

**Ce fait linguistique
n'existe pas,
je l'ai rencontré**

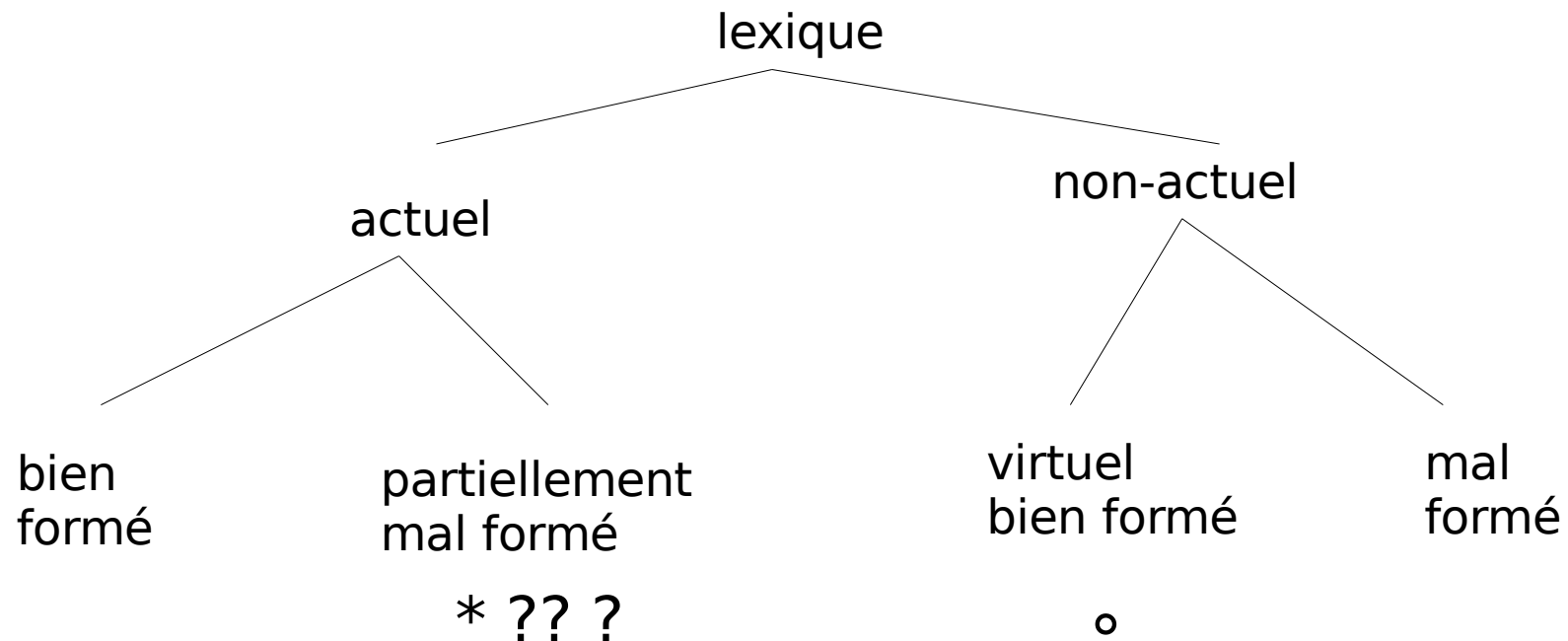
Benoît Habert

Quelques conventions

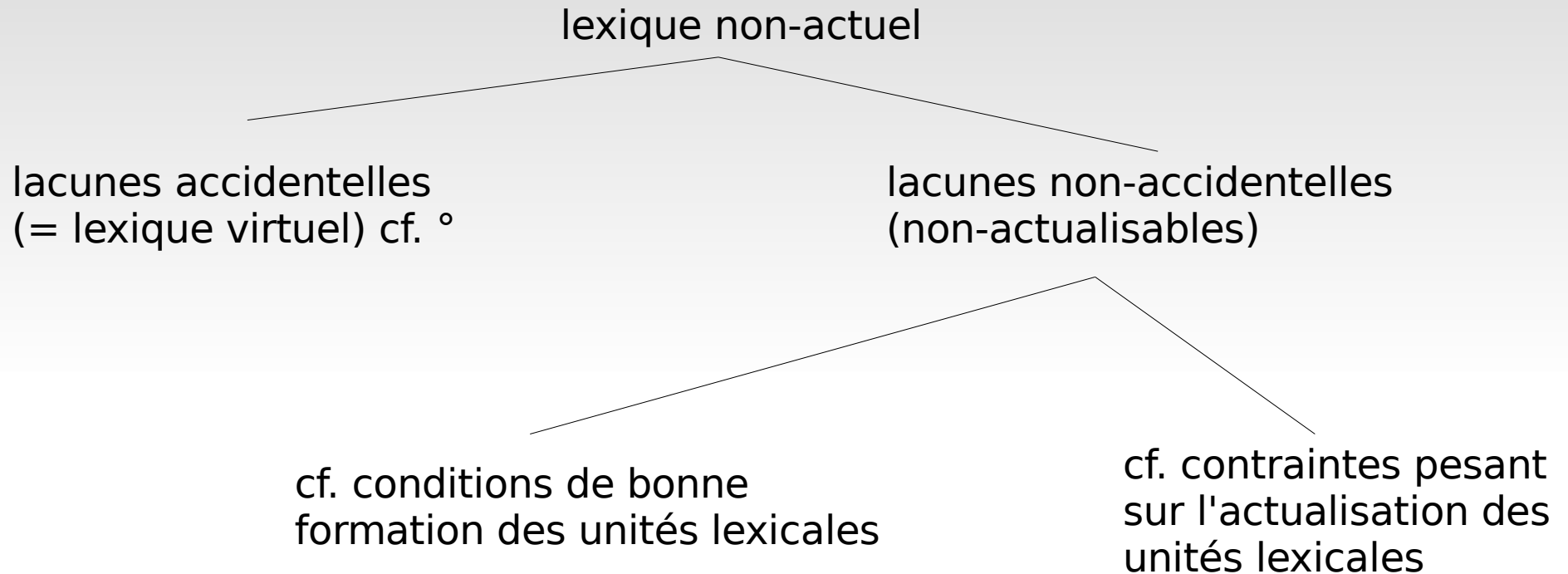
- * antéposé : réalisation impossible
(degré éventuel : ?, ??)
- ° antéposé : réalisation possible mais non actualisée
- [rien] : réalisation possible, actualisée ou non
- ⊗ : **réalisation ± impossible mais actualisée**

Lexique non-actuel

D. Corbin « Entre les mots possibles et les mots existants : les unités lexicales à faible probabilité d'apparition », *Silexicales*, 1997



Lacunes lexicales



Lacunes non accidentelles et conditions de bonne formation

D. Amiot « Prérromantisme vs *prépâleur », *Sillexicales*, 1997

- nom base : doit référer à une réalité temporellement bornée
 - Marc a (de l'humour | *du préhumour)
 - Une prédalle vs *une prétable
- ancrage dans un cadre temporel préconstruit (passé culturel, scénarios) : précambrien vs *préabus, *prépâleur – *préterreur vs pré-Terreur

Prétoile

- *préhumour* : 0, mais *pré-humour* : On peut observer une forme de \otimes *pré humour* vers l'âge d'un an ; *portrait au-pré humour*
- *pré-abus* : Même si l'oncle ne lui a rien fait... ben des propos déplacés c'est peut-être un comportement \otimes *pré-abus*
- *prépâleur* ou *pré-pâleur* : 0

Prétoile

- *Pré-terreur :*
 - *dans une situation de pré-terreur, Lyon se distingue comme un foyer contre la Révolution ; le film porte plus sur la « pré terreur » et de la début de la Terreur*
 - *Jean-Pierre Normand (dans ses jours pré-Terreur Crampante et pré-Jake Blues)*

Lacunes non accidentelles et contraintes d'actualisation

- Raisons pragmatiques : *chauffable, démodable, usable, violable*
- Raisons sémiotiques : *galactophore* pour '*milkman*'
- Raisons symboliques : *bras gauche de Npropre humain*
 - ... *si cette dénomination a peu de chances d'être retenue, c'est parce que la valeur symbolique qu'elle connote dévaloriserait son référent ...*

Démodable, usable...

- *chauffable* : Nounours ⊗ *Chauffable Au Four Micro-Onde* ; rouleau de calandre ⊗ *chauffable*
- *démodable*
 - *l'Octavia n'est pas démodable*
 - *Ordinateur utile mais ⊗ démodable rapidement...*
- **usable* : fréquent en anglais, peu visible en français
- *violable* : dans le TLFi

Galactophore

- nom donné aux canaux excréteurs de la glande mammaire (située dans le sein et sécrétant le lait).
- employé comme adjectif et pas comme substantif

*Le bras gauche de N_{Propre}

- *Collaborateur du Libre Journal de la France Courtoise et “⊗ bras gauche” de Serge de Beketch, Gofman l'assista aussi sur Radio Courtoisie... (www.wikio.fr/news/Patrick+Gofman)*
- *Seul Amato, le bras droit américano-vénézuélien de Strelski, restait de marbre... L'étonnant ⊗ bras gauche de Strelski était un Irlandais obèse.*

John Le Carré Le directeur de nuit, Laffont, 1993

Les imbéciles heureux qui sont nés quelque part

- Je suis d'Orléans
- Je suis (du Loiret | du Perche | du Trégor)
- Je suis (de la Région Centre | du Centre | de Vendée + de la Vendée + de l'Atlantique)
- Je suis (de Belgique + Lybie)
- ?Je suis de la Belgique
- Je suis de la Belgique wallonne

Je suis de la France

- *Moi, ⊗ je suis de la France. Je ne dis pas : je suis la France. ⊗ Je suis de la France. Toutes mes pensées, toutes mes façons d'être, toutes mes sensations, toutes mes vibrations, elles sont de la France.*

François Mitterrand

Interview radio-télévisée (2 mars 1986)

Sociolectes

- Génétique : *La glycoprotéine B (gB) est la glycoprotéine principale de l'enveloppe du cytomégalovirus humain, elle **code pour** le cadre de lecture ouvert UL55 ; en ce qu'elle **code pour** le variant de la protéine FMO2 humaine de séquence SEQ ID N° 3 qui présente un résidu lysine*
- Economie vs langue générale
 - *Le yuan **s'apprécie** avant la visite des Européens en Chine ; Le TSX **s'apprécie** légèrement.*
 - *L'ancienneté **s'apprécie** à la date d'envoi de la lettre de licenciement ; A quelle date **s'apprécie** la composition du patrimoine à déclarer au titre de l'impôt sur la fortune ?*

Confusions attestées

- **as least as* pour *at least as* ('du moins dans la mesure où)
- ... *Honus Whiting was beginning to understand... his wife's opinion, ⊗ as least as it pertained to Empire Falls* [dans un roman]
- [Manning, 2003] trouve 4 ex. dans le New York Times de 1995, 2 dans le NYT de 1996 et de nombreux sur le Web (Etats-Unis, Australie, Afrique du Sud)

Ténu et impossible 1/2

- Examen dans [Manning 2003] des jugements d'acceptabilité de [Pollard & Sag 1994] sur des verbes comme *to regard, to think, to consider...*
 - Echantillon (NYT) de 300 occurrences de *to regard*
 - 2 cadres données comme impossibles y figurent
 - [to regard NP NP]
 - [to regard NP Vpinf] *Conservatives argue that the Bible*
⊗ *regards homosexuality to be a sin*

Ténu et impossible 1/2

- frontière empirique d'acceptabilité : 1%
- « nivellement » des décalages de fréquence
 - $\simeq 75\%$ [to regard NP as NP] *We regard Kim as an acceptable candidate*
 - $\simeq 17\%$ [to regard NP as AdjP] *We regard Kim as quite acceptable*
 - $\simeq 4\%$ [to regard NP as Vping] *We regard Kim as being among the most acceptable candidates*
 - $\simeq 2\%$ [to regard NP as PP] *We regard Kim as among the most acceptable candidate*

Observation directe : apories

- Délimitation des « mots composés » au LADL
- Pierre de touche : degré de figement
 - restrictions plus ou moins nombreuses qu'une séquence connaît par rapport aux transformations dont elle serait théoriquement passible
- G. Gross (88, 96) : examen des N Adj

N°	transformation	séquence libre	séquence à figement
1	adjectif possible comme attribut.	<i>nous avons un climat froid – notre climat est froid</i>	<i>cette arme blanche - #cette arme est blanche</i>
2	nominalisation de l'adjectif	<i>Léa a le teint blanc. La blancheur du teint de Léa</i>	<i>L'examen blanc de Léa *La blancheur de l'examen de Léa</i>
3	rupture paradigmatique sur l'adjectif	<i>du papier bleu, violet, blanc...</i>	<i>une douche écossaise, #française</i>
4	variation libre en nombre		<i>*le denier public</i>
5	adjonction d'un adverbe	<i>une explication très vague</i>	<i>*un terrain très vague</i>
6	coordination d'un adjectif	<i>une explication longue et fastidieuse</i>	<i>*une étoile filante et brillante</i>
7	reprise par le nom seul	<i>j'ai lu un livre passionnant. Ce livre...</i>	<i>*Le bras droit de Luc a appelé. Ce bras...</i>
8	rupture paradigmatique sur le nom		<i>du beurre noir, #de la margarine noire</i>
9	adjectif remplaçable par un complément de nom	<i>une faute grammaticale, de grammaire</i>	<i>la médecine douce, *de douceur</i>

Observation directe : apories

N°	transf.	prédit	ex.	Google04	Google07	Yahoo07
1	adj. attribut	-	<i>*la guerre est froide</i>	5 (a)	694	73
2	adj. nominalisé	-	<i>*la froideur de la guerre</i>	4 (b)	9	10
3	commutation adj.	-	<i>* la guerre chaude</i>	1 (g)	1450	300
4	sing. / plur.	+	<i>les guerres froides</i>	383 (c)	950	86
5	adj. adj.	-	<i>*la guerre très froide</i>	39 (d)	81	14
6	adj. coordonné	-	<i>*guerres froide et chaude</i>	1 (e)	3	1
7	reprise par nom	+	<i>La guerre froide... cette guerre...</i>	?	?	?
8	commutation nom	-	<i>*la paix froide</i>	283 (f,g,h)	946	360
9	adj / de nom	-	<i>*la guerre de froideur</i>	0	0	1

Observation directe : apories

- a) ⊗ Cette guerre est froide car il n'y a pas d'affrontement direct entre les deux Grands...
- b) ⊗ La froideur de la guerre est peut-être en cause. Les acteurs sont bons, mais...
- c) Le pluriel renvoie en particulier à de nombreuses occurrences de deux titres de livres (⊗ *Les guerres froides*, d'Y. Raynaud, aux Editions Théâtrales et ⊗ *Guerres froides*, de Virginie Lou, chez Actes Sud).

Observation directe : apories

- d) Aux postes les plus sensibles de la CIA..., il est l'un des protagonistes de la ⊗ guerre très froide qui, de 1985 à la chute du rideau...
- e) Cette radicalisation de l'offensive contre l'informatique libre dessine une sorte de ⊗ nouvelle guerre froide et sans pitié...

Observation directe : apories

- f) Le pacte de Genève... pose les termes concrets d'une ⊗ paix froide mais réelle...
- g) ⊗ Guerre chaude et paix froide...
- h) [Petite histoire de la Guerre froide]... nouveau durcissement des relations est-ouest, si bien qu'on commence à parler de la « ⊗ paix froide ».

Observation directe : apories

- Google de 2004 à 2007
 - croissance de la Toile
 - évolution ou paramétrage du moteur
- Yahoo : résultats en retraits par rapport à Google, mais pas de manière proportionnelle
 - indexation plus faible
 - indexation différente

Observation directe : apories

- Propriété peu testable : reprise de *guerre froide* dans la même phrase ou le même paragraphe
- Confirmation de la commutation singulier/pluriel et du non passage de l'adjectif au nom correspondant
- Contradiction des autres prédictions, de manière plus ou moins affirmée (*guerre chaude / la froideur de la guerre*)

Modéliser sur données attestées

- Barkema (93, 94) : mesurer la flexibilité d'expressions toutes faites en corpus
- Occurrences de 450 expressions et de leurs variations (graphiques, morphologiques, syntaxiques) dans les 20 millions de mots « nus » du corpus de Birmingham
- *Cold war* : 111 occurrences + 13 variations

Modéliser sur données attestées

- 1. renewed Cold War
- 2. the melting Cold War
- 3. the world Cold War
- 4. continuing, ever-present 'cold' war
- 5. the Cold War won by Europeans who 'destalinizez' Eastern Europe

Modéliser sur données attestées

- 6. the cold war which threatened to divide the world into two ideological armed camps
- 7. a not-so-cold war against Khadafi
- 8. the awkward cold war thought up by the American paranoids, who should be back in the law offices of middlewestern towns
- 9. a period of cold and hot civil war which ended with Hitler's invasion of Austria

Modéliser sur données attestées

- 9. a period of cold and hot civil war which ended with Hitler's invasion of Austria
- 10. a kind of cold civil war
- 11. the cold war that existed between the two giants, the United States and...
- 12. the Cold War in Washington
- 13. the cold war between the Nature Conservancy Council and the farmers

Modéliser sur données attestées

dét. adj. <i>cold war</i>	3 (1, 2, 4)	<i>the melting cold war</i>
dét. <i>cold war</i> prop.	2 (6, 11)	<i>the Cold War that existed...</i>
dét. <i>cold war</i> SP	2 (12, 13)	<i>the cold war between</i>
dét. <i>cold war</i> part. passé	1 (5)	<i>the Cold War won by Europeans...</i>
dét. adj. <i>cold war</i> part. passé	1 (8)	<i>the awkward cold war thought up by...</i>
dét. adv. <i>cold war</i> SP	1 (7)	<i>a not-so-cold war against Khadafi</i>
dét.nom <i>cold war</i>	1 (3)	<i>the world Cold War</i>
dét. <i>cold</i> adj. <i>war</i>	1 (10)	<i>a kind of cold civil war</i>
dét. <i>cold</i> coord. adj. <i>war</i> prop.	1 (11)	<i>a period of cold and hot civil war</i>

Modéliser sur données attestées

- Les variations observées sont-elles plus ou moins importantes que ce à quoi on pouvait s'attendre
- Pierre de touche : corpus arboré de Nimègue (130 000 mots)
- Sur les 16 183 SN, 1 257 relèvent du schéma [{adjectif qualificatif} {nom commun masculin singulier}] et 3 171 de variations de ce schéma

Modéliser sur données attestées

- fréquence attendue de la réalisation A de x =
occurrences de x * (occurrences de la réalisation A
dans le corpus de Nimègue / occurrences de x et de
ses variations dans le corpus de Nimègue)
- Dans le cas présent :
 - $49,15 = ((111 + 13) * (1\ 257 / 3\ 171))$

Modéliser sur données attestées

Réalisation	% attendu	% réel	écart
dét. <i>cold war</i>	39,64	89,52	+49,88
dét. <i>cold war</i> prop.	6,12	1,60	-4,52
dét. <i>cold war</i> SP	15,52	1,60	-13,92
dét. <i>cold wars</i>	19,87	0,00	-19,87
dét <i>cold wars</i> SP	2,30	0,00	-2,30

Modéliser sur données attestées

- Biais de « peser » les variations dans un corpus avec une balance étalonnée sur une autre corpus
 - Nimègue : trop petit pour noter les variations
 - Birmingham : pas arboré et donc ne permet pas d'obtenir la répartition des variations selon leur schéma
- Ecart brut fruste entre fréquence attendue / observée
- Construction d'un modèle et expérimentation

Possible et attesté

	possible	impossible
attesté	abracadabrantésque	⊗ double-cliquer
non attesté	°requin-faucille	*déjoli

J.-C. Milner *Introduction à une science du langage*, Seuil 1989

- *On peut émettre un jugement différentiel concernant les données de langue. ... Autrement dit, le jugement concerne ce qui en matière de langue est possible ou impossible ; il suppose donc qu'il y a un impossible de langue.*
- *... une donnée de langue peut être possible matériellement, c'est-à-dire attestée, et impossible en langue, ou inversement.*
- *... l'intervention grammaticale majeure a lieu quand une donnée matériellement attestée est jugée linguistiquement impossible.*

Le langage : immédiat / instrumenté

*L'une des particularités des sciences du langage, [...] c'est que le langage est sans médiation à disposition du locuteur : je puis produire, à volonté, des phrases, les tronquer, y introduire tel élément que je choisis, etc. Il se pourrait que ce soit le seul exemple d'une manipulation sans instrument, du moins le seul qui se soit maintenu dans un état développé d'une discipline scientifique. S. Auroux *La raison, le langage et les normes* PUF 1998*

*Dans la science d'aujourd'hui, il est rare que l'on « voie » sans l'aide d'instruments I. Hacking *Concevoir et expérimenter* Christian Bourgeois 1989*

Outils linguistiques

... [les grammaires et les dictionnaires] *obéissent à une procédure d'accrétion : ils sont construits les uns à partir des autres, ajoutent des informations et de nouveaux éléments. C'est par là qu'ils atteignent à un moment ou à un autre, des dimensions et des caractéristiques qui font qu'ils ne correspondent à la compétence d'aucun locuteur réel.*

S. Auroux *La raison, le langage et les normes* PUF 1998

Représentativité du jugement linguistique

Les données qu'un linguiste locuteur-auditeur produit par introspection ... ne figurent le savoir linguistique d'aucun locuteur-auditeur particulier identifiable, pas même celui du linguiste qui les produit ; on peut les concevoir comme un fragment de la somme des savoirs linguistiques de l'ensemble des sujets constituant la communauté linguistique, fragment dont le statut sociologique est nécessairement indéterminé.

P. Corbin « De la production des données en linguistique introspective », *Théories linguistiques et traditions grammaticales*, Presses Universitaires de Lille, 1980

Les mondes possibles de l'introspection

... intérêt spécifique de l'introspection : la possibilité d'envisager d'autres énoncés que ceux qui sont attestés. L'introspection peut alors être conçue comme l'instrument privilégié d'une recherche sur les limites ultimes du possible prédictible à partir des observables.

[P. Corbin 1980]

Variation des intuitions

[Possibilité de] *considérer les intuitions contradictoires non pas systématiquement comme des obstacles au bon déroulement des obstacles qu'il convient à toute force de réduire, mais comme les indices d'éventuelles variantes dans les pratiques langagières de la communauté linguistique de référence que le linguiste devrait se contenter d'enregistrer sans chercher à les modifier ni à les hiérarchiser.*

[P. Corbin, 1980]

Décalages possible - attesté

fréquence	localisation	causes possibles
rare	isolement	lapsus variation intentionnelle
↕	↕	↕
commun	dissémination	variation interne changement