



Western Washington University
Western CEDAR

WWU Honors Program Senior Projects

WWU Graduate and Undergraduate Scholarship

Winter 2019

The Fuck is Going On? The Grammaticalization of Taboo Words in English

Laura E. Munger

Western Washington University, munger-laurae@gmail.com

Follow this and additional works at: https://cedar.wvu.edu/wwu_honors



Part of the [Linguistics Commons](#)

Recommended Citation

Munger, Laura E., "The Fuck is Going On? The Grammaticalization of Taboo Words in English" (2019).
WWU Honors Program Senior Projects. 108.
https://cedar.wvu.edu/wwu_honors/108

This Project is brought to you for free and open access by the WWU Graduate and Undergraduate Scholarship at Western CEDAR. It has been accepted for inclusion in WWU Honors Program Senior Projects by an authorized administrator of Western CEDAR. For more information, please contact westerncedar@wwu.edu.

The Fuck Is Going On?

The Grammaticalization of Taboo Words in English

Laura Munger

Western Washington University

Abstract

This study examines the grammaticalization of the verb “fuck” to the periphrastic interrogative “the fuck”. I use various forms of data from the Corpus of Contemporary American English and the Corpus of Historical American English to analyze (1) the evolution of “fuck” from a concrete lexical verb to an item diversified in meaning and part of speech, (2) the development of the phrase “what the X” from “what the devil”, and (3) the specific grammaticalization of “the fuck” as an interrogative. These processes are then discussed in terms of specific principles of grammaticalization which they illustrate.

The Fuck Is Going On?

The Grammaticalization of Taboo Words in English

Introduction

The intention of this study is to examine the grammaticalization of “fuck” to “the fuck”, looking at how a lexical verb referring to copulation developed into a periphrastic interrogative. This is accomplished through the quantitative analysis of data from Brigham Young University’s Corpus of Historical American English and Corpus of Contemporary American English. This paper will also explore how the evolution of “fuck” fits into the general framework of grammaticalization and illustrates principles of grammaticalization such as persistence, divergence, and layering.

In order to analyze this grammaticalization process, the study is divided into three sections of analysis. The first section traces the general development of “fuck” from a verb with very concrete meaning referring to sex, to a noun used to intensify other components of a sentence. The second section examines the phrase that I term for convenience “what the X”. This construction allows the intensification of any question word by the insertion of certain taboo words, primarily “devil”, “hell”, and “fuck”. The final section of analysis looks specifically at the reduced and grammaticalized form “the fuck”, which I argue has become a periphrastic interrogative, and which is able to serve similar functions as the word “what”.

Literature Review

Several studies have looked at constructions involving “the fuck”. These include Hoeksema and Napoli (2008) and den Dikken and Giannakidou (2002). Hoeksema and Napoli (2008) used corpus data to analyze the use of “the fuck” as an intensifier when placed

immediately after a verb. They argued that “the use of taboo terms as intensifiers spread from wh-constructions to these [verb] constructions and, finally, to degree intensifier constructions” (Hoeksema and Napoli, 2008, p.347). However, the Hoeksema and Napoli study did not actually examine the use of taboo terms with interrogatives in any detail.

Den Dikken and Giannakidou (2002) discussed interrogatives with taboo words as being polarity items, pointing out that their distribution in sentences reflected that of polarity items rather than conventional wh-elements. However, they did not discuss the further grammaticalized form of these in which the original wh-element is dropped. Additionally, unlike both this study and Hoeksema and Napoli, they did not utilize any corpus data.

Other relevant research includes the sociolinguistic studies by McEnery and Xiao (2004) and Howe (2012). Both of these studies looked at the occurrence of “fuck” in the British National Corpus and compared usage based on various sociolinguistic variables such as age, gender, social class, and education level. Additionally, McEnery and Xiao categorized different senses and ways in which “fuck” is used and compared usage of these particular categories across groups. This is discussed further below.

Methodology

In this study I relied on data from the Corpus of Historical American English, as well as the Corpus of Contemporary American English to a lesser degree. The first corpus contains 400 million words and draws from a variety of genres of written English. The second corpus contains 560 million words and draws from a variety of genres of written English as well as spoken, spanning from 1990 to 2017.

I gathered two different types of data from these corpora. The first was total frequency. For this type of data, I used the Corpus of Historical American English and found how often

particular items (“fuck” used as a noun, “fuck” as any part of speech, and the phrases “what the devil”, “what the hell”, and “what the fuck”) occurred across decades.

The second type of data were token sets. I took 50 token samples of “fuck” from the Corpus of Historical American English from every decade from 1960 to 2010, as well as all the tokens of “fuck” from earlier decades. I then coded these tokens with a system based on the categories used by McEnery and Xiao. In addition, I coded tokens from each decade based on part of speech and on whether or not they were used literally.

Unfortunately, due to its taboo nature “fuck” has historically not been very common in written materials. As such, the data for its use in this study does not begin until the 1930s. Additionally, from 1930 to 1959 each decade has less than 50 total instances of “fuck” as a noun. Because of this there is some potential difficulty in drawing conclusions about semantic trends in the usage of “fuck”.

Analysis of “Fuck”

The first line of analysis to pursue is to trace the evolution of “fuck”. According to Hughes (2006) the ultimate etymology of “fuck” is unclear. It is first attested to in English in 1503 as a verb referring to the act of having sex. It was later converted to a noun, attested to as early as 1663 according to the *Oxford English Dictionary*. This initial conversion retained most of the original semantic content, with the new noun form of “fuck” still referring to a sex act. Both the verb and noun forms of “fuck” later developed many nonsexual polysemies.

Looking at data from the Corpus of Historical American English provides some insight into the word’s usage in written English. It does not appear in the corpus until the 1930s. Table 1 shows the use of “fuck” as a noun per million words from every decade between 1930 and 2010.

Table 1

<i>Year</i>	<i>Uses of “fuck” as a noun per million words</i>
<i>1930-1939</i>	0.24
<i>1940-1949</i>	0
<i>1950-1959</i>	0.04
<i>1960-1969</i>	3.38
<i>1970-1979</i>	11.3
<i>1980-1989</i>	9.4
<i>1990-1999</i>	13.81
<i>2000-2009</i>	16.03

The table shows that while “fuck” as a noun did appear as early as the 1930s, it was not used with any prevalence in written English until the 1960s. After that its use continued to grow steadily (apart from a small drop in the 1980s), with it being about five times more common in the 2000s as in the 1960s.

Table 2

<i>Year</i>	<i>Total uses of “fuck” per million words</i>
<i>1930-1939</i>	0.65
<i>1940-1949</i>	0
<i>1950-1959</i>	0.57
<i>1960-1969</i>	8.63
<i>1970-1979</i>	33.8
<i>1980-1989</i>	23.07
<i>1990-1999</i>	34.46
<i>2000-2009</i>	38.83

Table 2 shows the use per million words of fuck regardless of part of speech for the same sections of time. It shows similar trends to those found in the noun specific data. The word saw some use from the 1930s to 1950s, became much more prevalent in the 1960s, dropped a little in the 1980s, but continued growing in use into the 2000s.

In order to study the development of different polysemous uses of the word “fuck”, I collected tokens from the Corpus of Historical American English. Tokens were collected from every decade from 1930 to 2009 and were coded according to the category scheme used in McEnery and Xiao (2004). The scheme divides taboo word use into nine different categories, shown in Table 3 (adapted from “Table 29: Category scheme for swear words” from McEnery and Xiao (2004)).

Table 3

<i>Category</i>	<i>Examples</i>
<i>General expletive</i>	(Oh) fuck!
<i>Personal insult referring to defined entity</i>	You fuck! / that fuck
<i>Cursing expletive</i>	Fuck you!/me!/him!/it!
<i>Destinational usage</i>	Fuck off! / he fucked off
<i>Literal usage denoting taboo referent</i>	He fucked her
<i>Emphatic intensifier</i>	Fucking marvelous! / in the fucking car
<i>Pronominal form</i>	Like fuck / fat as fuck
<i>Idiomatic set phrase</i>	Fuck all / give a fuck / thank fuck
<i>Metalinguistic or unclassifiable due to insufficient context</i>	The use of the word “fuck” / you never fucking

Table 4, below, shows the distribution of “fuck” across these categories in data from 1930 to 2009 as percentages. 50 tokens were taken from each decade, except for 1930-39 which had only 16 tokens in the corpus, 1940-49 which had no tokens, and 1950-59 which had 14 tokens. These tokens were coded according to the scheme outlined in Table 3. Metalinguistic and unclassifiable tokens were disregarded when calculating the percentages in Table 8. Additionally, it should be noted that none of the tokens analyzed here fell into the Pronominal Form category, so that category is omitted from Table 4.

Table 4

	<i>Personal Insult</i>		<i>Literal Usage</i>		<i>Emphatic Intensifier</i>		<i>Idiomatic Set Phrase</i>
	<i>General Expletive</i>	<i>Referring to Defined Entity</i>	<i>Cursing Expletive</i>	<i>Destinational Usage</i>	<i>Denoting Taboo Referent</i>		
<i>1930s</i>	0.0%	0.0%	6.25%	0.0%	68.75%	0.0%	25.00%
<i>1950s</i>	0.0%	7.14%	50.00%	0.0%	21.42%	0.0%	21.42%
<i>1960s</i>	11.11%	0.0%	15.55%	4.44%	35.55%	2.22%	31.11%
<i>1970s</i>	10.63%	0.0%	27.65%	4.25%	27.65%	2.12%	27.65%
<i>1980s</i>	10.20%	4.08%	22.44%	4.08%	18.36%	4.08%	36.73%
<i>1990s</i>	4.34%	0.0%	34.78%	4.34%	10.86%	8.69%	36.95%
<i>2000s</i>	14.28%	2.04%	34.69%	2.04%	10.20%	4.08%	32.65%

Looking at the data in Table 4, two main trends become apparent. The first is that literal use of “fuck” has decreased steadily since the 1960s. It is difficult to draw any definite conclusions from earlier decades because of the small number of tokens available, but it should be noted that in the 1930s more than half of the tokens were literal usages. This conforms with the idea that literal usage has been progressively decreasing. The second major trend is the diversification of use of “fuck” over time. In the 1930s there are only three different types of use. By the 2000s, there are seven. The decrease in use as denoting the taboo referent (i.e. sex) in favor of other taboo uses indicates that “fuck” is undergoing a process of semantic bleaching. As the word loses its original semantic content, it is able to be used more generally as a taboo word, fulfilling a variety of different taboo word roles. The change can be described by the following cline:

Sex (Taboo) > General Taboo Use

A useful illustration of this cline can be found in the use of “fuck” as a cursing expletive. When speakers use the word in this way, for instance with the sentence “Fuck you,” they do not do so with reference to sex. In fact, the sentence would not make syntactic sense if the verb “fuck” were being used here in a literal sense. However, despite not referring literally to sex, the word remains taboo in this usage.

Another way to look at the evolution of “fuck” is to analyze the parts of speech it is used as, as well as looking at whether the usage is literal or general. Table 5 displays data for just such a purpose:

Table 5

<i>Years</i>	<i>Verb (literal)</i>	<i>Verb (general)</i>	<i>Noun (literal)</i>	<i>Noun (general)</i>	<i>Interjection (always general)</i>
<i>1930-1939</i>	56%	12%	6%	25%	0%
<i>1940-1949</i>	X	X	X	X	X
<i>1950-1959</i>	21%	64%	0%	14%	0%
<i>1960-1969</i>	32%	26%	4%	26%	12%
<i>1970-1979</i>	24%	36%	2%	26%	12%
<i>1980-1989</i>	22%	40%	0%	24%	14%
<i>1990-1999</i>	10%	44%	4%	40%	2%
<i>2000-2009</i>	14%	36%	0%	36%	14%

The data is taken from the Corpus of Historical American English. The same sampling methodology was used as with the data in Table 4. I coded the data by part of speech and whether or not it referred to sex. The only parts of speech which surfaced were verbs, nouns, and interjections. In each decade the verbs were most common, followed by nouns, and then interjections. There is a potential trend in the data for nouns becoming more common and verbs becoming less, but its difficult to say for certain with these particular numbers. The data in Table 5 also further illustrates the shift towards general non-literal use discussed above. It is unclear whether the interjection form of the word “fuck” derives from the original verb or the noun variant.

Analysis of “What the X”

Having discussed the development of “fuck”, we can now turn to the development of “what the X”. The phrase originated with “what the devil”, which, according to the *Oxford English Dictionary*, was probably borrowed into English from the French equivalent. The phrase is attested to as early as 1393. It can be used with a variety of question words, primarily “what”, but also “why”, “when”, “where”, and “how”.

Since the original development of “what the devil” its use has expanded to allow the substitution of other taboo words and euphemisms for the “devil” component. These include terms such as “dickens”, “hell”, and “fuck”. The most prominent such substitutions are “hell” and “fuck”. The prevalence over time of these different versions of “what the X” are shown below in Table 6.

Table 6

	<i>“What the devil” per million words</i>	<i>“What the hell” per million words</i>	<i>“What the fuck” per million words</i>
1810-1819	12.7	0	0
1820-1829	3.61	0	0
1830-1839	2.47	0	0
1840-1849	3.12	0.06	0
1850-1859	1.64	0.06	0
1860-1869	1.64	0.12	0
1870-1879	1.56	0	0
1880-1889	1.53	0.05	0
1890-1899	1.26	0.05	0
1900-1909	2.26	0.36	0
1910-1919	2.55	0.88	0
1920-1929	3	2.69	0
1930-1939	2.68	12.03	0
1940-1949	2.75	15.48	0
1950-1959	2.61	13.65	0
1960-1969	1.58	21.06	0.75
1970-1979	1.01	23.01	2.44
1980-1989	1.07	19.32	2.8
1990-1999	0.32	20.36	2.72
2000-2009	0.74	23.37	4.48

As Table 6 illustrates, “what the devil” was relatively prevalent in the early 19th century, but subsequently declined greatly in usage. It is still used in contemporary English, but very rarely. We also see that “what the hell” first emerged in the mid-19th century. It came to prominence in the 20th century and remains the most prevalent form of “what the X”. Finally, there is “what the fuck”. This form first occurs in the corpus in the 1960s and has since grown steadily in usage, although it is nowhere near “what the hell” in its usage. It should be noted, however, that because of the taboo nature of these forms the written records probably do not reflect the points at which they entered into the language. In all likelihood they were used in spoken English for some time but were not put into written materials because they were considered too informal and impolite.

With the data from Table 6, we can begin to come to some conclusions regarding “what the X”. First, the development of “what the X” can be described by the following cline:

what the devil > what the hell > what the fuck

Taking the first shift in the cline, from “devil” to “hell”, it can be speculated that the change was facilitated by their shared religious domain, perhaps through a metonymical process. However, it is important to note that the “the” element of the original “what the devil” phrase is preserved in “what the hell”. (We can conclude that the “the” is in fact being preserved because of all these taboo terms “devil” is the only one generally accompanied by the definite article outside of constructions such as “what the X”.) The significance of this preservation is that it indicates a loss of the original semantic content of “devil” in this phrase. “Devil” has lost enough semantic content that the phrase as a whole has become essentially a single unit, expressing a question with intensified emotional content or emphasis. “The” is now part of that unified

phrase, rather than simply being there to accompany “devil”. Similarly, the fact that “hell” is able to be inserted next to the “the” indicates that it does not have enough of its original religious semantic content to make “the hell” seem nonsensical and illogical. Therefore, while the shared religious domain of the more concrete lexical forms of the “devil” and “hell” may in some part help the transition from “what the devil” to “what the hell”, it is probably a small part since both have evidently lost much of that semantic content by the time they are used in “what the X”.

Similar principles are at play in the second transition described in the “what the X” cline from “what the hell” to “what the fuck”. Examining them through the lens of their original lexical meanings, there would seem to be a leap occurring across semantic domains, forming the semantic cline religion > sex. However, for the same reason that we can conclude that “hell” is present in the phrase without all of its original religious semantic content (namely, the otherwise illogical presence of “the”) we can conclude that “fuck” has also entered “what the X” without all of its original sexual semantic content. This bleaching would make the apparent jump from religion to sex much more plausible.

Analysis of “The Fuck”

While both “fuck” and “what the X” have undergone obvious processes of change, the grammaticalization of “fuck” does not become completely apparent until the emergence of the interrogative “the fuck”. This form is illustrated in sentences such as:

(1) *The fuck happened here, man?*

(2) *The fuck does he know?*

Both of these sentences use “the fuck” in place of the interrogative pronoun “what”. (It should be noted that I found no instances in the corpora used in this study of “the fuck” being used for other question words, despite the existence of “why the X”, “how the X”, etc.) The

structure of the sentence is otherwise unchanged, as illustrated by the same sentences said using “what”:

(3) *What happened here, man?*

(4) *What does he know?*

Looking at sentences (1) and (2), one might think that “the fuck” is simply being used for the purpose of intensification. However, we can discount this idea because these sentences with neither “the fuck” nor “what” are either ungrammatical or ask an entirely different question:

(5) **Happened here, man?*

(6) *Does he know?*

Sentence (5) is ungrammatical without either a “what” or interrogative “the fuck”. And sentence (6) is asking a yes or no question about someone’s knowledge, rather than asking what that knowledge is. These sentences demonstrate that “the fuck” does indeed carry grammatical content equivalent to that of “what” in this context.

The form came into use as a shortened version of “what the fuck”. This reduction removed the element, “what”, that originally gave the phrase its interrogative function. Despite this removal, the phrase retained that interrogative function. This indicates that the original “what the fuck” has been reanalyzed such that it is no longer simply a “what” being modified by “the fuck”. Instead, the entire phrase carries the meaning and function of marking an intensified or urgent question. This reanalysis allowed “the fuck” to continue to be used for that same exact function after the reduction of the original phrase.

Unfortunately, both the Corpus of Historical American English and the Corpus of Contemporary American English contain very few instances of the interrogative “the fuck”. As such it is difficult to use corpus data to observe any trends in its usage. However, there are still

several observations to be made about the construction, particularly regarding specific principles of grammaticalization illustrated by “the fuck”.

Looking at the overall evolution of “fuck”, there is a clear case of decategorialization from a major category to a minor. This process is illustrated by the cline below:

Verb > Noun > Periphrastic Interrogative

As the cline shows, “fuck” began originally as a verb. It was subsequently converted to a noun. Later, after the noun form was inserted into the “what the X” construction, “fuck” was reanalyzed as part of a phrase indicating a question. The presence of the definite article with the interrogative “fuck” appears to indicate that it is still a noun, however its function in the sentence does not match that of a noun. Therefore, it may be best to consider “the fuck” as a periphrastic interrogative.

“Fuck” also exhibits divergence. According to Hopper and Traugott (2003), when divergence occurs “a less grammatical form may split in two, one variant maintaining its former characteristics, the other becoming more grammatical” (p. 115). In the case of fuck, the original lexical form referring to sex has remained in use in the language, even while a variant form underwent grammaticalization into an interrogative. This is clearly seen in some of the corpus data discussed above. For example, in the Corpus of Contemporary English 4% of all instances of “fuck” as a noun still express the original sexual meaning. The Corpus of Historical American English produces similar results, with 6% of all uses of “fuck” as a noun in the 1990s and 2000s referring to sex. From a certain perspective, this divergence seems almost necessary to the actual grammaticalization of “fuck”. The existence of the sexual meaning of the word is what “fuck” derives its taboo status from. This taboo status was necessary for it to be inserted into the “what the X” phrase and thereby begin grammaticalizing as an interrogative.

Another principle illustrated in this study is layering. Again drawing from Hopper and Traugott (2003), “layering is the synchronic result of successive grammaticalization of forms which contribute to the same domain” (p. 125). The relationship between “what” and the interrogative “the fuck” is a clear example of this. Both “what” and “the fuck” grammaticalized to fulfill very similar syntactic roles, i.e. to form questions. Despite the relative rarity of the interrogative “the fuck” the two forms do indeed exist simultaneously in English, forming two distinct layers of English interrogatives.

However, even though the two forms perform similar functions, there are still important differences between them. Some of these are syntactic, with “the fuck” being much more limited in its potential placement in a sentence. Of the instances of “the fuck” found in the corpus data for this study, 94% were sentence initial. There are also pragmatic and sociolinguistic limits on the usage of “the fuck” due to its taboo status. These differences between “what” and “the fuck” reflect the persistence that affects “the fuck”. The old characteristics of “fuck” as a sexual taboo word still affect the ways in which the grammaticalized form can be used.

Conclusions and Further Research

There are several ways in which this research could be expanded upon in future. I believe that further examination of the evolution of “fuck” leading up to its insertion into “what the X” is called for in order to gain a comprehensive perspective on the grammaticalization of “fuck”. Additionally, further syntactic analysis of the interrogative “the fuck” is necessary in order to better understand the nature of this periphrastic interrogative and how it compares to the more conventional “what”. Both of these directions of research would probably necessitate the use of different data sources containing more colloquial and spoken data, as well as greater chronological breadth.

That being said, despite its limitations this study does manage to shed some light on the grammaticalization of “fuck”. The conversion of “fuck” from a verb to a noun, as well as its semantic bleaching allow its use in the “what the X” construction. The reanalysis of this same construction subsequently allows the shortened form “the fuck” to retain its interrogative function, creating a new, periphrastic, alternative to the older interrogative “what”.

References

Corpus of Contemporary American English (2017) Available from

<https://corpus.byu.edu/coca/>

Corpus of Historical American English (2017) Available from <https://corpus.byu.edu/coha/>

Den Dikken, M., Giannakidou, A. (2002). From Hell to Polarity: “Aggressively Non-D-Linked” Wh-Phrases as Polarity Items. *Linguistic Inquiry*, 33, 31-61.

Fuck (n.d.). In *Oxford English Dictionary*. Retrieved from

<http://www.oed.com.ezproxy.library.wvu.edu/view/Entry/75196>

Hell (n.d.). In *Oxford English Dictionary*. Retrieved from

<http://www.oed.com.ezproxy.library.wvu.edu/view/Entry/51468>

Hoeksema, J., Napoli D. J. (2008). Just for the hell of it: A comparison of two taboo-term constructions. *Journal of Linguistics*, 44, 347-378.

Hopper, P.J., Traugott, E.C. (2003). *Grammaticalization*. New York, NY: Cambridge University Press.

Hughes, G. (2006). Fuck. In *An Encyclopedia of Swearing*. (pp. 188-194). Taylor and Francis Group.

McEnery, A., Xiao, Z. (2004). Swearing in modern British English: the case of *fuck* in the BNC. *Language and Literature*, 13, 253-268.