

Parlez-vous #hashtag ? Quelques éclairages sur l'anglicisme *hashtag* et ses substituts français *mot-dièse* et *mot-clic*

Melissa Schuring^{1,*}, et Anne Vanderheyden^{2,*}

¹QLVL (Quantitative Lexicology and Variational Linguistics) – KU Leuven – Belgique

²GaP (Grammar and Pragmatics) – Universiteit Antwerpen – Belgique

Mot-dièse ne veut malheureusement rien dire du tout pour la plupart d'entre nous... *Hashtag* est un emprunt nécessaire, il comble un manque que *mot-dièse* ne comblera certainement pas, car il s'éloigne par trop du génie de la langue.
(Etienne Jacob, 2017, *Le Figaro*)

Résumé. Cet article porte sur *hashtag*, un anglicisme, et ses substituts français *mot-dièse* et *mot-clic* à partir de l'observation d'un corpus d'occurrences. L'analyse quantitative montre que le mot *hashtag* s'emploie le plus fréquemment en France, alors qu'au Canada, c'est la variante *mot-clic* qui est la plus fréquente. Du point de vue sémantique, les trois mots s'avèrent être des synonymes ayant deux sens : (i) « mot-clé cliquable », et (ii) « signe hashtag » renvoyant au signe (#). L'analyse morpho-syntaxique de *hashtag* met en évidence que cet anglicisme s'est complètement intégré dans la langue française. La suite de l'article focalise sur la séquence hashtag. Celle-ci est construite par un procédé morphologique, que nous avons appelé le procédé de *hashtaguisation*. L'analyse des données montre, en plus, les fonctions diverses que prend la séquence hashtag et fait apparaître clairement que ses domaines d'emploi se sont élargis : réservée initialement aux réseaux sociaux, elle trouve désormais entrée dans la langue « hors ligne », voire dans la langue parlée. Sa fonction y est de captiver l'attention de l'interlocuteur.

Abstract. Do you speak #hashtag? Some Insights into the Anglicism *hashtag* and its French Substitutes *mot-dièse* and *mot-clic*. This article focuses on the Anglicism *hashtag* that, based on a corpus study, is

* Corresponding author : melissa.schuring@kuleuven.be

* Corresponding author : anne.vanderheyden@uantwerpen.be

compared to its French substitutes, *mot-dièse* and *mot-clic*. The quantitative analysis shows that the word *hashtag* is used most frequently in France, whereas in Canada *mot-clic* is most prominent. Semantically, the three words turn out to be synonyms with two meanings: (i) "clickable keyword" and (ii) "hashtag symbol", referring to the symbol (#). The morpho-syntactic analysis of *hashtag* shows that this Anglicism is fully integrated into the French language. The remainder of the article foregrounds the hashtag sequence. It is constructed by the morphological process we call *hashtaguisation*. The data analysis demonstrates, in addition, the various functions of the hashtag sequence and clearly shows that its contexts of use have expanded: initially reserved for social networks, it now finds entry in "offline", what is more, in spoken language, where it serves to capture the attention of the interlocutor.

1 Introduction

Le *hashtag* est un mot-clé précédé du symbole # qui est employé essentiellement sur les réseaux sociaux. Il permet de repérer facilement tous les messages qui mentionnent le même sujet (cf. *dictionnaire Orthodidacte*). Introduit en 2007 par Chris Messina¹, le phénomène du *hashtag*, ainsi que le terme même, deviennent courants à partir de 2009 grâce au réseau de microblogage Twitter. Aux Etats-Unis, en 2012, *hashtag* est même devenu le mot de l'année.

Pour « faire la chasse » à l'anglicisme *hashtag* en français, les organisations de francisation en France et au Canada ont recommandé deux substituts français, qui prennent chaque fois la forme d'un mot composé constitué de termes juxtaposés et reliés au moyen d'un trait d'union. Au Canada, où la politique de la langue est plus rigoureuse qu'en France (Vincent 2005), *L'Office québécois de la langue française* (OQLF) a proposé en janvier 2011 déjà un équivalent français pour l'anglicisme *hashtag* sous la forme du terme *mot-clic*. En France, deux ans plus tard, en janvier 2013, *la Commission générale de terminologie et de néologie* (CGTN) – qui s'occupe officiellement de l'aménagement lexical en France (Kim 2015 : 83) – a recommandé le substitut *mot-dièse*.

Jusqu'à présent, *hashtag* et ses substituts français n'ont pas encore fait l'objet d'une étude sur corpus approfondie. Les études existantes portent surtout sur l'analyse des données quantitatives visant à étudier la fréquence d'utilisation, avec ou sans comparaison entre la France et le Québec (Elchacar 2017 ; Kim 2015 ; Vincent 2015) ou portent seulement sur la séquence *hashtag* (voir e.a. Caleffi 2015 ; Scott 2015 ; Vidak et Jackiewicz 2016 ; Wikström 2014 ; Zappavinga 2015). L'objectif de cet article sera dès lors de décrire en détail le sémantisme et le fonctionnement des trois termes, *hashtag*, *mot-dièse* et *mot-clic*, à partir d'un corpus d'occurrences attestées de ces termes, prises dans les bases textuelles *frTenTen12* et *frTenTen17* (datant respectivement de 2012 et de 2017) de *SketchEngine*. Nous avons aléatoirement distillé un sous-corpus de 300 occurrences pour chaque mot.

L'article s'articule en deux grandes parties. La première partie est consacrée à l'étude de *hashtag*, *mot-clic* et *mot-dièse* en tant que « mots ». Nous y expliquerons, tout d'abord, l'étymologie de ces trois termes et nous vérifierons s'ils font déjà partie de la nomenclature des dictionnaires (§ 2). Puis, nous présenterons les résultats principaux de l'analyse linguistique de notre corpus (§ 3), en abordant les aspects suivants : la fréquence relative des trois termes étudiés (3.1), leur polysémie (3.2), la valeur du symbole # (3.3), le degré d'intégration de l'anglicisme *hashtag* dans la langue d'accueil (3.4), et le rapport entre le terme – anglais *versus* français – et la langue utilisée dans la séquence *hashtag* (3.5). La deuxième partie de l'article porte non sur le « mot » *hashtag*, mais sur la « séquence *hashtag* ». Nous y décrivons le procédé morphologique dit de « *hashtaguisation* » (§ 4) et

nous analyserons quelques contextes d'emploi et quelques fonctions communicatives du hashtag, tout en montrant qu'il y a là une évolution intéressante en cours (§ 5).

2 Les dictionnaires : étymologie et histoire des mots

Hashtag est une composition des mots *hash* et *tag*. *Hash*, attesté depuis 1967, a le sens de « symbole hashtag » (#) et est sans doute une abréviation de *hash mark*, qui s'employait aux États-Unis pour référer aux rayures d'une veste militaire (*Oxford English Dictionary*). *Tag*, provenant du monde du graffiti et désignant une « étiquette », trouve son origine dans le proto-germanique (**tag-*); en moyen anglais, on utilisait *tagge*, mot qui ressemble au mot suédois *tagg* et qui signifie « épine » (*Online Etymology Dictionary*). La composition *hashtag* existe depuis 2007, sa signification actuelle étant :

A word or phrase preceded by a hash sign (#), used on social media websites and applications, especially Twitter, to identify messages on a specific topic » (*Oxford English Dictionary*).

En France, le mot *hashtag* s'emploie tel quel à partir de 2009 et peut donc être considéré comme un anglicisme, un emprunt de sens et de forme à l'anglais.

Le terme proposé au Canada en 2011 pour remplacer *hashtag*, est *mot-clic* qui est un néologisme transparent. *Clic*, « onomatopée exprimant un bruit sec, un claquement » (*Wiktionnaire*) trouve son origine en 1460, dans l'interjection *clic, clac*. Au XVI^e siècle, s'y ajoute le sens de « bruissement » et un siècle plus tard on trouve attesté le sens actuel, mais sous la forme *klik*. Au début du XX^e siècle, la graphie change en *clic*.

Pour ce qui est du terme *mot-dièse*, recommandé par la *Commission générale de terminologie et de néologie* en France à partir de 2013, il est plus sujet à controverse. *Dièse*, du latin *diesis* (« quart de ton », ensuite en bas-latin « demi-ton »), réfère à un signe musical. Or, ce signe musical n'est pas identique au signe hashtag (*pound sign*), comme le montrent les dessins ci-dessous. La différence se trouve au niveau des deux barres transversales qui sont ou bien horizontales (*pound sign*) ou bien obliques (*dièse*). La dénomination de ce « pound sign » en français est officiellement *croisillon*.



Fig. 1. *Croisillon* (pound sign) versus *dièse* (signe musical).

En français, les trois mots – l'anglicisme *hashtag* et les substituts *mot-clic* et *mot-dièse* – sont introduits dans le *Larousse en ligne* en 2013, avec la définition suivante :

Mot-clé cliquable, précédé du signe dièse (#), permettant de faire du référencement sur les sites de microblogage : Le hashtag *#chien* regroupe les posts consacrés au chien sur Twitter. (Recommandation officielle : mot-dièse.) [Au Québec, on dit mot-clic.]

Plusieurs dictionnaires mentionnent en plus que les trois termes étudiés sont réservés ou limités au domaine de l'informatique, ce qui ne correspond pas tout à fait à la réalité, comme nous le verrons plus loin.

3 Les termes *hashtag*, *mot-dièse*, *mot-clic* : une étude de corpus

3.1 Fréquence absolue et relative d'utilisation

Notre corpus montre une grande différence de fréquence entre les trois termes selon la base lexicale utilisée *frTenTen12* ou *frTenTen17* – et donc une évolution dans le temps. Ainsi dans le tableau ci-dessous, nous constatons une forte augmentation des occurrences de *hashtag* (x 4). Pour *mot-clic*, le nombre d'occurrences augmente aussi, quoiqu'il s'agisse d'une augmentation moins significative (environ x 2,5). Rappelons aussi que *mot-dièse* n'a été recommandé qu'en 2013 par le CGTN, et que, par conséquent, il n'y a pas d'occurrences dans le corpus de 2012. En 2017, par contre, il y en a plus de quatre cents.

Tableau1. Nombre d'occurrences des trois termes dans les corpus

	occurrences dans <i>frTenTen12</i>	occurrences dans <i>frTenTen17</i>
<i>hashtag</i>	2121	9715
<i>mot-clic</i>	120	295
<i>mot-dièse</i>	0	417

Les figures suivantes représentent la proportion d'utilisation de *hashtag*, *mot-dièse* et *mot-clic* dans *frTenTen12* et *frTenTen17*. Nous constatons que *hashtag* représente 95 % et 93 % de l'utilisation totale, que *mot-clic* n'occupe que 5 % et 3% et que *mot-dièse* va de 0% à 4%.

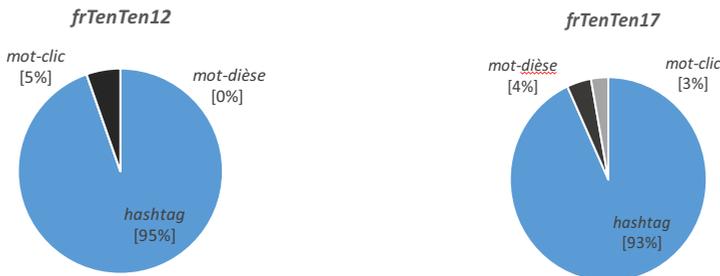


Fig. 2. Proportion d'utilisation de *hashtag*, *mot-dièse*, *mot-clic* dans *frTenTen12* et *frTenTen17*.

Comme notre corpus est constitué d'exemples qui ont été pris aléatoirement dans les bases textuelles de *SketchEngine* et que la source des occurrences dans *SketchEngine* est indiquée par des (hyper)liens se terminant souvent sur *.com* (au lieu de *.fr* ou *.ca*), il nous a été impossible de classer ces exemples en fonction de leur origine française ou canadienne. Pour ce qui est de l'attitude des locuteurs par rapport au choix lexical entre l'anglicisme et le substitut français, nous référons à Kim (2015) qui a étudié *hashtag* sur un corpus journalistique français et canadien. Ses conclusions sont qu'en France, on préfère utiliser l'anglicisme *hashtag* (82,86 % de l'utilisation totale), alors qu'au Québec, *mot-clic* est le terme le plus répandu (82,35 % de l'utilisation totale). Toujours selon Kim (2015 : 99), cette différence d'emploi tient à l'intervention gouvernementale tardive de la part de la France. Entre 2010 et 2013, les Français ont été contraints à employer l'anglicisme *hashtag*,

l'équivalent *mot-dièse* n'existant pas encore, alors que pour les Canadiens, la variante *mot-clic* était déjà disponible deux ans plus tôt. Plusieurs autres études sur *hashtag* confirment ces résultats (Elchacar 2017 ; Vincent 2015).

3.2 Polysémie de *hashtag*, *mot-dièse*, *mot-clic*

L'étude de notre corpus nous a montré que le mot *hashtag* est polysémique. Dans un premier sens *hashtag* peut être paraphrasé par « mot-clé cliquable » (*Larousse en ligne*). Il représente 83 % des occurrences de *hashtag* dans notre corpus. Le *hashtag* pouvant correspondre dans ce cas non seulement à un mot, mais aussi à un syntagme ou même à une phrase entière (voir § 4), nous appelons ce sens « *hashtag séquence* » ou *hashtag_{séq.}*. Les exemples (1) et (2) en sont une illustration. Dans un deuxième emploi, *hashtag* est synonyme de « signe » (#), nous désignons ce sens comme *hashtag_{signe.}* Ce deuxième sens (3)-(4) n'est toutefois pas fréquent avec une part de 6 % seulement dans le corpus².

- (1) Ou #vraimentPD. Depuis deux mois, ce **hashtag** se taille son petit succès sur le Twitter francophone. (*frTenTen12*)³
- (2) Créer un **hashtag** est donc un vrai métier, procédant d'une réelle réflexion en amont. (*frTenTen12*)
- (3) Et c'est bien sur cet aspect que se concentrent les améliorations de la deuxième version : filtrage du flux en fonction des personnes et des dossiers suivis, utilisation du **hashtag** (#) pour classer les commentaires par sujet [...] (*frTenTen12*)
- (4) Nous avons une structure exceptionnelle simplement en utilisant un **hashtag** (#). (*frTenTen12*)

Comme l'élément *mot* dans *mot-clic* et *mot-dièse* est difficilement compatible avec le sens *hashtag_{signe.}*, notre hypothèse initiale pour *mot-dièse* et *mot-clic* était que la polysémie de *hashtag* ne vaudrait pas pour les deux équivalents français et que ceux-ci ne présenteraient donc que le premier sens *hashtag_{séq.}* comme l'illustrent (5)-(6). Les exemples (7)-(8), toutefois, prouvent le contraire :

- (5) Il y a dix ans, un *hashtag* dépassait difficilement les 7 000 partages, alors qu'aujourd'hui, un **mot-dièse** populaire peut être mentionné des millions de fois. (*frTenTen17*)
- (6) Vous pouvez consulter toutes ces photos en utilisant le **mot-clic** "#MyVanityFairCover". (*frTenTen17*)
- (7) Au-delà de l'industrie du divertissement ou des nouvelles technologies, plus propices et plus promptes à adopter le langage des réseaux sociaux, leurs signes distinctifs, tels que le **mot-dièse** ou l'arobase de Twitter ou le "j'aime", le pouce et les "amis" de Facebook, sont de plus en plus utilisés par des marques qui s'adressent au grand public. (*frTenTen17*)
- (8) Les tweets sont publiés sur Twitter avec le **mot-clic** pertinent (le #) entre le 1er et le 10 octobre. (*frTenTen17*)

Dans (7), il est clair que *mot-dièse* a le sens de *hashtag_{signe.}* Le *mot-dièse* est même annoncé comme un « signe distinctif » et se met dans une liste avec l'arobase, le « j'aime » et le pouce de Facebook. Il en va de même dans (8), où *mot-clic* correspond aussi à la signification *hashtag_{signe.}*, vu que *mot-clic* y est suivi d'une explication rudimentaire entre guillemets (le #), # renvoyant au signe croisillon.

Nous concluons que *hashtag*, *mot-dièse* et *mot-clic* sont des synonymes, qu'ils sont, en plus, polysémiques et qu'ils peuvent avoir deux sens : *hashtag_{séq.}* et *hashtag_{signe.}*

3.3 La valeur du signe

Notre corpus a montré que les hashtags apparaissent aussi dans la communication en dehors de Twitter (Facebook, Instagram), contrairement donc à ce qui en est dit dans les dictionnaires (voir § 2). Considérons les exemples (9) à (11) :

- (9) Sur twitter, le **hashtag** [#Fast4Egypt](#) ("jeûne pour l'Egypte") est préconisé pour les jeûneurs qui veulent s'associer aux coptes à travers le monde. (*frTenTen12*)
- (10) Dans la journée de mercredi, le **mot-dièse** [#MuslimLivesMatter](#) ("la vie des musulmans compte", en français) a fait son apparition. (*frTenTen17*)
- (11) Les affiches, sur lesquelles est inscrit le **mot-clic** [#remigration](#), ne semblent pas destinées aux seuls " migrants illégaux ". (*frTenTen17*)

Ce que ces exemples ont de particulier, c'est qu'il y a répétition : il y a d'une part le mot (*hashtag*, *mot-dièse*, *mot-clic*), et d'autre part le signe *hashtag* suivi de la séquence *hashtag*. Strictement parlant, dans (9) par exemple, on devrait lire cela comme suit : *Sur twitter, le hashtag hashtag Fast4Egypt* [...]. Cette répétition de *hashtag* est présente dans plus de 90% des occurrences de notre corpus. Notre hypothèse est que, si on lisait les exemples (9) à (11) à haute voix, le signe (#) ne serait pas prononcé, ce qui revient à dire que '#' a le statut d'un signe graphique, purement matériel qui fait partie de la séquence *hashtag* et qui la délimite. Cela explique pourquoi il est nécessaire d'ajouter un mot, en l'occurrence *hashtag*, *mot-dièse* ou *mot-clic*, qui précise que la séquence qui suit est un *hashtag*_{séq.}. Dans cette interprétation, le croisillon (#) n'a donc d'autre fonction que d'indiquer la borne gauche de la séquence *hashtag*, comme le font les espaces blancs dans un texte écrit. Cela permet de le définir comme un signe de ponctuation (Zappavigna 2015).

3.4 L'insertion formelle de l'anglicisme *hashtag* dans la langue française

Bogaards (2008 : 34) distingue trois étapes dans le processus d'intégration d'un mot étranger dans la langue d'accueil. Dans une première étape, le mot étranger est accompagné d'une définition ou d'un synonyme. Bogaards qualifie ce mot, dans ce stade, de « xénisme ». Dans la deuxième étape, la définition du mot étranger n'est plus nécessaire, mais on voit apparaître souvent un commentaire « autonymique » du locuteur ; le mot étranger se trouve alors en italiques ou est placé entre guillemets. Les mots qui se trouvent à la troisième étape d'intégration, ne nécessitent plus de commentaire et s'emploient comme un mot de la langue d'accueil : ils sont alors complètement intégrés dans la langue.

Dans ce paragraphe, nous allons déterminer à quelle étape d'intégration on doit situer l'anglicisme *hashtag* en vérifiant trois critères : (i) l'intégration morphologique, (ii) la productivité morphologique et (iii) l'intégration dans la phrase.

3.4.1 Intégration morphologique

Selon *Larousse* et le *Petit Robert*, *hashtag* est masculin, bien qu'il soit difficile d'attribuer un genre aux anglicismes, comme il ressort des études portant sur ce sujet. Pour Pergnier (1989 : 39), les raisons de l'attribution d'un genre (masculin ou féminin) à un anglicisme sont complexes et « largement mystérieuses ». Surridge (1982) pose, sans vraiment donner les raisons, que le masculin est la règle.

Dans notre corpus, presque toutes les occurrences de *hashtag* sont du masculin (98,97 %) à quelques rares exceptions près ((12)-(13)) et qui sont sans doute des coquilles :

- (12) [#dailytraining](#) une nouvelle **hashtag**. (*frTenTen12*)

- (13) L'utilisateur sera montré toutes les tweets qui utilisent la même **hashtag**.
(frTenTen12)

Quant au pluriel de *hashtag*, il n'est pas mentionné par *le Petit Robert* ni par *Larousse*, mais les dictionnaires populaires tels que *Wikipédia* signalent *hashtags* comme forme du pluriel. Ceci a été confirmé par les données de notre corpus : 89,42 % des *hashtags* au pluriel s'écrivent avec s (*hashtags*), même si on constate là aussi une légère hésitation entre *hashtag* (14) et *hashtags* (15).

- (14) Les messages sur Twitter contenant **des hashtag** ne sont pas nécessairement bienvenus sur Facebook (frTenTen12)
(15) La troisième, est une recherche par **hashtags**. (frTenTen12)

Du point de vue morphologique donc, il y a encore des hésitations – fussent-elles minimales – par rapport au genre et par rapport au pluriel pour *hashtag*.

Signalons, en passant, que *mot-dièse* et *mot-clic* sont du genre masculin, conformément au genre de *mot*, et que leur pluriel - qui n'est pas attesté dans notre corpus -, serait *mots-dièse*⁴ et *mots-clics* selon *Wiktionnaire*.

3.4.2 Productivité morphologique

Ni les organisations de francisation en France et au Québec, ni les dictionnaires traditionnels ne proposent de dérivés pour *hashtag*. Notre corpus permet toutefois de poser que *hashtag* s'étend à d'autres catégories grammaticales, telles que le verbe à l'infinitif (16), le verbe conjugué (17) ou le participe passé (en emploi adjectival ou non) (18) :

- (16) **Hashtaguer** ou ne pas **hashtaguer** son mariage ? Telle est la question. On a pesé le pour et le contre, histoire de trancher ce cas de conscience 3.0. (*Marie-Claire*)
(17) On ne **hashtague** pas la photo que l'on envoie via la messagerie privée d'Instagram. (frTenTen12)
(18) Pour être vu, être suivi, être écouté, être visionné, être liké, être **hashtagué**, confiez nous la création de vos vidéos. (frTenTen12)

En plus, dans les exemples (19) à (20), *hashtag* est employé comme substantif épithète, fonctionnement qui pourrait être le premier pas vers un emploi adjectival de *hashtag* :

- (19) Le conférencier qui ouvre la journée d'interventions et d'échanges communique un code hashtag. (frTenTen12)
(20) Allez dans hashtag qui contribuera à un voyage, et même regarder tous les hashtags sur Flickr pour déterminer ces choses hashtag. (frTenTen12)

Un anglicisme qui élargit ses emplois à d'autres catégories grammaticales est un signe d'intégration dans la langue d'accueil (Vincent 2015 : 152). *Hashtag* fait preuve donc d'une productivité morphologique, productivité qui est inexistante dans le cas de *mot-dièse* et de *mot-clic*, probablement parce qu'il s'agit de mots composés et que le trait d'union complique le procédé de dérivation.

3.4.3 Intégration dans la phrase

Dans la majorité des occurrences du corpus, *hashtag* s'intègre tel quel dans la phrase, sans mention d'une définition ou d'une explication et sans italiques ou guillemets. En même temps, nous avons remarqué qu'une part importante des phrases avec *mot-dièse* et *mot-clic*

dans le corpus contiennent aussi *hashtag*. Pour *mot-dièse*, il s'agit de 24,16 % des phrases. Dans (21), par exemple, *mot-dièse* est mentionné comme traduction de *hashtag*, dans (22) *hashtag* se trouve entre parenthèses comme synonyme de *mot-dièse*. *Hashtag* apparaît donc là où le lecteur ne sait pas ce que c'est qu'un *mot-dièse*. Pour *mot-clic*, nous constatons le même phénomène (23), même si le nombre d'occurrences est alors plus bas (9,37 % des phrases) que dans le cas de *mot-dièse*. Cela s'explique sans doute par le fait que *mot-clic* est beaucoup plus répandu que *hashtag* au Québec.

- (21) Le **hashtag**, ou **mot-dièse** en français, a fêté ses 10 ans le 23 août dernier. (*frTenTen17*)
- (22) Sur les réseaux sociaux, le **mot-dièse (hashtag)** "article 102" est le plus partagé. (*frTenTen17*)
- (23) Le **hashtag** ou **mot-clic** est composé du signe # et d'un mot-clé (ou groupe de mots-clés non séparés par des espaces) qui, assemblés, permettent de "marquer" un contenu. (*frTenTen17*)

3.4.4 Conclusion

L'analyse de la productivité morphologique signale que *hashtag* est déjà bien intégré dans la langue française, ce qui est d'ailleurs confirmé par les fréquences d'utilisation de ce mot (cf. 3.1). De plus, il est presque toujours employé sans autres commentaires (synonymes, définitions) et sans marques autonymiques. Ces constatations permettent de ranger *hashtag* dans la troisième et dernière étape d'intégration. Bogaards (2008 : 34) fait remarquer aussi que les mots arrivés à la dernière étape ne sont pas nécessairement connus par tout le monde et que l'emploi initial est souvent restreint. *Hashtag* est utilisé en effet principalement par la population qualifiée de jeune, et surtout, mais pas exclusivement, dans le monde numérique.

3.5 Langue du terme versus langue de la séquence hashtag

Nous nous sommes demandé aussi si le terme employé (*hashtag*, *mot-dièse*, *mot-clic*) a une influence sur la langue utilisée dans la séquence hashtag qui suit. Pour vérifier cela, nous avons classé toutes les séquences hashtag de notre corpus par langue : français, anglais ou autre. Dans l'exemple (24) se trouve le mot anglais (*hashtag*) et une séquence hashtag en anglais (#MyDealBreaker), tandis que dans (25), le terme anglais (*hashtag*) est combiné avec une séquence hashtag française (#Sarkocasuffit). Les mêmes combinaisons existent pour *mot-clic* et *mot-dièse*.

- (24) Tweet @ HuffPostWomen avec **hashtag #MyDealBreaker** et dites-nous ce que vous ne défendons[sic] dans une relation. (*frTenTen12*)
- (25) Le **hashtag #Sarkocasuffit** était premier ce matin sur Twitter, juste devant #Franceforte et #Hervémorin (qui a annoncé qu'il se désistait au profit de Nicolas Sarkozy). (*frTenTen12*)

La figure 3 donne la fréquence relative de la combinaison 'langue du terme' versus 'langue de la séquence hashtag'. Le terme anglais (*hashtag*) est suivi d'une séquence hashtag anglaise dans 65 % des cas. Les termes français (*mot-clic* ou *mot-dièse*) apparaissent devant une séquence hashtag française dans 59 % des cas. Le test Chi carré ($\chi^2 = 53,1$, $p < 0,001$) montre que les termes français (*mot-clic*, *mot-dièse*) sont suivis plus souvent d'un hashtag français et que le terme anglais (*hashtag*) est suivi plus fréquemment d'un hashtag anglais. Le lien entre la langue du terme et la langue du hashtag est fort (V de Cramer : 0,35, effet important).

TERME ANGLAIS (HASHTAG)

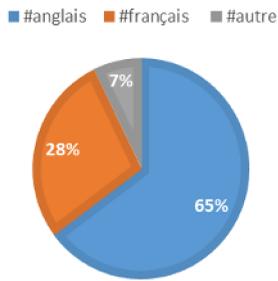


Fig. 3. Fréquence relative de la langue du terme versus la langue de la séquence *hashtag*.

4 La *hashtaguisation* : la construction du *hashtag*_{séq}

Dans les § 2 et § 3, nous avons analysé *hashtag*, *mot-dièse* et *mot-clic* en tant que mots. Dans les paragraphes suivants, nous allons étudier les séquences *hashtag*. Dans la phrase « Utilisez le hashtag #MoiAussi ! », par exemple, l'objet d'étude des paragraphes suivants sera la séquence qui suit le signe #, dans ce cas-ci donc : « MoiAussi ». Nous commençons par une analyse de la construction de ces séquences.

Twitter a développé une série de règles 'grammaticales' pour construire une séquence *hashtag*. Ainsi, les séquences (composées du signe # et d'une séquence de lettres par exemple), ne peuvent dépasser les 140 caractères ; ils peuvent contenir aussi des numéros⁵, mais pas des caractères spéciaux, à l'exception du tiret bas ; dans les séquences, il n'y a jamais d'espaces entre les mots, mais on met souvent une majuscule pour indiquer le début d'un mot (Caleffi 2015) ; il est permis, vu le nombre limité des caractères, d'abréger les mots... Ces quelques règles permettent, par exemple, aux utilisateurs de Twitter de construire des hashtags du type #MoiAussi, #toucheapasàmaconstitution ou #Sarkocasuffit. Mis à part les restrictions de Twitter, tout le monde est libre de construire ses propres hashtags et de les placer librement dans la phrase, ce qui crée une infinité de possibilités, comme cela se voit dans les exemples suivants :

(26) L'#Europe n'est plus un choix. #Sarkozy #Toulon (Twitter)

(27) Pas de #fondDePension car pas de #retraite car jamais aimé #travailler pour de l'#argent... (Twitter)

(28) Tous[sic] le monde a le droit d'être Heureux. #HétérosNousSoutenonsLeMariagePourTous (Twitter)

Dans (26), les hashtags #Sarkozy et #Toulon ne participent pas à la construction syntaxique : ils ne font pas partie de la phrase *L'#Europe n'est plus un choix*. Dans (27), par contre, les hashtags *#fondDePension*, *#retraite*, etc. sont intégrés à la phrase. Il est même possible que le hashtag constitue à lui seul une phrase entière, comme dans (28) (cf. aussi Vidak 2016 ; Zappavigna 2015).

Il faut remarquer aussi que les mots collés les uns aux autres et formant la séquence *hashtag* ne correspondent pas à des procédés de formation morphologiques connus tels que la dérivation ou la composition. Pour cela les possibilités de combinaison⁶ sont trop nombreuses. Nous sommes par conséquent d'accord avec Caleffi (2015 : 47,67) pour dire que la construction des hashtags est un procédé morphologique créant des unités linguistiques (des *hashtags*_{séq}) qui ne semblent appartenir à aucune classe lexicale. C'est ce

procédé de formation que nous désignons désormais par le néologisme inventé de *hashtagisation*.

5 Le hashtag_{séq} : des évolutions en cours

Pour terminer, signalons brièvement deux évolutions qui se profilent. La première évolution a trait au contexte d'emploi de la séquence hashtag. Apparu d'abord sur Twitter, le hashtag s'est étendu très vite, d'abord à d'autres réseaux sociaux comme Facebook ou Instagram. Mais, par après, on le voit apparaître aussi dans la communication « hors ligne » (Caleffi 2015 : 46). Les grandes entreprises, par exemple, les utilisent dans leurs campagnes publicitaires : elles les mettent sur les panneaux publicitaires, sur leurs produits et dans leurs spots télévisés. La figure 4 présente quelques-unes de ces campagnes. Parallèlement, on peut repérer des hashtag_{séq} dans les titres de journaux (figure 5) ou encore dans les slogans socio-politiques (figure 6).



Fig. 4. Les affiches des campagnes publicitaires du Thalys et de l'école Européenne de la communication.



Fig. 5. *Le Monde*, 20/12/2019.



Fig. 6. Rassemblement contre les violences faites aux femmes, à Marseille, en 2017, REUTERS/Jean-Paul Pelissie.

Enfin, on observe aujourd'hui une autre extension du contexte d'emploi, dans la mesure où le hashtag_{séq} n'est plus restreint à l'écrit, mais qu'il semble gagner de plus en plus l'oral, et plus précisément le langage des jeunes. Des phrases comme « je suis hashtag confus » ou « hashtag yolo » s'entendent de plus en plus. Il y a même des jeunes qui, en parlant, doublent leur propos d'un geste hashtag (figure 7) : ils mettent l'index et le majeur d'une main sur les mêmes doigts de l'autre, formant ainsi le signe physique du hashtag.



Fig. 7. Le geste « hashtag », (Caleffi 2015 : 66) - *socialmedium.co.uk*.

La seconde évolution concerne la fonction du hashtag_{séq}. Sa fonction première, sur Twitter, est d'organiser, de catégoriser des tweets portant sur un certain sujet⁷. Le hashtag « (indique) le ou les sujet(s) du message posté sur Twitter et (permet) de l'inscrire dans le fil thématique correspondant » (Vidak et Jackiewicz 2016 : 5). Ainsi, dans (29), le hashtag #moiaussi (et sa version anglaise) range le tweet dans les messages portant sur le sujet des comportements sexuels transgressifs :

(29) Le ministre de la Culture français a demandé une enquête concernant les allocations versées par le Centre national des lettres à l'écrivain. #moiaussi #metoo. (Twitter)

Or, comme il a été noté dans la littérature, cette fonction de catégorisation thématique, bien que première, n'est pas la seule. Les hashtags_{séq} ont obtenu une fonction supplémentaire, qui est dite « communicative » (Shapp 2014 : 4) ou « expressive » (Vidak et Jackiewicz 2016 : 8). Ces hashtags_{séq} « communicatifs » ne servent donc pas à relier le tweet à ceux qui contiennent le même mot-clé (fonction de catégorisation), mais à communiquer un contenu supplémentaire (Vidak et Jackiewicz 2016 ; Wilkström 2014 ;

Zappavigna 2015) : une émotion (30), un commentaire (31), la révélation du sujet principal du tweet (32) et (33), des félicitations (34) :

- (30) Prochaine étape, Ruan Pienaar dans une cave à vins. [#hilarant](#) (Twitter)
- (31) Ouais....,échouer un alcootest uniquement à cause de la consommation d'un médicament [#monoœil](#) (Twitter)
- (32) Pour info ce n'est pas moi qui sera là aujourd'hui mais mon fantôme ... [#teampasdormi](#) (Twitter)
- (33) Niveau circulation st pol/t cest la cata [#ratémontrain](#) 🙄🙄 (Twitter)
- (34) [#JoyeuxAnniversaire](#) @TalOfficial je t'aime ❤️ (Twitter)

Ce que ces quelques exemples illustrent, c'est que les hashtags_{séq} n'y servent plus à la catégorisation thématique, mais qu'ils ajoutent de l'information au contenu sémantique des tweets et qu'ils contribuent donc au sens du message.

Une dernière question qui subsiste concerne la raison d'être du croisillon, que ce soit sous la forme du signe #, à l'écrit, ou du geste, à l'oral : pourquoi s'en sert-on ? Comparons à ce sujet l'exemple suivant d'un tweet, avec (35) et sans croisillons (36) :

- (35) Un mirage ? Une illusion ? Un miracle ? Appelez cela comme vous le voulez, j'ai du mal à en croire mes yeux 😊 - le come back de la [#ligne13 #metro13](#) (il est 8h50 métro pas bondé du tout 🙄) [#greve2janvier](#) (Twitter)
- (36) Un mirage ? Une illusion ? Un miracle ? Appelez cela comme vous le voulez, j'ai du mal à en croire mes yeux 😊 - le come back de la ligne13, metro13 (il est 8h50 métro pas bondé du tout 🙄🙄) grève 2 janvier.

Il est clair que les séquences hashtaguées dans (35) sont ressenties comme plus importantes et que leur force illocutoire est beaucoup plus grande. Mettant en valeur les séquences hashtaguées, le but principal du croisillon est donc d'attirer l'attention de l'interlocuteur, ce qui est confirmé par plusieurs auteurs. Scott (2015 : 14) parle du hashtag comme « a highlighting device ». Vidak et Jackiewicz (2016 : 10) le définissent comme un des moyens privilégiés de l'emphase.

6 Conclusion

L'objectif de cette étude était d'analyser le sémantisme et le fonctionnement de trois mots qui n'ont guère fait l'objet d'études approfondies : *hashtag*, *mot-dièse* et *mot-clic*. L'analyse d'un corpus d'occurrences contemporaines dans la première partie de cet article, nous a permis de tirer les conclusions suivantes. (i) *Hashtag* s'emploie beaucoup plus fréquemment (93% de l'utilisation totale) que *mot-clic* et *mot-dièse*. (ii) *Hashtag* est polysémique : le premier sens étant *hashtags_{séq}* signifiant « mot-clé cliquable », le deuxième, *hashtag_{signe}*, renvoyant au signe (#). Contrairement à notre hypothèse initiale, *mot-dièse* et *mot-clic* présentent la même polysémie. (iii) Dans la majorité des occurrences, le croisillon (#) ne se comprend plus comme un signe, mais fait partie de la séquence *hashtag*. (iv) Du point de vue intégration morphologique, *hashtag* appartient à la troisième et dernière étape d'intégration dans la langue d'accueil (Bogaards 2008), vu sa fréquence élevée, sa productivité morphologique et son emploi autonome (sans définition). (v) Le terme employé (*hashtag* versus *mot-dièse* et *mot-clic*) a une influence sur la langue dans laquelle la séquence hashtag qui suit est formulée.

Dans la deuxième partie de notre étude, nous avons, d'une part, étudié le phénomène de *hashtagisation* pour décrire le procédé morphologique par lequel se créent les *hashtags_{séq}*. Selon nous, les *hashtags_{séq}* sont des unités linguistiques à part entière. Nous avons observé,

d'autre part, que le hashtag_{séq} est un phénomène en évolution de plusieurs façons. Réserve initialement à Twitter, il apparaît aujourd'hui aussi dans la communication hors ligne, tant écrite qu'orale : dans des campagnes publicitaires, dans des journaux, dans des slogans politiques, voire même dans la langue parlée des jeunes. En plus, sa fonction a changé elle-même : d'un élément permettant de catégoriser des posts sur un même sujet à l'origine, il est devenu un élément du contenu à proprement parler, qui permet d'évaluer le message – en ajoutant une émotion, un commentaire, une opinion à un message – et qui attire l'emphase. Des recherches supplémentaires, entre autres sur d'autres types de corpus, sont toutefois nécessaires pour inventorier et préciser ces différentes fonctions.

Références bibliographiques

- Bogaards, P. (2008). *On ne parle pas français: la langue française face à l'anglais*. Bruxelles : De Boeck Supérieur.
- Caleffi, P-M. (2015). The 'hashtag': a new word or a new rule ?. *SKASE journal of theoretical linguistics*, 12,2, 46-69.
- Elchacar, M. (2017). Le traitement lexicographique des anglicismes au vu de la variation géographique : l'exemple de deux outils en ligne. *Repères DoRiF*, 14, 1- 27.
- Kim, M. (2015). L'anglicisme et l'intervention linguistique gouvernementale: étude de l'implantation terminologique en francophonie. *Elis*, 3,3, 79-102.
- Pergnier, M. (1989). *Les anglicismes*. Paris : PUF.
- Noailly, M. (1990). *Le substantif épithète*, Paris, PUF.
- Page, R. (2012). The linguistics of self-branding and micro-celebrity in Twitter: the role of hashtags. *Discourse & Communication* 6, 2, 181-201.
- Scott, K. (2015). The pragmatics of hashtags: Inference and conversational style on Twitter. *Journal of Pragmatics*, 81, 8-20.
- Shapp, A. (2014). *Variation in the use of Twitter hashtags*, dissertation. New York University.
- Surrige, M. E. (1982). L'attribution du genre grammatical aux emprunts anglais en français canadien: le rôle des homologues et des monosyllabes., *Glossa*, 16,1, 28-39.
- Vidak, M., A. (2016). Le mot-dièse (hashtag): émergence d'une nouvelle forme de figement dans une diachronie très courte. *Language design*, Special Issue, 217-234.
- Vidak, M., A. Jackiewicz (2016). Les outils multimodaux de Twitter comme moyens d'expression des émotions et des prises de position. *Cahiers de praxématique*, 66.
- Vincent, N. (2005). Perception dans les journaux français de la lutte québécoise aux anglicismes. *Circula*, 2, 76-96.
- Wikström, P. (2014). #srynotfunny: Communicative functions of hashtags on Twitter. *SKY Journal of linguistics*, 27, 127-152.
- Zappavigna, M. (2015). Searchable talk: the linguistic functions of hashtags. *Social Semiotics*, 25, 3, 274-291.

¹ En 2007, Messina, un programmeur anglais, envoie un tweet à l'organisation de Twitter : « How do you feel about using # (pound) for groups. As in #barcamp [msg]? ». La réponse initiale de Twitter était négative. Mais depuis l'incendie à San Diego de 2008, le hashtag #SanDiego apparaît partout sur Twitter. Ce sont en fait les utilisateurs de Twitter qui ont introduit et entériné le procédé.

² Dans les 11 % restants des occurrences de *hashtag* de notre corpus, il était impossible de déterminer le sens par manque de contexte. Nous avons rangé ces occurrences dans la classe des « ambigus ».

³ Nous n'avons pas corrigé les éventuelles erreurs dans les exemples cités.

⁴ Dans un autre sous-corpus de *SketchEngine (Timestamped French*, que nous venons de découvrir), nous avons pu repérer toutefois deux formes pour le pluriel de *mot-dièse* : *mots-dièse* (189 occ) et *mots-dièses* (134 occ).

⁵ Cependant, une séquence contenant uniquement des numéros ou une séquence qui commence par un numéro ne sera pas acceptée comme hashtag.

⁶ Par exemple, l'ajout de chiffres, de tirets bas, de majuscules...

⁷ cf. Kehoe et Gee 2011, cité par Zappavigna (2015 : 282) : « hashtags (...) having a classificatory function of indicating the topic of a post or its “aboutness” ». Wikström (2014 : 130) parle de *topic tags*. Cf. aussi Scott (2015 : 12) : «The original and most straightforward function of the hashtag is to make the content in the tweets searchable ». Zappavigna (2015) réfère au *hashtag* en le caractérisant de “*searchable talk*”.