

Le mot-dièse (hashtag): émergence d'une nouvelle forme de figement dans une diachronie très courte

MARKO VIDAK

STIH, Université Paris-Sorbonne

Résumé: L'utilisation du mot-dièse, création récente issue des réseaux sociaux numériques, est devenue particulièrement importante au point que celui-ci apparaît en dehors de son cadre d'origine. Ses propriétés formelle, syntaxique, sémantique et fonctionnelle ainsi que son utilisation laissent à penser qu'il s'agit d'une séquence figée en termes de phraséologie. Sur la base de l'étude d'un vaste corpus de tweets, nous explorons ici les traits caractéristiques du mot-dièse pour vérifier s'il peut être rapproché des séquences figées, voire défini comme telle. Le signe #, qui signale le figement, pourrait être interprété comme l'indice du moule locutionnel. L'analogie entre le mot-dièse et les séquences figées attestées en langue permet de considérer celui-ci comme un lieu spécifique entre la néologie et le figement.

Mots-clés: mot-dièse; figement; création lexicale; moule locutionnel; réseaux sociaux numériques

Abstract: Recently appeared in the discourse of online social networks, the hashtag has become particularly frequent and is henceforth used outside its original frame. Its formal, syntactic, semantic and functional properties along with its use suggest that it can be considered as a set phrase in terms of phraseology. Based on the study of a vast corpus of tweets, this research aims to explore the properties of the hashtag in order to verify whether it can be considered as a set phrase. The # sign, which marks the fixation, could be seen as an indicator of the phrasal mould. The analogy between the hashtag and the attested set phrases suggests that it can be considered as a result of a peculiar combination of both neology and fixation.

Keywords: hashtag; fixation; lexical creation; idiomatic pattern; digital social networks

Introduction

Le mot-dièse (*hashtag* en anglais) est une séquence spécifique, apparue récemment sur Twitter à côté des autres outils multimodaux employés dans la communication sur ce réseau social numérique (RSN). Depuis sa

création, il connaît une popularité croissante au point de conquérir d'autres RSN et même de sortir de leurs limites pour s'insérer petit à petit dans d'autres univers discursifs. Avec son fonctionnement linguistique particulier et innovant, le mot-dièse (MD) suscite un important intérêt pour la recherche, notamment du point de vue phraséologique. Nous explorons ici les traits caractéristiques du MD pour vérifier s'il peut être rapproché des séquences figées (SF), voire défini comme telle.

Notre objectif est d'analyser et de décrire le mot-dièse en tant que SF. Nous proposons un examen linguistique de son statut de SF selon les critères formels, sémantiques et pragmatiques. Les spécificités formelle et fonctionnelle du MD nous amènent à postuler que le signe dièse, qui en est un trait caractéristique et définitoire, peut s'apparenter à l'indice du moule locutionnel observé en phraséologie.

Cette recherche est fondée sur l'examen d'un vaste corpus de tweets¹ récoltés depuis deux ans. Il s'agit d'un corpus initial de messages émis autour de la loi sur le mariage pour tous, récoltés pour les besoins de notre recherche sur les propriétés linguistiques du MD (Jackiewicz et Vidak, 2014) et qui compte plus de 500 000 messages soit presque 14 millions de formes. À cela s'ajoute un corpus de plusieurs milliers de tweets récoltés depuis, au fil de l'eau et sans restriction thématique.

L'étude du mot-dièse en tant que séquence figée s'inscrit dans le cadre de nos travaux de recherche menés sur le MD (Jackiewicz et Vidak, 2014; Vidak et Jackiewicz, sous presse). Sur le plan théorique, nous nous sommes inspirés de travaux en analyse du discours sur Twitter et son discours (Cervulle et Pailler, 2014; Longhi, 2012, 2013, 2015; Paveau, 2012, 2013a, 2013b, 2013c), ainsi que sur la notion de formule en discours (Krieg-Planque, 2009, 2012). Dans le domaine de la phraséologie, nous nous appuyons notamment sur les travaux de recherche sur le figement (Gross, 1996; Mejri, 2004, 2005a, 2005b) ainsi que sur l'étude des phrasèmes (Mel'čuk, 2003, 2011).

L'étude du corpus a permis de dégager un certain nombre de réflexions concernant un rapprochement possible entre le mot-dièse et les séquences figées. Nous l'avons structuré selon trois axes principaux : (i) rappeler les principales propriétés du mot-dièse, notamment linguistiques, nécessaires à l'étude phraséologique ; (ii) décrire son fonctionnement en tant que séquence figée ; (iii) vérifier son statut de SF à travers les principaux tests communément admis.

¹ Tweet : message émis sur Twitter.

1. Les propriétés linguistiques du mot-dièse

Pour saisir les propriétés linguistiques du mot-dièse il est nécessaire de rappeler les particularités de la communication sur Twitter car elles jouent un rôle important dans la création des MD et dans leur exploitation.

1.1. Twitter, un lieu d'écriture à part

Twitter est un RSN de microblogage créé en 2006 et qui n'a de cesse d'évoluer notamment dans l'interface d'écriture et de lecture des messages ainsi que dans les formes d'écriture. Il permet de fédérer une communauté importante d'utilisateurs (plus de 500 millions dans le monde dont plus de 2 millions en France). Les auteurs des messages peuvent être des individus ou des groupes de personnes, des institutions, des entreprises etc., qu'ils aient ou non une notoriété en dehors de ce réseau. Le nombre d'utilisateurs et de messages postés est en croissance permanente, ce qui fait de ce RSN une plateforme internationale segmentée en communautés linguistiques.

L'unité de discours sur Twitter est le tweet, un message dont la longueur ne peut dépasser 140 caractères, espaces compris. La forme courte appelle à la concision, ce qui engendre des conséquences linguistiques. D'une part, de nombreux utilisateurs adoptent un style particulier fondé notamment sur l'abréviation au sens large ainsi que sur l'insertion de formes d'art graphique issues de la combinaison de caractères alphanumériques (frimousses ou ASCII art par exemple). De l'autre, les différents outils multimodaux permettent de dépasser les limites fixées : mot-dièse, lien internet, formules d'adressage ou contenus multimédia incorporés dans le message. Ces outils multimodaux, qui permettent l'interconnexion avec d'autres contenus extérieurs au message, impliquent la délinéarisation de celui-ci en faisant ainsi l'impasse sur la limite formelle tout en maintenant le nombre de caractères en dessous de celle-ci. Le contexte d'énonciation et le cotexte jouent ainsi un rôle important dans l'interprétation des messages postés. Le style d'écriture sur Twitter est à la fois une forme d'identité et une solution de reconnaissance entre utilisateurs et groupes d'utilisateurs : un choix lexical, grammatical ou orthographique peut y servir de signe de reconnaissance.

Dans ce RSN qui revendique l'absence de toute modération, l'expression peut être considérée comme libre tout comme elle l'est dans un échange verbal spontané entre personnes en contact direct. En dehors

de quelques cas de censure pour des propos racistes, l'usage de la langue ne connaît pas d'autres limites que les techniques.

1.2. Le mot-dièse

Nous rappelons ici les principales propriétés du mot-dièse (voir (Jackiewicz et Vidak, 2014)), nécessaires à prendre en considération pour son examen en tant que séquence figée.

Le mot-dièse est une séquence multimodale créée à l'origine dans Twitter. Il s'agit d'un ensemble constitué du symbole dièse² (#) qui précède une suite de caractères alphanumériques. L'ensemble ainsi constitué est typographiquement soudé car aucun blanc ne peut s'insérer à l'intérieur du MD. Celui-ci est donc délimité par des blancs et commence toujours par le signe #, ce qui le rend facilement repérable dans le texte. À cela s'ajoute une couleur différente de la police des caractères indiquant que le bloc est cliquable en tant qu'hyperlien. Sa taille n'est pas limitée et peut théoriquement occuper la totalité du tweet, soit 139 caractères alphanumériques précédés du signe #. La taille moyenne d'un mot-dièse observée dans notre corpus est de 19 caractères. Son usage est facultatif – il apparaît pourtant dans plus de la moitié des messages du corpus (64 %). Le nombre de MD utilisés dans le tweet est également libre. Enfin, aucune règle d'écriture spécifique ne régit sa création ce qui en fait un élément dont le nombre théorique de réalisations est infini.

Dès les débuts de Twitter, le mot-dièse a été proposé comme solution pour créer des groupes de discussion. Il a également été utilisé de manière spontanée pour indiquer la thématique du message et faciliter la recherche et la navigation dans le site. Il suffit de cliquer sur un mot-dièse pour accéder à l'ensemble des messages qui le contiennent. Un message peut ainsi être indexé dans un seul ou dans plusieurs fils thématiques en fonction du nombre de mots-dièse qu'il contient. Rapidement, les MD se sont émancipés de leur fonction initiale d'indexation en développant diverses fonctions discursives.

Dans une analyse de ce RSN, Zappavigna (2011) indique que le mot-dièse permet une affiliation diffuse de la communauté autour d'un sujet. En d'autres termes, différentes personnes peuvent n'avoir en commun que l'intérêt qu'ils portent à un sujet donné et qui se manifeste à travers une recherche effectuée grâce au mot-dièse. L'auteure explique

² Bien que la dénomination *croisillon* soit plus adaptée pour désigner le signe #, nous adoptons ici le terme *dièse* pour ne pas brouiller les pistes par rapport au terme *mot-dièse*.

que le MD rend ainsi le tweet investigable. La création des MD étant libre, plusieurs formes concurrentes peuvent coexister pour référer à une même thématique. Tout l'intérêt des usagers réside dans la reprise du même mot-dièse s'ils souhaitent voir leurs messages apparaître dans un seul et même fil et augmenter ainsi la visibilité de leurs messages. Deux mots-dièse différents qui indexent un même sujet sont ainsi en concurrence et renvoient à deux fils différents (ex : *#Sarko*, *#Sarkozy*, *#NSarkozy*, *#NicolasSarkozy* etc.)³. Notons que la moitié des MD dans notre corpus sont des hapax. Au vu de son efficacité et de sa popularité, ce type d'étiquetage collaboratif spontané a peu à peu été adopté dans d'autres RSN.

Pour indiquer qu'il participe à la fois du linguistique et du technologique, M.-A. Paveau (2013a) appelle le mot-dièse « technomot » ou « technomorphème ». La Commission générale de terminologie et de néologie conseille l'utilisation du terme *mot-dièse* à la place de l'anglais *hashtag* et précise qu'il « signale un sujet d'intérêt et est insérée dans un message par son rédacteur afin d'en faciliter le repérage. »

Dans (Jackiewicz et Vidak, 2014) nous avons décrit les propriétés linguistiques du MD et dans (Vidak et Jackiewicz, sous presse) nous avons exploré la manière dont il est utilisé pour exprimer l'émotion, pour véhiculer des opinions, des partis-pris ou bien pour créer et entretenir la polémique. Nous avons pu constater qu'il permet d'affiner la dimension dialogique dans Twitter et qu'il participe à l'ouverture du contexte immédiat du tweet en le connectant à un cotexte sous-jacent. Il en découle qu'il s'interprète en fonction du contexte immédiat tout comme du contexte adjacent auquel il renvoie. Il véhicule ainsi un contenu informationnel doublement contextualisé.

Du point de vue formel, le MD peut être constitué d'un lexème seul (*#joli*, *#aujourd'hui*) ou de plusieurs lexèmes agglutinés (*#mariagepourtous*, *#RemetsToiEnQuestionSi*). La majorité des MD sont des noms ou des GN, cependant toutes les autres catégories grammaticales peuvent également fonctionner comme des mots-dièse (*#nier*, *#jamais*, *#aïe...*). Du point de vue syntaxique, le mot-dièse peut être adjoint au message (au début ou à la fin) sans participer à la construction syntaxique de celui-ci. Il est alors le plus proche d'un usage purement indexical. Il peut être intégré à la phrase et assumer une fonction syntaxique (souvent

³ La discrimination entre les majuscules et les minuscules ainsi qu'entre les lettres accentuées ou non n'est pas pertinente dans Twitter. Nous reproduisons ici les MD tels quels.

les fonctions les plus usuelles du GN) ou bien il peut constituer à lui seul une phrase à part entière, qu'elle soit minimale (*#jattends*) ou développée (*#HétérosNousSoutenonsLeMariagePourTous*).

1.3. Les fonctions du mot-dièse

Outre sa fonction d'origine (indexer et indiquer le sujet du message), le mot-dièse a également développé des fonctions pragmatiques. Il peut s'agir de MD médiatifs qui indiquent la source de l'information fournie (souvent des sigles des titres d'émissions de télévision ou d'institutions) ou cadratifs (*#departementales2015*, *#Lyon*). Un MD peut être utilisé comme marqueur métalinguistique (*#pourrire*, *#sérieux*). Il peut servir à recueillir des opinions ou des témoignages, à constituer une communauté d'intérêt et maintenir le dialogue, à exprimer les émotions, des jugements ou des partis-pris (*#désolé*, *#pitoyable*, *#Jaimepas...*), à entretenir la polémique dans un débat, etc.

En créant un mot-dièse, les usagers de Twitter manifestent leur volonté de créer un contenu interactif que d'autres personnes pourront reprendre dans leurs propres messages. Sachant qu'un peu plus d'un tiers des messages ne comporte pas MD, le fait d'en employer un ou plusieurs relève d'un choix délibéré et conscient. Outre que pour indexer leurs messages et indiquer leur thématique, les auteurs s'en servent aussi pour interpeller les autres usagers et établir la communication, même indirecte. L'une des motivations peut également être la volonté de devenir auteur d'un MD qui serait repris par une communauté large. On assiste ainsi sur Twitter à une production foisonnante qui excelle en inventivité linguistique dans la recherche d'originalité. Il faut rendre le MD visible, attrayant et efficace. Cependant, nombreux sont ceux qui ne sont guère repris, notamment parmi les MD construits sur le patron *#Xpourtous* à partir de *#mariagepourtous* (*#ConnardPourTous*, *#mariageCIVILpourtous* etc.). Certaines réalisations suggèrent que le mot-dièse a été créé dans le seul but d'attirer l'attention du lecteur et exploitent alors l'effet d'emphase ainsi obtenu (*#BizzareVousAvezDitBizzareCommeCestBizzare*, *#CARTONROUGEALHOMOPHOBIE* etc.).

L'inventivité des usagers est dans d'autres cas couronnée de succès, ce qui se vérifie à travers les différentes variantes de la formule « mariage pour tous ». Certaines d'entre elles ont été largement plébiscitées à travers les reprises (*#floppourtous*, *#concertpourtous* ou *#égalitépourtous*). L'écho de la formule d'origine, suffisamment forte et ancrée dans les

usages discursifs dans et en dehors de Twitter, se ressent ainsi à travers se formes de réfection qui confirment ainsi le statut de formule du motif de départ. Pour un examen détaillé du mot-dièse en tant que formule discursive au sens de Krieg-Planque (2009, 2012) nous renvoyons à nos deux articles (2014, sous presse). Le fait qu'un mot-dièse puisse être reconnu en tant que formule et subir des défigements est un des critères nécessaires pour lui reconnaître le statut de SF.

Les MD apparaissent comme nécessaires voire indispensables à une communication de qualité sur les réseaux sociaux. Ils sont utilisés non seulement pour indexer le contenu et indiquer la thématique du message mais également pour rendre le discours plus expressif et sémantiquement plus riche et pour augmenter la visibilité du message. Pour cela, le MD est comme une sorte de « particule communicative », un élément linguistique multimodal qui permet d'établir et d'entretenir la communication entre usagers et également de créer les liens cotextuels entre différents messages.

2. Du mot-dièse à la séquence figée

2.1. Les différents types de mot-dièse – séquence figée

Avant tout examen phraséologique, deux couples d'oppositions doivent être distingués. D'une part, l'on constate globalement deux usages du MD en fonction de son intégration syntaxique dans le message : il peut être soit intégré à la structure syntaxique de la phrase et assumer une fonction, soit adjoint et sans rapports syntaxiques avec le reste de l'énoncé. Les cas intermédiaires sont assez rares. D'autre part, il s'agit de distinguer deux types formels de mots-dièse : les MD monolexicaux et les MD polylexicaux, les seconds étant constitués du signe dièse et d'au moins deux lexèmes.

Du point de vue phraséologique, l'examen des MD du corpus fait ressortir à ce stade trois cas de figures distincts : le MD est un lexème simple ou un syntagme libre ; le MD est créé à partir des phrasèmes existants en langue ; le MD permet la création d'un nouveau phrasème. Il semble que le MD lexème simple ne soit pas éligible au statut de séquence figée. Cependant, il pourrait être considéré comme pragmatème (voir *infra*, 3.1). Signalons également parmi les MD polylexicaux ceux qui sont soit des hapax ou bien ne connaissent pas une popularité suffisante pour être considérés comme des SF.

Les phrasèmes préexistants sont de bons candidats à la transformation en MD. Leur statut de séquence figée facilement reconnaissable par les locuteurs, la fréquence de leur usage ainsi que les contenus connotatifs qu'ils véhiculent permettent leur utilisation aisée dans les messages de Twitter aussi (*#mieuxvauttardquejamais*, *#lechangementsmaintenant*, *#mèreporteuse*, *#JeSuisCharlie*). Cette transformation en MD peut alors être considérée comme une reconnaissance de leur figement avancé. Ils subissent une transformation formelle (soudure et adjonction du signe #) qui ne semble pas altérer leurs propriétés phraséologiques. Leur emploi sur Twitter est sensiblement le même que celui en dehors des RSN. Il peut arriver qu'un phrasème devenu MD subisse certaines transformations (insertion, changement paradigmatique, ajout de déterminants etc.) sans que cela rompe le lien avec le phrasème d'origine – il s'agit alors de cas classiques de défigement. Ces mots-dièse peuvent occuper toutes les positions dans le message, leur intégration ou non dans celui-ci et le type de rapport qu'ils entretiennent avec celui-ci pouvant être analysés du point de vue pragmatique.

Le troisième cas concerne les mots-dièse qui accèdent au statut de phrasème en raison de la nature même du MD⁴, de leur fréquence d'usage et de leurs éventuels défigements. En voici quelques exemples : *#AnnulezLaMarcheDu11Janvier*, *#OnFaitLaPaix*, *#NousSommesEnsemble*⁵, *#UnePhraseDeProf* etc. Dans cet univers d'échanges où l'immédiateté et la réactivité sont de mise, il est encore difficile d'observer des MD-phrasèmes qui se stabilisent de manière particulièrement durable. Ces cas sont présents notamment sur Twitter, mais leur reprise en dehors de ce RSN n'est pas à exclure étant donné les usages du mot-dièse en dehors de son média d'origine. Néanmoins, il semble pertinent de comparer ces réalisations avec les séquences figées en langue. D'autre part, l'observation des phrasèmes dans Twitter pourrait permettre de donner une indication sur l'état de figement des phrasèmes préexistants : plus il est figé et répandu dans l'usage, plus il a de chances d'être repris dans les RSN.

⁴ Il s'agit d'un figement formel de fait, qui s'accompagne d'un indice typographique facilement reconnaissable et permettant son identification.

⁵ Trois mots-dièse créés sur Twitter à la suite des attentats de janvier 2015 à Paris et devenus particulièrement populaires.

2.2. Les propriétés formelles du mot-dièse qui en font une séquence figée

Le MD polylexical peut être considéré soit comme une séquence non contrainte à laquelle est ajouté le symbole #, soit comme une séquence figée. Les considérer comme des séquences libres revient à négliger les conséquences syntaxiques, sémantiques et pragmatiques qui découlent de cette réunion de lexèmes.

Les propriétés formelles du mot-dièse polylexical en font une SF de fait car la nouvelle séquence ainsi faite fonctionne comme un tout et est traitée comme telle par les usagers et par les moteurs du RSN. Ce figement formel est soutenu par le fait que chaque structure ainsi figée renvoie à un seul fil de messages et que tout éventuel changement de structure résulte par la rupture de ce lien cotextuel, entravant ainsi la fonction originelle du MD. La solidarité entre constituants à l'intérieur d'un MD est ainsi très forte. D'autre part, un MD-séquence figée est irréductible à l'un de ses constituants. Un fois formé, il acquiert un sens qui est, la plupart du temps, compositionnel et transparent. La transparence des mots-dièse est une caractéristique qui leur est inhérente d'une part à cause du caractère très récent du phénomène en comparaison avec les SF opaques dont le sens évolue depuis bien plus longtemps. D'autre part, le caractère transparent des MD polylexicaux semble pour l'instant être une nécessité dans un univers où les échanges sont rapides et où la possibilité de reprise du MD dépend directement de sa lisibilité immédiate.

D'un point de vue syntaxique on observe également un blocage total dû, lui aussi, à la soudure formelle et fonctionnelle du MD. Par ailleurs, un mot-dièse polylexical peut, comme nous l'avons observé précédemment, n'avoir aucune fonction syntaxique dans l'énoncé et être tout simplement adjoind à celui-ci pour indiquer la thématique et permettre l'indexation du message. Son rapport avec le reste de l'énoncé peut être alors interprété comme une sorte de commentaire métadiscursif. Ce qui change ici par rapport aux SF dans l'usage hors du discours des RSN, c'est la possibilité que des MD-SF remplissent des fonctions syntaxiques de tous genres, notamment celles du GN. Néanmoins, nous observons une mobilité assez faible de ces MD dans la phrase puisqu'elle dépend de l'intégration syntaxique de celui-ci et que la majorité des MD est utilisée en fin de message, en tant que commentaire métadiscursif ou simple élément d'indexation.

#SmsTesMonCrush c'est beaucoup moins intelligent que le
#OnFaitLaPaix quand même...

#onfaitquoi #OnFaitLaPaix ? On lit Véliocas sur la nouvelle
Collaboration #UMP #PS et les islamistes, en librairie
<http://www.youtube.com/watch?v=sf52hihhfkM> ...

#OnFaitLaPaix "à jamais" merci cher inconnu de mettre tant de
sympathie dans ces messages de paix

Les deux derniers exemples illustrent la compositionnalité du sens et la proximité que les MD entretiennent avec leurs constituants pris séparément. On remarque que le mot-dièse peut être traité soit comme un tout figé, soit comme un agglomérat de constituants dont le sens ne s'est pas effacé, ce qui se justifie par une évolution diachroniquement très courte. Néanmoins, reste le fait que, de par leur figement formel, les MD accèdent à un autre sens, un sens très fortement ancré dans le contexte de leur usage, ce qui les sépare des séquences libres constituées des mêmes éléments. Ceci dit, au vu de leur compositionnalité, il semblerait que les mots-dièse qui ne reprennent pas les séquences déjà figées et attestées ailleurs que sur les RSN ne soient pas éligibles *stricto sensu* au statut de séquences figées. Ce point précis sera discuté dans la troisième partie de cet article.

2.3. La fonction du signe

Comme nous l'avons vu, l'adjonction du signe dièse est nécessaire pour la création d'un mot-dièse. Le signe dièse a tout d'abord une fonction technique propre à l'univers des RSN : il permet au système informatique de reconnaître et de traiter la séquence en tant qu'élément classant et interactif. Outre son caractère technique, le dièse joue également le rôle de signe distinctif permettant l'emphase. Nombre de MD hapax sont ainsi créés pour des besoins de mise en valeur de la portion textuelle à laquelle ils sont adjoints, pour attirer l'attention du lecteur sur le contenu ainsi marqué :

#hemoroïdeFiscale

#humonestbienserrédanscettelanif

#JeComprendLes82PourcentsDeFrançaisQuiEnOnRienAfoutre

#paysdesdroitsdelhommeetdelafemmeetdesenviesfechacun

D'un point de vue linguistique, le dièse est signe du figement pour tout usager ou lecteur du message qui contient des MD. Il indique que le contenu qui le suit doit être pris en compte comme un tout et interprété comme tel, ce dont témoignent la possibilité de l'intégration syntaxique

du mot-dièse, la possibilité de recevoir un déterminant ou son emploi métalinguistique. L'on peut donc considérer le signe dièse comme un élément formel qui sert à figer la séquence à tous points de vue. Ce signe a aussi l'avantage de faciliter le traitement automatique des langues.

Du moment où il est considéré comme signe du figement, le dièse ne peut-il pas alors correspondre à l'indice formel du moule locutionnel, une sorte de marqueur de matrice lexicale permettant la création d'une séquence figée ? Dans les exemples précédemment cités on constate la facilité avec laquelle on peut créer une séquence formellement figée sur Twitter grâce au # sans que cela relève uniquement de sa dimension technique. Cela se vérifie également à travers les nombreux cas de défigement présents dans notre corpus, notamment à partir de la formule « mariagepourtous ». Non seulement celle-ci est perçue comme figée en discours en dehors de Twitter et reprise comme telle sous forme de mot-dièse mais elle y est également transformée pour donner des variations originales parfois largement plebiscitées par la communauté: *#floppourtous*, *#mathématiquespourtous*, *#manifantimariagepourtous*, *#1millionpourtous* etc.

En créant le mot-dièse on a offert aux locuteurs la possibilité de marquer de manière formelle et explicite le phénomène de figement. Il ne serait pas étonnant de voir, dans un futur plus ou moins prochain, le signe dièse adjoint aux SF en dehors des RSN tout comme on pourrait imaginer que cet usage se popularise au point de devenir un fait de discours et d'accéder à terme au niveau de langue.

3. Examen phraséologique du mot-dièse

3.1. Le statut de phrasème

Examiné du point de vue phraséologique selon Mel'čuk (2011), le mot-dièse apparaît à l'origine comme un syntagme libre (non contraint et compositionnel) :

D'un point de vue paradigmatique, un signe complexe **AB** est non contraint si et seulement si [=ssi] le signe **AB** lui-même et ses signes constituants **A** et **B** sont sélectionnés par le locuteur strictement pour leur sens : 1) **AB** est sélectionné indépendamment de la situation extralinguistique impliquée dans son énonciation et 2) **A** est sélectionné indépendamment de **B** et vice versa. Autrement dit, **AB**, **A** et **B** sont sélectionnés d'après les règles générales de sélection de la langue. (Mel'čuk, 2011, p. 42)

Il semblerait ainsi qu'il ne soit pas éligible au statut de phrasème et qu'il ne puisse être considéré que comme une séquence formellement figée. Or, ceci est valable pour la toute première création de chaque MD. En effet, dès l'instant où il est repris par un autre usager, les constituants du MD ne sont plus indépendants de la situation extralinguistique car le choix de reprise du MD est fait en fonction de l'écologie de Twitter et de l'intention de l'usager. Si l'on admet alors que le MD (repris) n'est pas une séquence libre et que son sens est compositionnel, il accède alors au statut de phrasème, notamment dans le groupe des pragmatèmes, clichés et collocations proposé par l'auteur.

Dans le même article, l'auteur définit le pragmatème comme suit : « Un phrasème est un pragmatème, ou un phrasème pragmatique, ssi il est contraint par rapport à une situation extralinguistique. Un pragmatème est un syntagme non libre fixé par rapport à la situation SIT où il est utilisé. » (Mel'čuk, 2011, p. 43) Si l'on considère le mot-dièse de ce point de vue, celui-ci apparaît comme un parfait pragmatème, étant contraint par la situation discursive dans laquelle il est employé. Ceci pourrait être signifié en termes de Mel'čuk comme suit : [SIT : dans Twitter, accompagné du signe dièse]. Cela équivaldrait à considérer tous les mots-dièse comme pragmatèmes et à les faire ainsi accéder au statut de phrasèmes, qu'ils soient monolexicaux ou polylexicaux. Compte tenu des propriétés du MD dans Twitter et des affordances de ce RSN ceci ne serait pas faux.

Or, que faire de tous les cas de MD-hapax que l'on rencontre dans l'usage ? La prise en compte de la reprise des MD par d'autres usagers semble être essentielle pour déterminer si celui-ci peut être jugé comme un phrasème ou non. Pour reprendre un exemple de l'auteur, songeons à l'écriteau « Défense de stationner » [SIT : sur un panneau de circulation] : il est reconnaissable et compréhensible par une grande majorité et ne présente guère de problème d'interprétation. En effet, toute séquence n'est pas forcément figée si elle n'est pas fixée dans l'usage, ce qui, dans Twitter, se traduit par la reprise du mot-dièse, ne serait-ce qu'à une petite échelle. Théoriquement, il suffirait ainsi qu'un mot-dièse soit repris par un autre usager pour devenir de fait un phrasème. Ceci est bien évidemment valable uniquement dans ce RSN où la pérennité d'un usage lexical n'est pas de même nature qu'en dehors de celui-ci. Nous pourrions donc écarter de la catégorie des phrasèmes tous les mots-dièse à occurrence unique. Il reste cependant un obstacle à franchir pour que le MD se fige à long terme: sa reprise doit être continue et s'effectuer même en dehors du RSN où il a été créé. Nous constatons aujourd'hui de plus en plus l'usage des MD hors Twitter et autres RSN, notamment dans les médias électroniques

mais également dans la presse et dans la publicité, et même dans les minimessages (SMS), bien que ces reprises ne soient pas toutes de même ordre. Une reprise et une recirculation large en dehors des limites des RSN paraissent nécessaires pour pouvoir reconnaître le statut de phrasème à un MD. Or, cette reprise n'est pas encore assez importante.

En suivant la catégorisation des phrasèmes par Mel'čuk (2011), il convient de vérifier si le MD peut correspondre à d'autres types de phrasèmes. Chaque MD est complètement contraint dans tous ses constituants et est, dans la grande majorité des cas, sémantiquement compositionnel. Il devrait donc correspondre au cliché. Or, une différence essentielle apparaît lorsqu'on compare un MD comme *#OnFaitLaPaix* avec *Mieux vaut tard que jamais*, exemple donné par l'auteur. Les deux sont complètement contraints à cette différence près que le MD cité est surtout contraint localement, dans Twitter, de manière formelle et technique, alors que le cliché cité est contraint dans l'usage et peut être invariablement repris dans tout autre univers discursif, ce qui ne serait pas le cas du MD. Ce type de contrainte constitue la différence entre les phrasèmes communément admis et attestés et les MD de manière générale. L'on peut donc difficilement se fonder sur ce seul critère pour définir un MD comme phrasème. Celui-ci ne correspond pas d'emblée ni à un cliché, ni à une collocation ou une locution. Les comparer de ce point de vue précis serait une erreur car le statut de la contrainte qui opère au niveau des phrasèmes attestés n'est pas le même que pour les MD. Rien n'interdit cependant de transformer un cliché en MD ou bien d'imaginer qu'un MD devienne un jour un cliché. La question est de savoir s'il sera alors repris tel quel, soudé et accompagné du signe dièse.

Toujours est-il que le MD peut être considéré comme un phrasème dans les RSN du fait de son statut sémantique spécifique qui en fait une unité de sens indivisible et irréductible à un seul de ses constituants, en plus des contraintes techniques et pragmatiques à l'œuvre. L'on peut ainsi dire que les mots-dièse répondent au critère de globalisation. La synthèse sémantique s'opère dès l'instant où le lecteur repère un MD dans le texte : le signe # et l'aspect visuel (couleur différente et agglutination des constituants) signalent alors la nécessité de prendre en compte le tout ainsi formé comme un ensemble ayant son propre sens.

2.4. Tests phraséologiques

Le MD n'a de sens et ne remplit ses fonctions de base que s'il est repris et diffusé à travers les discours de Twitter. Dans le cas contraire, il se voit

relégué au statut d'élément discursif à vocation stylistique. Autrement dit, le mot-dièse ne peut pleinement fonctionner et remplir ses fonctions essentielles s'il n'est pas attesté par les autres utilisateurs. L'on peut dire qu'il s'agit d'un cas de figement dans une diachronie très courte compte tenu de l'écologie spécifique de Twitter. Sa reprise de plus en plus observée en dehors de ce RSN joue en faveur de son statut de séquence figée. Cependant, il paraît difficile de parler de séquence figée en langue pour un type d'unité et non pour une occurrence donnée.

En tant que séquence figée par définition, le MD polylexical répond aux principaux critères de Gross (1996), repris par Mejri (2005a) au sujet du degré de figement, car aucune variation ni transformation ne peut être admise sous peine de compromettre son fonctionnement dans Twitter. À ce stade, il serait donc inutile de soumettre les MD à ce type de tests. Les contraintes formelle et fonctionnelle semblent être en contradiction avec son statut de séquence figée en comparaison avec les SF attestées en langue. Au vu des précédentes observations et de celles de Mejri à propos du figement et de la néologie (2005b), nous serions tenté de refuser le statut de séquence figée aux mots-dièse. Cependant, comme nous l'avons vu, des caractéristiques formelles, sémantiques, syntaxiques et des observations de son usage nous permettent de dire que le MD n'est pas une séquence dont le figement serait contraint uniquement par son fonctionnement à l'intérieur du discours de Twitter et des autres RSN. La manière dont les locuteurs s'approprient les MD existants et leur reprise observée en dehors de ces RSN, même s'il ne s'agit pour l'instant que des balbutiements, indiquent que l'on ne doit pas d'emblée écarter cette nouvelle solution de figement qui s'offre ainsi à la langue. Ce qui ressort ici c'est que le mot-dièse s'apparente à un mécanisme particulièrement dynamique de formation des séquences figées dont seules certaines réalisations pourraient accéder au statut de SF attestée. Il participe ainsi à la fois de la néologie et du figement en permettant de créer de nouvelles unités qui sont reprises telles quelles.

2.5. Où situer le mot-dièse dans le lexique d'une langue ?

Du fait de leur statut de séquences figées et de leur fonctionnement linguistique propre, les mots-dièse peuvent être considérés comme des unités à part entière. Leur place dans le lexique d'une langue ne relève pas simplement de l'assemblage compositionnel des sens des unités qui les constituent. D'autre part, leur sens global est, pour l'heure, lié à leur usage dans le contexte des RSN dans lesquels ils apparaissent. C'est leur reprise

en dehors de ces réseaux qui nous incite à envisager la possibilité de leur classification par rapport aux unités lexicales connues et décrites en langue.

En reprenant la distinction proposée par Mel'čuk dans son article sur les phrasèmes dans le dictionnaire (2011, p. 41) et au vu de nos précédentes observations, nous pouvons constater que le choix de la classe dont ils relèveraient est délicat. Les MD monolexicaux posent notamment un problème car chacune de leurs occurrences peut être identifiée en tant qu'unité lexicale qui la constitue. Ce faire reviendrait à supprimer la dimension même du mot-dièse et à décrire ainsi le lexème en question comme une réalisation possible en contexte. Cette solution peut sembler satisfaisante si l'on prend en compte les occurrences uniquement dans le cadre de Twitter et des autres RSN qui utilisent les mots-dièse. D'autre part et selon le même principe, les MD polylexicaux pourraient être assimilés, selon le cas, aux phrasèmes lorsqu'ils en reprennent un déjà attesté en langue ou aux syntagmes libres dans le cas contraire. Cependant, ce serait négliger leur sens global, leur dimension multimodale et leur fonctionnement en tant que pragmatèmes au sens large.

La solution pourrait être trouvée en considérant les mots-dièse comme des sous-classes spécifiques des lexèmes pour les MD monolexicaux et des phrasèmes pour les MD polylexicaux. Ce faisant, on admet d'emblée le statut de phrasème à un nombre non négligeable de mots-dièse, ce qui, dans l'état actuel des connaissances sur ces nouvelles formes linguistiques, peut être remis en cause au vu du caractère encore trop récent et non pérennisé du phénomène. L'autre solution pourrait être trouvée dans la prise en compte des mots-dièse en tant qu'unités lexicales d'un nouveau genre en leur conférant une place à part entière, une sorte de troisième classe à côté des lexèmes et des phrasèmes. Ce choix est facilement critiquable au vu de leur proximité avec ces deux classes et du manque de recul que nous avons par rapport à ces réalisations spécifiques. L'on serait alors tenté de considérer chacune des réalisations du mot-dièse comme une spécialisation de sens d'une unité correspondante à laquelle on a simplement ajouté le signe # pour pouvoir l'utiliser dans les RSN. Cela reviendrait ainsi à admettre que chaque mot-dièse possède sa paire lexicale dépourvue de dièse et donc de faire une place parmi les entrées du dictionnaire, aussi riche et extensible soit-il, aux formes comme *flop pour tous*, *on fait la paix* ou encore *annulez la marche du 11 janvier*. On voit bien que considérer ainsi ces ensembles qui n'ont pas de signification propre en dehors de leur usage en tant que mot-dièse serait erroné. Une

autre solution pourrait être de considérer les mots-dièse comme des agglutinations de lexèmes auxquels on a ajouté le signe dièse afin de les transformer en unités techniques opérables dans un univers discursif spécifique et dont le sens est complètement compositionnel. Cela reviendrait à nier et rejeter toutes les observations que nous et d'autres chercheurs avant nous avons apportées. Ce serait de fait négliger une réalité observable dans les discours à travers les différents usages toujours plus riches et imprévisibles qui sont faits des mots-dièse et qui débordent désormais du cadre des RSN.

La solution n'est pas commode à trouver et il n'est probablement pas encore arrivé le moment où ce genre de considération mérite d'être fait. Néanmoins, ces observations permettent de constater que le statut du mot-dièse (lexical, syntaxique, sémantique et pragmatique) est spécifique et qu'il mérite d'être pris en compte en tant que tel, ce qui revient à accorder au mot-dièse un statut linguistique à part entière. L'utilisation croissante des mots-dièse et l'adhésion qu'il rencontre auprès des locuteurs témoignent d'un nouveau phénomène linguistique encore très récent sur l'échelle de l'évolution des langues naturelles. Son apparition récente et son évolution observée jusqu'à ce jour appellent à la fois à un suivi rapproché et à une prise de distance critique.

Conclusion

Les propriétés et le statut spécifique du mot-dièse en font un élément de langage spécifique dont l'analyse et la description ne sont pas aisées. Il s'agissait à l'origine d'une séquence artificiellement figée et de portée locale dans Twitter, où elle avait été créée. Les différents travaux de recherche menés sur son sujet et notre étude des réalisations dans un vaste corpus nous permettent de dire qu'il s'agit d'une séquence qui est figée et contrainte tout d'abord par sa fonction indexicale. La nécessité qu'un mot-dièse soit repris pour pouvoir pleinement fonctionner est un des principaux critères permettant de le classer parmi les séquences figées. Bien qu'il s'agisse de séquences totalement bloquées du point de vue syntaxique et que leur sens soit global, en raison d'une diversité trop importante des réalisations rencontrées il serait difficile d'identifier tous les mots-dièse comme phrasèmes. Comparé à d'autres séquences figées en langue, le mot-dièse est doté de fonctions supplémentaires liées au contexte de son utilisation et qui lui confèrent un trait sémique spécifique. En étant intimement liés au contexte de leur production et contraints par celui-ci, les MD se rapprochent le plus des pragmatèmes au sens large.

Au vu de leur nature et des particularités de leur fonctionnement, il est malaisé de leur trouver une place dans le lexique à ce stade de leur évolution. Cependant, étant plébiscités par l'usage croissant qui en est fait et par leur reprise en dehors des RSN de leur origine, les mots-dièse s'approchent de plus en plus des phrasèmes attestés en langue. Pour le moment, leur caractère trop récent et de fait pas encore stabilisé, leur évolution rapide comparativement à l'ontogénèse des autres éléments langagiers et notamment des séquences figées attestées, ainsi que la nature de leur reprise en dehors des RSN, encore balbutiante, imposent la prudence et la prise de distance critique dans leur analyse et leur description.

Cette étude a également permis de postuler que l'adjonction du signe dièse et le rôle qu'il joue au sein du mot-dièse permettent de considérer ce signe comme indice formel du moule locutionnel.

La surveillance rapprochée de l'évolution des mots-dièse et de leur reprise en dehors des contextes de leur origine est nécessaire pour poursuivre leur étude du point de vue phraséologique. Leur émergence et leur évolution rapide pourraient permettre d'observer la création de nouveaux phrasèmes. Il conviendrait pour cela de continuer à les confronter aux phrasèmes déjà attestés en langue ainsi que d'étudier leur degré de figement tel qu'il a été décrit dans les travaux de S. Mejri.

Bibliographie

- Cervulle, M. et Pailler, F. (2014) #mariagepourtous : Twitter et la politique affective des hashtags. *Revue française des sciences de l'information et de la communication*, 4. DOI:10.4000/rfsic.717
- Gross, G. (1996) *Les expressions figées en français: noms composés et autres locutions*. Paris: Ophrys.
- Jackiewicz, A. et Vidak, M. (2014) Étude sur les mots-dièse. *SHS Web of Conferences*, 8, 2033-2050. DOI:10.1051/shsconf/20140801198
- Krieg-Planque, A. (2009) *La notion de formule en analyse du discours: cadre théorique et méthodologique*. Besançon: Presses universitaires de Franche-Comté.
- Krieg-Planque, A. (2012) *Analyser les discours institutionnels*. Paris: Armand Colin.
- Longhi, J. (2012). Discours, style, format : contraintes et niveaux de structuration de la textualité des Tweets de Mouloud. *SHS Web of Conferences*, 1, 1127-1141. DOI:10.1051/shsconf/20120100022
- Longhi, J. (2013) Essai de caractérisation du tweet politique. *L'information grammaticale*, 136, 25-32. Repéré à <https://halshs.archives-ouvertes.fr/halshs-00940202/document>

- Longhi, J. (2015) Pratiquer la twittérature à travers la twittécriture. Position théorique, mise en pratique et retours d'expérience. Dans *Numérique et écriture littéraire. Mutations des pratiques*. Paris: Hermann.
- Mejri, S. (2004) Introduction : Polysémie et polylexicalité. *Syntaxe et sémantique*, N° 5(1), 13-30. Repéré à http://www.cairn.info/resume.php?ID_ARTICLE=SS_005_0013
- Mejri, S. (2005a) Figement absolu ou relatif : la notion de degré de figement. *Linx. Revue des linguistes de l'université Paris X Nanterre*, (53), 183-196. doi:10.4000/linx.283
- Mejri, S. (2005b) Figement, néologie et renouvellement du lexique. *Linx. Revue des linguistes de l'université Paris X Nanterre*, (52), 163-174. DOI:10.4000/linx.231
- Mel'čuk, I. (2003) Collocations dans le dictionnaire. Dans T. Szende (dir.), *Les écarts culturels dans les dictionnaires bilingues* (pp. 19-64). Paris: H. Champion.
- Mel'čuk, I. (2011) Phrasèmes dans le dictionnaire. Dans *Le figement linguistique la parole entravée* (pp. 41-61). Paris: H. Champion.
- Paveau, M.-A. (2012) Activités langagières et technologie discursive. L'exemple de Twitter. *La Pensée du discours*. Hypothèses.org. Repéré à <http://penseedudiscours.hypotheses.org/8338>
- Paveau, M.-A. (2013a) Genre de discours et technologie discursive. Tweet, twittécriture et twittérature. *Pratiques. Linguistique, littérature, didactique, Théories et pratiques des genres* (n° 157-158), 7-30. Repéré à <https://hal.archives-ouvertes.fr/hal-00824817>
- Paveau, M.-A. (2013b) Hashtag [Dictionnaire]. Billet. Repéré à <http://technodiscours.hypotheses.org/488>
- Paveau, M.-A. (2013c) Tweet [Dictionnaire]. Billet. Repéré à <http://technodiscours.hypotheses.org/385>
- Vidak, M. et Jackiewicz, A. (sous presse) Les outils multimodaux de Twitter comme moyens d'expression des émotions et des prises de position. *Langage et Société*.
- Zappavigna, M. (2011) Ambient affiliation: A linguistic perspective on Twitter. *New Media & Society*, 13(5), 788-806. DOI:10.1177/1461444810385097

VIDAK, Marko, Chercheur associé,
STIH, Université Paris-Sorbonne
<marko.vidak@yahoo.fr>