

ARTICLE

# Prescribing French: A corpus-linguistic approach to official terminology in French newspapers

Gyula Zsombok

Middlebury College

Email: [gzsombok@middlebury.edu](mailto:gzsombok@middlebury.edu)

(Received 11 February 2020; revised 08 June 2020; accepted 14 June 2020; first published online 09 November 2020)

## ABSTRACT

In France, English is often perceived as a negative influence on the language in the eyes of purist institutions like the French Academy. Terminological commissions have been established to replace foreign expressions with French terminology that is regularly published in the *Journal officiel de la République française*. Although the Toubon Law of 1994 prescribes the use of this terminology in government publications, speakers are merely encouraged to do so. This article investigates the variation between English lexical borrowings and their prescribed equivalents in a large newspaper corpus containing articles from 2000 to 2017 in order to see whether formal written language complies with the purist recommendations. Time is treated with a new dynamic approach: the probability of using a prescribed term is estimated three years before and three years after official prescription. Fifty-four target terms are selected from the lexical fields of computer science, entertainment industry and telecommunication, including emblematic prescribed words such as *courriel* and *mot-dièse*. The analysis reveals that prescription is only effective when it follows already attested use. Furthermore, conservative newspapers show higher proportions of recommended terminology, especially as compared to newspapers specializing in technology.

## 1. INTRODUCTION

Language planning institutions in France and Québec have traditionally exerted strong top-down influence on the French lexicon and continue to intervene in the usage of the language whenever they deem it necessary (Candel, 2015, 2017). In France, government-issued decrees since 1972 outline legal support for the work of so-called terminological commissions, aimed at replacing incoming foreign terms and expressions with French terminology. In Québec, the Charter of the French Language (1977) grants similar powers to the *Office québécois de la langue française*. While the long consultations and complex decision-making processes underlying the work of these professional lexical innovators remain mostly unknown to the public, the successes and pitfalls of some of their widely used lexical creations often elicit considerable public interest. New terms that made it

into the mainstream, or seem likely to catch on eventually, such as *ordinateur* for 'computer', *courriel* for 'email', and not long ago *infox* for 'fake news', are recognized as successful acts of enrichment of the French language. Others are derided as failures that call into question the efficacy of the very same top-down terminological creation process; after all, *accès sans fil* (and its proposed abbreviation *ASFI*) and *terminal de poche* or *ordiphone* fell into oblivion in French linguistic history shortly after their official release as recommended alternatives for 'wireless internet access (wifi)' and 'smartphone' (Kim, 2015, 2018). However, rather than conducting 'a discussion of change in the lexicon [that] soon descends the citation of individual examples' (Posner, 1997: 143), it seems much more appropriate to search for general processes and criteria that might propel some lexical innovations to thrive where others tend to fail.

While studies of historical, especially semantic, change in the lexicon have been abundant, previous scholarship on the systematic longitudinal effects of terminology creation in French is nonexistent. The few quantitative and attitudinal studies of the effects of prescriptivism seem to agree on one point: there are broad discrepancies between prescription and practice. Attempts at regulating the variable use of complex grammatical structures, such as the subjunctive mood and future tenses in formal written registers, for instance, have been blatantly ineffective due to the fact that prescriptivism tries to impose a strict form-function symmetry on an inherently variable linguistic system and, predictably, fails to influence vernacular usage at all (Poplack et al., 2013; Poplack and Dion, 2009). It has also been shown that the French and the Québécois tend to demonstrate purist attitudes towards the influence of English, but speakers in Québec tend to be more inclined to use prescribed terminology than those in France (Walsh, 2016). With the increasing use, control and consumption of social media, official and spontaneous word creations can be expected to become more frequent and widely diffused (Blackwood, 2013), which makes it particularly relevant to find out what could make some newly created terms thrive and others fall quickly into disuse.

In French today, English is perceived as the most productive donor language, making borrowings from English the most relevant source to test the efficacy of terminological prescription. The type of corpus selected for such a study also has to control for media and topic, since the probability of the entrenchment of English loanwords in French depends greatly on their frequency and dispersion across written registers (Chesley and Baayen, 2010), and borrowings are more numerous in topics associated with the prestige of English in domains such as sport (Bernard-Béziade and Attali, 2012; Martin, 2007), music, entertainment, media (Ostyn et al., 1997), computer science and politics (Varga et al., 2011), especially American politics (Chaput, 2010). While spoken language could provide abundant lexical borrowings, only their recurrence in widely diffused written media, among them newspapers, can be taken as indication of their use and entrenchment in the language (Saugera, 2017). Newspapers also provide a good test case for the longitudinal effects of prescription in written registers thanks to their frequency of publication and the strict editorial constraints to comply with standard language norms. Although newspapers are expected to comply with the rules of standard French, lexical borrowings have been shown to occur even in widely distributed, international journalistic writing.

This article proposes to measure the efficacy of top-down intervention on the lexicon from a single source: the publications of the *Commission d'enrichissement de la langue française* (CELFL, 'Commission for the Enrichment of the French Language') in *Journal officiel*. Previous studies investigated terminology implantation drawing on manuals by multinational IT companies (Saint, 2015), specialized communications in the domain of nanotechnology (Quirion, 2011), or newspaper articles from French-speaking territories (Kim, 2015, 2018) and they confirmed striking variation pertaining to the relative frequency of official recommendations: publications in Québec are more likely to reject Anglicisms when French terms are at their disposal. Rather than following previous studies that focused on a variety of prescriptive inputs during a given time period, this study applies precise data-analytical methods to a single, large and publicly available newspaper corpus to examine the variable use of English lexical borrowings and their prescribed French equivalents before and after the official date of publication of the new French term in *Journal officiel*, the legal source of laws, decrees and other official information published by the French Parliament and the national government. Using a single source as input leads to less contextual variation than in previous studies, which examined a wide variety of prescriptive literature or sources of opinion. It has the advantage of controlling for the type of innovation (single lexical items) and the onset and timeline of prescription (three years prior and after the official date of recommendation). Although prescription has had little effect on the use of complex grammatical structures, such as the subjunctive, it is expected to affect the lexicon more easily, since lexical change precedes grammatical change in contact situations (Poplack, 2018; Thomason and Kaufman, 1991). Target words examined in this study come from technology and entertainment industries, as they are among the most productive sources of lexical creation in contemporary French. Section 2 describes the sources of prescriptive terminology and English borrowings used in this study and explains the methods of data collection and analysis. Section 3 reports on the results of a generalized additive model and reveals that prescription does affect the probability of the use of some terms. The analysis will shed light on the proportion of successful terminological innovations and offer some insights into the possible role of the scope and readership of the French newspapers included in the corpus.

## 2. TERMINOLOGICAL COMMISSIONS IN FRANCE

Motivated by prescriptivism and purism (Thomas, 1991), the CELFL in France puts in place extensive procedures to find the most suitable terms for objects, tools and concepts in the fields of science, technology, media and culture. In order to achieve this goal, the CELFL coordinates the work of multiple universities, industry partners, ministry commissions and other language institutions such as the *Académie française* ('French Academy'). The process of accepting and prescribing a new French term to replace a foreign expression also involves multiple steps of revisions. First, the new term is created by a group of experts including linguists and scientists, but also members of the community where the term is intended to be used. When the group of experts proposes a neologism, they are instructed to examine if creating that neologism is necessary, if the term is transparent (i.e. it

is directly associated with the notion it covers), and if it complies with the morphological and syntactic rules of French. They can consult other francophone partner institutions, the *Association française de normalisation* (Afnor, 'French Standardization Association') and academic linguistic laboratories. Afterwards, the *Académie française* evaluates the proposal and decides whether to approve or to return it for further revisions. When a term is finalized, the proposal is submitted for ministerial approval. At the end, the term and its definition are published in the *Journal officiel*, as well as on *FranceTerme*, a website dedicated to the work of terminological commissions (*Rapport annuel 2017 de la Commission d'enrichissement de la langue française*, 2019).

The prescription procedure is informed and motivated by purism. The essence of purism lies in the attempt to rid the language of foreign elements by using two processes: saying not only what linguistic features ought to be included, but also what features are to be excluded (Langer and Nesse, 2012). Purism is directed primarily at the lexicon, but it affects the codification, cultivation and planning of all aspects of standard languages. It is often associated with the notions of 'decadence' and 'corruption' of a language and its lexical stock, and the responsibility to remedy this situation (Chansou, 2003). While purism is considered an abstract ideology motivating negative attitudes towards foreign elements in the language, prescriptivism is very much concrete: it is the materialization of purism. The quintessential trait that differentiates prescriptivism from purism lies in authority: while every speaker has the power to show purist attitudes towards a presumably ideal form of language, only a limited number of speakers with authority have the power to act by prescribing the use of that, purportedly ideal, form. Language laws are only one of the many forms of prescriptivism. Dictionary editors, grammar book authors and language instructors all have the means to demonstrate prescriptive tendencies. Thus, contrary to purism, prescriptivism is not inherently conservative; its norms can change (Walsh, 2014).

The goal of prescriptivism is to achieve a desired form, while language planning outlines how it can be done (Calvet, 2017). The French government's decree of 7 January 1972 calls for terminological commissions to propose new terms to denominate new realities, or to replace lexical borrowings, and to publish them in the government gazette *Journal officiel de la République française*. The Bas-Lauriol Law of 31 December 1975 prohibits all foreign expressions in administrative publications, if a localized terminology is available. The Toubon Law of 4 August 1994 establishes that all public information should be accessible in French, including advertisements, contracts and communications. Even though the Toubon Law applies solely to the government, and it does not concern private entities, it has caused a controversy in France. In an attempt to harmonize the work of multiple terminological commissions, the decree of 3 July 1996 brought to life the *Commission générale de terminologie et de néologie* ('General Commission of Terminology and Neology'), which was subsequently renamed *Commission d'enrichissement de la langue française* in the decree of 25 March 2015, while also altering the functions of the terminological commissions. The new decree specified that official terminology mandatorily replaced equivalent foreign expressions in state-issued publications, correspondences and documents. Additionally, its use is also obligatory in trade contracts dealing with public

missions and services in France. And yet, to what extent mandatory terminology becomes mainstream remains an open question. Official language laws, including words produced by the CELF and published in *Journal officiel*, are merely recommendations for the public; they are not mandatory. Therefore, it is crucial to examine under what circumstances prescriptive efforts succeed, if at all, and when and how they are able to replace foreign borrowings with French-only terminology. The present article uses quantitative methodology borrowed from corpus linguistics to try to arrive at such generalizations.

### 3. METHODOLOGY

#### 3.1. Corpus

The effect of prescription was examined using a combined variationist and corpus linguistic approach. The lexical items selected for this study were considered variables with two variants: the original English lexical borrowings and the French terms prescribed by the CELF. Occasionally, a third variant was also available, which was either a colloquial term or a prescribed term in Québec, but not in France. This third term was not included in the study. Measuring the use of both lexical items allows us to observe the recommended term's relative frequency as opposed to its absolute frequency, which produces the implantation coefficient (Quirion, 2003). These lexical variables all belonged to the semantic fields of computer science, information technology, audiovisuals or entertainment industry, which have been shown to be productive domains in the late twentieth and early twenty-first century in previous scholarship. A total of 54 terms were selected as target variables in this study from the official list of more than 7,600 recommended terms published on *FranceTerme*. The selected terms were distributed among the following domains: 12 from audiovisuals, 2 from electronics, 26 from computer science and 14 from telecommunication. In the domain of audiovisuals, the lexical items denoted concepts related to television and cinema (e.g. *blockbuster*, *prequel*, *trash television*, *trailer*) as well as video games (e.g. *gameplay*, *game designer*). Computer science was represented by words related to computers (e.g. *freeware*, *toner*), the Internet (e.g. *pop-up window*, *webcam*, *cloud computing*), services offered on the web (e.g. *triple play*, *bundle*), and electronic devices and their contents (e.g. *e-reader*, *e-book*). The domain of telecommunication included terms that belong to traditional telephone (e.g. *call center*, *SIM card*, *femtocell*) and online communication (e.g. *email*, *chat*). Two items came from the domain of electronics: *OLED* (screen technology) and *multitask*. There was considerable overlap between the domains of these terms on *FranceTerme*. For instance, *smartphone* was listed under computer science, but it could just as well have been part of telecommunication. Similarly, *tablet* figured among audiovisuals, but it could have also been grouped within computer science. This overlap was due to the prescription procedure: officially recommended translations can be requested by different terminological commissions at the same time, which means that the definition of a term and its grouping into a given domain might very well depend on the commission that reviews it first. For this reason, these domains should be considered informative rather than deterministic groupings. They were not designated as factors in subsequent analyses.

To translate *e-book*, two terms were recommended by the CELF at two different times. On 18 January 2005, *livre électronique* was proposed, but on 4 April 2012 two additional recommendations were announced to make semantic nuances more explicit: *livre numérique* to designate the electronic version of a book and *liseuse* to refer to 'e-reader'. Previously, these two meanings were both expressed by *livre électronique*; consequently, the 4 April 2012 decision rescinded the 18 January 2005 recommendation. Both terms were part of the present corpus, distinguished by the times of their recommendation. Such modifications underscore the flexibility of prescriptivism: in cases such as these, the terminological commissions decided to annul a previously prescribed term and recommend two new ones that they found better suited for multiple contexts of use.

Two types of newspaper sources were used based on their audiences: *La Croix*, *Le Figaro*, *Le Monde*, *Le Parisien*, *Les Echos* as newspapers with general audience and *01net* and *Journal du Net* as newspapers targeting audiences specialized in new Internet technologies. The first category comprised five daily national newspapers from France of wide distribution, all covering a variety of topics. All five titles are published in France, but most of them are also accessible internationally, which means that a significant portion of the French-speaking population around the world can read them. As mentioned previously, lexical borrowings can occur even in widely distributed, international journalistic writing. The newspapers *01net* and *Journal du Net* are published for specialized audiences, well-versed in new media and technologies, thus in these texts more English lexical borrowings were expected to occur. While all general audience newspapers are originally printed, it is essential to note that the specialized ones are primarily online publications. At the same time, being published primarily online would not affect the results as specialized professionals are assumed to regularly use the Internet.

The data were collected from Nexis Uni, an online database storing primarily journal articles, web-based publications and legal documents. Nexis Uni has limited access to certain selected newspaper titles in this corpus, which helped narrow the number of variables. Since the goal of this study was to observe the effect of top-down intervention on the probability of use of the prescribed terms, time was treated with a new approach, hitherto unattested in quantitative sociolinguistic studies. The frequencies of each term in the corpus (henceforth, term frequencies) were examined three years before and three years after the time of prescription to see whether a term had been attested before its official recommendation, and also whether prescription did or did not affect its probability of use after prescription. Thus, time was treated as a dynamic component: for each variable, time 0 marked the month of prescription and observation points expanded up to three years before and after time 0. For instance, *tutoriel* was prescribed for 'tutorial' in April 2007, thus the observed time period for this term stretched from April 2004 to April 2010.

### 3.2. Method

In the corpus, frequency equals the number of articles, even if the target term itself was repeated multiple times in the article. This approach allowed to control for articles that repeated a term and thus could have caused distortions in frequency

counts: it might have looked like a given variant is abundantly used, when in reality, it was only used in one publication. This counting method aligns well with measurements of dispersion intended in this study. A total of 47,241 articles were collected for the 54 variables (Table 1, Appendix). Prescribed variants occurred in 16,111 articles (34%), while non-prescribed variants figured in 31,130 articles (66%). The five most frequent variants were English terms *smartphone*, *email*, *big data*, and prescribed terms *courriel* ‘email’, *centre d’appels* ‘call center’. There were six more variants used in more than 1,000 articles: English terms *cloud computing* and *chat* and prescribed terms *vidéo à la demande* ‘video on demand’, *boîtier multiservice* ‘box’, *dégroupage* ‘unbundling’, and *livre numérique* ‘e-book’. Nine variants had zero frequency: English terms *core network*, *multitask*, *SIM card* and prescribed terms *microblogue* ‘microblog’, *capture de jeu* ‘performance capture’, *filoutage* ‘phishing’, *présuite* ‘prequel’, *scénariste-conseil* ‘script doctor’ and *écran fragmenté* ‘split screen’. Even though these nine terms did not occur in the corpus, their English or prescribed equivalents did – they were thus included in the model.

Furthermore, in order to examine the effect of prescriptivism over time, generalized additive models (GAM) were implemented, using GAM in the *mgcv* (Mixed GAM Computation Vehicle) package in R (Wood, 2017). Unlike traditional linear models that are fit to explore linear relationships between dependent and continuous independent variables, GAMs are preferred when linear relationships between variables cannot be assumed, for instance, due to nonlinear change over time and nested dependencies between variables (Winter and Wieling, 2016). In such cases, including the present corpus, rather than yielding ‘inferences’ based on raw frequency counts, GAMs provide ‘estimates’ about the ‘probability of use’ of dependent variables, and express those estimates as logistic values between 0 and 1. Thus GAM are considerably more flexible in uncovering hidden patterns in the data (Larsen, 2015). In this study, GAMs model the probability of use of each lexical item or variant, with the application value as the frequency count of the prescribed variant, measured against all other variants combined. The output of relative frequencies therefore corresponds to a six-year-long estimate of implantation coefficients. If prescriptivism does indeed have an effect on the probability of use of a prescribed variant after its prescription time, one would expect to see an increasing S-curve after prescription at time 0. As we shall see in the next section, some of the lexical variants examined in this study show such trajectories.

Several GAMs were tested by adding so-called ‘interactions’ and ‘smooths’ to the model. In the GAM evaluated as the best fit for the data, the logit probability of the frequency count of the prescribed variant is measured against the frequency counts of the other variants, with time added as a non-linear smooth term multiplied by the lexical items, while also adding lexical items and the newspaper source as main effects. Lexical items were included as *by* variables with the time smooth term in order to display individual patterns for the effect of prescription for each lexical item. This would not have been possible if lexical items had been used as a main effect. This method allowed a better understanding of how prescription affects each term over time to be gained. At the same time, since GAM can only take one *by* variable, newspapers were included as main effect, using mean estimates for each of them.

## 4. RESULTS

### 4.1. Summary tables

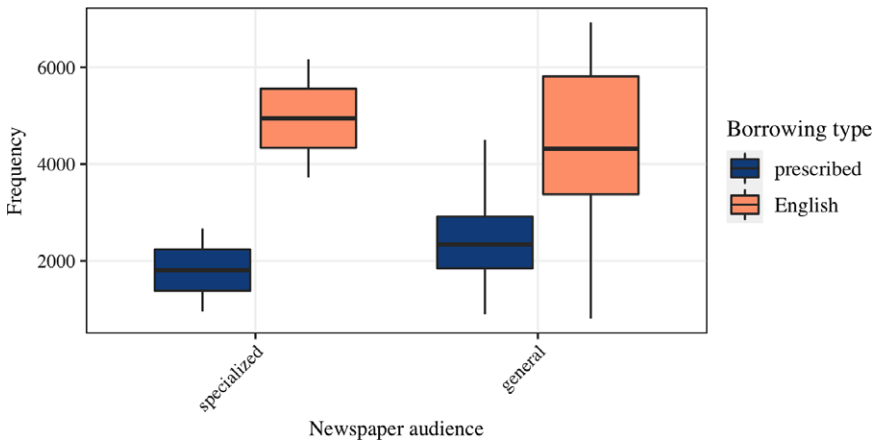
Summaries of parametric coefficients and approximate significance of smooth terms of the generalized additive models (GAM) for all 54 lexical variables and seven newspaper titles are shown in Tables 2 and 3 (Appendix). Parametric coefficients represent the main effects in the model for all lexical items and the estimated degrees of freedom in the summaries of the approximate significance of smooth terms indicate whether the smooth term is linear (close to 1), or non-linear (farther from 1). The measures ‘deviance explained’ (Table 3) is 62.7% and the adjusted  $r^2$  is 0.67, which suggests that the GAM used in this analysis represents a good model for the data. Although these summary tables contain p-values that traditionally indicate significance, in this study they are approximate and neglect smoothing parameter uncertainty. For this reason, a better interpretation of the results through the visualizations of probabilities of use patterns over time will be displayed in Figures 2 to 5 in the next sections, using the *visreg* (Breheny and Burchett, 2017) and the *ggplot2* packages in R (Wickham, 2009). The effects of initial use and of prescription from very low ( $> .0$ ) to very high ( $> .75$ ) for each term are summarized in Table 4 (Appendix).

### 4.2. Variation among newspapers

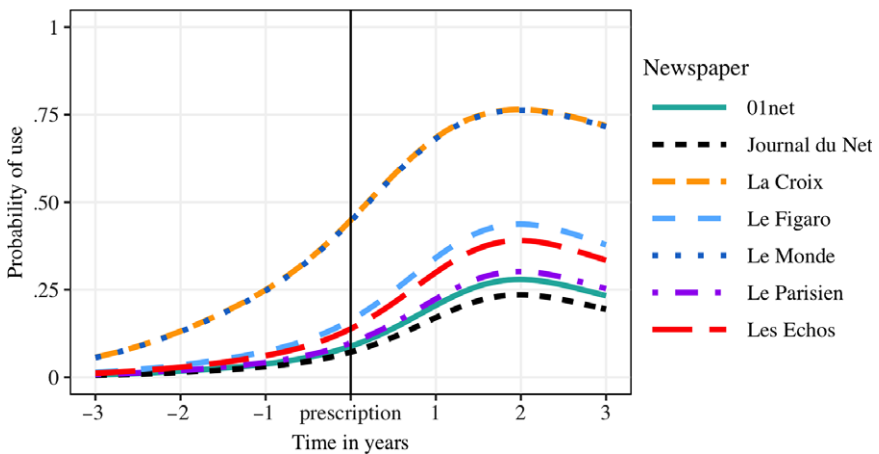
The newspaper sources show considerable variation in the use of the prescribed terms and their English equivalents (Figure 1). In the corpus of 47,241 articles, 33,732 are from general newspapers and 13,509 articles are from specialized newspapers. Prescribed variants occurred less frequently in the newspapers specialized in new Internet technologies, which thus confirms the hypothesis that these sources would be likely to choose the original lexical borrowing because their audience is exposed to the new realities of the tech industry earlier than the audience of the general newspapers. Figure 1 also suggests that prescription has little effect on terminology implantation in *Journal du Net* and *01net*. In the case of the general newspapers, prescribed terminology also figures less frequently, although with increased variability among general newspaper titles. The distribution of the variables in these publications indicates that the conservative leaning Christian newspaper *La Croix* is particularly apt to use official terminology, surpassing slightly the use of English terms. *Le Monde* is also close to an even distribution, which is particularly interesting because it has the highest international readership among French national newspapers (BOOK Presse Grand Public 2016, 2017). One would assume that international expectations to comply with the standard written norms affects the use of prescribed terminology in a newspaper such as *Le Monde*, but this is not the case. Other national newspaper titles clearly prefer English variants in over two-thirds of all cases, with *Le Parisien* using the least prescribed terminology.

Figure 2 depicts the results of GAMs applied to all prescribed terminology combined over time in each newspaper. In this and all subsequent GAM visualizations, estimated probability values are shown on the vertical axis and time (in years) before and after the time of prescription appears on the horizontal axis. Since in the GAMs, newspaper sources were included as main effect and not a *by*





**Figure 1.** Distribution of prescribed vs. English lexical variants in general and specialized newspapers.



**Figure 2.** Estimated probabilities of use of prescribed terminology three years before and three years after prescription in each newspaper.

variable of the smooth term, curves represent probabilities of use rather than frequency counts. The initial and the final probabilities of use across the observed time period correspond to the starting points and endpoints of each line shown in different colour lines. They allow us to evaluate whether a prescribed term had been used or not before its official prescription by the CELF (high or low values on the bottom left) and whether its probability of use increased or decreased over time (high or low values on the top right).

In Figure 2, the first observation to make is the distinct patterns of *La Croix* and *Le Monde* from the other newspapers, overlapping yellow and blue curves in the top of the figure, indicating that the estimated probabilities of use of prescribed terminology in both newspapers reach 0.72 by the end of the third year after

prescription. Furthermore, even though prescribed terms are more or less frequent before prescription, official recommendation does boost their probability of use for at least a few years afterwards. The result, in each case, is a S-curve-like diffusion that is slow at first, but accelerated after the time of prescription and throughout most of the observed time period. There is a stable increase in the estimated probability of use reaching 0.38 in *Le Figaro*, 0.33 in *Les Echos* and 0.25 in *Le Parisien*. *Le Figaro* has a comparable size of readership to *Le Monde*, but it is considered more conservative, which might explain its tendencies to adopt prescribed terminology at a higher probability than other news sources, such as *Les Echos*.

The lowest increase among national newspapers with a general audience in Figure 2 is shown by *Le Parisien*. This could be due to its more colloquial news profile that would entail that lexical choices in *Le Parisien* tend to mirror everyday speech more closely, including the recurring use of newly adopted foreign borrowings for certain concepts. Although their estimated probabilities of use are on comparable paths to the general newspapers, the two specialized newspapers, *01net* and 0.19 in *Journal du Net*, show more moderate uptakes over time. By the end of the third year after prescription, the model estimates a probability of use of officially recommended variants at 0.23 in *01net* and 0.19 in *Journal du Net*. This suggests that these sources also employ prescribed variants, but less frequently than the other sources.

### 4.3. Positive effects of prescription

Figure 3 comprises terms that had the strongest increase in probability of use at one point after prescription. The term *zone d'accès sans fil*, recommended in 2005 for *hotspot*, shows one of the lowest estimated initial and final probabilities of use. Interestingly, its use in newspapers increased slightly at the time of prescription, but real uptake happened only at the end of the observed time period. *Courriel*, another term with low initial probability of use, on the other hand, shows a classic S-curve like diffusion, with just a slight decline at the end of the observed time period. The uptake of this term might have been boosted by its official recommendation for 'email' in Québec prior to its prescription in France. Among lowest initial uses on this figure belong to *rétrolien* and *multiplexeur*. The former, recommended for *trackback* and denoting a linking system between websites, is a loan creation based on *rétro-* and *lien* from 2010, while the latter, prescribed in 2003, is an assimilated loanword echoing the verb *multiplexer*, where the suffix *-er* was adapted as *-eur*. Both terms show a nearly identical and almost linear increase from low to very high probabilities of use in six years.

Among the terms that show strong effects of prescription, *i.e.* a boost of more than 0.50 in probability of use over time (Table 4, Appendix), *liseuse* and *tablette* had medium initial usage. *Liseuse* was prescribed in 2012 when the CELF revised its recommendation for *e-book*. Originally, in 2005, *livre électronique* was recommended as both the device and the digital copy of the book, which was annulled by the 2012 decision specifying that *e-reader* is ought to be translated as *liseuse*, a semantic loan that borrows the meaning of *electronic reader* to the already existing word *liseuse*. *Tablette* is a semantic loan from 2011 that contributes new meaning to the existing French word for 'tablet' (as in 'a shelf', 'pharmaceutical tablet' or 'clay tablet'), and now it also refers to 'tablet computer' as in

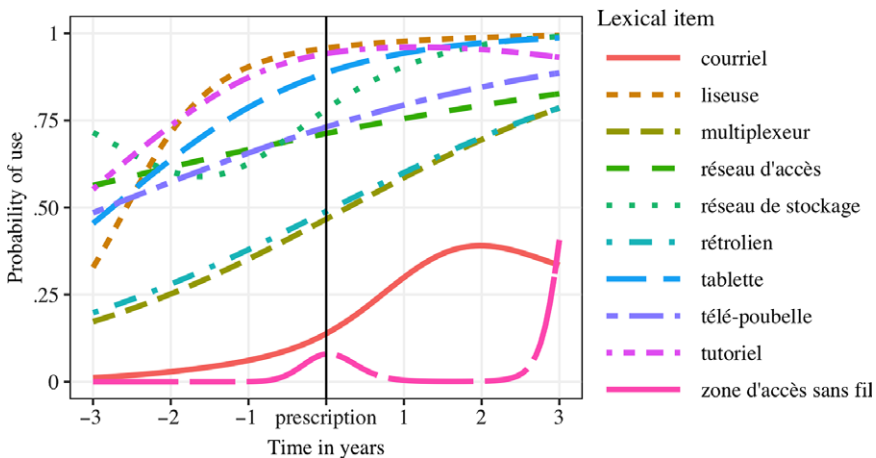


Figure 3. The estimated probability of use for terms showing strong and medium effects of prescription.

English. Similarities in form and word length observed for *tablette* ‘tablet’ might also have contributed to the strong increase in the probability of use of the prescribed term *tutoriel*, recommended in 2007. Denoting to the training section of a game or a software, the term is based on the English *tutorial*, which derives from the English verb *to tutor*. Borrowing the adjectival suffix *-el* to transform the English borrowing into a quasi-cognate in French, *tutoriel* attests to the creative use of morphophonology in new terminology creation.

Six prescribed terms demonstrated medium effect of prescription, indicating an at least 0.25 increase by the end of the observed time period: *courriel*, *réseau d'accès*, *réseau de stockage*, *télé-poubelle*, *tutoriel* and *zone d'accès sans fil*. The highest initial and final usage is associated with *réseau de stockage*, prescribed for *storage area network* in 2010. This loan rendition does not include every element of the original English compound word because it omits *area*. Additionally, even though both *réseau* and *stockage* correspond to *network* and *storage*, respectively, *stockage* itself is an assimilated loanword from the English *stock*. Thus, it seems that *stockage* has already been so entrenched in French, that the CELF did not object to the use of this assimilated loan hence calquing a new term. *Réseau d'accès* and *télé-poubelle*, two other terms with high estimated initial and final probabilities of use, are a loan translation for *access network* from 2003 and a loan creation for *trash TV* from 2010. The latter is not considered a loan translation because *télé-* is truncated from *télévision*, making it more of a neologism than a translation, as the prefix itself means *far* in English. Similar to other terms with high probabilities of use, *réseau d'accès* and *télé-poubelle* also show a steady linear increase over time.

#### 4.4. Weak to no effects of prescription

Prescribed terms for which prescription had only a weak effect, i.e., an increase of <0.25 in the probability of use over time (Table 4, Appendix), are shown in Figure 4.

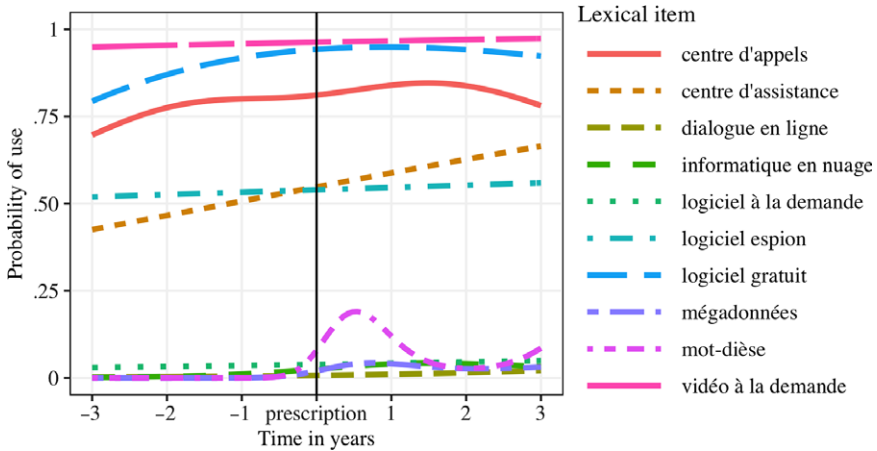
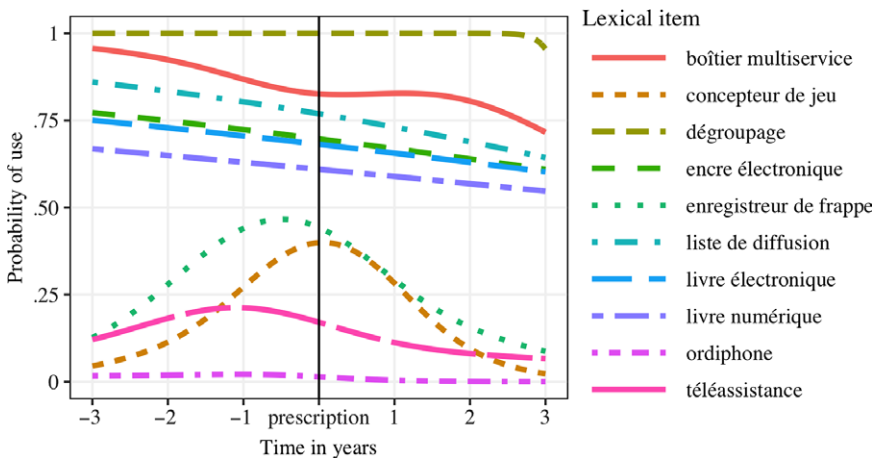


Figure 4. The estimated probability of use of the prescribed terminology three years before and three years after prescription with weak effects of prescription.

*Centre d'appels*, *centre d'assistance*, *logiciel espion*, *logiciel gratuit* and *vidéo à la demande*, all with  $>0.4$  initial probabilities of use, the display medium to high initial usage. *Centre d'appels* is a loan translation prescribed in 2003 to replace *call center*. Perhaps the relative success of the prescribed variant lies in its form: it clearly refers to the same concept as the English borrowing, with the added advantage of being short as well. While the loan rendition *centre d'assistance*, prescribed in 2007, differs semantically from its English equivalent, *help desk*, it can help users link its referential meaning to a family of prescribed terms denoting locations rather than means of assistance. The jury is still out on *logiciel espion*, a compound loan translation from 2007, that is relatively unlikely to replace *spyware* based on its probability of use over the six-year period.

Semantic similarities with the already entrenched *logiciel* might be the reason for the high probability of use of *logiciel espion*, recommended for *spyware*. The same type of similarity might be at play for *logiciel gratuit* that appears to render well its English variant *freeware*, a portmanteau composed of *free* and *software*. The last term of this group is the prescribed variant *vidéo à la demande*, a loan translation of *video on demand*. This term, recommended in 2006, mirrors the original English compound, very likely contributing to its success in probability of use.

Among the terms with the weakest effects of prescription and very low initial use (see Table 4, Appendix), *mot-dièse*, is emblematic on social media. It is the prescribed term for *hashtag*, the pound symbol, used mainly on Twitter to tag words or expressions with a searchable link to express belonging to a topic. The word *hashtag* is never explicitly written out when the term is used: only the symbol #. However, when the discourse is about hashtag, the word is spelled out. *Mot-dièse* was recommended in 2013 to replace the original English word, but based on Figure 4, its probability of use did not have a remarkable uptake.



**Figure 5.** The estimated probability of use of the prescribed terminology three years before and three years after prescription with weak negative effects of prescription.

The official term is a neologism, a loan rendition of *hash* by *dièse* (as in the musical symbol for ‘sharp’) and *tag* by *mot*. Even though it clearly refers to the symbol #, and it is as short as *hashtag*, its only uptake in France seems to have been due to the initial discussions that it sparked for a short time after its prescription. Its probability of use at year 3 of the observed period was among the lowest.

Terms for which final usage displayed no change or ended up even lower than their initial estimated probability of use are shown in Figure 5. One prescribed term in this group that has had virtually no attested initial usage and prescription did not affect this outcome either is *ordiphone* for ‘smartphone’. Other recommended variants, such as *concepteur de jeu* ‘game designer’, *téléassistance* ‘hotline’, and *enregistreur de frappe* ‘keylogger’ show medium to high estimated probability of use before prescription followed by decrease after prescription. Thus, even though the CELF adopted an already attested term as official terminology, prescription did not seem to help sustain the use of these promising terms. It is possible that newspapers reacted to the newly available terminology, but as time went by, they reverted to the English variants. Although the French *concepteur de jeu*, officially recommended by the CELF in 2011 for *game designer*, is a semantically transparent loan translation, *game* and *designer*, are well-attested Anglicisms in French, which might have contributed to the apparent success of their English compound over the officially preferred French term. *Smartphone* represents a particular case because the word is composed of the suffix *-phone*, which is natively used in French just as much as in English to denote ideas related to the telephone. The recommended term from 2009 itself, *ordiphone*, contains the same suffix. The prefix *smart-*, which refers to the ‘intelligence’ of the device compared to its predecessors, is quite productive in similar expressions, such as *smart TV*, *smart fridge*, *smart home*, etc. It is replaced by the prefix *ordi-*, which originates from the well-established prescribed term: *ordinateur*. And yet, despite the identical length and the similarity with other nativized expressions, *ordiphone* was clearly not preferred over *smartphone*.

*Boîtier multiservice* ‘box’, *dégroupage* ‘unbundling’, *encre électronique* ‘e-ink’, *liste de diffusion* ‘mailing list’, *livre électronique* ‘e-book’ and *livre numérique* ‘liseuse’ is among prescribed terminology with high (>.5) to very high (>.75) initial usage, whose probability of use either has not changed decreased over time. Yet again, as suggested in the visualization in Figure 5, the CELF adopted terms that were attested in previous usage, but prescription had no effect on their estimated probability of use over time. The term *dégroupage*, referring to the opening-up of a local telephone network to competition and widely used prior to its prescription as preferred to ‘unbundling’, was the only widely adopted variant in the corpus whose 100% probability of use was sustained at such levels for the observed period.

*Boîtier multiservice* was recommended in 2007 to replace ‘box’. This term attributes new meaning to *boîtier*, referring to a device through which multiple telecommunication services (e.g. telephone line, internet, television) can be accessed rather than to a particular service in the cloud. The recommendation by the CELF explicitly adds to the definition of *boîtier multiservice* the prohibition of using the English word *box*, which appears frequently in brand names. The English variant ‘mailing list’ differs somewhat from its prescribed counterpart *liste de diffusion*, which was created through loan rendition: while *liste* is equivalent to *list* in English, *diffusion* diverges from *mailing*. Nonetheless, recommended in 2007, *liste de diffusion* is successfully implanted probably thanks to its semantic similarities shared with ‘mailing list’.

The last term in this group unsuccessful terms, *livre numérique*, is a particular case because it represents only one of the official terms replacing ‘e-book’. In 2005, the CELF recommended *livre électronique*, a loan translation of *electronic book* subsequently shortened to ‘e-book’, as replacements for both the electronic device and the digital edition of a book. However, to further clarify the distinction between *e-book* and the device *e-reader*, the CELF decided to change its recommendation in 2012 and propose *livre numérique* for the book, and *liseuse* for the device. This proves the flexibility of prescriptivism, as opposed to purism: while the latter aims at preserving a state of the language perceived as prestigious, prescriptivism endorses change in order to provide language users with forms based on established consensus.

## 5. CONCLUSION

Table 4. (Appendix) presents the summary of initial use and effect of prescription for all 54 prescribed terms examined in this study. The diffusion of words with medium or low initial usage shows the largest effects of official recommendation, which suggests that top-down lexical change is the most successful when prescription follows usage. However, this general tendency has to be nuanced. Recommended terms with high probability of use throughout the observed period revealed that whenever the CELF endorsed a variant already well attested in usage, prescription did nothing more than endorsing a *fait accompli*: lexical change through the replacement of the English borrowing had already taken place and its statement by an official act seemed superfluous. However, when prescription followed the incipient uptake of a recommended term, i.e., it took

place during the initial phases of diffusion, official recommendation could have boosted the probability of use of the new term. Such strong increases (that is, a high probable effect of prescription) characterized variants with previously attested medium initial usage. These variables included *tablette* 'tablet', *liseuse* 'e-reader', *rétrolien* 'trackback' and *multiplexeur* 'multiplexer', and with a moderate effect *courriel* 'email', *réseau d'accès* 'access network', *télé-poubelle* 'trash tv', *tutoriel* 'tutorial'. What do these variants have in common? The answer seems straightforward: their length compares favorably to that of their English equivalents, the two share obvious semantic similarities or the recommended terms are direct translations of the English borrowings, with a clear reference to meaning. And yet, there are many prescribed terms characterized by the same similarities that do not enjoy a comparable increase in probability of use. Among the examples was *ordiphone* that could have indeed replaced 'smartphone', as well as *microblogue* 'microblog', *femtocellule* 'femtocell', and so on.

Other terms with weak effect of prescription display merely noticeable increase, yet they constitute change compared to initial usage. These included *dialogue en ligne*, *informatique en nuage*, *logiciel à la demande*, *mégadonnées* and *mot-dièse*. *Big data* is a relatively recent buzzword to characterize a great amount of available information, and it was meant to be replaced by *mégadonnées* in 2014. Even though the official term possesses the same elements of the compound word with the prefix *méga-* equaling *big*, *big data* seems to be too well entrenched in journalistic texts. *Cloud computing* has also gained popularity thanks to cloud services such as *Dropbox* or *iCloud*. While *nuage* is a semantic loan prescribed in 2010 and providing an already existing word with new meaning based on its English borrowing, it does not compare with the use of *cloud*. In spite of the slight uptake of *logiciel à la demande* after its official recommendation in 2013 for 'software as a service', its probability did not significantly increase. In fact, 'software as a service' was affected by word formation processes that seem to have propelled its abbreviated form *SaaS* to greater use, becoming brand-like attributes to its respective service. The prescribed variant was not able to achieve the same level of recognition.

Certain lexical terms seemed to have been utterly unaffected by prescription, following their own trajectory, no matter how popular they were before they were officially recommended. Among them were prescribed variants used with high initial probability in specialized contexts, such as *dégroupage* 'unbundling' and *encre électronique* 'e-ink' and *enregistreur de frappe* 'keylogger' that did not demonstrate an uptake soon after the prescription time.

As far as the actual frequency counts of the two lexical variants were concerned, overall all journalistic publications demonstrated some uptake in probability of use of prescribed terms over time, but in varying degrees. Generally speaking, specialized newspapers presented the lowest probability of use of prescribed terminology: their lexical preferences, even if initially affected by prescription, remained favorable to English. General newspapers, on the other hand, gladly complied, in particular *La Croix* and *Le Monde* that displayed the strongest compliance with the new prescriptive lexical norms. There was no overwhelming evidence from this corpus that political preferences could influence the use of prescribed terminology. Future studies of more extensive newspaper sources

could focus on the publication and editorial processes rather than generalizations based on political affiliation.

Although this study has identified several promising general tendencies, it still remains unclear what led to increase in probability of use. In earlier versions of the generalized additive model the type of borrowings (e.g. calque, loan translation, loan rendition) was used as a main effect, but it did not yield any significant effect. Additionally, the time period based on the year of prescription was also included to investigate why terms such as *ordinateur* for ‘computer’ and *écran* for ‘screen’ were successful, and *ordiphone* for ‘smartphone’ and *écran fragmenté* for ‘split screen’ were not. However, the model’s estimates were not improved by the time period, which means that post-2000 official recommendations behaved in a similar fashion than earlier recommendations.

In conclusion, targeted frequency counts and generalized additive models revealed that there is considerable variation among the prescribed variants with regards to the uptake of previously more or less well-entrenched terminology after prescription time. At this point, the question is whether we can accept that each prescribed variant has a unique story to tell, and there is no recipe for their successful implantation in use. If that were the case, one would not have advanced any more in our discoveries of ‘general processes of change in the lexicon’ (see Posner’s comment in the Introduction). However, it seems that the results of this study are encouraging enough to make three general observations. The first is that prescription is most effective when it endorses the incipient change of a previously attested term. The second is that for lexical variants that already enjoy high probability of use, prescription is largely superfluous, although it can conceivably provide legitimacy for continued use. For lexical variants with low or medium probability of use, success is not foreseeable; their individual histories could be preferable, indeed.

## REFERENCES

- Bernard-Béziade, M., and Attali, M. (2012). L’évidence des anglicismes dans les discours journalistiques français: Le cas du sport. *Language Problems and Language Planning*, 36.1: 120–135. <https://doi.org/10.1075/lplp.36.2.02ber>
- Blackwood, R. (2013). French, language policy and new media/Französisch, Sprachpolitiken und neue Medien/Le français, la politique linguistique et les nouveaux média. *Sociolinguistica*, 27.1. <https://doi.org/10.1515/soci.2013.27.1.37>
- BOOK *Presse Grand Public 2016* (BOOKS ACPM/OJD). (2017). L’Alliance pour les chiffres de la presse et des médias. [https://www.acpm.fr/var/ojd/storage/files/books\\_pdf/A/OJD\\_BOOK\\_presse-grand-public-2016\\_195.pdf](https://www.acpm.fr/var/ojd/storage/files/books_pdf/A/OJD_BOOK_presse-grand-public-2016_195.pdf)
- Breheny, P., and Burchett, W. (2017). Visualization of regression models using visreg. *The R Journal*, 9.2: 56. <https://doi.org/10.32614/RJ-2017-046>
- Calvet, L.-J. (2017). *La sociolinguistique*. Paris: Que sais-je ?
- Candel, D. (2015). Vers une didactique de l’enrichissement du français (le processus du vocabulaire officiel). In: M. Lebreton (ed.), *La didactique des langues et ses multiples facettes: Mélanges offerts à Jacqueline Feuillet professeur émérite, université de Nantes*. Paris: Riveneuve.
- Candel, D. (2017). Prescription and tradition: From the French Dictionnaire de l’Académie to the Official French Language Enrichment Process (1996-2014). In: I. Tiekens-Boon van Ostade and C. Percy (eds), *Prescription and Tradition in Language: Establishing Standards across Time and Space*. Clevedon: Multilingual Matters, pp. 273–287.



- Chansou, M.** (2003). *L'aménagement lexical en France pendant la période contemporaine: (1950—1994) ; étude de sociolexicologie*. Paris: Champion.
- Chaput, L.** (2010). La variation stylistique en tant que procédé discursif dans les blogues journalistiques. In M. Burger (ed.), *Les médias et le politique: Actes du colloque "Le français parlé dans les médias", Lausanne, 1-4 septembre 2009*. Centre de linguistique et des sciences du langage.
- Chesley, P. and Baayen, R. H.** (2010). Predicting new words from newer words: Lexical borrowings in French. *Linguistics*, **48.6**. <https://doi.org/10.1515/ling.2010.043>
- Kim, M.** (2015). L'anglicisme et l'intervention linguistique gouvernementale: Étude de l'implantation terminologique en Francophonie. *ELIS - Échanges Linguistiques En Sorbonne*, **3.3**: 79–102.
- Kim, M.** (2018). *Variation terminologique en francophonie. Élaboration d'un modèle d'analyse des facteurs d'implantation terminologique* [Paris IV Sorbonne]. <https://tel.archives-ouvertes.fr/tel-01859665>
- Langer, N. and Nesse, A.** (2012). Linguistic purism. In: J. M. Hernández Campoy and J. C. Conde Silvestre (eds), *The Handbook of Historical Sociolinguistics*. Oxford: Wiley-Blackwell.
- Larsen, K.** (30 July 2015). *GAM: The Predictive Modeling Silver Bullet*. <https://multithreaded.stitchfix.com/blog/2015/07/30/gam/>
- Martin, E.** (2007). "Frenghish" for sale: Multilingual discourses for addressing today's global consumer. *World Englishes*, **26.2**: 170–190. <https://doi.org/10.1111/j.1467-971X.2007.00500.x>
- Ostyn, S., Lienart, N. and Klein, J.-R.** (1997). *L'anglicisme et la presse: Enquête et analyse à travers quatre quotidiens français et belges*. <https://doi.org/10.5169/SEALS-399970>
- Poplack, S.** (2018). *Borrowing: Loanwords in the Speech Community and in the Grammar*. Oxford: Oxford University Press.
- Poplack, S. and Dion, N.** (2009). Prescription vs. praxis: The evolution of future temporal reference in French. *Language*, **85.3**: 557–587. <https://doi.org/10.1353/lan.0.0149>
- Poplack, S., Lealless, A. and Dion, N.** (2013). The evolving grammar of the French subjunctive. *Probus*, **25.1**: 139–195. <https://doi.org/10.1515/probus-2013-0005>
- Posner, R.** (1997). *Linguistic Change in French*. Oxford: Clarendon Press.
- Quirion, J.** (2003). Methodology for the design of a standard research protocol for measuring terminology usage. *Terminology*, **9.1**: 29–49. <https://doi.org/10.1075/term.9.1.03qui>
- Quirion, J.** (2011). Dynamique terminologique et terminométrie: Une complémentarité nécessaire pour le suivi de l'évolution de l'usage des termes et des connaissances. *Terminology*, **17.1**: 113–133. <https://doi.org/10.1075/term.17.1.07qui>
- Rapport annuel 2017 de la Commission d'enrichissement de la langue française** (p. 156). (2019). Délégation générale à la langue française et aux langues de France. <https://www.culture.gouv.fr/Sites-thematiques/Langue-francaise-et-langues-de-France/Politiques-de-la-langue/Enrichissement-de-la-langue-francaise/Le-dispositif/Rapport-annuel-de-la-Commission-d-enrichissement-de-la-langue-francaise-2017>
- Saint, E. C.** (2015). La terminologie de référence au service de la langue française au Québec et son usage dans le domaine de l'informatique. *Revue de l'Université de Moncton*, **44.2**: 167–198. <https://doi.org/10.7202/1031004ar>
- Saugera, V.** (2017). *Remade in France: Anglicisms in the lexicon and morphology of French*. Oxford: Oxford University Press.
- Thomas, G.** (1991). *Linguistic Purism*. London: Longman.
- Thomason, S. G. and Kaufman, T.** (1991). *Language Contact, Creolization, and Genetic Linguistics*. Berkeley: University of California Press.
- Varga, D., Orešković Dvorski, L. and Bjelobaba, S.** (2011). English loanwords in French and Italian daily newspapers. *Studia Romanica et Anglica Zagabiensia*, **56**: 71–84.
- Walsh, O.** (2014). 'Les anglicismes polluent la langue française'. Purist attitudes in France and Quebec. *Journal of French Language Studies*, **24.3**: 423–449. <https://doi.org/10.1017/S0959269513000227>
- Walsh, O.** (2016). *Linguistic Purism: Language Attitudes in France and Quebec*. Amsterdam: John Benjamins.
- Wickham, H.** (2009). *Ggplot2: Elegant graphics for data analysis*. Springer.
- Winter, B. and Wieling, M.** (2016). How to analyze linguistic change using mixed models, Growth Curve Analysis and Generalized Additive Modeling. *Journal of Language Evolution*, **1.1**: 7–18. <https://doi.org/10.1093/jole/lzv003>
- Wood, S. N.** (2017). *Generalized Additive Models: An Introduction with R* (2nd edition). Boca Raton: CRC Press.

**Table 1.** Summary of article counts by lexical item and type used for the newspaper analysis

Lexical item	Type	n	Lexical item	Type	n	Lexical item	Type	n
access network	English	21	gameplay	English	308	script doctor	English	10
<i>réseau d'accès</i>	prescribed	49	<i>jouabilité</i>	prescribed	78	<i>scénariste-conseil</i>	prescribed	0
banner	English	2	hashtag	English	655	scripted reality	English	42
<i>bannière</i>	prescribed	753	<i>mot-dièse</i>	prescribed	77	<i>réalité scénarisée</i>	prescribed	5
big data	English	2,676	helpdesk	English	55	SIM card	English	0
<i>mégadonnées</i>	prescribed	99	<i>centre d'assistance</i>	prescribed	55	<i>carte SIM</i>	prescribed	667
blockbuster	English	866	hotline	English	699	smartphone	English	9,657
<i>grosse machine</i>	prescribed	175	<i>téléassistance</i>	prescribed	107	<i>ordiphone</i>	prescribed	47
box	English	328	hotspot	English	296	software as a service	English	242
<i>boîtier multiservice</i>	prescribed	1,451	<i>zone d'accès sans fil</i>	prescribed	3	<i>logiciel à la demande</i>	prescribed	8
bundle	English	48	keylogger	English	22	split screen	English	47
<i>offre groupée</i>	prescribed	69	<i>enregistreur de frappe</i>	prescribed	7	<i>écran fragmenté</i>	prescribed	0
call center	English	499	mailing list	English	23	spyware	English	249
<i>centre d'appels</i>	prescribed	2,011	<i>liste de diffusion</i>	prescribed	73	<i>logiciel espion</i>	prescribed	212
chat	English	1,071	microblog	English	200	storage area network	English	15
<i>dialogue en ligne</i>	prescribed	17	<i>microblogue</i>	prescribed	0	<i>réseau de stockage</i>	prescribed	29
cloud computing	English	1,912	movie bootleg	English	1	tablet	English	70
<i>informatique en nuage</i>	prescribed	64	<i>film piraté</i>	prescribed	76	<i>tablette</i>	prescribed	650
core network	English	0	multiplexer	English	16	toner	English	46
<i>réseau d'infrastructure</i>	prescribed	31	<i>multiplexeur</i>	prescribed	13	<i>encre en poudre</i>	prescribed	4

(Continued)

Table 1. (Continued.)

Lexical item	Type	n	Lexical item	Type	n	Lexical item	Type	n
e-book	English	50	multiplexing	English	30	trackback	English	7
<i>livre électronique</i>	prescribed	133	<i>multiplexage</i>	prescribed	101	<i>retrolien</i>	prescribed	6
e-book (2)	English	622	multitask	English	0	trailer	English	149
<i>livre numérique</i>	prescribed	1,093	<i>multitâche</i>	prescribed	298	<i>bande-annonce</i>	prescribed	773
e-ink	English	38	OLED	English	240	trash TV	English	10
<i>encre électronique</i>	prescribed	73	<i>DELO</i>	prescribed	8	<i>télé-poubelle</i>	prescribed	27
e-mail	English	7,322	performance capture	English	40	triple play	English	858
<i>courriel</i>	prescribed	2,586	<i>capture de jeu</i>	prescribed	0	<i>triple service</i>	prescribed	2
e-reader	English	72	phishing	English	354	tutorial	English	8
<i>liseuse</i>	prescribed	890	<i>filoutage</i>	prescribed	0	<i>tutoriel</i>	prescribed	58
femtocell	English	32	pop-up (window)	English	139	unbundling	English	2
<i>femtocellule</i>	prescribed	3	<i>fenêtre intrusive</i>	prescribed	1	<i>dégroupage</i>	prescribed	1,416
freeware	English	34	prequel	English	60	video on demand	English	59
<i>logiciel gratuit</i>	prescribed	305	<i>présuite</i>	prescribed	0	<i>vidéo à la demande</i>	prescribed	1,480
game designer	English	31	reverse engineering	English	13	webcam	English	884
<i>concepteur de jeu</i>	prescribed	10	<i>retro-ingénierie</i>	prescribed	12	<i>cybercaméra</i>	prescribed	6
TOTAL							English	31,130
							prescribed	16,111
							N	47,241

**Table 2.** Summary of the parametric coefficients in the GAM for the newspaper corpus

	Estimate	Standard Error	z-value	p-value	
Intercept (réseau d'accès,01net)	0.51	0.34	1.5	0.14	
bannière	5.51	0.97	5.68	< 0.001	***
mégadonnées	-7.17	1.03	-6.94	< 0.001	***
grosse machine	-3.08	0.35	-8.76	< 0.001	***
boîtier multiservice	0.67	0.34	1.94	0.05	.
offre groupée	-0.8	0.4	-1.99	0.05	*
centre d'appels	0.45	0.34	1.32	0.19	
dialogue en ligne	-5.75	0.42	-13.55	< 0.001	***
informatique en nuage	-4.87	0.47	-10.47	< 0.001	***
réseau d'infrastructure	704.07	12976370	0	1	
livre électronique	-0.31	0.38	-0.81	0.42	
livre numérique	-0.61	0.34	-1.78	0.08	.
encre électronique	-0.24	0.41	-0.59	0.55	
courriel	-2.63	0.34	-7.72	< 0.001	***
liseuse	2.19	0.43	5.14	< 0.001	***
femtocellule	-25.93	35.56	-0.73	0.47	
logiciel gratuit	1.56	0.39	4	< 0.001	***
concepteur de jeu	-2.68	0.58	-4.63	< 0.001	***
jouabilité	-2.28	0.38	-6.06	< 0.001	***
mot-dièse	-7.53	2.1	-3.59	< 0.001	***
centre d'assistance	-0.74	0.41	-1.82	0.07	.
téléassistance	-2.97	0.36	-8.33	< 0.001	***
zone d'accès sans fil	-8.49	3.69	-2.3	0.02	*
enregistreur de frappe	-2.08	0.65	-3.18	< 0.001	***
liste de diffusion	0.09	0.43	0.21	0.83	
microblogue	-706.91	7931469	0	1	
film piraté	6449.25	25961.3	0.25	0.8	
multiplexeur	-0.91	0.54	-1.7	0.09	.
multiplexage	0.16	0.42	0.39	0.7	
multitâche	704.21	3916310	0	1	
DELO	-4.66	0.5	-9.34	< 0.001	***
capture de jeu	-706.44	11833607	0	1	
filoutage	-705.88	3593203	0	1	

(Continued)

Table 2. (Continued.)

	Estimate	Standard Error	z-value	p-value	
fenêtre intrusive	-4058.87	24121.78	-0.17	0.87	
présuite	-706.7	10063858	0	1	
rétro-ingénierie	-1.56	0.64	-2.46	0.01	*
scénariste-conseil	-707.15	25451365	0	1	
réalité scénarisée	-8.89	2.34	-3.81	< 0.001	***
carte SIM	704.11	2978006	0	1	
ordiphone	-6.27	0.38	-16.52	< 0.001	***
logiciel à la demande	-4.18	0.52	-8.01	< 0.001	***
écran fragmenté	-706.99	10384034	0	1	
logiciel espion	-0.84	0.36	-2.36	0.02	*
réseau de stockage	1.07	0.83	1.28	0.2	
tablette	1.4	0.37	3.81	< 0.001	***
encre en poudre	-388.83	1116.53	-0.35	0.73	
rétrolien	-0.84	0.71	-1.18	0.24	
bande-annonce	1.2	0.36	3.28	< 0.001	***
télé-poubelle	0.15	0.64	0.24	0.81	
triple service	-5051.96	7488.82	-0.67	0.5	
tutoriel	1.46	0.58	2.53	0.01	*
dégroupage	20.13	11.46	1.76	0.08	.
vidéo à la demande	2.3	0.36	6.31	< 0.001	***
cybercaméra	-7.44	1.49	-4.98	< 0.001	***
Journal du Net	-0.23	0.08	-2.84	0.01	**
La Croix	2.13	0.08	25.32	< 0.001	***
Le Figaro	0.7	0.07	10.31	< 0.001	***
Le Monde	2.12	0.07	30.48	< 0.001	***
Le Parisien	0.11	0.07	1.67	0.1	.
Les Echos	0.5	0.06	7.83	< 0.001	***

Significance codes: 0 (\*\*\*\*) 0.001 (\*\*\*). 0.01 (\*\*). 0.05 (\*). 0.1 (.) 1

**Table 3.** Approximate significance of smooth terms in the GAM for the newspaper

	Effective Degree of Freedom	Reference Degrees of Freedom	$\chi^2$	p-value	
s(lexical: réseau d'accès)	1	1	1.44	0.23	
s(lexical: bannière)	1.8	2.22	1.98	0.44	
s(lexical: mégadonnées)	2.94	3.14	12.02	0.02	*
s(lexical: grosse machine)	1.02	1.04	0	0.98	
s(lexical: boîtier multiservice)	3.09	3.57	47.05	< 0.001	***
s(lexical: offre groupée)	2.15	2.63	11.01	0.01	**
s(lexical: centre d'appels)	3.22	3.67	19.22	< 0.001	***
s(lexical: dialogue en ligne)	1	1	4.89	0.03	*
s(lexical: informatique en nuage)	2.13	2.55	4.35	0.15	
s(lexical: réseau d'infrastructure)	1	1	0	1	
s(lexical: livre électronique)	1	1.01	2.05	0.15	
s(lexical: livre numérique)	1	1.01	6.17	0.01	*
s(lexical: encre électronique)	1	1.01	1.59	0.21	
s(lexical: courriel)	3.77	3.97	1232.56	< 0.001	***
s(lexical: liseuse)	2.2	2.69	76.88	< 0.001	***
s(lexical: femtocellule)	1.86	2.04	0.53	0.75	
s(lexical: logiciel gratuit)	2.01	2.47	7.3	0.05	*
s(lexical: concepteur de jeu)	2.21	2.71	4.51	0.2	
s(lexical: jouabilité)	2.98	3.47	13.54	0.01	**
s(lexical: mot-dièse)	3	3.15	13.94	0.01	**
s(lexical: centre d'assistance)	1	1.01	1.99	0.16	
s(lexical: téléassistance)	2.68	3.2	17.41	< 0.001	***
s(lexical: zone d'accès sans fil)	2.73	3.02	2.26	0.48	
s(lexical: enregistreur de frappe)	1.94	2.35	2.14	0.38	
s(lexical: liste de diffusion)	1	1	2.26	0.13	
s(lexical: microblogue)	1	1	0	1	
s(lexical: film piraté)	1.85	1.98	0.07	0.97	
s(lexical: multiplexeur)	1	1.01	2.78	0.1	.
s(lexical: multiplexage)	1.54	1.91	0.64	0.65	
s(lexical: multitâche)	1	1	0	1	
s(lexical: DELO)	1.01	1.02	0.01	0.92	
s(lexical: capture de jeu)	1	1	0	1	
s(lexical: filoutage)	1	1	0	1	

(Continued)

Table 3. (Continued.)

	Effective Degree of Freedom	Reference Degrees of Freedom	$\chi^2$	p-value	
s(lexical: fenêtre intrusive)	1.86	1.98	0.03	0.98	
s(lexical: pré suite)	1	1	0	1	
s(lexical: rétro-ingénierie)	1	1.01	4.13	0.04	*
s(lexical: scénariste-conseil)	1	1	0	1	
s(lexical: réalité scénarisée)	1	1.01	6.17	0.01	*
s(lexical: carte SIM)	1	1	0	1	
s(lexical: ordiphone)	2.75	3.27	48.71	< 0.001	***
s(lexical: logiciel à la demande)	1	1	0.13	0.72	
s(lexical: écran fragmenté)	1	1	0	1	
s(lexical: logiciel espion)	1	1	0.2	0.66	
s(lexical: réseau de stockage)	1.97	2.42	3.47	0.27	
s(lexical: tablette)	1	1	43.26	< 0.001	***
s(lexical: encre en poudre)	1.39	1.63	0.62	0.72	
s(lexical: rétrolien)	1	1.01	0.67	0.42	
s(lexical: bande-annonce)	3.77	3.97	56.08	< 0.001	***
s(lexical: télé-poubelle)	1	1	1.96	0.16	
s(lexical: triple service)	1.48	1.72	0.64	0.71	
s(lexical: tutoriel)	1.86	2.26	7.78	0.04	*
s(lexical: dégroupage)	1	1	2.22	0.14	
s(lexical: vidéo à la demande)	1	1	1.42	0.24	
s(lexical: cybercaméra)	2.44	2.83	1.76	0.59	
	Adjusted $r^2$	Deviance explained			
	0.676	62.6%			

Significance codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

**Table 4.** The effect of prescription for the 54 prescribed terms in light of the initial usage. The strongest increase (i.e. the highest effect) is found for terms with medium and low initial usage

initial use → effect ↓	very low > .0	low > .05	medium > .25	high > .50	very high > .75
strong increase > .50		multiplexeur, rétrolien	liseuse, tablette		
medium increase .25 - .49	courriel, zone d'accès sans fil		télé-poubelle	réseau d'accès, réseau de stockage, tutoriel	
weak increase .01 - .24			centre d'assistance	centre d'appels, logiciel espion	logiciel gratuit, vidéo à la demande
no change 0	capture de jeu, cybercaméra, DELO, écran fragmenté, encre en poudre, femtocellule, fenêtre intrusive, filoutage, microblogue, présuite, scénariste-conseil, triple service	grosse machine			bannière, carte SIM, film piraté, multitâche, réseau d'infrastructure
weak decrease .01 - .24	concepteur de jeu, dialogue en ligne, informatique en nuage, logiciel à la demande, mégadonnées, mot-dièse, ordiphone	enregistreur de frappe, téléassistance		livre numérique	boîtier multiservice, dégroupage, encre électronique, liste de diffusion, livre électronique, multiplexage
medium decrease .25 - .49					bande-annonce, offre groupée
strong decrease > .50				jouabilité	réalité scénarisée, rétro-ingénierie

**Cite this article:** Zsombok G (2021). Prescribing French: A corpus-linguistic approach to official terminology in French newspapers. *Journal of French Language Studies* 31, 270–293. <https://doi.org/10.1017/S0959269520000204>