‘LOL’, ‘OMG’ and Other Acronyms and Abbreviations

A study in the creation of initialisms

Ida Lundell
Abstract

Marchand (1969) claims that abbreviations and acronyms, which are also known as ‘initialisms’, are used to create “names of new scientific discoveries, trade-names, names of organizations, new foundations or offices, but occasionally, and chiefly in American English, personal and geographical names are also coined in this way” (Marchand, 1969: 452). However, initialisms that originate from netspeak, such as ‘LOL’, are different from the initialisms Marchand (1969) describes. These initialisms are not names of organizations or scientific discoveries; rather, they describe everyday things or phases. This kind of initialism is a new phenomenon that seemingly did not exist before the Internet, and the aim of this essay is thus to examine whether Internet has given us increased opportunities for this type of word formation. A corpus of informal English was created for this study and ten initialisms were extracted from that corpus. These initialisms were then examined in the Corpus of Contemporary American English. The data from the COCA showed that people do form and use this kind of initialism, albeit they do it very infrequently. Furthermore, the corpus data showed that these items are used in all different genres that exist in the COCA, except for academic writing. The data found in this study indicates that people seem create a new kind of initialism, and that people therefore have new opportunities to use this kind of word formation. However, the data is too sparse to draw any definite conclusions from it.

Keywords
Morphology, acronyms, abbreviations, initialisms, slang, corpus linguistics, the Internet, netspeak.
Contents

1. Introduction.................................................................................................................. 4

2. Background .................................................................................................................. 5
   2.1 Theories on general morphology and initialisms................................. 5
   2.2 Theories on slang in general................................................................. 6
   2.3 The Internet as a means of communication and its role in language change 7
   2.4 Netspeak and initialisms................................................................. 8
   2.5 Previous studies ........................................................................ 9

3. Method......................................................................................................................... 12
   3.1 Corpus of Informal English................................................................. 12
   3.2 Corpus of Contemporary American English........................................ 13

4. Results.......................................................................................................................... 14
   4.1 Distribution across different genres in the COCA................................. 16
   4.2 Distribution over time in the COCA......................................................... 18

5. Discussion...................................................................................................................... 21
   5.1 Discussion of results ........................................................................ 21
   5.2 Limitations............................................................................................. 23
   5.3 Future research...................................................................................... 23
   5.4 Conclusion .............................................................................................. 24

References ......................................................................................................................... 26
1. Introduction

Ever since the Internet was introduced as a means of communication, researchers have been interested in how it affects the way people use language. The way language is used online generally differs from both the way it is used in more formal writing, and also the way it is used in speech. This ‘online’ language is based on features found both in written and spoken language, and it crosses over the boundaries that usually divide spoken and written language (Crystal, 2001). Language online tends to contain significantly more slang than written language in other contexts (Coleman, 2012). Slang in itself is not something brand new. However, some features of the language used in instant messaging conversations are completely new. The language online, which Crystal (2001) calls ‘netspeak’, contains a new kind of initialism, which is the general term for acronyms, abbreviations, and multi-word phrases that are compressed into a single word.

Marchand (1969) describes initialisms as a type of word formation used mainly to form names of organizations and also names of scientific discoveries (Marchand, 1969: 452). However, the new initialisms that are used online are not abbreviations or acronyms that function as names of organizations or names of scientific discoveries; initialisms are now being formed through taking clauses that deal with more general, everyday things, and turning these into abbreviations and acronyms. These new initialisms are different from the ones Marchand (1969) describes, and people therefore appear to have begun creating a completely new kind of initialisms. Previous studies have included this kind of initialism into the larger concept of ‘netspeak’ or ‘Internet speech’. Furthermore, previous studies in this field have mostly been concerned with who uses these initialisms, and also whether netspeak has influenced adolescents’ literacy and ability to tell slang words and expressions apart from more formal ones. However, no studies have looked at these new initialisms as an isolated feature, and no studies have looked at them outside of the Internet. If this kind of initialism is only found online, then it might just be a part of netspeak. However, if people use this kind of initialism outside the Internet, it might indicate that a change in general in how people form and use initialisms has taken place.

The purpose of this essay is thus to examine whether Internet has given us increased opportunities for this type of word formation.
2. Background

2.1 Theories on general morphology and initialisms

Marchand (1969) calls the process of forming acronyms and abbreviations ‘word-manufacturing’ and describe these initialisms as “more or less arbitrary parts of words that have been welded into an artificial new word” (Marchand, 1969: 452). As previously mentioned, this type of word formation is mostly used to create names of organizations and sometimes also scientific discoveries (Marchand, 1969: 452).

Bauer (1988) defines acronyms as “words formed from the initial letters in a name, title or phrase” (Bauer, 1988: 39). Furthermore, acronyms should be distinguished from abbreviations because acronyms are pronounced as new words, and not just a sequence of letters (Bauer, 1988).

Ljung (2003) describes acronyms and abbreviations as a “product of irregular word formation” (Ljung, 2003: 153). Ljung (2003) claims that there are two main kinds of word formations. One kind is called regular word formation, which concerns the process of adding inflections to words in order to create words that belong to another grammatical category, for example through adding a suffix to a noun and thereby creating an adverb or adjective; and the other kind is called irregular word formation, which creates new words and enriches the vocabulary.

Whereas Bauer (1988), only identified two types of initialisms, Ljung (2003) recognizes four different types. Ljung (2003) defines abbreviations as a “string of letters, most of which are spelled with capital letters” (Ljung, 2003: 157). Ljung (2003) furthermore lists ‘DNA’ as an example of abbreviations. The only difference between abbreviations and acronyms is that acronyms “have to be pronounceable” and that acronyms “can only contain letter sequences permitted in ordinary English words” (Ljung, 2003: 158). Furthermore, Ljung lists ‘NATO’ (North Atlantic Treaty Organization) as an example of an acronym. The third kind of initialisms according to Ljung’s theory comprises words such as ‘RAF’, which can be treated either as acronyms or abbreviations (Ljung, 2003: 157). The fourth kind of initialism contains blends of acronyms and abbreviations, where one part of the word is pronounced as an actual word and another is pronounced a letter at a time. Words such as ‘CDROM’ and ‘email’ are examples of this category (Ljung, 2003: 157). Abbreviations and acronyms are most often used to create names for different organizations, although they do sometimes refer to concepts.
such as ‘DNA’ or ‘laser’ (Ljung, 2003: 157).

Harley (2006) describes abbreviations and acronyms as “behaving like single words but they are actually several words, as far as meaning is concerned” (Harley, 2006: 8). Initialisms are a form of extreme clipping, and while this kind of word creation has been used since the 1800s, it did not become a popular form of creating words until the second half of the 20th century (Harley, 2006: 96). The process of creating initialisms was born “as a natural outgrowth of a proliferation of bureaucratic institutions named with long, unwieldy compounds and phrases (Harley, 2006: 97). Harley (2006) mentions only three different categories of initialisms; abbreviations, acronyms, and words that originally were initialisms that now have undergone a change and are nowadays more like blends. Humvee, originally HMMWV; and jeep, originally G.P, for General Purpose, are examples of the third category (Harley, 2006: 97).

Denning, Kessler, and Leber (1995) point out that initialisms are created because people want to “reduce potentially long names to something manageable” (Denning et al, 1995: 59). In speech, initialisms will often become more frequent than the full form, and will sometimes even replace the full form completely (Denning et al, 1995: 59). According to Denning et al (1995), this occurs both to abbreviations and acronyms.

Of all authors mentioned above, only Ljung (2003) and Harley (2006) acknowledge that there seems to be a new kind of initialisms currently in the process of becoming popularized. Ljung (2003) describes them as “a smaller group of abbreviations and acronyms that stand for more common phrases…examples include ‘ASAP’ for ‘as soon as possible’” (Ljung, 2003: 158). Harley describes these as “initialisms that have come into use primarily in electronic communications of one kind or another; email, instant messaging and chatrooms” (Harley: 2006, 97). Furthermore, “people who are typing have a particular impetus to economize on frequently used phrases, or phrases inserted to maximize communicative flow rather than convey actual information” (Harley: 2006: 97). Harley (2006) acknowledges that these new initialisms are a product of the Internet, while Ljung (2003) does not. However, both these authors published after the introduction of the Internet; before that, this kind of initialism was not acknowledged by these scholars, which indicates that this kind of initialisms may indicate a new morphological process.

### 2.2 Theories on slang in general

The Oxford English Dictionary Online defines slang as “language of a highly colloquial type,
considered as below the level of standard educated speech, and consisting either of new words or of current words employed in some special sense”.

Slang changes rapidly, and this change is mostly caused by teenagers and young adults. Older people also use slang but they are generally not bothered with keeping their language fashionable, and thus opt for using words that already exist in their vocabularies (Andersson & Trudgill, 1992). According to Andersson and Trudgill, “when we get older and more secure in our own language, more is needed to change it. Linguistically we become more conservative as years pass by” (Andersson & Trudgill, 1992: 47).

As quite often happens with new inventions, slang has been frowned upon and disliked by some people, especially older speakers. There has long been a fear that slang will destroy children’s speech, and that they will be unable to distinguish slang expressions from neutral expressions (Andersson & Trudgill, 1992: 38). Furthermore, the notion of correct English is very “integrated” in speakers (Burridge, 2004: 158).

2.3 The Internet as a means of communication and its role in language change

The introduction of the Internet and computer technology has given birth to several completely new words and expressions. According to Crystal, “the rate with which [Internet users] have been coming up with new terms has no parallel in contemporary language use” (Crystal, 2001: 67). Since the Internet has given birth to many new expressions, there is no surprise that these expressions will be found in other domains as well (Dent, 2007: 3). Therefore, according to Dent, it would not be surprising to find slang that originated online in contexts outside the Internet. Coleman (2012) theorizes that slang in general is changing more quickly than ever before because of the widespread use of the Internet. When online, you are exposed to a large number of slang expressions, many of which may not have been encountered before. Internet users tend to use slang expressions before they are completely familiar with their meanings if they are exposed to them frequently enough. According to Coleman (2012), this could lead to a chain reaction between Internet users and a spike in the frequency with which an expression is used. When a word is used a lot, it automatically acquires a shorter lifespan because the high frequency with which it is used, makes the expression lose some of its meaning (Coleman, 2012).
Coleman (2012) credits role-playing gaming online as being the birthplace of netspeak. Multi-user dungeon games, commonly called ‘MUD’s, are games where several users play together. According to Coleman (2012), “because of the speed of play and the relatively limited range of likely topics in MUD’s, acronyms and abbreviations became commonplace among the users of these games” (Coleman, 2012: 270).

Crystal (2001) claims that it will be increasingly difficult to continue typing only using initialisms, as was the norm in the early days of texting, since it allows for too much ambiguity. This type of language that only uses initialisms will probably die out, or at least become less common (Crystal, 2001: 230).

The norms and conventions applied when writing online is vastly different from both written and spoken language. The Internet offers completely novel dimensions of stylistic variation (Crystal, 2004). Dent describes the phenomenon of Internet language as a ‘blurring of boundaries between spoken and written language” (Dent, 2007: 4). For example, the grammar is characterized by highly colloquial constructions and non-standard use (Crystal, 2001). Orthographic features have been affected, for example through the removal of capitalization and punctuation (Crystal, 2001). Netspeak “is better seen as written language that has been pulled some way in the direction of speech than as spoken language which has been written down” (Crystal, 2004: 79).

2.4 Netspeak and initialisms

Crystal (2001) claims that instant messaging has “motivated a whole new genre of abbreviated forms. The acronyms are no longer restricted to words or short phrases, but can be sentence-length” (Crystal, 2001: 86).

Coleman (2012) discusses the process of using initialisms to create slang expressions. Coleman (2012) notes ‘LOL’, which is an initialism created from the phrase ‘laughing out loud’, as being a good example of a slang initialism. It can either be an abbreviation or an acronym, depending on whether you choose to spell the word out or pronouncing it like words typically are pronounced (Coleman, 2012: 39). Coleman furthermore notes that several of these initialisms concern laughing. Apart from ‘LOL’, there are initialisms such as ‘ROFL, which stands for ‘rolling on floor laughing’; and ‘PMSL’, which stands for ‘pissing myself
laughing’ (Coleman, 2012: 95). Many of these initialisms are used online as a substitute for laughing, in an attempt to convey feelings in a way that resembles that of spoken language (Coleman, 2012: 95).

Kinsella (2010) claims that in some cases, these initialisms are not to be interpreted literally; the initialism is simply there to convey the tone of the utterance (Kinsella, 2010: 65). Netspeak has become a code which parents cannot decipher, which makes teenagers use it whenever they want to distinguish themselves from their parents and want to create an in-group feeling among themselves (Kinsella, 2010: 67).

Baron (2011) looked at previous studies that focused on netspeak, and concluded that teenagers in general seem to be able to distinguish between formal and informal writing, and that netspeak seemingly does not influence teenagers’ literacy (Baron, 2011: 122). However, it is probably premature to claim that netspeak has not influenced language in more formal situations (Baron, 2011: 123). Baron claims that “historically, there are ample cases in which language judged to be somehow improper works its way into more formal everyday speech and writing” (Baron, 2011: 123). However, Baron (2011) also notes that “there is little evidence that online and mobile devices will be radically reshaping offline language anytime soon” (Baron, 2011: 123). Given their relative highly frequency in netspeak, it would not be surprising if a few elements such as ‘lol’ and ‘brb’ eventually found their way into offline language as well (Baron, 2011: 123).

2.5 Previous studies

Varnhagen, McFall, Pugh, Routledge, Sumida-MacDonald, and Kwong (2009) carried out a study to find out whether there were any differences between how different types of informal language were used in IM conversations. This study found that the most common features of IM language included typing completely in lower case, followed by abbreviations such as ‘omg’, and acronyms such as ‘lol’ (Varnhagen et al, 2009).

Tagliamonte and Denis (2008) studied language in use in IM conversations among Canadian teenagers. These scholars noticed that there were extremely few IM abbreviations and acronyms in use, and that their test subjects were more prone to use both slang and lower case typing in their IM conversations than initialisms. These authors further noted that ‘LOL’ was “used by the participants in the flow of conversation as a signal of interlocutor
involvement, just as one might say mm-hm in the course of a conversation” (Tagliamonte and Denis, 2008: 11). ‘LOL’ was the most frequently occurring initialism, followed by ‘OMG’ and ‘BTW’ (Tagliamonte and Denis; 2008: 11). ‘LOL’ was most commonly used by younger teenagers, and the use of ‘LOL’ declined quickly as the teenagers got older (Tagliamonte and Denis, 2008: 13).

Similar to Tagliamonte and Denis (2008), Baron (2011) carried out a study on how ‘special language’ is used in IM conversations. ‘Special language’ refers to acronyms, abbreviations and emoticons (Baron, 2011: 120). Baron found that the use of abbreviations and acronyms was generally low. This study excluded initialisms that were not specific to the Internet; furthermore, this study also excluded shortenings that were common in ‘offline writing’ such as ‘hrs’ for ‘hours’ and ‘cuz’ for ‘because’. Baron created an IM corpus, containing 11,718 words, and extracted all initialisms present in that particular corpus. In total, there were 121 initialisms such as ‘cya’ for ‘see you’, and ‘lol’ for ‘laughing out loud’ (Baron, 2011: 120). Around one percent of all findings were netspeak items.

Kinsella (2010) studied how netspeak has affected spoken language. The study was carried out on campus at an Australian university. A questionnaire was sent out to 95 participants between the ages of 15 and 25. It showed that approximately 85% of the respondents used Internet slang in everyday spoken conversations (Kinsella, 2010: 69). Furthermore, about 75% reported that they have accidentally used netspeak in spoken conversations (Kinsella, 2010: 70). Kinsella concluded that netspeak appears to have found its way into spoken language, at least among this particular group of young people.

To sum it up, the previous studies that this investigation is based on, have included abbreviations and acronyms in the larger concept of netspeak, and consequently have not observed this new kind of initialisms as a separate phenomenon. Several studies on netspeak, e.g. Baron (2011) and Varnhagen et al (2009), have established that Internet users create and use this type of initialism online. Some studies, e.g. Kinsella (2010), have also looked at whether these initialisms are used in spoken language. Authors such as Coleman (2012) and Crystal (2001), whose works deal mainly with slang and the effects of the Internet on language, mention this new way of creating initialisms. Some works on general morphology, such as Ljung (2003) and Harley (2006), mention this kind of initialisms, albeit very briefly. However, authors such as Bauer (1988) and Marchand (1969) do not mention this way of creating initialisms at all, which suggests that this may be a relatively new linguistic process. Previous studies that mention these initialisms, have only talked about them in relation to the Internet, and thus may have missed the view of them as a more general process and an
indication of a change in how people create initialisms on a broader scale. If this way of creating initialisms is only found online, it might indicate that it is just a part of a particular group’s jargon, and not a change in general and outside of that group. However, if these types of initialisms are found in contexts outside the Internet, then it might be that there has been a change in how people form and use initialisms. The aim of this essay is thus to examine whether Internet has given us increased opportunities for this type of word formation.
3. Method

This study was carried out in three different steps. Step one included creating a corpus of informal English with data retrieved from a bulletin board chat room called Ohnotheydidnt, since corpus of informal English was readymade and available online. Step two consisted of sorting the items collected in step one according to the frequency with which they appeared in the Corpus of Informal English. This was done with the help of an online program called ‘Word Frequency Counter’. Thereafter, the ten most frequently occurring initialisms in the corpus data were selected for the actual analysis. Step three included running the ten initialisms chosen in step two through the Corpus of Contemporary American English.

3.1 Corpus of Informal English and the analysis of the informal data

In order to examine how initialisms are used online, a corpus of informal English was created for this study. It contains 31 000 words and the data was collected from a bulletin board site (ohnotheydidnt.livejournal.com) that allows for asynchronous chatting, referred to henceforth as ‘ONTD’. Members of the site discuss a wide variety of topics by leaving comments on posts posted on the site. Most posts on ONTD deal with a specific topic; the topic of the post thus affects what is said in the comments. The vast majority of posts on ONTD are connected to popular culture in at least one way, although on some occasions, the users are free to discuss anything they like. These less controlled posts are called ‘Friday Free For All’, and the data collected for the corpus used in this study were gathered from one of these posts. These posts were chosen specifically because of their lack of topic-specific discussions. When the users are free to discuss anything they like, the language they use tends to resemble the language previous studies (eg. Denis and Tagliamonte, 2008) have identified as occurring in instant messaging conversations.

The data collected from the corpus of informal English was analyzed with help of an online program called ‘Word Frequency Counter’, which analyzes the frequency of words that appear in a text. The items were then sorted into a list based on the frequency with which they appeared in the corpus. The results generated from the frequency count were examined and the ten most frequently occurring abbreviations and acronyms that were found in the Corpus of Informal English were chosen for further examination. These items are presented in the
As shown in Table 1, these initialisms made up 546 of the 31000 words in the corpus, or 1.72% of all the words in the corpus. These numbers are in general agreement with those of Tagliamonte and Denis (2008), and Baron (2011).

### 3.2 Corpus of Contemporary American English

The Corpus of Contemporary American English (henceforth referred to as ‘COCA’) was chosen for this study because it offers a large collection of contemporary data from a wide variety of sources and also a wide variety of genres; for example, different disciplines of writing and spoken material. This wide variety of sources allows for comparisons between different genres of writing, and also allows for comparisons between spoken and written language. Furthermore, while the COCA represents quite a big number of different genres, it
is still somewhat limited. A big portion of the text that is found in the corpus is fictional writing. Fiction may not be an actual reflection of language use, but rather reflects only what authors believe is language use.

When possible, the frequency expressed as per million words was chosen for examination rather than the actual number of words, since the distribution across different genres or time periods in the COCA might not be equal. Furthermore, items that generated less than twenty hits in the corpus were not included in this study. All results were then controlled to see whether the hits actually were related to the topic of this essay or not. As became apparent in the results generated from the COCA, the initialism ‘OMG’ can also denote a special type of cholesterol called ‘Omg cholesterol’, and these hits were subsequently left out of this study.
4. Results

Of the ten items chosen for this study, only seven of these items were used in the COCA. Out of those seven items, only five items were used frequently enough to be analyzed further. The frequencies of the initialisms are presented in the table below. Several of these initialisms have multiple meanings, and ‘number of relevant occurrences’ refers to the number of occurrences where the chosen initialisms have been used in a way which is relevant for this study.

Table 2: The number of occurrences of the chosen initialisms in the COCA

<table>
<thead>
<tr>
<th>Initialism</th>
<th>Number of occurrences in the COCA</th>
<th>Number of relevant occurrences in the COCA</th>
</tr>
</thead>
<tbody>
<tr>
<td>‘LOL’</td>
<td>82</td>
<td>76</td>
</tr>
<tr>
<td>‘OMG’</td>
<td>387</td>
<td>72</td>
</tr>
<tr>
<td>‘WTF’</td>
<td>35</td>
<td>31</td>
</tr>
<tr>
<td>‘BTW’</td>
<td>23</td>
<td>20</td>
</tr>
<tr>
<td>‘TBH’</td>
<td>11</td>
<td>0</td>
</tr>
<tr>
<td>‘IMHO’</td>
<td>7</td>
<td>7</td>
</tr>
<tr>
<td>‘IKR’</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>‘IRL’</td>
<td>198</td>
<td>7</td>
</tr>
<tr>
<td>‘OFC’</td>
<td>68</td>
<td>0</td>
</tr>
<tr>
<td>‘FYI’</td>
<td>382</td>
<td>372</td>
</tr>
</tbody>
</table>

As shown in Table 2, only three of these initialisms generated no hits at all, viz., ‘TBH’, ‘IKR’, and ‘OFC’. ‘IMHO’ and ‘IRL’ both occurred seven times each in the corpus but due to their very low frequency, they will not be analyzed in this study. Of the initialisms examined, only five fit the criteria of occurring more than twenty times in the corpus, and these were selected for further examination: ‘LOL’, ‘OMG’, ‘WTF’, ‘BTW’, and ‘FYI’. ‘FYI’ was the most commonly occurring initialism with 372 hits in the COCA. It occurred approximately two hundred times more than any other initialism. ‘LOL’ and ‘OMG’ generated similar frequencies, at 76 and 72 hits in the COCA. ‘WTF’ occurred slightly less often, at 31 hits in total. Out of the initialisms chosen for this study, ‘BTW’ was the least common.
4.1 Distribution across different genres in the COCA

The results from the ‘academic writing’ category in the COCA were left out of this study because there were no actual occurrences of these initialisms there. Whenever the selected items occurred in that category, they were always used in quotations. In other words, there were no new hits there, and no real examples of people using the initialisms. To facilitate the reading of the figures below, the results are split up into multiple figures.

Figure 1: the frequency of ‘LOL’ and ‘OMG’ per million spread across different genres

As can be seen in Figure 1 ‘LOL’ and ‘OMG’ have similar frequencies. ‘LOL’ appears to be used more evenly between genres, while ‘OMG’ is used very little in fiction and much more in speech. ‘LOL’ is most commonly used in newspapers, but there are no big differences between its frequencies in magazines, newspapers and spoken data. Both these items have quite low frequencies, with frequencies of 0.25 per million being the largest number.
As shown in Figure 2, ‘WTF’ and ‘BTW’ were all used less than ‘LOL’ and ‘OMG’. ‘WTF’ has its peak in magazines, with a frequency of 0.19 occurrences per million words, but, less than half as often in newspapers, and even more infrequently in spoken data and in fiction. The use of ‘BTW’ did not differ as greatly between genres as ‘WTF’. ‘BTW’ was most commonly used in fiction, followed by magazines. It was used less than ‘WTF’ in spoken data, and in newspapers as well.

Out of all items examined in this study, ‘FYI’ is the item with the highest frequency, as shown in Figure 3. It peaks at approximately 2.75 hits per million, which is much higher than any other item. It is used quite unevenly in different genres. It is used the most in magazines, and much less frequently in all other genres.
4.2 Distribution over time in the COCA

Figure 4: The frequency per million of ‘LOL’ over time

Figure 4 shows that the frequency of ‘LOL’ has generally been very low, although it did experience a sharp increase during the 2005-2009 period. It appeared for the first time during the 1990-1994 period but it was not used more frequently until much later.

Figure 5: The frequency per million of ‘OMG’ over time

As shown in Figure 5, ‘OMG’ was not used at all until the 2000-2004 period. It then experienced a peak which began during the 2005-2009 period and continued throughout the 2010-2012 period.
As shown in Figure 6, the way ‘WTF’ has been used over time resembles that of ‘OMG’. It was not used at all until the 1995-1999 period. Its usage increased slightly between the 2000-2004 and 2005-2009 time period. From there on, it has increased sharply, reaching a level of almost 0.5 occurrences per million words in the COCA.

Figure 7 reveals that ‘BTW’ has been used ever since the beginning of the corpus, although very sparsely. Its usage has increased quite evenly throughout the COCA. As of 2012, although its usage has increased, it is still used infrequently, as indicated by the low number of occurrences per million words.
‘FYI’, shown in Figure 8, like the other items in this study, has been increasing since the beginning of the COCA. However, it is the only item in this study that seems to be declining. Its usage increased rapidly between the 2000-2004 time periods, and then declined rapidly thereafter. However, compared to the other items in this study, it is still experiencing a relatively high frequency per million.
5. Discussion

5.1 Discussion of results

The results indicate that the items analyzed in this study appear to be new inventions in mainstream texts. The corpus data show that they did not exist until the mid-1990s, and most of them were not used more frequently until several years later. Furthermore, while these items are found in the COCA, none of them occur very frequently. This low frequency of initialisms corresponds with the findings of previous studies, such as Tagliamonte and Denis (2008); and Baron (2011). The initialisms selected for this analysis generated hits across all four genres that were included in this study. Although they do not appear in spoken data as often as in written data, these initialisms exist in spoken data. This corresponds with the results of Kinsella’s (2010) study, vis., that these initialisms are used in spoken language. Of the items analyzed in this study, ‘FYI’ appears most frequently in the COCA, followed by ‘LOL’ and ‘OMG’, and then ‘WTF’ and ‘BTW’. In contrast, ‘LOL’ and ‘BTW’ were the most commonly used initialisms in Tagliamonte and Denis’s (2008) study. ‘LOL’ was the most frequently used initialism in the Corpus of Informal English, and ‘BTW’ was the fourth most commonly occurring one. ‘FYI’ was the least commonly occurring initialism in the Corpus of Informal English, but it is the most commonly used initialism in the COCA corpus. Through analyzing the data in the COCA, ‘FYI’ appears to co-occur together with ‘FAQ’, which stands for “Frequently Asked Questions” (Oxford English Dictionary Online). These two types of initialisms are seemingly preferred by authors who write magazine articles, as ‘FYI’ is used much more in magazine articles than in any other type of writing. Of the initialism chosen for this study, it is the only one that is used more in the COCA than in the Corpus of Informal English.

Using Marchand’s (1969) definition that initialisms are used to create names of organizations or new scientific discoveries, it is possible to see that initialisms such as ‘LOL’ are different than the ones Marchand describes. These initialisms are not used for forming names of scientific discoveries or companies, but are used to form less formal, more low-style initialisms of everyday phrases regarding, for example, the actions or the emotional state of a person. Furthermore, these initialisms often appear to be formed on phrases like ‘oh my god’ and ‘by the way’, and these are phrases one would expect to see perhaps in spoken language
rather than written language. The results from the COCA indicate that people appear to be forming a new kind of initialism that did not exist before the Internet. People also use these initialisms outside of the Internet, which makes it possible to argue that they are not just jargon used by a particular group. While it appears that people are creating a new kind of initialisms that did not exist before, the data are too sparse to be used to draw any definite conclusions from it. For example, not all of the ten selected initialisms appear in the corpus. ‘IMHO’ and ‘IRL’ only appear seven times each, and seven times in a corpus that contains 450 million items is too little to make any real difference in a study like this. Furthermore, some of the items that are used frequently online are not found in the corpus at all, which could indicate that people have not yet completely taken to using this kind of initialism more generally. The items that are used in the corpus, apart from ‘FYI’, appear infrequently, which makes it hard to draw any kind of meaningful conclusions from the data. If this kind of initialism was an established and frequent phenomenon, then the initialisms would most likely appear more often in the COCA than what is currently the case. However, while it is very sparse, the data from the COCA show that there are some items in use which are formed in this new fashion. This shows that while there appears to be a new kind of initialisms in use, because of the low number of occurrences in the corpus it is difficult to draw any definite conclusions regarding these initialisms. To sum it up, the corpus evidence suggests that new kinds of initialisms are first appearing on the Internet and have begun to appear outside of the Internet, providing more opportunities to use this kind of word formation. However, the data collected in this study are not sufficient to warrant any strong claims concerning these initialisms.

There are many possible explanations for why this kind of initialism is not used more frequently in the COCA. As seen in Denis and Tagliamonte (2008) and Baron’s (2011) studies, these items are not as common in netspeak as one might expect. Since these items are not very frequent in online written language, it is not very surprising that they appear infrequently in other contexts. Another explanation could simply be that this kind of initialism originates from the Internet, as Crystal (2001) and Coleman (2012) claim these items do. It can be assumed that many of those who spend a lot of time online using this kind of language are likely teenagers. It is furthermore also possible to assume that people who use this kind of language online, are also more prone to use these initialisms in other contexts as well. Andersson and Trudgill (1992) claim that as we get older, we become more conservative linguistically. This means that older people are less likely than young people to change their vocabulary. It could be one of the reasons why these initialisms are not used very much in the
COCA corpus, and could also mean that these new initialisms are not going to be used more frequently until the next generation of Internet users become a majority. If you are used to netspeak because you have been exposed to the language used online throughout your life, then this kind of initialisms might not seem so strange or foreign or even new to you, and you might thus be more inclined to use them generally. Another explanation why these initialisms are not used much in the COCA might be connected to the notion of correct language, as Burridge (2004) explains. The notion of correctness when it comes to using the English language is something you learn at a very early age in school. These initialisms are, according to Coleman (2012), slang expressions, and are also seen as a feature of netspeak. If people perceive these initialisms as netspeak, and thus also as being slang expressions, then they might refrain from using them in more formal circumstances.

5.2 Limitations

Corpus methodology was chosen for this study because it allows for a broad overview when it concerns language in use in different contexts and provides an overview of how certain aspects of language are used in everyday speech and writing. However, the COCA creates a number of limitations when it comes to this study. For example, the data in the COCA are quite formal; which could serve as an explanation as to why these items rarely appear. These initialisms are, as previously stated, quite informal in their nature, and that might limit the usage of them in a context that may not allow for informal expressions. This could mean that these initialisms will become more frequent in informal language in time, but that they will never truly be used in more formal language. Another limitation of basing the study on the COCA corpus is the lack of control of the sources used to create the corpus data. It can be assumed that these initialisms are mostly used in magazines aimed at teenagers, but the proportion of such texts may be disproportionately in the corpus. Previous studies, such as Tagliamonte and Denis (2008), have established that the main users of these initialisms are teenagers, and would thus be the target audience for texts making use of such initialisms. Lastly, this study has only focused on American English and may not give an accurate view of how these items are used in the English speaking world as a whole.
5.3 Future Research

Whether or not these new initialisms are a sign of an ongoing change in word formation processes is probably something that will take time to unravel. It may require several decades of data before it is possible for researchers to investigate this phenomenon thoroughly. A useful starting point for such study would be the question of how these initialisms are used online, and the extent to which they occur in netspeak. Since different types of communication are directed towards different audiences, future research in this field could be focused on how these initialisms are used in regards to different target audiences. In addition to the initialisms used for this study, the Corpus of Informal English contained several initialisms that were strictly abbreviated items, such as ‘bf’ for boyfriend. These items were not included in this study because they are formed differently. These initialisms are simply single words that have been shortened. Future research could focus on investigating these abbreviations and the extent to which they are used in non-netspeak contexts. Due to the time restrictions given for this project, working with several corpora was not possible. Future studies could therefore include corpora of informal spoken English to see the extent to which these initialisms are used in spoken language. Presumably, much language change starts out in informal language. Informal written language, which nowadays mostly takes place through texting or in chatting online, has already been investigated since netspeak is part of it, but informal spoken language has not been investigated to the same extent.

5.4 Conclusion

Marchand (1969) claims that initialisms are mostly used to create names of new scientific discoveries, organizations or brands. However, initialisms that originate from netspeak, such as ‘LOL’, are different from the kind of initialisms that Marchand describes. This new kind of initialism is not used as names of organizations or scientific discoveries; rather, they denote phrases concerning everyday things. Out of the ten initialisms chosen for this study, only seven of them appeared at all in the COCA. Only five of them appeared in the COCA frequently enough to be used in this analysis. Judging by the data gathered from the COCA, people do form and use this new kind of initialisms, albeit very infrequently. Furthermore, the corpus data show that these items are used in the different genres that exist in the COCA, except for academic writing. The results also indicate that people use this kind of initialisms,
and it may be that a change is currently taking place, but any strong conclusions cannot be
drawn from the limited amount of data gathered for this study.
References


