

Mémo Systèmes Lexicaux

Exemples lexicographiques *

Sandrine Ollinger[†] et Alain Polguère[‡]

[†]ATILF, CNRS, Université de Lorraine [‡]Université de Lorraine, CNRS, ATILF

25 juin 2020

Table des matières

1	Principes de sélection des exemples lexicographiques	2
2	Contenu de la zone d'exemples EX d'un article lexicographique	4
2.1	Entrée des citations	5
2.2	Exploitation des citations comme exemples lexicographiques	5
3	Modalité de référencement des citations selon leur source	6
3.1	Citations de corpus type Frantext (obligatoires)	7
3.2	Citations de corpus type FrWac (obligatoires)	8
3.3	Citations de corpus journalistiques hors Frantext (obligatoires)	9
3.4	Citations de corpus oraux type TCOF	10
3.5	Exemples hors bases	10
3.5.1	Exemples hors bases type <i>illustration du sens</i> (obligatoires)	10
3.5.2	Exemples hors bases type <i>publication</i>	10
3.5.3	Exemples hors bases type <i>audiovisuel</i>	11
3.5.4	Exemples hors bases type <i>conversation</i>	11
3.5.5	Exemples hors bases type <i>Web</i>	12
3.5.6	Exemples hors bases type <i>courrier</i>	13
3.5.7	Exemples hors bases type <i>fabrication</i>	14
3.5.8	Exemples hors bases type <i>temporaire</i>	14
3.6	Citations hors-norme	14
4	Citations adaptées	15
	Bibliographie	15

Résumé

Nous présentons le contenu de la zone d'exemples d'un article de *Système Lexical* du type RL- (Polguère, 2014), en illustrant notre propos sur le cas spécifique du *Réseau Lexical du Français RL-fr*. Il s'agit principalement d'identifier les différentes classes d'exemples potentiellement présents dans les articles et d'indiquer comment ces exemples doivent être sélectionnés et présentés par la lexicographe.

*Ce texte a été élaboré en s'inspirant de Benoit et al. (2011).

1 Principes de sélection des exemples lexicographiques

Un texte ou segment de texte est introduit comme *exemple lexicographique*¹ dans la *zone d'exemples (EX)* d'un article de lexie car :

- i) il contient au moins une occurrence d'emploi du *mot-vedette* ;
- ii) il permet d'illustrer une ou plusieurs propriétés (sémantiques, formelles, combinatoires) du mot-vedette.

Les onze principes ci-dessous contrôlent la sélection des exemples.

Principe 1 L'exemple doit permettre de circonscrire avec suffisamment de précision le sens de la lexie-vedette, tout particulièrement lorsque cette dernière appartient à un vocable polysémique.

Principe 2 Les exemples doivent nécessairement provenir d'une source apparaissant dans la liste des sources présentée par l'éditeur *Dicet* (Gader et al., 2012). Les principaux types de sources utilisés pour la construction du RL-fr sont détaillés dans la section 2 *Contenu de la zone d'exemples EX d'un article lexicographique*.

Principe 3 Autant que possible, un exemple est une phrase syntaxique (1a). Il peut cependant s'agir d'un énoncé non phrastique, comme un syntagme nominal (1b) ou d'une suite d'énoncés formant un très court texte (1c).

Important Les exemples et les sources apparaissent avec le format de présentation utilisé dans les vues-articles de *Dicet*. Chaque exemple contient en gras le mot-clé dont il illustre l'emploi.

- (1) a. Franchomme, lui, **comptait** ses œufs, ses noix et ses pommes.
Frantext PERRY Jacques, *Vie d'un païen*, 1965, p. 31
- b. Municipales : **vent** de panique à droite
FrWac février 2008, <http://municipales.actu.orange.fr/articles/aquitaine/Bayrou-90-engagements-pour-Pau-dont-tramway-leger-et-centre-culturel.html> [titre]
- c. Ni commode, ni table, ni chaise, ni bergère autour de nous, ne restaient plus que le lit, qui est un droit légal, et, comme un crapaud sur la moquette jaune, le téléphone. . .
– On vous **a cambriolé!** criai-je.
Frantext BILLETDOUX Marie, *Un peu de désir sinon je meurs*, 2006, p. 108

Principe 4 Si l'exemple est une citation, il doit provenir d'un texte qui n'est pas antérieur à 1950.

On peut exceptionnellement être amené à transgresser ce principe, par exemple, pour illustrer une lexie dite *inactive*, qui existait dans un état antérieur de la langue et n'est présente en synchronie qu'au sein de locutions, proverbes, etc. Tel est le cas de la lexie FLEURETTE : l'exemple (2a) illustre son emploi comme lexie autonome dans un état antérieur du français et (2b) son emploi actuel au sein de la locution 「CONTER FLEURETTE」.

- (2) a. Les hommes se sont fait une habitude de débiter la **fleurette**, et les femmes de s'en repaître avidement.
Frantext CAYLUS (DE) Anne-Claude, *Histoire de Monsieur Guillaume, cocher*, 1737, p. 56

1. Nous dirons dorénavant simplement *exemple*, sauf si le contexte requiert d'être plus explicite.

- b. Quelques semaines plus tard, en effet, déjà lassé par l'attente, le beau maréchal se mit à **conter fleurette** à sa jeune pupille rousse, Bonne de Pons.

Frantext CHANDERNAGOR Françoise, *L'Allée du Roi*, 1981, p. 198

Bien entendu, les exemples de type (2b) doivent apparaître dans l'article de la locution et non dans celui de la lexie inactive.

Principe 5 Les exemples ne doivent pas provenir de textes possédant une syntaxe particulière. Les extraits de poèmes, chansons, etc., sont admissibles pour peu que leur formulation relève d'un vocabulaire et d'une syntaxe « ordinaires ».

Principe 6 Les troncations de phrases ou de syntagmes – *cf.* l'exemple extrait du TLFi en (3) ci-dessous – sont totalement proscrites.

- (3) Le bâtiment (...) évoquait l'idée d'un bateau de plaisance (...) assurément capable de tenir la haute mer (...). Le **tableau** d'arrière ne portait ni nom ni trace aucune de l'immatriculation nautique. [GRACQ Julien, *Le rivage des Syrtes*, 1951, p. 77]

Dans le cas d'un exemple trop long, se référer à la section 4, page 15, qui porte sur les exemples adaptés.

Principe 7 Selon qu'il s'agit d'un vocable traité en profondeur ou simplement ébauché, le lexicographe renseignera de manière différente le champ des exemples.

- Dans le cas d'un vocable traité en profondeur, ce dernier devra minimalement contenir pour chaque acception : 3 exemples de Frantext et 3 exemples d'autres sources (Web, etc.).
- Dans le cas d'un vocable ébauché, seul 1 exemple est nécessaire pour chaque acception créée, quelle que soit la source.

Principe 8 Il est préférable de ne pas prendre d'exemple contenant une erreur de français. Cependant, la lexicographe peut être amené à utiliser une citation qui illustre parfaitement l'acception désirée, mais qui contient une erreur de français. Dans une telle situation, les erreurs de français doivent être corrigées par la lexicographe, qui note dans le champ **Commentaires** de l'exemple les modifications effectuées sous le format suivant :

Correction : <Forme erronée> -> <Forme corrigée>

Par exemple :

Correction : petits -> petites

La normalisation des caractères (*cf.* Principe 10 plus bas) n'est pas considérée comme une correction véritable et ne donne pas lieu à un commentaire.

Si la citation contient nombre d'abréviations non conventionnelles et d'éléments de langage SMS, il s'agit vraisemblablement d'une citation non adaptée. Plutôt que de corriger/réécrire ces éléments linguistiques, on cherchera un autre exemple. Bien entendu, ce type de citation est tout à fait approprié si l'on cherche justement à illustrer une lexie relevant du style SMS.

Principe 9 Si le nom d'une personne est important pour comprendre le sens de l'exemple, alors c'est probablement un mauvais exemple. En cas de doute on ne conserve pas l'exemple. Il est préférable d'avoir un exemple du Web clair et sans nom propre. Toutefois, si l'exemple est vraiment bon, on l'adapte (*cf.* citations adaptées section 4, page 15).

Dans tous les cas, il vaut mieux éviter de prendre des exemples avec des noms propres ou des pronoms qui ne permettent pas de savoir s’il s’agit d’un individu, d’un animal ou bien d’un objet.

La lexicographe n’anonymise pas les exemples de Frantext. À l’inverse, pour tous les autres corpus, les noms d’individus – sauf les personnages historiques (Louis XVI, Armand Jean du Plessis de Richelieu) ou les personnes publiques (artiste, politicien, etc.) – sont à anonymiser par la majuscule du nom suivie d’un point **lorsqu’il apparaît nécessaire de préserver la vie personnelle des personnes citées**. Par exemple, Paul Durant devient Paul D., M. Smith devient M. S.

Rappelons que deux points ne peuvent se suivre. Il est donc exclu de conserver le point de ponctuation après le point d’anonymisation (*M. S.).

Consigne de bon sens. Il est impératif d’**utiliser son bon jugement** dans la procédure d’anonymisation et de n’anonymiser que lorsque cela est nécessaire.

Principe 10 Le formatage d’une citation – italique, gras, saut de ligne, etc. – doit toujours être restitué. De plus, pour éviter les problèmes d’affichage posés par les caractères fantômes, il est fortement conseillé à la lexicographe de passer les exemples dans un bon éditeur de fichiers textes *.txt* (un simple copier-coller suffit).

Il faut veiller à la normalisation systématique des caractères suivants :

- usage d’apostrophes courbes (‘...’, et non ‘...’);
- usage de guillemets dits *français* (« ... », et non “...”);
- usage approprié des tirets : « — » et non « - » devant les tours de parole, citations et éléments insérés;
- introduction d’espaces insécables selon les normes typographiques du français de France;
- mise des accents sur les majuscules;
- mise en indice ou exposant de ce qui doit l’être : « M^{me} » et non « Mme », « 3^e » et non « 3e », etc.²;
- usage des caractères spéciaux : « œ » au lieu de « oe », « æ » au lieu de « ae », etc.

Note. En cas de doute sur les normes typographiques, se reporter au guide « Typographie » de la suite logicielle Antidote³.

Principe 11 Les exemples doivent être ordonnés selon leur pertinence linguistique propre. Ainsi, un exemple qui illustre un emploi « standard » du mot-clé doit apparaître avant un exemple qui illustre un emploi marginal ou quasi-déviant. Cela implique que l’on mélange éventuellement dans l’ordonnement linéaire de la zone EX des articles lexicographiques (voir prochaine section) des exemples provenant de sources distinctes.

2 Contenu de la zone d’exemples EX d’un article lexicographique

Il convient de distinguer deux opérations lexicographiques relativement à l’exemplification :

1. **l’entrée** d’une *citation* dans la base de données;
2. **l’exploitation** de cette citation en tant qu’*exemple lexicographique* pour une lexie donnée via la création d’une association entre cette lexie et la citation en question – un exemple est donc une association lexie → citation.

Ces deux opérations sont effectuées au moyen de l’éditeur Dicet.

2. Rappelons, à ce propos, qu’il faut écrire « 1^{er} » et « 2^e », et non « 1^{ier} » et « 2^{ième} ».

3. <https://www.antidote.info/fr/>

2.1 Entrée des citations

Nous énumérons ci-dessous une série de directives « génériques » concernant l'entrée de citations dans la base lexicale : ces instructions valent pour toute citation, quelle que soit sa source. Nous détaillerons dans la section 3 les instructions propres à chaque type de source.

- (a) Il ne faut surtout pas intégrer à la base une citation qui en chevauche ou en contient une autre. La lexicographe doit donc s'assurer qu'une citation qu'elle désire entrer dans la base ne possède pas d'intersection avec une citation déjà entrée.
- (b) L'éditeur permet à la lexicographe de laisser un commentaire dans la fiche d'édition de la citation (*cf.* plus bas le champ **Commentaires** dans la Figure 13, page 13). C'est notamment de cette façon que doivent être signalées les modifications opérées dans la citation (corrections orthographique, anonymisations, etc.).
- (c) Si la citation est un titre, la lexicographe ajoute [titre] à la suite des informations entrées dans **Emplacement**, par exemple :
 - (4) Rencontre avec Emmanuelle R., **flûte** solo de l'ONL
FrWac février 2008, http://www.onl.fr/onl_info/pages_saison0708/conferences.htm [titre]
- (d) Les noms de mois s'écrivent en chiffres dans le champ **Parution** (jour/mois/année)⁴. Les chiffres de mois et de jour ne doivent pas être précédés du 0 : février se note 2 et non 02 – *cf.* plus bas, Figure 4, page 8.

2.2 Exploitation des citations comme exemples lexicographiques

La base lexicographique réunit toutes les citations utilisées comme exemples dans le RL-fr. Lorsqu'une lexicographe a besoin d'un exemple, elle doit commencer par rechercher des exemples illustratifs de sa lexie parmi les citations déjà présentes dans la base (*cf.* étape lexicographique décrite ci-dessus). Pour cela, elle dispose de deux possibilités :

1. utiliser les fonctionnalités de recherche de citations offertes par l'éditeur Dicet ;
2. faire une recherche dans les exports de citations régulièrement effectués par S. Ollinger ; cette recherche s'effectue au moyen de la plateforme *TXM* (Heiden, 2010; Heiden et al., 2010)⁵, qui présente l'avantage de permettre des requêtes complexes et d'effectuer un affichage des résultats plus lisible (concordances).

Remarque. Il est demandé à la lexicographe de profiter du fait qu'elle exploite une citation préexistant dans la base pour vérifier que celle-ci a été correctement entrée et référencée.

Ci-dessous, les directives générales qui concernent la mise en œuvre d'une association lexie → citation.

- (a) Une fois que la lexicographe a créé une association lexie → citation elle effectue les deux opérations suivantes :
 - (a) identifier l'occurrence de la lexie dans cette citation ;
 - (b) effectuer un clic droit dans le texte de la citation pour sélectionner **Marquer une occurrence** dans la *pop-up* qui s'ouvre alors – *cf.* Figure 1.

4. Les feuilles de style de l'éditeur Dicet se chargent de calculer le nom du mois à partir du numéro correspondant, pour affichage dans les fenêtres lexicographiques.

5. <http://textometrie.ens-lyon.fr/>

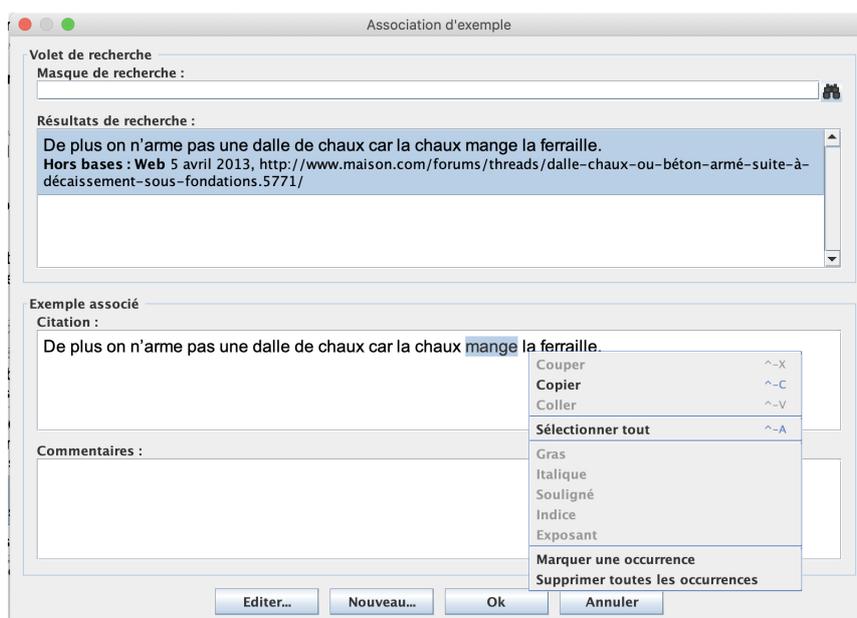


FIGURE 1 – Identification d’une occurrence de lexie dans une citation

Bien évidemment, dans le cas où l’exemple contient plusieurs occurrences de la lexie, toutes les occurrences doivent être marquées. La lexicographe doit être vigilante quant à la sélection de la forme fléchie complète de la lexie – par exemple :

- (5) Les boulons qui le fixaient **avaient** un peu **mangé** le bois.
- (b) L’éditeur permet à la lexicographe de laisser un commentaire dans la fenêtre d’association lexie → citation.
- (c) Après avoir créé une association lexie → citation, la lexicographe doit veiller à ce que l’exemple résultant soit coché. Cocher un exemple signifie qu’il est jugé correct et validé par la lexicographe. Toutefois, il est possible de laisser un exemple non coché afin de rendre compte de certains emplois jugés « limites » et sur lesquels la lexicographe veut apporter un commentaire. Il apparaît alors en grisé dans les articles lexicographiques visualisés par Dicet – cf. partie *vue-article* de la Figure 2.

3 Modalité de référencement des citations selon leur source

Les citations qui servent d’exemples dans le RL-fr sont des cinq types suivants :

1. « Citations de corpus type Frantext » ;
2. « Citations de corpus type FrWac » ;
3. « Citations de corpus journalistiques hors Frantext » ;
4. « Citations de corpus oraux type TCOF » ;
5. « Exemples hors bases ».

L’ordre d’énumération ci-dessus correspond à l’ordre de présentation par défaut des exemples dans le champ EX des articles lexicographiques.

Nous détaillons maintenant le mode de référencement propre à chaque type de citation, tel qu’il doit être appliqué lors de l’entrée de la citation dans la base.

De plus on n'arme pas une dalle de chaux car la chaux **mange** la ferraille.
Hors bases : Web 5 avril 2013, <http://www.maison.com/forums/threads/dalle-chaux-ou-beeton-arme-suite-a-decaissement-sous-fondations.5771/>

En fait en forçant le serrage des vis de fixation du support de transmission j'**ai mangé** le pas de vis du carter moteur.
Hors bases : Web 9 avril 2013, <http://www.fgmarder.net/t358-demonter-un-embayage>

La rouille **a mangé** toute la carrosserie. Celle-ci est désormais à refaire.
Illustration du sens

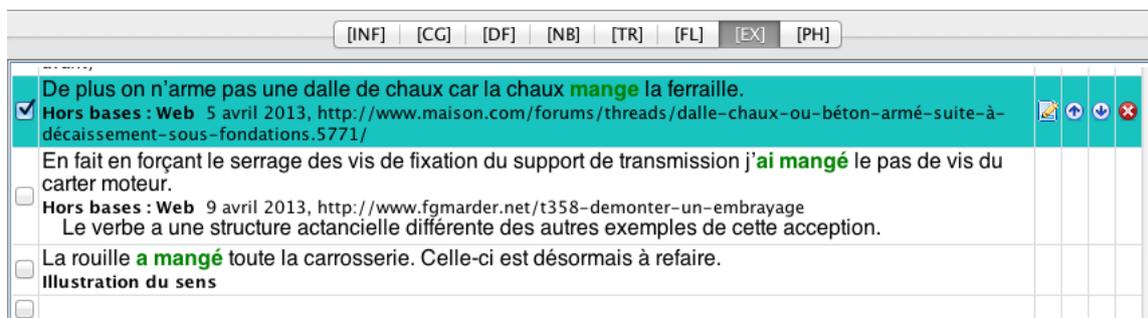


FIGURE 2 – Sélection et visualisation des exemples

3.1 Citations de corpus type Frantext (obligatoires)

Frantext (Bernard et al., 2002) est interrogeable en ligne⁶ avec la plateforme *Allegro*⁷. Il doit être systématiquement consulté, notamment pour les textes littéraires qu'il contient. Il faut cependant veiller à respecter le Principe 4 (page 2, ci-dessus) et donc, n'utiliser que des citations de textes dont la publication n'est pas antérieure à 1950⁸.

Rappelons que le formatage d'une citation – italique, gras, saut de ligne, etc. – doit toujours être restitué.

Pour entrer un exemple provenant de Frantext, il faut, dans la page d'édition d'exemple :

- sélectionner la section « Citations de corpus type Frantext », la source « Frantext » est sélectionnée automatiquement ;
- saisir le texte de l'exemple dans la case « Citation » ;
- noter « Nom, Prénom » de l'auteur ou « Nom, Prénom ; Nom, Prénom » s'il y a plusieurs auteurs (*cf.* Figure 8, page 11) ;
- entrer le titre de l'ouvrage tel qu'il est présenté dans Frantext ;
- inscrire l'année de parution ;
- marquer le numéro de page – « p. <Numéro> », ou « pp. <Numéro1>-<Numéro2> » si l'exemple se trouve sur deux pages – dans la case « Emplacement », **en prenant soin de faire suivre p. et pp. d'une espace insécable.**

Pour les noms d'auteurs à particule, la particule se place après le nom et entre parenthèses comme : Beauvoir (de), Simone. Attention à ne bien appliquer cette règle que dans le cas de particules véritables. Ainsi, il faut écrire De Gaulle, Charles et non *Gaulle (de), Charles⁹.

Le référencement d'une citation de Frantext est affiché sous l'exemple, de la façon suivante :

- (6) Le **basket-ball** est le sport principal en Amérique du nord.
Frantext COLLECTIF sous la dir. de R. Caillois, *Jeux et sports*, 1967, p. 1349

Remarque En cas d'indication de plusieurs dates dans Frantext, retenir l'édition la plus ancienne. Par exemple, pour des œuvres d'Alphonse Daudet qui ont été rééditées en 2011, on retiendra le texte original ainsi que sa date de parution.

6. <http://www.frantext.fr>

7. <https://halshs.archives-ouvertes.fr/halshs-02381605>

8. Nous avons actuellement accès à 1 324 textes, correspondant à près de 73,5 millions de mots-occurrences.

9. On peut dire *Beauvoir a écrit ce livre*, mais non **Gaulle a écrit ce livre*.

Section :		Source :	
Citations de corpus type Frantext		Frantext	
Citation :			
Le basket-ball est le sport principal en Amérique du nord.			
Auteurs (nom, prénom; ...):			
Collectif, sous la dir. de R. Caillois			
Titre :			
Jeux et sports			
Parution (jour/mois/année) :		Emplacement :	
		1967	p. 1349

FIGURE 3 – Entrée d'un exemple provenant de Frantext

3.2 Citations de corpus type FrWac (obligatoires)

FrWac¹⁰ – construit dans le contexte du projet *WaCky wide web* (Baroni et al., 2009)¹¹ – est une image du contenu des pages Internet du domaine `.fr` construite en février 2008 et comportant environ 1,6 milliard de mots.

Le référencement d'une citation de FrWac se fait, outre la mention du corpus source, par la mention de la date de création du corpus (février 2008), ainsi que l'URL de la page web source. Pour entrer un exemple provenant de FrWac, il faut, dans la page d'édition d'exemple :

- sélectionner la section « Citations de corpus type FrWac », la source « FrWac » est sélectionnée automatiquement ;
- saisir le texte de l'exemple dans la case « Citation » ;
- noter la date dans les cases correspondantes : le mois de février par un 2 et l'année 2008 ;
- entrer l'adresse internet complète dans la case « Emplacement ».

Section :		Source :	
Citations de corpus type FrWac		FrWac	
Citation :			
Pour s'assurer que Jake lui remboursera bien tout l'argent qu'il lui doit, Johnny et sa bande kidnappent son petit frère, Zack.			
Auteurs (nom, prénom; ...):			
Titre :			
Parution (jour/mois/année) :		Emplacement :	
	2	2008	http://tetelle43.blogs.allocine.fr/?blog=tetelle43&tool=\&page=18&f=1

FIGURE 4 – Entrée d'un exemple provenant de FrWac

Un exemple FrWac est visualisé dans l'éditeur de la façon suivante :

- (7) Pour s'assurer que Jake lui remboursera bien tout l'argent qu'il lui doit, Johnny et sa bande **kidnappent** son petit frère, Zack.
FrWac février 2008, <http://tetelle43.blogs.allocine.fr/?blog=tetelle43&tool=&page=18&f=1>

Important Les textes de FrWac, qui sont issus du Web, sont souvent fautifs et les citations FrWac doivent impérativement être corrigées pour ce qui est de l'orthographe, de la ponctuation, etc. Cette remarque s'applique bien évidemment aussi aux citations hors bases provenant du Web (section 3.5.5, p. 12).

10. https://www.clarin.si/noske/run.cgi/first_form?corpname=frwac;align=

11. <http://wacky.sslmit.unibo.it/doku.php>

À noter que certaines caractéristiques des textes du Web – provenant des blogs notamment – peuvent tout à fait être préservées dans les citations RL-fr : mise en majuscules pour l’emphase, points d’exclamation ou d’interrogation répétés, etc.

3.3 Citations de corpus journalistiques hors Frantext (obligatoires)

Pour entrer un exemple provenant de *L’Est Républicain*, il faut, dans la page d’édition d’exemple :

- sélectionner la section « Citations de corpus journalistiques hors Frantext »;
- veiller à ce que la source « L’Est Républicain » soit sélectionnée;
- saisir le texte de l’exemple dans la case « Citation »;
- noter dans les cases correspondantes la date de parution en jour, mois et année en chiffres;
- entrer l’identifiant de l’article dans la case « Emplacement ».

Attention à ce que la référence soit bel et bien celle de l’article et non du titre, tel qu’indiqué dans la Figure 5 ci-dessous.

```
toutes les violences racistes en France , notamment celles qui visent des Juifs et des Arabes » </p> </article> <article id="id3327882_2002-04-11"
auteur="non-specifie"> <titre id="id3327887_2002-04-11"> A la barre </titre> <division auteur="non-specifie"> <titre id="id3327892_2002-04-11">
```

FIGURE 5 – Identifiant article et identifiant titre dans *l’Est Républicain*

Le référencement se fait tel qu’indiqué dans la Figure 6 ci-dessous.

Section :		Source :	
Citations de corpus journalistiques hors Frantext		L’Est Républicain	
Citation :			
Gabriel, 21 ans, venait de cambrioler la cave d’un retraité et d’y dérober 17 bouteilles.			
Auteurs (nom, prénom; ...):			
Titre :			
Parution (jour/mois/année):		Emplacement :	
11	4	2002	3327882

FIGURE 6 – Entrée d’un exemple de *l’Est Républicain*

Finalement, la citation apparaît comme suit dans la vue-article :

- (8) Gabriel, 21 ans, venait de **cambrioler** la cave d’un retraité et d’y dérober 17 bouteilles.
L’Est Républicain 11 avril 2002, 3327882

Remarque 1 Nous avons renoncé à mentionner les numéros de pages et de colonnes pour les citations journalistiques, car cette information n’est pas nécessairement disponible dans les versions électroniques des journaux.

Remarque 2 Si l’auteur de l’article journalistique est mentionné dans les données, la lexicographe peut l’entrer dans la fiche RL-fr de la citation dans le cas où il s’agit d’un auteur connu et où il peut être pertinent de considérer qu’il s’agit d’une citation de cet auteur particulier.

3.4 Citations de corpus oraux type TCOF

Nous ne disposons pas à l'heure actuelle de corpus d'envergure de productions orales en français, interrogeables au moyen d'un concordancier. C'est pourquoi nous n'envisageons pas le recours à de tels corpus dans le cadre du projet RELIEF.

Mentionnons que des corpus de productions orales sont tout de même disponibles à l'ATILF. Il s'agit des corpus du projet *TCOF* (Traitement de Corpus Oraux en Français)¹², en accès libre sur le site du CNRTL : <http://www.cnrtl.fr/corpus/tcof/>.

3.5 Exemples hors bases

En raison de leur grande valeur lexicographique potentielle, il est essentiel pour ces **citations** hors bases de reproduire exactement ce qui a été lu ou entendu, afin de maintenir leur nature de citation. Si le lexicographe ne les a pas notées et n'en a qu'un souvenir approximatif, la retranscription qu'il donne relèvera d'une classe particulière d'exemples hors bases : les exemples fabriqués (section 3.5.7 ci-dessous).

3.5.1 Exemples hors bases type *illustration du sens* (obligatoires)

Pour chaque acception de vocable travaillé, il est demandé au lexicographe d'entrer un exemple inventé ou simplifié qui illustre très clairement le sens de l'acception. Les exemples de ce type pourront notamment être utilisés lorsque l'on générera les *Tableaux synoptiques* des vocables.

Ces exemples sont à intégrer dans la rubrique « illustration du sens », sans information additionnelle de source. Ils doivent rester grisés dans les vues-articles et ne sont donc pas cochés dans les associations lexie-exemple, tel qu'illustré dans la Figure 7 ci-dessous.

Section :	Source :
Exemples hors bases	Illustration du sens
Citation :	
Je mange équilibré à chaque repas.	
Auteurs (nom, prénom; ...) :	
Titre :	
Parution (jour/mois/année) :	Emplacement :

FIGURE 7 – Entrée d'un exemple Hors bases : Illustration du sens

3.5.2 Exemples hors bases type *publication*

Il s'agit de citations de livres, de journaux, etc., qui doivent avoir le même référencement que les citations de Frantext, comme par exemple :

Remarque Les citations extraites d'ouvrages ou documents mis à disposition sur le Web – sur Google Books, notamment – entrent dans cette classe et doivent être référencées relativement à l'ouvrage dont elles proviennent. Elles ne correspondent pas à la classe 3.5.5 ci-dessous.

12. Cf. André et Canut (2010).

Section :		Source :	
Exemples hors bases		Hors bases : publication	
Citation :			
Les croissances ralenties des économies industrielles installées concourent plus aux échanges mondiaux que les économies en démarrage rapide. En ce sens, le freinage sensible des investissements directs à l'étranger dans les années 2000 a pu paraître à certains le signe d'une rupture dans la mondialisation ; cette question ne pourrait être vraiment éclairée qu'au travers d'une problématique de recherche approchant continuellement les échanges.			
Auteurs (nom, prénom; ...) :			
Hugon, Philippe; Michalet, Charles-Albert			
Titre :			
Les nouvelles régulations de l'économie mondiale			
Parution (jour/mois/année) :		Emplacement :	
		2005	p. 255

FIGURE 8 – Entrée d'un exemple Hors bases : publication

3.5.3 Exemples hors bases type *audiovisuel*

Ce sont des citations d'énoncés radiophoniques, télévisuels ou cinématographiques. Elles sont référencées de la façon suivante.

- Radio/télévision

<Nom de l'émission de radio/télévision>, <Nom de la chaîne de radio/télévision>, <Date de diffusion (jour, mois et année)>

Section :		Source :	
Exemples hors bases		Hors bases : audiovisuel	
Citation :			
On a la découpe des montagnes. Là, on a une dent acérée.			
Auteurs (nom, prénom; ...) :			
Titre :			
Des racines et des ailes			
Parution (jour/mois/année) :		Emplacement :	
6	2	2013	France3

FIGURE 9 – Entrée d'un exemple Hors bases : audiovisuel (radio/télévision)

- (9) On a la découpe des montagnes. Là, on a une **dent** acérée.
Hors bases : audiovisuel *Des racines et des ailes*, 6 février 2013, France3

- Cinéma

Réplique de <Nom du personnage> (joué(e) par <Nom de l'acteur>), <Titre du film>, <Année de réalisation>

- (10) — Cal, mon **cigare** est devant ta tronche depuis 20 minutes. Si ça te gêne pas, c'est pire que ce que je pensais !
Hors bases : audiovisuel RÉPLIQUE DE JACOB PALMER (JOUÉ PAR RYAN GOSLING), *Crazy, Stupid, Love*, 2011

3.5.4 Exemples hors bases type *conversation*

Ce sont des citations d'énoncés entendus dans des conversations, présentées selon le patron de référencement suivant :

Section :		Source :	
Exemples hors bases		Hors bases : audiovisuel	
Quotation :			
— Cal, mon cigare est devant ta tronche depuis 20 minutes. Si ça te gêne pas, c'est pire que ce que je pensais !			
Auteurs (last name, first name; ...):			
Réplique de Jacob Palmer (joué par Ryan Gosling)			
Titre :			
Crazy, Stupid, Love			
Publication (day/month/year):		Location :	
		2011	

FIGURE 10 – Entrée d'un exemple Hors bases : audiovisuel (cinéma)

<Date (jour, mois et année)> <Très brève caractérisation du contexte d'énonciation>

- (11) J'ai bien fait de prendre mon gilet, je ne me **déshabille** pas, il fait trop froid !
Hors bases : conversation 10 septembre 2011, entendu dans le TGV Metz-Nice à 6h55

Section :		Source :	
Exemples hors bases		Hors bases : conversation	
Citation :			
— J'ai bien fait de prendre mon gilet, je ne me déshabille pas, il fait trop froid !			
Auteurs (nom, prénom; ...):			
Titre :			
Parution (jour/mois/année):		Emplacement :	
10	9	2011	entendu dans le TGV Metz-Nice à 6h55

FIGURE 11 – Entrée d'un exemple Hors bases : conversation dans Dicot

C'est au lexicographe d'évaluer le degré de précision nécessaire pour la bonne caractérisation du contexte d'énonciation, **en veillant à être le plus économe possible**. En règle générale, le lieu d'énonciation devra être spécifié, mais il peut être utile, dans certains cas, de caractériser aussi de façon approximative le locuteur et/ou le destinataire.

3.5.5 Exemples hors bases type *Web*

Le lexicographe doit indiquer la date de consultation du site dans les cases correspondantes ainsi que l'adresse complète du site dans « Emplacement ».

Les citations du web suivent le patron de référencement suivant :

<Date de consultation de la page (jour, mois et année)> <URL de la page web>

Ils apparaissent donc de la façon suivante :

- (12) Ma fille était **propre** à l'âge de 2 ans et demi, mais mon fils, qui a presque 3 ans, ne s'intéresse pas du tout à l'apprentissage du pot que je lui enseigne.
Hors bases : Web 11 septembre 2012, <http://www.babycenter.fr/toddler/apprentissage-pot/plus-long-garcons-expert/>

Section :		Source :	
Exemples hors bases		Hors bases : Web	
Citation :			
Ma fille était propre à l'âge de 2 ans et demi, mais mon fils, qui a presque 3 ans, ne s'intéresse pas du tout à l'apprentissage du pot que je lui enseigne.			
Auteurs (nom, prénom; ...) :			
Titre :			
Parution (jour/mois/année) :		Emplacement :	
11	9	2012	.babycenter.fr/toddler/apprentissage-pot/plus-long-garcons-expert,

FIGURE 12 – Entrée d'un exemple Hors bases : Web

Il est important de vérifier que l'URL est correctement entrée dans le champ correspondant lorsqu'on en fait le copier-coller. Il peut arriver que des caractères spéciaux (accentués, notamment) soient mal retranscrits par Dicot lors de cette opération.

Il serait en principe pertinent de donner également la date de rédaction/modification de la page consultée, mais cette information est trop souvent soit non disponible, soit non fiable (mises à jour faites sans en indiquer la date). On doit se résoudre à donner une information incomplète. C'est justement une des faiblesses du Web en tant que corpus et du Web dans sa réalité textuelle : ce n'est pas un corpus contrôlé ou contrôlable.

Rappel Tel que mentionné plus haut à propos de FrWac (section 3.2, p. 8), les citations provenant du Web doivent être soigneusement révisées pour la qualité du français (orthographe et ponctuation, notamment).

3.5.6 Exemples hors bases type *courrier*

Les citations de type courrier peuvent provenir notamment :

- de courrier/courriel personnel ;
- de courrier/courriel administratif ;
- de courrier/courriel professionnel ;
- de courrier/courriel commercial.

Il faut suivre le patron de référencement suivant :

<Très brève caractérisation du contexte si nécessaire>,
 <Date (jour, mois et année)> <type de courrier ou courriel>

Section :		Source :	
Exemples hors bases		Hors bases : courrier	
Citation :			
Je cours après mes retards.			
Auteurs (nom, prénom; ...) :			
Professeur d'université			
Titre :			
Parution (jour/mois/année) :		Emplacement :	
17	7	2013	courriel professionnel
Commentaires :			
Registre peu formel.			

FIGURE 13 – Entrée d'un exemple Hors bases : courrier

- (13) Je **cours** après mes retards.
Hors bases : courrier Professeur d'université, 17 juillet 2013, courriel professionnel

3.5.7 Exemples hors bases type *fabrication*

Une section de la zone d'exemples du RL-fr est réservée aux exemples fabriqués, comme (14) :

- (14) Gronder un enfant de trois mois, c'est comme **pisser dans un violon** : ça ne sert à rien.
Hors bases : fabrication

Ce type d'exemple ne contient aucun référencement autre que la section « Exemples hors bases » et la source « Hors bases : fabrication ».

3.5.8 Exemples hors bases type *temporaire*

Certains exemples sont introduits par le lexicographe uniquement pour fixer le sens d'une lexie *créée à la volée*, comme cible d'un lien de fonction lexicale contrôlé par un mot-vedette d'article que le lexicographe est en train de construire. Ces exemples sont généralement récupérés sur le Web et sont présents de façon temporaire dans la base. Ils seront remplacés par de véritables exemples lexicographiques lorsque la lexie créée à la volée deviendra l'objet d'une description lexicographique.

On peut référencer ces exemples de façon très simple. Ainsi, le lexicographe introduit l'exemple avec la seule mention **Temporaire**, via la section « Exemples hors bases » et la source « Hors bases : temporaire », comme l'ensemble de données de la Figure 14, qui apparaît dans la vue-article tel qu'en (15).

FIGURE 14 – Entrée d'un exemple Hors bases : temporaire

- (15) Le **titulaire** du poste en congé maternité prend un temps partiel de droit.
Hors bases : temporaire

Les lexicographe doivent éviter d'utiliser cette source d'exemples vu la diversité de nos sources déjà disponibles, donc de lui préférer un exemple référencé.

Note Les exemples temporaires doivent toujours restés non cochés (et donc apparaître en grisé dans les vues-articles).

3.6 Citations hors-norme

Un article RL-fr peut contenir des citations relevant d'un français « hors-norme ». Ces citations peuvent être de plusieurs natures, notamment :

- *greffes collocationnelles* (Polguère, 2007) du type *rongé de peur* (sur le modèle de *rongé d'angoisse/de remords/...*), qui nous semblent être atypiques, mais dont nous pensons qu'elles ont une fréquence d'emploi significative ;
- emplois d'une lexie qui nous semblent relever d'un glissement sémantique assez bien attesté, mais dont nous ne jugeons pas la lexicalisation effective ;

— emplois d’une lexie avec un régime « anormal », mais qui méritent pour diverses raisons d’être relevés.

De telles citations doivent être référencées selon les patrons proposés plus haut, mais laissées grisées. Elles représentent ce que l’on pourrait appeler une *zone de turbulence lexicale* dans l’utilisation de la lexie décrite. Dans tous les cas, de telles citations doivent être accompagnées d’un commentaire qui les analyse et justifie leur inclusion dans un article de RL-fr.

4 Citations adaptées

Il peut arriver que le lexicographe veuille introduire dans un article un exemple qui est en fait une *citation adaptée*. Ainsi, (16b) est une citation adaptée de (16a), extrait de l’*Est Républicain*.

Les exemples adaptés sont référencés d’après leur source d’origine et différenciés des autres par la mention entre crochets [adapté] dans « Emplacement ».

- (16) a. Le **nombre** de places étant limité pour permettre un bon fonctionnement des différents ateliers inscrivez -vous sans tarder auprès de MM. Jacques A. 36 rue du Lt Clerc à Allamps (tél. 03.83.25.44.39), Gérard C. 2 rue Pasteur à Colombey (03.83.52.00.66), Christian D. 4 rue de la Prairie à Allamps (03.83.25.46.24), Yves D. 2 rue Gambetta à Blénod-les-Toul (03.83.62.58.67) ou Maurice M. 32 rue du Lt Clerc à Allamps (03.83.25.40.28).
- b. Le **nombre** de places étant limité pour permettre un bon fonctionnement des différents ateliers, inscrivez-vous sans tarder.
L’EST RÉPUBLICAIN 21 mai 1999, 369395 [adapté]

Comme on le voit, l’adaptation en (16b) de (16a) consiste principalement en la simplification du texte par troncation.

Les exemples adaptés peuvent provenir de toute source précédemment mentionnée (y compris Frantext). L’adaptation des exemples est plus souvent utilisée pour rendre plus « élégant » ou compact un énoncé d’une certaine lourdeur. Elle peut également consister en la correction d’un texte contenant trop de fautes d’orthographe, de syntaxe, etc. Elle sert aussi à rendre plus autonome un énoncé qui, sinon, pourrait paraître étrange ou peu clair une fois tiré de son contexte. En bref, elle permet de récupérer des exemples demandant trop de modifications et ne pouvant être conservés en l’état par l’usage d’un simple « [sic] ». Le lexicographe entre l’exemple selon sa source, puis ajoute [adapté] dans « Emplacement », à la suite des informations éventuellement présentes.

Section :	Source :
Citations de corpus journalistiques hors Frantext	L’Est Républicain
Quotation :	
Le nombre de places étant limité pour permettre un bon fonctionnement des différents ateliers, inscrivez-vous sans tarder.	
Authors (last name, first name; ...):	
Title :	
Publication (day/month/year) :	Location :
21 5 1999	369395 [adapté]

FIGURE 15 – Entrée d’un exemple Hors bases : citation adaptée

Bibliographie

- Virginie André et Emmanuelle Canut. Mise à disposition de corpus oraux interactifs : le projet TCOF (Traitement de Corpus Oraux en Français). *Pratiques : linguistique, littérature, didactique*, 147/148:35–51, 2010.
- Marco Baroni, Silvia Bernardini, Adriano Ferraresi et Eros Zanchetta. The WaCky wide web: a collection of very large linguistically processed web-crawled corpora. *Language Resources and Evaluation*, 43(3):209–226, 2009.
- Jean-Luc Benoit, Veronika Lux-Pogodalla et Alain Polguère. Mémo RLF. Exemples lexicographiques. Rapport technique, Projet RELIEF/RLF, ATILF CNRS, 8 pages, Nancy, 2011.
- Pascale Bernard, Josette Lecomte, Jacques Dendien et Jean-Marie Pierrel. Un ensemble de ressources informatisées et intégrées pour l'étude du français : FRANTEXT, TLFi, Dictionnaires de l'Académie et logiciel Stella, présentation et apprentissage de leurs exploitations. Dans *Actes de TALN 2002*, volume 2, pages 3–36, Nancy, 24–27 juin 2002.
- Nabil Gader, Veronika Lux-Pogodalla et Alain Polguère. Hand-Crafting a Lexical Network With a Knowledge-Based Graph Editor. Dans *Proceedings of the Third Workshop on Cognitive Aspects of the Lexicon (CogALex III)*, pages 109–125, Mumbai, 15 décembre 2012. The COLING 2012 Organizing Committee.
- Andrew Hardie. CQPweb – combining power, flexibility and usability in a corpus analysis tool. *International Journal of Corpus Linguistics*, 17(3):380–409, 2012.
- Serge Heiden. The TXM Platform: Building Open-Source Textual Analysis Software Compatible with the TEI Encoding Scheme. Dans Ryo Ootoguro, Kiyoshi Ishikawa, Hiroshi Umemoto, Kei Yoshimoto et Yasunari Harada, dir., *Proceedings of the 24th Pacific Asia Conference on Language, Information and Computation*, pages 389–398, Sendai, Japan, 2010. Institute for Digital Enhancement of Cognitive Development, Waseda University. URL <http://halshs.archives-ouvertes.fr/halshs-00549764>. Région Rhône-Alpes Cluster 13; Université Lyon 3.
- Serge Heiden, Jean-Philippe Magué et Bénédicte Pincemin. TXM : Une plateforme logicielle open-source pour la textométrie – conception et développement. Dans Sergio Bolasco, Isabella Chiari et Luca Giuliano, dir., *Statistical Analysis of Textual Data – Proceedings of 10th International Conference Journées d'Analyse statistique des Données Textuelles*, volume 2-3, pages 1021–1032, Rome, 2010. Edizioni Universitarie di Lettere Economia Diritto. URL <http://halshs.archives-ouvertes.fr/halshs-00549779>.
- Alain Polguère. *Soleil insoutenable et chaleur de plomb* : le statut linguistique des greffes collocationnelles. Dans Marie-Claude L'Homme et Sylvie Vandaele, dir., *Lexicographie et terminologie : compatibilité des modèles et des méthodes*, pages 247–291. Les Presses de l'Université d'Ottawa, Ottawa, 2007.
- Alain Polguère. From Writing Dictionaries to Weaving Lexical Networks. *International Journal of Lexicography*, 27(4):396–418, 2014. URL <https://academic.oup.com/ijl/article/27/4/396/932994/From-Writing-Dictionaries-to-Weaving-Lexical>. Special Issue: Dictionaries and the Digital Revolution: A Focus on Users and Lexical Databases.