

# ***Digging Into A Perfume Lover's Life***

An analysis of Conor McTeague's fragrance interests

**By Chloe Gordon**

June 2021

## Background part 1

When Conor McTeague passed away in April 2021, he left all of his perfumes to the Institute of Art and Olfaction. The IAO is a non-profit organization primarily known for their awards program for the artisan/niche perfume industry.



Because Conor was a massive collector, the IAO needed volunteers to help with data collection. While cataloging his perfume collection and reading through his blog, ScentHurdle, I learned a lot about a person I had never met.

## *Background part 2*

Some bonus work came about when his friends were having trouble compiling his blog posts. Last month I pulled his blog into text files and had some fun reading his engaging style in the process.

I developed some assumptions about Conor after cataloging under 300 perfume bottles with the project lead, Daniel.

In the project here, I use various exploratory analysis techniques to see if the assumptions I formed about Conor after going through half of his collection are consistent with what he wrote about.

# *Presentation Outline*

- > Introduction
- > Is there a preference for vintage fragrances?
- > Is there a preference for certain perfume notes?
- > A quick Chanel briefing

**Q&A and code breaks after each section**

***Assumption 1:  
He LOVED Vintage Collecting***

## Section 1 Speaker Notes

While we came across many interesting contemporary perfumes, it felt like many of the 550 bottles were vintage. Did his writing reflect a substantial interest in vintage (pre-2000) perfumes?

In the post tags (next slide) I used the earliest year listed to classify the focus of that post.

There were a total of 630 posts on Conor's site, 563 of which had a year tag.

The conclusion in this assumption is that he did not write more, either in number of posts or number of words, about vintage fragrances.

2016 HERMÈS JEAN-CLAUDE ELLENA

## HERMÈS EAU DE NÉROLI DORÉ, 2016

06/15/2016 BY JTD

OTHER

## WHAT INTERESTS YOU IN PERFUME?

07/07/2016 BY JTD

1924 2007 CHANEL ERNEST BEAUX JACQUES POLGE

## CHANEL CUIR DE RUSSIE, 1924 (2007)

01/31/2015 BY JTD

1972 CHRISTIAN DIOR EDMOND ROUDNITSKA

## DIGGING (INTO) VINTAGE: CHRISTIAN DIOR DIORELLA, 1972

01/04/2012 BY JTD

# Did Conor Post More on Vintage Fragrances?

Post Frequency by Perfume Launch Decade



**Chanel N.5:**  
1921



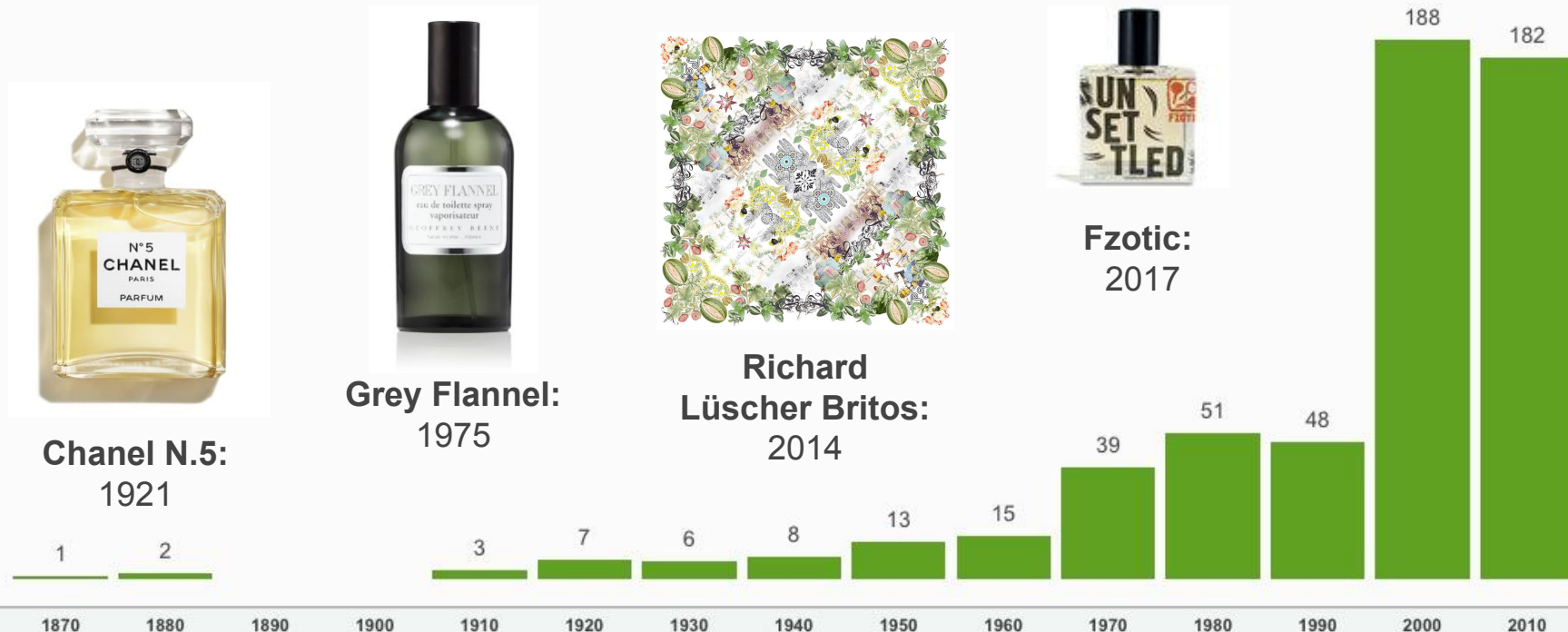
**Grey Flannel:**  
1975



**Richard  
Lüscher Britos:**  
2014



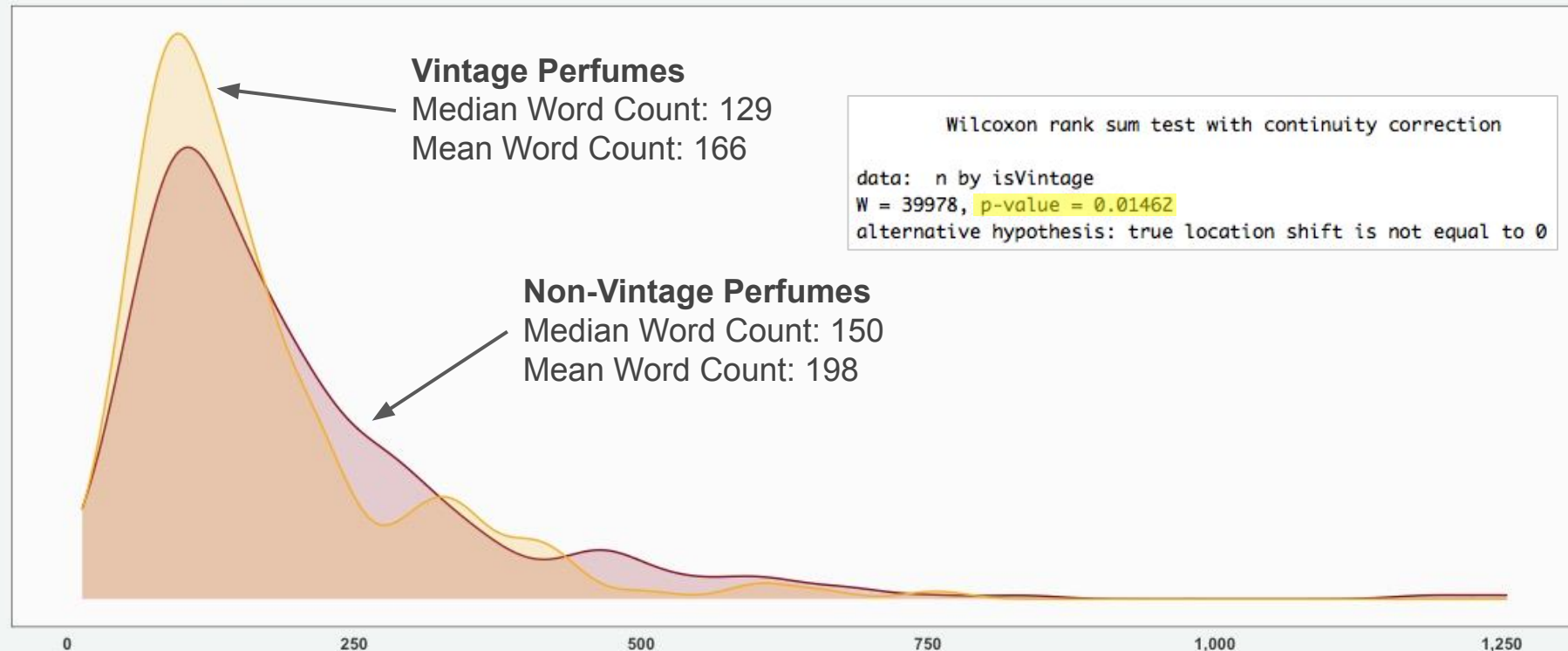
**Fzotic:**  
2017





# Perhaps He Spent More Time Writing about Vintage Perfumes?

Word Count per Post



***Assumption 2:***  
***He Didn't Like White Florals as Much as He Let On***

## Section 2 Speaker Notes

Here I open with the assumptions I learned from working with Daniel while sorting through Conor's most "loved" bottles. As we sifted through the bottles, he kept getting astonished at the number of deep, earthy fragrances that were almost empty. Daniel was under the impression that Conor loved light, white florals.

This analysis could use more cleaning, because I may not have captured all of the plurals, related aroma chemicals, and translations in French, the lingua franca of perfumery, or other relevant languages.

An issue with the TF-IDF section is that notes (like iris) can be used in the perfume name, inflating that term's calculated "importance" in a post.

## Popularity of Certain Notes part 1

How often do the lighter notes (**rose, iris, lily, jasmine, narcissus, galbanum**) appear in Conor's posts versus heavier notes (**chypre, musk, leather, vanilla**)?

| Note      | Note Mentions | Note Mentions /<br>Total Posts |
|-----------|---------------|--------------------------------|
| Chypre    | 692           | 1.10                           |
| Rose      | 504           | 0.80                           |
| Musk      | 401           | 0.64                           |
| Leather   | 297           | 0.47                           |
| Vanilla   | 265           | 0.42                           |
| Iris      | 159           | 0.25                           |
| Lily      | 132           | 0.21                           |
| Jasmine   | 99            | 0.16                           |
| Narcissus | 25            | 0.04                           |
| Galbanum  | 21            | 0.03                           |

| Note      | searchTerms          |
|-----------|----------------------|
| <chr>     | <chr>                |
| iris      | iris                 |
| rose      | rose                 |
| chypre    | (chypre oakmoss)     |
| lily      | (lily muguet lilies) |
| vanilla   | vanill               |
| musk      | (musk musc)          |
| jasmine   | jasmin               |
| galbanum  | galbanum             |
| narcissus | narcissus            |
| leather   | leather              |

```
# Create function to count mentions of a word
countFunction <- compose(~ sum(.x),
                          ~str_count(postsDF1$postText,
                                     pattern = regex(paste0("\\s", .x),
                                                       ignore_case = TRUE)))

# Note Mentions to CSV
noteCountDF <- perfumeNotes %>%
  mutate(`Note Mentions` = map_dbl(searchTerms, countFunction)) %>%
  select(-searchTerms)
```

## Popularity of Certain Notes part 2

We can look in which posts these same notes are used more frequently using Term Frequency-Inverse Document Frequency.

| Post Title   | Note    | Count | Term Frequency | Inverse Document Frequency | TF-IDF |
|--|---------|-------|----------------|----------------------------|--------|
| le Labo Iris 39, 2006  | Iris    | 12    | 0.128          | 2.615                      | 0.334  |
| digging (into) vintage: Paloma Picasso mon Parfum, 1984                  | Jasmine | 5     | 0.094          | 2.333                      | 0.22   |
| Amouage Lyric Woman, 2008  | Rose    | 17    | 0.137          | 1.592                      | 0.218  |
| digging (into) vintage: Maitre Parfumeur et Gantier Iris Bleu Gris, 1988 | Iris    | 11    | 0.08           | 2.615                      | 0.21   |
| Annick Goutal Musc Nomade, 2008  | Musk    | 17    | 0.142          | 1.394                      | 0.198  |
| Donna Karan Gold, 2006   | Lily    | 5     | 0.075          | 2.594                      | 0.194  |
| Cuir de Lancome, 2007  | Leather | 10    | 0.123          | 1.561                      | 0.193  |
| Indult Reve en Cuir, 2008  | Leather | 18    | 0.118          | 1.561                      | 0.185  |
| Etro Gomma, 1989   | Leather | 11    | 0.113          | 1.561                      | 0.177  |
| Aftelier Perfumes Vanilla Smoke (edp), 2015                              | Vanilla | 20    | 0.072          | 2.181                      | 0.156  |

```
# Note Term Freq - Inverse Document Freq to CSV
postsDF1 %>%
  select(postTitle, postText) %>%
  unnest_tokens(word, postText) %>%
  filter(!word %in% tidytext::stop_words$word) %>%
  filter(!word %in% proustr::stop_words$word) %>%
  mutate(word = ifelse(word %in% c("lily", "muguet", "lilies"),
                        "lily",
                        word),
         word = ifelse(word %in% c("jasmin", "jasmines"),
                        "jasmine",
                        word),
         word = ifelse(word %in% c("cuir", "leathers", "cuirs", "leathery"),
                        "leather",
                        word),
         word = ifelse(word %in% c("musc", "musks", "musky"),
                        "musk",
                        word),
         word = ifelse(word %in% c("chypres", "oakmoss"),
                        "chypre",
                        word),
         word = ifelse(word == "roses",
                        "rose",
                        word)) %>%
```

***Assumption 3:***  
***He Preferred Chanel N.19 to N.5***



## Section 3 Speaker Notes

From perusing a few posts, I saw quite a few warm mentions of Chanel N.19. Of course this could be my own bias, since the N.19 extrait is among my favorites. At any rate, my judgement tells me he prefers N.19, despite it being mentioned fewer times than N.5. This is mostly because Chanel N.5 is **the** reference for modern or feminine or aldehydic floral perfumes.

A good step for further analysis would be to rank the perfumes mentioned in Conor's writings to commercial popularity. There are two limitations I can think of here. First, it will be difficult to automate picking out the perfume names from the post text. Looking for capitalized words may be a good start, but I don't think that will solve all of the problem. Second, using a site like basenotes.net or fragrancesoftheworld.info would require using their intellectual property.

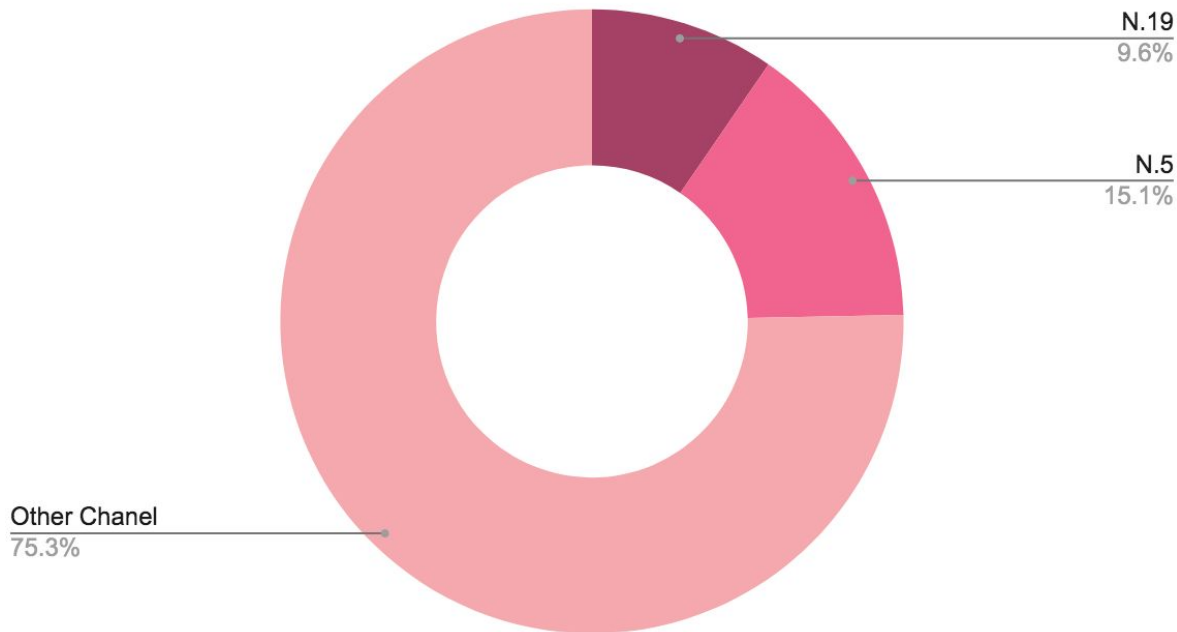
# The Battle of Chanel

Conor didn't write an entire post about the most popular perfume of all time, Chanel N.5. But he did reference it 33 times in 19 posts.

Chanel N.19, a stern green-iris scent, is referenced 21 times, including a dedicated post.

There were a total of 165 mentions of "Chanel".

Chanel N.5 Is Mentioned More than N.19



```
##### Chanel
# (I made the Chanel graphic in Google Sheets)
# Chanel Post Counts
postsDF1 %>%
  filter(str_detect(postText, "Chanel")) %>%
  select(postUrl, postTitle) %>%
  nrow() #95

# N.5
postsDF1 %>%
  filter(str_detect(postText, regex("chanel\\s+[no0]*.{0,2}5", ignore_case = TRUE))) %>%
  select(postUrl, postTitle) %>%
  nrow() #19

# N.19
postsDF1 %>%
  filter(str_detect(postText, regex("chanel\\s+[no0]*.{0,2}19", ignore_case = TRUE))) %>%
  select(postUrl, postTitle) %>%
  nrow() #21

# Chanel Term Counts
str_count(postsDF1$postText, pattern = regex("chanel", ignore_case = TRUE)) %>%
  sum() #165

# N.5
str_count(postsDF1$postText, pattern = regex("chanel\\s+[no0]*.{0,2}5", ignore_case = TRUE)) %>%
  sum() #33
# N.19
str_count(postsDF1$postText, pattern = regex("chanel\\s+[no0]*.{0,2}19", ignore_case = TRUE)) %>%
  sum() # 21
```