

# Analyse du corpus français

Nous allons procéder à l'analyse du corpus français à l'aide de 6 documents (voir ci-dessous) obtenus après application du script bash, et du logiciel iTrameur (outils d'analyse textométrique créé par notre enseignant).

Après avoir appliqué le script bash aux URLS français (64 urls), nous obtenons 6 documents :

- 1 fichier contenant l'ensemble des données textuelles des pages web : *dumpglobal*
- 1 fichier contenant les contextes (2 lignes au-dessus et en-dessous) des occurrences du mot « geek » : *contexteglobal*
- 2 index : un du *dumpglobal*, un du *contexteglobal*
- 2 listes de bigrammes (couples de mots qui se suivent) : un du *dumpglobal*, un du contexte global

Pour l'analyse du corpus français nous analyserons dans un premier temps les cooccurrents de *geek*, puis dans un second le lexème *geek* et ses dérivés. Enfin nous étudierons les noms d'agents ayant pour référent une personne versée dans la culture geek. À chaque fois, nous débiterons par une explication de la méthodologie d'analyse, puis l'analyse du corpus et enfin l'interprétation des données.

## Les co-occurrents de *geek* et leur contexte :

Afin de comprendre et de définir ce qu'est un geek, nous allons nous intéresser aux unités textuelles qui accompagnent l'emploi du mot « geek ». Les co-occurrents qui ressortiront du document *dumpglobal* représenteront une vision globale des thèmes abordés dans un article se référant aux « geeks », tandis que les cooccurrents présents dans *contexteglobal* nous en apprendrons plus sur le « geek » lui-même. L'analyse des deux fichiers semble donc complémentaire pour aborder ce que représente un « geek » en français.

## Méthodologie d'analyse

Les résultats de l'analyse du corpus seront représentés sous forme de tableaux, un pour le *dumpglobal* et un autre pour le *contexteglobal*. Les lignes ont été triées par thèmes pour une vision des données plus dynamique.

Première colonne : à partir de l'index nous avons relevé les mots lexicaux présents parmi les 200 unités textuelles les plus employées dans chaque document.

Deuxième colonne : le nombre d'occurrences par mot

Troisième colonne : à partir des bigrammes, nous avons fait une recherche dans le document pour trouver les co-occurrents directs les plus utilisés pour chaque mot de la première colonne.

Quatrième colonne : nous avons utilisé certains outils d'iTrameur : le concordancier et les cooccurrents du pôle visé (les mots de la première colonne) avec un contexte gauche et droit de 10 unités textuelles ; afin de trouver les cooccurrents dans un contexte phrastique et d'explicitier l'emploi de certains bigrammes.

Cinquième colonne : une analyse rapide des occurrences.

Après ce premier travail, certains mots de la première colonne n'étaient pas du tout pertinents pour notre analyse : structures des pages web (onglets de redirection vers un autre thème de contenu journalistique), nom du journal. Exemple : *politique, économie, social, série, bouton, article, réponse, commentaire, la gazette du geek*.

Certains ont été gardés car apportant une nouvelle nuance à notre analyse, tandis que les autres ont été supprimés pour ne pas surcharger le tableau de données inutiles pour la présente analyse.

## Analyse du corpus dumpglobal

[insérer tableau fr\_dump]

Ce qui ressort : il y a une culture geek et beaucoup d'évènements à destination des geeks partout en France mais majoritairement à Paris (capitale de la France).

Les principaux intérêts des geeks sont les jeux vidéo et les films/ séries. On y retrouve des thèmes communs tels que la science-fiction, l'action et le fantastique. Une série de films semble plutôt iconique : Star Wars.

## Analyse du corpus contexteglobal

[insérer tableau fr\_contexte]

Ce qui ressort : il y a une culture geek aussi appelé culture pop / populaire. Des évènements geek ont lieu partout en France : conventions, festivals ; parmi lesquels on retrouve la Comic Con, la Paris Games Week et la Japan Expo.

Un genre semble représentatif des geeks : la science-fiction ; que l'on retrouve dans la littérature, les films et les jeux vidéo. Dans les séries, on retrouve The Big Bang Theory et Star Trek, et dans les films Star Wars et le Seigneur des Anneaux.

Les geeks semblent se passionner pour les nouvelles technologies. On peut observer le développement d'un art geek. La communauté geek existe et les femmes en font partie. La culture geek est pour le grand public.

Le terme *nerd* apparaît, bien que peu utilisé.

## Interprétation des résultats

De nombreux évènements – conventions, festivals, expositions – sont organisés autour de la culture geek partout en France et notamment dans la capitale, Paris. On relèvera notamment la Comic Con, la Japan Expo et la Paris Games Week. Ces évènements ont lieu depuis plusieurs années et sont de plus en plus populaires. Lors de ces rassemblements, un grand nombre de geek semblent adeptes du cosplay. Le cosplay consiste à jouer le rôle de ses personnages préférés en imitant leur apparence (costume, maquillage). Le geek est une personne passionnée, et durant ces évènements « geeks », on retrouve un grand nombre de leurs passions qui semblent centrées sur les jeux vidéo, mais également axées sur ce qui touche au domaine de la science-fiction et du fantastique.

Concernant les jeux auxquels s'adonnent les geeks, une grande part est consacrée aux jeux vidéo et on retrouve même quelques mentions de l'eSport – compétition de jeux vidéo en réseau local. Cela peut être mis en lien avec le goût des geeks pour la technologie et l'informatique : internet, programmation, nouvelles technologies, technologies mobiles, ordinateur, console de jeu. Les autres jeux que pratiquent les geeks sont les jeux de rôle et les jeux de société tirés des univers de science-fiction qui passionnent les geeks tels que Star Wars et Le Seigneur des Anneaux.

La science-fiction et le fantastique sont des genres narratifs principalement littéraires et cinématographiques. On retrouve ces genres aussi bien dans la littérature – mangas, BD, comics (DC) – que dans les films et séries geek. Dans le septième art geek, on retrouve des films de la série Star Wars qui s'insèrent dans un univers de fantasy et de science-fiction spatiale. Mais aussi des films des enseignes Marvel et DC qui relatent des histoires de super-héros, et des films d'animation produits par Disney et Pixar. Il est plus que probable que les personnages dont les cosplayers prennent l'apparence lors des évènements geeks soient tirés des univers cinématographiques et littéraires geek.

La « culture geek » est un terme qui revient souvent dans notre corpus. Mais jusque-là il était difficile de cerner en quoi consisterait une « culture » qui serait dite « geek ». Ici, « culture » n'est pas

utilisé dans son emploi agricole mais bien dans l'acception d'un ensemble de connaissances propres à un groupe social – ici les geeks – et qui rassemble et soude les membres de cette communauté. Au sens sociologique, la culture est considérée comme *l'ensemble des traits distinctifs, spirituels et matériels, intellectuels et affectifs, qui caractérisent une société ou un groupe social* (définition de l'UNESCO). En ce sens il s'agit bien d'une culture que l'on observe chez les geeks. Ils sont passionnés de jeux vidéo et par l'univers de la science-fiction (littérature, cinéma). Ils partagent un goût commun pour un certain nombre de films au point que certains deviennent des figures de proue du genre geek – ce qui est le cas pour les films Star Wars.

Il existe une communauté qui se revendique « geek » et qui possède et pratique cette culture geek. La communauté geek est mixte, on y retrouve aussi bien des hommes que des femmes. La culture que véhicule la communauté geek n'est pas réservée à un certain type de personne mais est disponible au grand public.

Il est à noter que le premier film Star Wars créé par Georges Lucas est sorti en 1977, et le dernier en date, en décembre 2019. L'univers geek n'est donc pas très récent mais existe bien depuis au moins plus de 40 ans. Et depuis tout ce temps de nouvelles œuvres cinématographiques et littéraires sont venues alimenter l'univers geek. Sans oublier les nouvelles technologies dont semble éprise la communauté geek qui n'ont fait que progresser et permis la création de plus en plus de support de jeux vidéo – portable, ordinateur, console – et donc de jeux. En prenant en compte ce perpétuel renouvellement de biens matériels à destination principale d'un public geek, et la création d'événements totalement axés sur la culture geek, on peut deviner que la communauté geek est un public que cherchent à atteindre un certain nombre d'entreprises. De plus, un aussi grand nombre d'événements annuels partout en France ne seraient pas mis en place pour un petit groupe. De cela on peut admettre que la communauté geek n'est pas simplement un petit groupe d'individus mais bien une véritable communauté largement représentée partout en France.

Toutefois, il y a un questionnement sur ce qu'est réellement un geek : certains cherchent à en obtenir une définition – tout comme nous. Certain s'interrogent sûrement sur ce qui fait un geek : consommer du contenu dit « geek » ferait-il de quelqu'un un geek ? Pourrait-on être geek sans le savoir ? Voilà bien des questions qui ne trouvent pas encore de réponses incontestables. Cependant, notre corpus nous a permis de mettre en relief un certain nombre de particularités propres à la communauté geek, telles que sa passion pour les jeux vidéo, les nouvelles technologies et la science-fiction, le tout appartenant à la culture geek.

## L'emploi du mot « geek »

Le mot « geek » est un emprunt à l'anglais *geek* /gɪk/ « 1. a person who is considered to be different from others, esp. a teenager who is socially awkward or spends too much time studying. 2. an expert in computers (a term of pride as self-reference, but often considered offensive when used by outsiders.) » [<https://www.wordreference.com/definition/geek>]

Avant d'aller plus loin on peut déjà observer qu'il y a eu un glissement de sens par rapport à l'emploi original. De ce qu'on a pu tirer de l'analyse des cooccurrents, un geek est un passionné de jeux vidéo, de technologie et de science-fiction. Le tout faisant partie de la culture geek, partagée par une large communauté de personnes qui vivent ensemble leurs passions lors d'événements partout en France.

Nous souhaitons savoir si le mot geek a été largement introduit (lexicalisé) dans la langue française. Pour ce faire nous allons observer son utilisation dans le corpus et ses dérivés.

## Méthodologie

Utilisation du dictionnaire d'iTrameur pour retrouver tous les mots construits à partir de *geek*, et du concordancier pour observer leur utilisation dans la phrase. Présentation des résultats

sous forme de tableau qui reprend l'unité textuelle, le nombre d'occurrences, le procédé de construction morphologique et le niveau de lexicalisation du néologisme dans la langue française.

Durant le relevé des données, nous nous sommes permis de supprimer les noms de journaux construits à partir de *geek*, exemple : *lagazettedugeek*.

### Analyse des unités textuelles dérivées de « geek »

Unité lexicale (néologisme)	Nombre d'occurrences	Procédé de construction morphologique	Lexicalisation du néologisme
geek	1213	Emprunt	+++++
geeks	350	Emprunt	+++++
geekette	18	Suffixation : féminin du geek	+++
geekerie	9	Suffixation : comportement de geek ou objet pour geek	++
geekettes	8	Suffixation : féminin du geek	+++
geekbuster	5	Jeux de mot : gosthbuster > geekbuster (chasseur de geek)	-
geekitude	5	Suffixation : état de celui ou de celle qui est geek	++
geeker	3	Conversion : agir comme un geek	++++
geekosphère	1	Préfixation (geek+sphère): domaine où les geeks sont présents	+
geekifier	1	Suffixation : rendre geek	+
geekisation	1	Suffixation : rendre geek	+
geekopedia	1	guide du geek : mot-valise (geek + wikipédia)	-

### Interprétation des résultats

Tout d'abord, il est bon de revenir sur la raison de la création de néologismes – de nouveaux lexèmes (unités lexicales) – dans les langues : le plus souvent l'apparition de néologismes dans une langue répond à des besoins : de dénominations de concepts, de cohérence lexicale... Dans notre cas la création de néologismes répond à des besoins de dénomination. Comme nous l'a montré l'analyse des cooccurrents, l'univers geek est assez récent – d'un point de vue linguistique. Et la langue française ne possédant pas de terme permettant de caractériser ce « genre », ajouté au fait que de nombreux contenus geek proviennent des États-Unis au début de l'apparition de ce « genre », il n'est pas difficile de concevoir que pour pallier un manque, la langue française ait fait un emprunt de lexème à l'anglais : « geek ».

Toutefois, pour qu'un néologisme devienne lexème d'une langue, il est nécessaire que ce dernier subisse un processus de lexicalisation – intégration au lexique d'une langue. Un bon indicateur de lexicalisation d'un néologisme est l'emploi de ce dernier par des institutions participant à l'établissement de la norme linguistique telles que les médias, les dictionnaires, l'école et l'Académie Française. Lors de notre analyse, nous nous sommes basés sur cet indicateur pour déterminer la lexicalisation des unités lexicales.

En français, l'emploi de « geek » est très largement lexicalisé : *geek* est présent dans les dictionnaires et les médias, *geek* est le lexème base de plusieurs dérivations morphologiques, et il s'accorde en nombre. La conversion du lexème nominal *geek* en adjectif ou en verbe – *geeker* – est

un procédé plutôt courant dans la langue française. L'adjectif *geek* possède les propriétés du geek et va les transmettre au nom qui le gouverne, comme dans « culture geek », « salon geek », « convention geek »... Le verbe « geeker » démontre que le geek n'est pas défini seulement par la culture qu'il partage avec sa communauté mais également ses actions.

Le néologisme « *geekette* » semble en cours de lexicalisation et nous confirme la présence de femmes parmi la communauté geek. La lexicalisation moins avancée de *geekette* peut être attribuée à plusieurs facteurs – un néologisme construit sur un néologisme demande plus de temps pour la lexicalisation, moins de femmes que d'hommes parmi les geeks, le genre grammatical de *geek* déplait et certains veulent une version féminine du nom... . C'est là un large sujet qui n'est pas le nôtre dans ce projet.

D'autres néologismes sont en début de lexicalisation : *geekerie*, *geekitude*, *geekifier*, *geekisation* et *geekosphère*. La « *geekerie* » et la « *geekitude* » suggèrent un état, un comportement geek, tandis que « *geekifier* » et « *geekisation* » appellent à une transformation progressive en quelque chose de geek. On ne naît pas geek, on le devient, et ce n'est pas quelque chose d'immédiat mais qui demande un certain investissement émotionnelle (le geek est un passionné). La *geekosphère* confirme l'existence d'une communauté geek.

Le néologisme « *geekodepia* » qui se rapport à un guide pour tout ce qui touche au geek est la preuve que la culture geek est une source d'interrogation pour les non-initiés, mais est également ouverte au grand public en proposant des outils qui permettent de rentrer plus facilement dans cet univers.

En dehors d'un cadre purement linguistique, le fait que le lexème *geek* ait été lexicalisé dans la langue française démontre une certaine reconnaissance de l'existence de cette communauté par l'ensemble des français, mais également une grande communauté de geeks.

## Les autres appellations :

Comme nous pouvons le voir dans l'analyse des corpus des autres langues, il y a différentes manières de se référer aux personnes intéressées par la culture geek. Nous allons nous efforcer, dans cette partie, de trouver les caractéristiques de chaque appellation.

## Méthodologie :

Nous nous sommes basés sur nos connaissances personnelles et les noms d'agent relevés dans l'analyse des cooccurrents de « geek » pour faire un premier relevé de différentes dénominations correspondant à des individus passionnés par la culture geek. Nous avons ensuite utilisé des outils d'iTrameur :

- Dictionnaire : pour déterminer si les appellations sont présentes dans le corpus
- Concordancier et Cooccurrences avec un contexte de 15 autour du pôle visé : pour déterminer les cooccurrents des appellations dans un contexte phrastique.

## Analyse du corpus français :

Nom	Nombres d'occurrences	iTrameur concordancier et cooccurrences (contexte 15)
Geek / geeks	1213/ 350	[voir analyse des cooccurrents]
Gamer/ gamers	31/ 29	jeux vidéo (27) ; culture geek (9) ; manga (0) ; référence au japon (0) ; technologie (0) ; informatique (0) ; différence (7)

Nerd/ nerds	41/ 18	culture geek (14) ; culture <i>nerd</i> (11) ; informatique (6) ; technologie (4) ; jeux vidéo (2) ; manga (0) ; référence au japon (0) ; différence (7)
Otaku/ otakus	19/ 4	référence au japon : japonais, pays du soleil levant, nipponnes (7) ; manga (4) ; anime (1) ; culture <i>otaku</i> (5) ; culture geek (3) ; jeux vidéo (0) ; différence (1)

Ce qui ressort : chacun porte un intérêt à la culture geek

- le *gamer* est plutôt tourné vers les jeux vidéo
- le *nerd* est quant à lui axé sur l'informatique et les technologies
- l'*otaku* est lui orienté vers des références japonaises.

### Interprétation :

Il y a d'autres manières de se référer à une personne portant un intérêt à la culture geek. Toutefois, on remarque une concordance entre une appellation particulière et une passion geek. On pourrait presque parler de spécialisation.

Le terme « *gamer* » est un nom d'agent, construit par suffixation (personne qui) sur l'emprunt du lexème « game » au lexique anglais. Le *gamer* désigne donc littéralement « une personne qui joue aux jeux vidéo ». Et c'est ce que l'on retrouve après analyse des co-occurents de *gamer* : près de la moitié des occurrences de *gamer* ont un lien avec les jeux vidéo. Pour étayer notre hypothèse de spécialisation d'une branche de la culture geek : la mention des jeux vidéo est pratiquement inexistante au contact de *nerd* et *otaku*.

Le terme « *nerd* » est un emprunt au jargon familier américain des années 1950. Ses cooccurents le caractérisent comme une personne portant un intérêt particulier pour l'informatique et la technologie. On remarque une coexistence des unités textuelles « culture geek » et « culture nerd ». Cela pourrait refléter une spécialisation du *nerd* dans une branche spécifique geek – l'informatique et les nouvelles technologies. Ainsi la culture nerd serait une culture se référant plus spécifiquement à l'informatique que ne le fait la culture geek.

Le terme « *otaku* » est un emprunt à la langue japonaise. Ses cooccurents le présentent comme une personne ayant un engouement certain pour la culture japonaise et notamment les mangas et les animes. On peut ici aussi constater une certaine spécialisation dans une des branches geek : la culture otaku qui se réfère aux mangas, aux animes provenant du Japon.

### Conclusion

On peut convenir que dans la communauté geek, tous partagent une culture commune : la culture geek. Culture qui regroupe de nombreuses passions telles que les jeux vidéo, l'informatique et la technologie, sans oublier la science-fiction qui se traduit par de nombreux contenus cinématographiques – Star Wars, les superhéros, les animes japonais, The Big Bang Theory – et littéraire – les comics, les BD, les mangas, les romans de science-fiction (Jules Verne). La communauté geek est mixte, ouverte au grand public et rassemble de plus en plus de monde. Il existe des événements annuels devenus incontournables partout en France qui réunissent de plus en plus de « geeks ».

Parmi les geeks, on peut convenir qu'il existe différentes spécialisations qui ont donné naissance à de nouvelles appellations : les gamers pour les passionnés de jeux vidéo, les nerds pour ceux portant un intérêt plus spécifique à l'informatique et à la technologie, et les otakus pour ceux

qui sont fans de la culture japonaise. Il est bon de spécifier que malgré leurs différences, tous partagent une passion commune pour la culture geek, peu importe la manière dont ils l'abordent.