</td>
<td>62</td>
<td>9</td>
<td>11</td>
</tr>
<tr>
<td></td>
<td>3,03%</td>
<td>46.85%</td>
<td>28.53%</td>
<td>18.62%</td>
<td>2.7%</td>
<td>3.3%</td>
</tr>
<tr>
<td>DEEPL</td>
<td>261</td>
<td>110</td>
<td>55</td>
<td>67</td>
<td>15</td>
<td>14</td>
</tr>
<tr>
<td></td>
<td>2.38%</td>
<td>42.15%</td>
<td>21.07%</td>
<td>25.67%</td>
<td>5.75%</td>
<td>5.36%</td>
</tr>
<tr>
<td>SMARTCAT</td>
<td>340</td>
<td>157</td>
<td>94</td>
<td>63</td>
<td>14</td>
<td>12</td>
</tr>
<tr>
<td></td>
<td>3.1%</td>
<td>46.18%</td>
<td>27.65%</td>
<td>18.53%</td>
<td>4.12%</td>
<td>3.53%</td>
</tr>
<tr>
<td>TOTAL OF SEGMENTS</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>10980</td>
</tr>
</tbody>
</table>

- **Blue**: Female character translated as male
- **Green**: Feminine object translated as masculine
- **Violet**: Non-binary character
- **Red**: Male character translated as female
- **Orange**: Masculine object translated as feminine
Distribution of types of errors per tool

- Blue: Female character translated as male
- Green: Feminine object translated as masculine
- Violet: Non-binary character
- Red: Male character translated as female
- Orange: Masculine object translated as feminine

**GOOGLE: 333 ERRORS IN TOTAL**
- Blue: 46.85%
- Green: 28.53%
- Violet: 18.62%
- Red: 3.30%
- Orange: 2.70%

**DEEPL: 261 ERRORS IN TOTAL**
- Blue: 42.15%
- Green: 25.67%
- Violet: 21.07%
- Red: 5.75%
- Orange: 5.36%

**SMARTCAT: 340 ERRORS IN TOTAL**
- Blue: 46.18%
- Green: 27.85%
- Violet: 18.53%
- Red: 4.12%
- Orange: 3.53%
RESULTS: corpus-based analysis
Blue: Female character translated as male
Green: Feminine object translated as masculine
Violet: Non-binary character
Red: Male character translated as female
Orange: Masculine object translated as feminine
Blue: Female character translated as male
Green: Feminine object translated as masculine
Violet: Non-binary character
Red: Male character translated as female
Orange: Masculine object translated as feminine

« Ana »

(Ana) - I can't today, sorry.
(Ana) - The part where I'm supposed to listen to you.
(Ana) - I've come with a proposal.
(Ana) - I've come to take back my town!

(Ana) - Je ne peux pas aujourd'hui, désolé.
(Ana) - La partie où je suis censée t'écouter.
(Ana) - Je suis venue avec une proposition.
(Ana) - Je suis venu pour reprendre ma ville!
« You »

- Blue: Female character translated as male
- Green: Feminine object translated as masculine
- Violet: Non-binary character
- Red: Male character translated as female
- Orange: Masculine object translated as feminine

Do you consider me a friend?

- Vous me considérez comme un ami?
- Me considéres-tu comme une amie?
« Kid & Kiddo »

- **Blue**: Female character translated as male
- **Green**: Feminine object translated as masculine
- **Violet**: Non-binary character
- **Red**: Male character translated as female
- **Orange**: Masculine object translated as feminine

Transcriptions:

(Pilmayken) - Salut, gamin.  
(Pilmayken) - Merci, gamin.  
(Pilmayken) - Tu m’as apporté ce que je t’ai dit, gamin?  
(Pilmayken) - Merci, petit.  
(Pilmayken) - Tu m’as apporté ce que je t’ai dit, petit?  
(Pilmayken) - Merci pour le masque, petit.  
(Pilmayken) - Petit, tu dois t’échapper.  
(Pilmayken) - Le prendre, le gosse.  
(Pilmayken) - Utilise le masque, petit.  
(Pilmayken) - Utilise le masque, gamin.  

(Awkan) - Salut, gamin!  
(Awkan) - Salut, mon petit!  
(Awkan) - C’est bon de te voir, gamin.  
(Awkan) - C’est bon de te voir, mon petit.  
(Awkan) - Tu t’es surpassé, petit.  
(Awkan) - Tu t’es surpassé, gamin.  

(Awkan) - Salut, biquet.  
(Awkan) - Il est bon de vous voir, kiddo.  

(Awkan) - Vous vous êtes surpassé, gamin.  

(Awkan) - Vous surpassez vous-même, kiddo.
Blue: Female character translated as male
Green: Feminine object translated as masculine
Violet: Non-binary character
Red: Male character translated as female
Orange: Masculine object translated as feminine
« They »

- Blue: Female character translated as male
- Green: Feminine object translated as masculine
- Violet: Non-binary character
- Red: Male character translated as female
- Orange: Masculine object translated as feminine
CONCLUSIONS AND FUTURE WORK
Blue: Female character translated as male
Green: Feminine object translated as masculine
Violet: Non-binary character
Red: Male character translated as female
Orange: Masculine object translated as feminine

<< You're really meanful, aren't you, Kris? >>
<< Kris, you are aware it's just a dummy, right? >>
<< Kris, I know you were a here the moment I saw you. >>
<< Furthermore only Kris can seal the DARK FOUNTAIN. >>
<< Kris, if you're jealous you can be a bad guy too. >>
<< Kris isn't jealous of your evildoing lifestyle. ! >>
<< Umm, you're a real go-getter, Kris! >>
<< Kris, are you sure you know what you're doing...? >>

<< Tu es vraiment méridioïde, n'est-ce pas, Kris? >>
<< Kris, tu sais que c'est juste un mannequin, non...? >>
<< Kris, tu sais que c'est juste un mannequin, n'est-ce pas...? >>
<< Kris, tu sais que c'est juste un mannequin, n'est-ce pas...? >>
<< Kris, tu est un émeutier formidable, Kris! >>
<< Kris, je sais que tu es un héros au moment où je t'ai vu. >>
<< De plus, seul Kris peut sceller la FONTAINE NOIRE... >>
<< Kris, si tu es jaloux, tu peux aussi être un méchant... >>
<< Kris n'est pas jaloux de ton style de vie malfaisant... ! >>
<< Umm, tu es un vrai... forceur, Kris! >>
<< Kris, tu es sûr de savoir ce que tu fais...? >>

<< Tu es vraiment méridioïde, n'est-ce pas, Kris? >>
<< Kris, tu sais que c'est juste un mannequin, non...? >>
<< Kris, tu sais que c'est juste un mannequin, n'est-ce pas...? >>
<< Kris, tu est un émeutier formidable, Kris! >>
<< Kris, je sais que tu es un héros au moment où je t'ai vu. >>
<< De plus, seul Kris peut sceller la FONTAINE SOMBRE... >>
<< Kris, si tu es jaloux, tu peux aussi être un méchant... >>
<< Kris n'est pas jaloux de ton style de vie malfaisant... ! >>
<< Umm, tu es un vrai... forceur, Kris! >>
<< Kris, tu es sûr de savoir ce que tu fais...? >>

<< Kris, je sais que tu es un héros dès que je t'ai vu. >>
<< De plus, seul Kris peut sceller la FONTAINE SOMBRE... >>
<< Kris, si tu es jaloux, tu peux aussi être un méchant... >>
<< Kris n'est pas jaloux de ton style de vie malfaisant... ! >>
<< Umm, tu es un vrai... forceur, Kris! >>
<< Kris, tu es sûr de savoir ce que tu fais...? >>

<< Eh, vous êtes un vrai... forceur, Kris! >>
<< Kris, tu es sûr de savoir ce que tu fais...? >>
SmartCat’s hall of fame

<s> (Veronica) - Thanks for convincing me to come. </s>
<s> (Veronica) - Merci de me convaincre de venir. </s>
<s> family, but that this one had rejected me, tossed me away... </s>
<s> famille, mais que celui-ci m'avait rejeté, m'a jeté loin... </s>
<s> -Do you consider me a friend? </s>
<s> -Ne vous considérez-vous comme un ami? </s>
<s> you watered me. </s>
<s> vous arrosées moi. </s>
<s> Damn, I was SURE he went this way... </s>
<s> Merde, j'étais SÛR qu'il est allé de cette façon... </s>
<s> We have to find her, Kris... </s>
<s> Nous devons trouver son, Kris... </s>
<s> If it picked you up and bit your face off. </s>
<s> Si ça vous a pris et mordu votre visage. </s>
<s> Oh, you mixed me up all nice~ </s>
<s> Oh, tu m'as mélangé tout gentil~ </s>
<s> Do I look pretty, Susie? </s>
<s> Puis-je regarder assez, Susie? </s>
<s> (Kuruf) - I am sorry this is all the involvement I am allowed. </s>
<s> (Kuruf) - Je suis désolé que ce soit toute l'implication que je suis autorisé. </s>
REFERENCES


