

The ‘hashtag’: A new word or a new rule?

Paola-Maria Caleffi

Born as a topic marker to make tweets searchable, today the pound symbol is gaining popularity not only on other social platforms, but also offline, with hashtags increasingly appearing in the linguistic landscape (e.g. headlines, advertising, political slogans). This paper analyzes hashtagging as a productive process of word-formation in English and Italian, both online and offline. The analysis is based on samples of hashtags from a corpus of tweets and samples appearing in the offline world. The study proposes a tentative taxonomy of hashtag types and poses questions on the nature of the ‘products’ of hashtagging. It also comments on the pragmatic exploitation of the collected samples.

Keywords: *Twitter; hashtag; hashtagging; word-formation*

1. Aim and scope of the study

The use of hashtags was introduced on Twitter as a way to classify messages (*tweets*) according to the topic, thus allowing users to easily search for specific content and share information related to it. Social designer Chris Messina is credited to have produced the very first Twitter hashtag, when, back in August 2007, he posted a tweet reading: “how do you feel about using # (pound) for groups. As in #barcamp [msg]?”.

A hashtag typically consists in a string of characters (possibly including numerical digits) preceded by the *pound* symbol # (also called *hash*). This combination serves as a label for the message itself and is “linked to a search for those characters” (Sagolla 2009: 167), which allows the retrieval of all tweets dealing with the labeled topic. By making it easier for users to find tweets related to a specific subject, the hashtag has developed into a “community building linguistic activity” (Zappavigna 2011: 2), fostering the creation of communities of people bound by their being interested in the same content and wishing to read and share information about it (Kricfalusi 2013). In fact, hashtags are no longer simply used as a categorization method, but they are specifically created by individual online users to comment on, praise or criticize ideas (#*democracyisbetter*) or people (#*celochiedebepp*), to promote brands (Coca Cola, #*AmericaIsBeautiful*) or events (#*Wimbledon*), to spread and provide updates on breaking news items (#*earthquake*), just to mention a few examples. Moreover, they are created by several social media experts, educators, institutions and major companies from all around the world to bring in more followers and keep them involved. For these reasons, hashtags show up continuously on Twitter, together with the evolving stream of content. Some of them have success and propagate, while others die immediately after birth and are restricted to a few messages.

Born on Twitter, the use of hashtags has now spread to other social media sites (like Instagram and Google+ in 2001, or Facebook in 2013) and it is largely and regularly employed by millions of social media users. What is more, hashtags are no longer restricted to the online world, since they are now starting to appear also in the linguistic landscape (Landry and Bourhis 1997) and in a variety of offline contexts, such as advertising boards, commercial shop signs, street manifestation posters and banners, television spots and

commercials, printed magazines and newspaper headlines, political slogans and speeches, to mention but a few.

This paper deals with such new phenomenon and tries to shed lights on its characteristics from a linguistic perspective. As a matter of fact, in this work I examine hashtags as ‘linguistic items’ produced through a potentially new word-formation mechanism, namely *hashtagging*, which would add to already existing morphological processes such as blending, clipping, compounding, etc. Specifically, this study seeks to explore (1) whether or not ‘hashtag’ is (only) a new word or (also) the specific product of a new morphological process, that is, hashtagging; (2) what the very nature of the ‘linguistic items’ produced through hashtagging is; (3) how hashtagging as a (potential) morphological rule works and (4) how hashtags are pragmatically exploited, both online and offline.

2. Background

Due to their potential as a form of social tagging – or “folksonomy” (Vander Wal 2007)¹ – and as a tool that facilitates the spread of information, Twitter hashtags have been the subject of several scholarly investigations.

Potts et al. (2011) have investigated the usage of hashtags by Twitter users exchanging information on natural disasters. Their findings show that inconsistency in content-tagging due to differences in conventions and usage among participants e.g. inconsistent formats, spellings, and word ordering) may be an issue when it comes to making such content more easily searchable and findable across the social web. Romero et al. (2011) have studied the mechanics of information diffusion on Twitter, finding “significant variation in how most-widely used hashtags on different topics” (Romero et al. 2011: 1) propagate within Twitter user population. Their study reveals that the sources of variation involve differences in *stickiness* (the probability of adoption of one hashtag based on the number of exposures), and in *persistence* (the extent to which “repeated exposures to a hashtag continue to have significant marginal effect” (Romero et al 2011: 1)). It also reveals that, variation in hashtag propagation seems to be aligned with the topic of the hashtag itself, with hashtags on politically controversial topics showing a particularly high degree of *persistence*. Hill & Benton (2012) have analyzed the exploitation of both general and TV-program dedicated hashtags during the broadcasting of a show (the TV program in their case being the American reality singing show *The Voice*). According to their conclusions, it seems that displaying hashtags on the screen during a program increases the number of dedicated tweets, and therefore viewers’ Twitter activity overall, as well as their engagement with the program during commercial breaks. Dickinson (2013) has explored the function of formulaic language in interactional discourse on Twitter, including the use of hashtags, which he refers to as “a Twitter-specific feature that may be considered formulaic” (Dickinson 2013: 24). His study highlights the fact that, besides having a practical facilitative purpose, hashtags foster affiliation with the related values and communities, making it possible for Twitter users “to interact with others on terms that other modes of communication cannot provide” (Dickinson 2013: 25).

¹ “Folksonomy is the result of personal free tagging of info and objects (anything with a URL) for one’s own retrieval. The tagging is done in a social environment (usually shared and open to others). Folksonomy is created from the act of tagging by the person consuming the information” (Vander Wal 2007, vanderwal.net/folksonomy.html).

However, hashtags have hardly been analyzed from a purely linguistic perspective. A language-based approach has been adopted by Cunha et al. (2011), who have studied the propagation of innovative hashtags on Twitter based on models for the analysis of linguistic innovation in speech communities. They have focused on the characteristics that distinguish a hashtag which spreads widely from one that fails to attract attention. Their findings reveal that the most used hashtags get very high frequencies of use, and that the most popular hashtags are simple, direct and short. This behavior resembles that of natural language, where people's linguistic choices within a speech community largely depend on the choices made by other people (Easley and Kleinberg 2010, quoted in Cunha et al. 2011) and where longer words tend to be avoided, presumably because they are uneconomic (Sigurd et al. 2004, quoted in Cunha et al. 2011).

Yet, to my knowledge, hashtags have never been looked at, neither investigated, as the result of a morphological process leading to the creation of 'linguistic items' whose identity does not match the definition of any part of speech in the traditional sense of the term. This is what the following sections are going to focus on.

3. The Twitter hashtag: Conventions and use(s)

According to the tweeting typographic conventions, a hashtag is a string of characters preceded by the symbol # (*hash*, or *pound*), as in *#barcamp*. It is not case-sensitive, so that, for instance, *#GRAVITY*, *#gravity* or any alternations of lower and upper case (*#GRaViTY*, *#gRaViTy* etc.) are possible and detected as the same unit. Instead, no whitespace characters are allowed. In a hashtag including two or more (recognizable) words, the initial letter of each word may be capitalized, as in *#SouthAfrica*, rather than *#southafrica*, or even *#South Africa*, which would lead the system to search for the topic "South" rather than "South Africa". Still according to the Twitter convention, a hashtag can contain numbers but cannot be made up entirely of numerical digits (e.g. *#123*), neither can it start with a number (e.g. *#123yo*). Special characters are not allowed either (e.g. *!*, *%*, ***, *\$* etc), except for the underscore (*_*). Finally, according to Twitter experts' recommendations, the best hashtags are those made up of a maximum of six characters, and three hashtags in one tweet should be the highest number of occurrences. Used excessively, hashtags could cause confusion and frustration. Therefore, they should be employed "sparingly and respectfully."²

Yet, since hashtags are user-initiated, Twitter users are free to develop their own ones, irrespective of the conventions, both in terms of hashtag shape and length, and in terms of their numerical presence within a single tweet. Indeed, hashtags have become some sort of a 'digital tic', occurring wildly and at times randomly online, with tweets and other social media posts often containing a high number of hashtags (or even only hashtags), as in:

- (1) *I love those people who feel like hash tagging 20 words #really #yourgay #hashtag #unicorns #whydoyouhashtag #randomhashtag #lasthashtag*

Moreover, although the Twitter hashtag was born as a categorization method, besides serving as metadata (namely, what the tweet is about), hashtags are now used in a number of ways

² Sources: www.twitter.pbworks.com, www.hashtag.org.

and for a number of purposes, very often as a contextual aside to comment on, give more depth to, or somehow emphasize what has been said, as in:

- (2) *Sarah Palin for President?? #Iwouldratherhaveamoose*
- (3) *My bestie has the best Instagram. Would it be weird if I started having her edit all my photos? #kidding #butnotreallykidding*
- (4) *My arms are getting darker by the minute. #toomuchfaketan*

but also as a disclaimer:

- (5) *BREAKING: US GDP growth is back! #kidding*

as a (seemingly) accidental remark or naming:

- (6) *Ahahahah Jack comunque ti tradisce... con mio fratello #ops*

to express personal feelings and emotions:

- (7) *#angry*

to support events or movements:

- (8) *#PrayforBoston*

for self-mockery:

- (9) *Feeling great about myself till I met an old friend who now races at the Master's level. Yup, there's today's #lessoninhumility*

for brand promotion:

- (10) *#ShareACoke*

for chat/conference participation:

- (11) *#ESSEconference*

These are just some of the main functions hashtags have been serving since the first *#barcamp* by Chris Messina in 2007, and it is exactly for this expansion of both their purposes and presence that they have aroused interest in many fields, especially those connected with communication practices.

4. Corpus and methodology

In order to answer my research questions, I collected a set of hashtags – some appearing online (Twitter) and others in contexts offline – in two different languages, namely English and Italian, and compared the way these hashtags are created and exploited in either language. To this aim, I specifically compiled a corpus made up of samples of tweets containing hashtags as well as examples of hashtags used in offline contexts (e.g. advertisements and advertising boards, headlines, political slogans). The Twitter part of the corpus includes 10000 tweets retrieved through Twitter API³ over a period of 6 months (March–August 2014) and is divided into two components, an English and an Italian one. Each component contains 5000 tweets posted from two different geographical areas within the relevant country, that is, 2500 tweets posted from London and 2500 tweets posted from Manchester for the English component; 2500 tweets posted from Milan and another 2500 posted from Rome for the Italian one. I decided to fetch my tweets from these areas as I deemed them to be somehow, though inevitably limitedly, representative of the way the practice is employed in the two countries.

Table 1 shows the details concerning the Twitter part of my corpus:

Table 1 *Details of the Twitter Corpus*

Subcorpus	Total tweets	Total tweets with #s	Percentage of tweets with #s	Percentage of #s per tweet
London	2500	606	24%	1.8%
Manchester	2500	554	22%	1.5%
Milan	2500	684	27%	2.0%
Rome	2500	717	29%	1.7%
Whole Corpus	10000	2561	26%	1.8%

Similarly, with reference to the examples of hashtags retrieved from the offline world, I collected samples of posters, banners, advertisements, headlines etc., both in English and in Italian. These will be shown in section 5.3.

5. Findings

5.1 *The word ‘hashtag’*

An investigation into the entries of the most authoritative dictionaries of both English and Italian has revealed the presence of the entry ‘hashtag’ beyond the specific environment of Twitter glossaries, although there is no full agreement on its definition, especially in English, as we can see from the following quotations from some English dictionaries:

³ The Interface automatically retrieved only tweets containing hashtags and being posted from the selected geographical areas.

“A hashtag is any **word**⁴ or **phrase** immediately preceded by the # symbol. When you click on a hashtag, you'll see other Tweets containing the same keyword or topic” (The Twitter Glossary);⁵

“The # **symbol**, called a hashtag, is used to mark keywords or topics in a Tweet. It was created organically by Twitter users as a way to categorize messages” (Twittonary, the Twitter Dictionary);⁶

“The **symbol** # on a phone or computer keyboard, used on Twitter for describing the general subject of a tweet (= message)” (Cambridge Dictionaries online);⁷

“The **word** or **phrase** preceded by a hash sign (#), used on social media sites as Twitter to identify messages on a specific topic”. Also, “the hash **sign** (#)” (Oxford Dictionaries online);⁸

“(on the Twitter website) A **word** or **phrase** preceded by a hash mark, used to denote the topic of a post” (Collins Dictionaries online);⁹

“A **word** with the symbol # in front of it, used especially in social media and microblogging to identify or search for subjects of interest”. Also, “SPOKEN: used to make a **humorous point** when speaking by pretending that what you are saying is a hashtag” (Macmillan Dictionary online).¹⁰

As for Italian dictionaries, the definitions seem to converge more unanimously towards ‘parola o frase’, where the meaning of the Italian word ‘frase’ specifically corresponds to the English words ‘clause’ (or ‘sentence’).¹¹ The following quotations provide some examples of this:

“s. m. inv. In alcuni motori di ricerca e, in particolare, in siti di microblogging, **parola o frase (composta da più parole scritte unite)**, preceduta dal simbolo cancelletto (#), che serve per etichettare e rintracciare soggetti di interesse” (Treccani Vocabolario Online);¹² [*noun, masculine, invariant. In some search engines, and in particular on microblogging sites, word or ‘frase’¹³ (made up of more words spelt without whitespace characters in between), preceded by the symbol #, which serves as a label enabling the retrieval of topics of interest*] (my translation);

⁴ All bold characters and translations from Italian into English which are present in the text are mine.

⁵ <https://support.twitter.com/entries/166337-the-twitter-glossary>

⁶ twittonary.com

⁷ dictionary.cambridge.org

⁸ www.oxforddictionaries.com

⁹ www.collinsdictionary.com

¹⁰ www.macmillandictionary.com

¹¹ A question here is whether to include the concept of ‘phrase’ when translating ‘frase’ into English. Italian actually has its own term for ‘phrase’ (‘sintagma’), and the English equivalent for ‘frase’ is ‘clause’. Since in the present study the hashtag symbol is more often followed by a word or a phrase rather than a clause, I decided not to translate the Italian word ‘frase’.

¹² www.treccani.it/vocabolario

¹³ For the translation of ‘frase’ see note 11. The word is deliberately left in Italian.

“n. m. o f. invar. (Internet) in alcuni social network, **parola o frase** preceduta dal simbolo cancelletto (#), che permette di contrassegnare i messaggi con una parola chiave utile a classificarli, rendendoli facilmente reperibili agli utenti interessati all’argomento” (Dizionario Garzanti Online).¹⁴ [*noun, masculine or feminine, invariant. (Internet) in some social networks, word or ‘frase’¹⁵ preceded by the symbol # which classifies messages by means of a keyword that makes them easily retrievable for users interested in the topic*].

In Italy, the word ‘hashtag’ has also been recognized by the Accademia della Crusca (a leading institution in the field of research on the Italian language) on whose site, in the section “New Words”, one can read:

parola (o **sequenza continua di parole**) preceduta dal simbolo #, usata nell'ambito dei *social network* per categorizzare e rendere ricercabili contenuti correlati. [...] La parola all'inizio indicava solamente il simbolo del cancelletto impiegato per creare chiavi di ricerca. [...] Nell'uso comune, tuttavia, il significato si è esteso a tutta la sequenza, che nella terminologia ufficiale è chiamata **hashtagged word**.¹⁶ [*word (or any unbroken sequence of words) preceded by the symbol #, used in the context of social networks as a means to categorize correlated topics and make them searchable. (...) Initially, the word referred only to the hash symbol which was used to create search keys. (...) In the vernacular, however, the meaning has extended to the whole sequence, which in the official jargon is called hashtagged word*].

Moreover, in January 2013 ‘hashtag’ was voted as the word of the year 2012 by the American Dialect Society, on whose website one can find the following definition: “*Hashtag* refers to the **practice** used on Twitter for marking topics or making commentary by means of a hash symbol (#) followed by a word or phrase.”¹⁷

Based on what illustrated above, there seems to be no doubt that the word ‘hashtag’ has definitely become part of both the English and the Italian vocabulary, despite the apparent slight discrepancy in its definition.

5.2 Hashtaging and its ‘products’

In this paper, I argue that *hashtaging* could be looked at as a new morphological mechanism producing items, called hashtags, whose linguistic nature may be difficult to identify and relate to any traditional part of speech. In order to try and shed light on the way this user-initiated mechanism possibly operates, I have considered all the hashtags in my corpus, from the ones included in the samples collected via Twitter, to those retrieved in the offline world.

¹⁴ www.garzantilinguistica.it

¹⁵ See notes 11 and 13.

¹⁶ www.accademiadellacrusca.it/it/lingua-italiana/parole-nuove

¹⁷ www.americandialect.org

5.2.1 Hashtagging online

Taxonomies

For the purpose of my investigation, I analyzed the 2561 hashtags contained in the 10000 tweets included in the Twitter part of my corpus, and I tried to make a taxonomy of the type of hashtags produced in both the English and the Italian components. Such (tentative) taxonomy is based on what follows the # symbol, for example acronyms, abbreviations, combinations of letters and numerical digits, and of course words and phrases. As far as words are concerned, I decided to include in my taxonomies hashtags made up of the # symbol + 1, 2, 3, 4 and 5 or more words. I decided to consider the # + 5 or more words type because hashtags with more than 5 words are definitely less frequent in my corpus (for example, in the English component the longest hashtag recorded – *#LifesIsAChallengeNdItsUpToUsToPlayAndWin* – is formed by 13 words and it is the only occurrence; in Italian, the longest hashtag in my corpus – *#nelsognodiunestatechevorreipotessenonfiniremai* – includes 11 words and, again, this length occurs only once). Moreover, it seemed more important to me to highlight the fact that there is variation in the number of words a hashtag may be made up of, rather than the actual number of words in it.

Tables 2 and 3 show the taxonomies concerning respectively the English and the Italian hashtags included in my Twitter corpus:

Table 2 *Taxonomy of English hashtags*

Type of hashtag	Example
# + acronym/abbreviation	<i>#ootd</i>
# + 1 word	<i>#marathon</i>
# + 2 words	<i>#prettyplace</i>
# + 3 words	<i>#ThingsNobodySays</i>
# + 4 words	<i>#fromwhereistand</i>
# + 5 or more words	<i>#IAmElyarsBillionthGirl</i>
# + letters and numbers	<i>#b2bhour</i>
# + ??	<i>#duhDumduhDumDuhDumDuhDumDuhmdduhm</i>

Table 3 *Taxonomy of Italian hashtags*

Type of hashtag	Example
# + acronym/abbreviation	<i>#sbam</i>
# + 1 word	<i>#cultura</i>
# + 2 words	<i>#Tortosalate</i>

# + 3 words	#riderefabene
# + 4 words	#NelTelefilmCheVorrei
# + 5 or more words	#èstataunagiornatadura
# + letters and numbers	#SS3
# + ??	#sboccinlikemotomorrow
# + blends	#autunnestate
# + code-switching	#milanobybye
# + swearword	#FerragostoDiMerda
# + dialectal expression	#gnapossofa
# + idiomatic expression	#gallinavecchiafabuonbrodo
# + reduplication	#maimai

5.2.2 Analysis of the hashtag types and of their pragmatic exploitation

+ acronym/abbreviation and # + letters and numbers

As we can see from Table 1 and Table 2 above, in both components there are # symbols followed by acronyms/abbreviations and by a mixture of letters and numbers. These two types are far more frequent in English. Their main characteristic is that they are not always semantically transparent (e.g. #pm; #sbam). Sometimes, however, they refer to easily recognizable shows (e.g. #BGT, Britain's Got Talent), events (e.g. #vmlm14, Vergin Money London Marathon 2014), or institutions (e.g. #BHITA, British Healthcare Trade Association), but they may also be “idiomatic” (Romero et al. 2011) (e.g. #ff, a short form for #FollowFriday),¹⁸ suggesting a strong sense of community building (Zappavigna 2011).

+ 1 word

The # symbol followed by 1 word is the most frequent in both components. With only few exceptions (e.g. #golfhalism), the word is usually an unmarked content word (e.g. #lorry; #moto) in the singular form and does not seem to carry any particular semantic/pragmatic load. When an adjective occurs, this normally expresses a feeling (e.g. #proud) or a comment (e.g. #pointless), acting as a meta-comment on the whole message. Many times, separate one-word hashtags are created for expressions where the two words appear to be semantically related with reference to the topic of the tweet (e.g. #gorgeous #food instead of #gorgeousfood). Names (not necessarily capitalized) of either a place (e.g. #Milano) or a person (e.g. #Renzi) are widely used in Italian, with more semantic/pragmatic load than the unmarked lexical items. Interestingly, in the Italian component, many of the items forming

¹⁸ A Twitter Trend created back in 2009 that has since then become a customary Friday activity. Each Friday you recommend Twitter profiles that you enjoy and appreciate to all your followers. Your #FF recommendations will encourage others to check out those profiles, generating more followers for them.

the # +1 word type are English words (e.g. *#brain*), even though an Italian equivalent would be available. This seems to be related to the strong pragmatic function of the use of Anglicisms in Italian, which is becoming increasingly trendy also offline. Spelling mistakes do not seem to affect the searchability of the hashtag (e.g. *#typogrpahy*), and one may wonder whether or not they are meant to make the hashtag itself more ‘appealing’ (e.g. *#gombloddo*).

+ 2 words

This type of hashtag is usually a noun phrase of the type attributive adjective+noun (e.g. *#shortfilm*; *#azionarioattivo*) or noun adjunct+noun (e.g. *#agegap*; *#videocollaborazione*). Possessive noun phrases used as determiners are also present in the English component (e.g. *#mensfashion*). As with the # + 1 word type, occurrences of English noun-phrases in the Italian tweets have been recorded (e.g. *#laughterlab*; *#personalstylist*, *#behappy*). The initial letter of each word is not generally capitalized to make the two words more easily recognizable, especially in Italian, (e.g. *#numeriprimi*). Two-word hashtags seem to be more pragmatically exploited than one-word ones. Some of them are used as greetings or best wishes expressions (e.g. *#GoodFriday*; *#buonadomenica*), to complain (e.g. *#vitaingrata*), to recommend (e.g. *#liveit*), to encourage (e.g. *#forzaAzzurri*), to express emotions/feelings (e.g. *#lovinglife*; *#chetristezza*), or to comment (e.g. *#bellaserata*; *#summertimesadeness*).

+ 3 word

Three-word hashtags are more varied in form. In English, they may be verb phrases (e.g. *#studyinlondon*), noun phrases (e.g. *#YouAndI*), prepositional phrases (e.g. *#withmylion*), but also clauses (e.g. *#godsnnotdead*). In Italian, there are different patterns of three-word hashtags, but most frequently they are noun phrases (e.g. *#yogadellarisata*). The glued words are rarely highlighted through capitalization of the first letter, especially in Italian (e.g. *#italianibravimorti*). Many three-word hashtags are used as slogans (e.g. *#supportindiefilms*; *#digitaliperescere*), or to make suggestions (e.g. *#justdoit*; *#riderefabene*). Some also express emotions/feelings (e.g. *#lovemydogs*) or comment on something (e.g. *#VeryWeirdDay*). In the Italian component, some three-word hashtags are used as exclamatory expressions (e.g. *#evvivaaisogni*) – at times sarcastically (e.g. *#certocomeno*) – warnings (e.g. *#questivoglionofregaci*), or emotions, and very often they are used as a code-switching from Italian into in English within the Italian tweet (e.g. *Tortillas, sangria & more...prima del pranzo...Non sono ancora arrivato al tavolo successivo #Iwannadie*).

+ 4 words

Four-word hashtags are more frequently entire clauses (e.g. *#WeAreTheBest*; *#sognounmondocosi*). However, both in the English and in the Italian components I have found examples of noun phrases (e.g. *#quoteoftheday*) or prepositional phrases (e.g. *#damammaamamma*). Capitalization to mark word borders are used more than with other types of hashtags, especially in English (e.g. *#KindlyFollowBack_Thanks*), but less frequently in Italian, which makes the reading quite difficult at times (e.g. *#sabbiaavulcanicache passione*). Pragmatically speaking, this hashtag type often serves as a comment or a suggestion (e.g. *#Make_ _ NotWar*; *#comediretiamo*). In general, four-word hashtags seem to be less catchy than shorter ones, with only some exceptions (e.g. *#KindlyFollowBack_Thanks*; *#Make_ _ NotWar*). Their specific pragmatic load appears lower, but they can still be used to comment (e.g. *#toomuchfaketan*).

+ 5 or more words

Long hashtags in the Italian component may be clauses (e.g. *#melhadettounNobel*), but also noun phrases (e.g. *#supereroicontrolamunicipale*). Once again, capitalization to signal word borders is generally neglected in Italian. Instead, it is present in the English hashtags, which, like in Italian, are often an entire clause (e.g. *#MentionPeopleYouReallyLove*). In this category, examples of ‘aphorisms’ e.g. (*#LifeIsACallangeNdItsUpToUsToPlayAndWin*), comments (e.g. *#itsnotthatfunnythough*) and also pragmatically strong exclamatory expressions (e.g. *#chelevacanze2014abbianoinizio*; *#chesenonbestemmioguarda*) have been recorded.

+ ??

Anything could belong here (especially interjections, onomatopoeic expressions, non-linguistic sounds, both in English (e.g. *#aaaaaaaaaaaaaarghhh*; *#wiiiiiiiiiiiiiiiiiiiiii*; *#duhDumduhDumduhDumduhDumduhmdduhm*; *#MattToIMit*), and in Italian (e.g. *#uff*; *#uscitofuoro*). I have also found bizarre neologisms (e.g. *#lindieanata*) and occurrences of misspellings (e.g. *#tipogarphy*; *#uncincertoalgiorno*) which sometimes generate confusion on the actual elements making up the hashtag (e.g. *#aspettirenzie*). The illocutionary force is particularly high with interjections and onomatopoeic items like *#aaaaaaaaaaaaaarghhh* and *#uff*, as they convey the tone of the message (angry, bored, etc.).

Interestingly, in this section I also included hashtagged grammar word (e.g. *#most*; *#me*; *#e*). These hashtags do not seem to serve as a categorization method. Rather, the *#*+grammar word type suggests the idea that on Twitter any word can be hashtagged and generate a link. A power that grammar words do not have offline.

+ blends, code-switching, swearwords, dialectal expressions, idioms and reduplication

These hashtags are quite frequent in the Italian component (e.g. *#selfistiskyarte*; *#votnostress*; *#coglioniututti*; *#stucorenubulassamai*; *#giuringiuretta*; *#sisi*), while no occurrences of this type of hashtags have been detected in the English one. We may regard them as representative of a high degree of creativity displayed by hashtagging in Italian.

To sum up, what the different types of English and Italian hashtags show is that hashtagging seems to be more productive and to display a higher level of creativity in the Italian component, with the inclusion of non-standard varieties of the language. Moreover, my Italian hashtags tend to be slightly longer than the English ones (although hardly ever exceeding 5 or 6 words) and to make use of acronyms and abbreviations less frequently. Word borders are signaled through capitalization far more in the English hashtags than in the Italian ones. Finally, the tweets in Italian often include hashtags in English, reflecting the widespread use of Anglicisms in the Italian language.

5.3 Hashtagging offline

Besides selecting examples of hashtags from the online world (specifically, from Twitter), I also collected examples of hashtags making their way out of the (social network) online world, and *physically* appearing in the offline environment. Some of these examples are shown in Figures 1–28 hereafter.



Figure 1 *Hashtags in the Press. The cover of the December 31, 2012 issue of Newsweek, the last United States print issue of the magazine (source: huffingtonpost.com)*



Figure 2 *Hashtags in sport. London 2012 Olympics. The Twitter hashtag #SAVETHE SURPRISE appeared on giant screens around the stadium (source: skipedia.co.uk)*



Figures 3 Hashtags used within the so called 'Hashtag Activism' (Protests outside the Nigeria Consulate in Johannesburg (source: theguardian.com)



Figure 4 Hashtags to support protests. Enbridge opponents in Vancouver (source: [Vancouver Observer Facebook](https://www.facebook.com/VancouverObserver))



Figure 5 Hashtags in social campaigns. UN Security Council pledges support for the 'Children, Not Soldiers' Campaign (source: childrenandarmedconflict.un.org)



Figure 6 *Hashtags in sport. Southampton displays its Twitter hashtag on seats at its stadium (source: ITV News)*



Figure 7 *Hashtags on shop windows (my photograph)*



Figure 8 *Hashtags in political comics (source: Caglecartoons.com)*



Figure 9 *Hashtag gifts (source: wheretoget.it)*



Figure 10 *Hashtags in advertisements (my photograph)*



Figure 11 *Hashtags in political campaigns. European Elections campaign of the Italian Partito Democratico (source: leuromanoncadeldalcielo.wordpress.com)*



Figure 12 *Hashtags in political campaigns. Opponent political parties using the same hashtag (source: ilfattoquotidiano.it)*



Figure 13 *Hashtags in political campaigns (extramuralactivity.com)*



Figure 14 *Hashtags used in 'Hashtag Activism'. Protesters against hidden corruption in the Church (source: estense.com)*



Figure 15 *Hashtags against drug addiction. Posters all around the streets of Rome (source: actroma.it)*



Figure 16 *Hashtag for the promotion of brands (my photograph)*



Figure 17 Hashtags for the promotion of brands (my photograph)



Figure 18 Hashtags in sport (source: ilgiornaleditalia.org)



Figures 19 *Hashtags in sport* (source: giornalettismo.com)



Figure 20 *Hashtags to promote professionalization. A campaign by young Italian Physicians* (source: catania.sisms.org)



Figure 21 *Hashtags used to emphasize specific words* (source: socialcasehistoryforum.com)



Figures 22 and 23 *Hashtags in support of research and instruction* (my photograph)

HASHTAG
#LAEFFEPOP

Figure 24 *'Hashtag' used as a proper name for a TV program* (source: Italian TV Channel LaEffe)



Figure 25 Hashtags on TV screens (source: flickr.com, paz.ca)



Figure 26 'Hashtag' as the title of a book (my photograph)



Figure 27 Hashtags on food (source: flicker.com, Michael Coté)

As the examples above show, hashtags are present offline as well. There, the categorization function has no reason to exist, and one may wonder for what purpose they are created. The few examples displayed in this section would lead us to think that the main aim of using hashtags outside the world of the web is that of emphasizing the message (be it an advertisement, a political slogan, the promotion of a social initiative etc.) by highlighting it through hashtagged words. The message takes a prepackaged and condensed form that one can read as a whole, the preceding # symbol functioning as a pre-positioned exclamation mark, possibly aiming at producing catchy formulations.

The examples of hashtags I have selected to represent their use in the *real world* are never too long (e.g. *#freshpolitics*). Indeed, the longest one in my corpus is formed by six words (*#leuromanoncadedalcielo*). Sometimes the hashtag is printed (*#SAVETHESURPRISE*), other times it is hand-written (e.g. *#PapaAscoltaErik*), showing an appropriation of the symbol and of the relevant morphological practice by language users outside the Web, also in everyday contexts of writing (e.g. *#ComprateLaBari*). In these particular cases, it seems that the typographic conventions recommended by Twitter experts (that is, within the digital setting where the hashtag was born) are largely disregarded, with whitespace characters being added and words being even written on different lines (as in the latter example).

As within the online world, it seems that Italian hashtags tend to be slightly more creative also offline, especially in terms of characters combination, or typographical shape (e.g. *#svoltiAMOLA SANITÀ*), as if to create more emphasis within something which is in itself already emphatic.

5.4 The spoken hashtag

Hashtags have also made their way into everyday spoken conversation, with people (especially teens) actually saying the word ‘hashtag’ out loud before speaking a word or phrase, as a verbal exclamation:

(12) *I’m hashtag confused!*

(13) *Hashtag are you kidding me?*

(14) *Hashtag, YOLO!*

There are also people who flash one another the hashtag sign — crossing their index and middle finger of one hand over the same two fingers of their other hand to create a physical hashtag, as Figure 29 shows.



Figure 28 A hashtag hand sign (source: socialmedium.co.uk)

This is confirmed by the number of articles,¹⁹ blogs²⁰ and videos²¹ on the topic that have been appearing over the last years, most of which look sarcastically at the new phenomenon. Although I would argue that this is just a fad, a sign of the *Zeitgeist*, what is interesting however is that it somehow confirms the trend of writing-the-way-we-talk and talking-the-way-we-write which is being brought about by Computer-Mediated-Communication in general (e.g. Baron 2008).

6. Conclusions

My investigation seems to confirm that ‘hashtag’ is definitely a new word in its own right which has entered both the English and the Italian vocabulary, in the written as well as in the spoken modes.

It further suggests that the hashtag can also be regarded as the product of a new morphological process, that is, *hashtagging*, a mechanism that can generate an almost unlimited variety of forms by lumping words together, and thus generating new ‘linguistic items’. If we consider the hashtag as the product of such mechanism, a question arises about the nature of these ‘linguistic items’, which are both words and yet not words. Indeed, a classification of hashtags according to the traditional categorization of parts of speech seems to be a critical issue. Hashtags may take such a variety of forms that a part-of-speech tagging based on natural language would hardly be possible. However, considering the spread of hashtagging practices also in the ‘real world’, it may not seem unlikely to expect the hashtag itself to become a new part of speech altogether in the natural language of the so called digital natives, and the # symbol to add to the list of affixes already available in the morphology of both English and Italian. To put it in Niola’s terms, hashtags are “iperparole” [*augmented words*], the whole of which “trasforma il mondo in un ipertesto, ovvero gli dà un nuovo significato” [*turns the world into a hypertext, that is, it gives the world a new meaning*] (Niola 2014: 9).

Similarly, *hashtagging* does not correspond to any of the already existing morphological processes (like compounding, blending, agglutination). At the moment, there does not seem to be a recognized purely linguistic term for these glued-together clusters, neither for the mechanism that produces them. According to my findings, *hashtagging* seems to be a morphological process allowing an extremely high degree of freedom and creativity, with only minor technological constraints on the Twitter platform (mainly linked to the limited number of characters a tweet may be made up of). In my corpus, such freedom and creativity can be recorded both in English and in Italian, with a slightly higher level of creativity in the latter, and an increasing tendency in both languages for hashtags to appear more and more frequently in contexts offline, where *hashtagging* seems to have become a very productive ‘(socio)morphological’ practice.

As for the pragmatic exploitation of hashtags, the findings of this study suggest that hashtags seem to serve more than the dual purpose of both referring to a topic and creating communities of people interested in that topic. As seen, hashtags have moved far beyond. In addition to serving as metadata (what the tweet is about), they are now used in a number of

¹⁹ e.g. www.theguardian.com/technology/shortcuts/2012/aug/01/how-to-say-hashtag-fingers

²⁰ e.g. www.cnet.com/news/the-spoken-hashtag-must-die-heres-how/

²¹ e.g. www.youtube.com/watch?v=57dzaMaouXA

ways and for a number of purposes. As ‘linguistic products’ of the mechanism of *hashtagging*, hashtags are making language lexically richer in catchy phrases (e.g. #senzapaura), idiomatic expressions (e.g. #MusicMonday), and also formulas (e.g. #lol). Moreover, in their compact and pre-packaged format, ‘hashtagged utterances’ appear to have, both online and offline, a stronger illocutionary force than their untagged counterparts in natural language, sometimes making the rest of the text (if any) within which they appear as totally marginal in conveying the meaning of the message. Indeed, ‘hashtagged utterances’ seem to be more effective when performing speech acts like, for example, commenting on something (e.g. #freshpolitics), greeting (e.g. #HappyBirthdayCapaldo), informing (e.g. #brazilwins), warning (e.g. #ignaziostaisereno), ordering (e.g. #boicottabeppegrillo), complaining (e.g. #Ihate), expressing solidarity (e.g. #weareallmonkeyes) and, in general, they seem to represent a new means through which language functions. Similarly, it appears that the # symbol provides the tagged word or phrase with more emphasis (e.g. *raccontare storie di #successo*).

Finally, it also appears that *hashtagging* is somehow enriching grammar words with a new dignity, in that they can be tagged just like content words (e.g. #and), thus not only becoming potential topic markers and community-building facilitators, but also playing their part in the construction of meaning. In this respect, the hashtag seems to be the product of a morphological mechanism that is providing hashtagged morphemes with a new role.

On the whole, it would appear that the reasons why people create hashtags by putting the # symbol before one or more words seem to be, both online and offline, more of a social and pragmatic nature, possibly simply a way of following one of the latest social-media related fashions. A classification of hashtags based on purely linguistic criteria may be too ambitious at this stage, if possible at all, due to the extremely flexible – and basically unpredictable – way with which they are created, producing a practically unlimited variability of forms.

Yet, in the age of Computer-Mediated Communication, the product of this fashionable social practice is worth a merely linguistic investigation. The presence of hashtags in the written and spoken modes of certain offline world contexts on the one hand, and the uncontrollable spread of new ways of communicating brought about by new digital media on the other hand, might not only change the notion of what it means to write and to speak – possibly leading to a more and more condensed style in both modes – but also potentially imply the redefinition of traditional linguistic concepts such as ‘word’ and ‘part of speech’. This, in turn, is likely to change the way of carrying out parsing and syntactic analysis, by first having to agree on what the relevant unit for syntactic analysis is.

References

- Baron, Naomi. 2008. *Always on: Language in an Online and Mobile World*. Oxford: Oxford University Press.
- Cunha, Evandro & Magno, Gabriel & Comarela, Giovanni & Almeida, Virgilio & Goncalves, Marcos A. & Benvenuto, Fabrício. 2011. Analyzing the dynamic Evolution of Hashtags on Twitter: A Language-Based Approach. In *Proceedings of the Workshop on Language in Social Media (LSM 2011)*, 58–65. Portland, Oregon. 23 June 2011.

- Dickinson, Paul. 2013. 'B/w U & me': The Functions of Formulaic Language in Interactional Discourse on Twitter. *The Linguistic Journal* 7(1). July 2013. 7–38.
- Easley, David & Kleinberg, Jon. 2010. *Networks, crowds, and markets: Reasoning about a highly connected world*. Cambridge: Cambridge University Press.
- Hill, Shwandra & Benton, Adrian. 2012. Social TV: Linking TV content to Buzz and Sales. *International Conference on Information Systems*. December 2012.
- Kriefalusi, Elizabeth. 2013. *The Twitter Hashtag: What Is It and How Do You Use It?* (<http://www.techforluddites.com>) (Accessed 01-08-2014).
- Landry, R. and Bourhis, R. (1997). Linguistic Landscape and Ethnolinguistic Vitality. An Empirical Study. *Journal of Language and Psychology* 6(1), pp. 23-49.
- Niola, Marino. 2014. *HASHTAG. Cronache di un paese connesso*. Milano: Bompiani.
- Potts, Liza & Seitzinger, Joyce & Jones, Dave & Harrison, Angela. 2011. Tweeting disasters: Hashtag Constructions and Collisions. *Proceedings SIGDOC '11, Proceedings of the 29th ACM International Conference on Design of Communication*, 235–240. ACM, New York, USA.
- Romero, Daniel, M. & Meeder, Brendan & Kleinberg, Jon. 2011. Differences in the Mechanics of Information Diffusion Across Topics: Idioms, Political Hashtags, and Complex Contagion on Twitter. *International World Wide Web Conference (WWW 2011)*, March 28–April 1, 2011. Hyderabad, India.
- Sagolla, Dom. 2009. *140 Characters: A Style Guide for the Short Form*. Hoboken: John Wiley & Son.
- Sigurd, Bent & Eeg-Olofsson, Mats & Van De Weijer, Joost. 2004. Word length, sentence length and frequency: Zipf's law revisited. *Studia Linguistica* 58(1). 37–52. Oxford: Blackwell.
- Vander Wal, Thomas. 2007. "Folksonomy". *Vanderwal.net*. (<http://vanderwal.net/folksonomy.html>) (Accessed 01-08-2014).
- Zappavign, Michele. 2011. Ambient Affiliation: A Linguistic Perspective on Twitter. *New Media & Society*, 13(5). 788–806. (<http://nms.sagepub.com/content/early/2011/05/26/14611444810385097.full.pdf+html>) (Accessed 01-08-2014).

Paola-Maria Caleffi
University of Verona
Italy
paolamaria.caleffi@univr.it

In SKASE Journal of Theoretical Linguistics [online]. 2015, vol. 12, no.2 [cit. 2014-06-25]. Available on web page http://www.skase.sk/Volumes/JTL28/pdf_doc/05.pdf ISSN 1336- 782X.