

## EXEMPLES DE COMBINATOIRES LEXICALES ÉTABLIS POUR L'ÉCRIT ET L'ORAL À LISBONNE

**Maria Fernanda Bacelar do Nascimento**

**Centro de Linguística da Universidade de Lisboa ( CLUL )**

Le *Corpus de Référence du Portugais Contemporain* du Centre de Linguistique de Lisbonne constitue, aujourd'hui, la base de plusieurs recherches linguistiques: des dissertations universitaires sur différents sujets, la construction de lexiques, des études sur le portugais parlé et des travaux lexicographiques.

En ce qui concerne les travaux lexicographiques, le *corpus* est utilisé comme source d'informations du *Nouveau Dictionnaire de La Langue Portugaise*, en cours de réalisation à l'Académie des Sciences de Lisbonne; il est, aussi, la base du *Dictionnaire de Combinatoires du Portugais* du Centre de Linguistique de Lisbonne. Pour ce qui est du Dictionnaire de l'Académie, le recours au *corpus* se limite au choix des citations à partir des concordanciers et à l'actualisation et validation de sa nomenclature. Pourtant, quoique d'une façon assystématique, l'observation des concordances permet aux rédacteurs d'améliorer les articles car elles leur fournissent, entre autres, des informations sur les constructions plus usuelles, sur les emplois caractéristiques d'un genre de discours donné, sur les emplois nouveaux, ou sur des emplois brésiliens, africains ou régionaux et même des informations encyclopédiques. Cela veut dire que le dépouillement des concordances pour la sélection des citations aide aussi les rédacteurs à rédiger les articles, à construire des exemples plus riches et plus naturels et à fournir des

renseignements pertinents sur les entrées. Mais, faute de temps, ce dépouillement n'a pas été systématique.

Quant au Dictionnaire des Combinatoires, il s'agit d'un projet de recherche sur les associations lexicales du portugais européen. La recherche porte sur un sous-corpus du CRPC de 12 millions de mots de la langue écrite et de près d'1 million de mots de la langue orale, avec la constitution présentée dans le tableau 1.

Tableau 1 - *Corpus DCP*: 12.282.392 mots (*sub-corpus* oral: 856.195, *sub-corpus* écrit: 11.426.197)

<b>pourcentage par type de discours</b>		
<b>oral</b>		
	+ espontâneo	80%
	+ formal	20%
<b>escrito</b>		
	jornalístico	60%
	literário	24%
	técnico, científico, didáctico	10%
	político	4%
	jurídico	2%

Pour ce projet notre objectif est d'observer, mesurer et interpréter les associations lexicales dans leurs environnements.

Nous avons, maintenant, accès aux informations suivantes:

- la liste de toutes les formes lexicales du *corpus* avec les fréquences à l'écrit et à l'oral.
- la liste des lèmes théoriques de ces formes.
- toutes les paires de mots de fréquence égale ou supérieure à 2 ( 2 millions et demi de paires ) et qui ont une distance entre les éléments de 1 à 5 places à gauche et de 1 à 5 places à droite du noeud.
- la liste des lèmes cooccurrents du noeud.
- la fréquence du noeud dans le *corpus* et dans les paires.
- les indices combinatoires des paires (résultat obtenu à partir de l'information mutuelle).
- les listes brutes des concordances des paires avec un contexte restreint ( 1 ligne ) ou avec un contexte élargi. Ces listes sont organisés par lèmes ( du noeud et des cooccurrents ) et contiennent les données de fréquence et les indices combinatoires.
- les associations de 5, 4, 3, et 2 mots contigus avec la fréquence et les contextes.
- les références bibliographiques de chaque contexte.

En partant de ces résultats, nous voulons étudier les mots:

- dans leurs relations avec d'autres mots avec lesquels ils cooccurrent plus au moins systématiquement, de façon continue ou discontinue.
- dans leurs relations avec les traits grammaticaux qui se dégagent de la cooccurrence lexicale.
- dans leurs relations énonciatives et contextuelles.

Ce genre de travail est devenue classique dans les études linguistique sur *corpus* et une vaste bibliographie sur le thème rend compte de nombreuses exploitations de *corpus* dont l'objet est l'étude des associations lexicales.

En ce qui concerne notre projet nous croyons que les concordances des paires ou des groupes de mots peuvent constituer la base de recherches pour l'établissement de méthodologies d'annotation morphosyntaxique, syntaxique et sémantique qui remplaceraient l'annotation mot à mot que nous ne considérons pas satisfaisante. Ils constitueront aussi une excellente base d'informations pour l'établissement de typologies textuelles fondées sur des critères internes puisque nous pourrons établir la relation entre le texte (oral, écrit, scientifique, politique, etc.) et des phénomènes d'association à la fois lexicaux et grammaticaux.

Mais, pour le moment, nous sommes dans une phase préliminaire d'analyse et d'interprétation des donnés en utilisant ce que nous avons appelé de Dictionnaire, plutôt comme un dictionnaire des dictionnaires puisqu'il est plein d'informations qui nous semblent, à plusieurs titres, être d'un très grand intérêt pour les travaux lexicographiques.

L'observation des résultats, soit en ce qui concerne les combinatoires soit en ce qui concerne les listes de formes ou de lèmes, soit en ce qui concerne la quantification des donnés, nous fait remarquer que les dictionnaires portugais ne tiennent pas compte d'un certain nombre de phénomènes que nous considérons indispensables à la caractérisation des mots, si on veut décrire leurs usages réels et pas simplement leurs virtualités. Ces phénomènes sont mis en évidence par les

analyses automatiques sur *corpus*.

En commençant par les listes lématisées, on peut s'apercevoir de quelques particularités concernant la morphologie flexionnelle. Par exemple, la lématisation du *corpus* oral révèle un assez grand nombre de verbes avec seulement 1, 2, 3 ou 4 formes qui, dans la plupart des cas appartiennent au Participe Passé ou à l'Infinitif. Nous avons observé, ensuite, que dans le *corpus* écrit qui est presque 12 fois plus grand que l'oral, quelques-uns de ces verbes se comportent de la même façon.

Le tableau 2 nous montre 3 exemples:

Le verbe *arreigar-se* (s'accrocher), *graduar* (grader) et *sofisticar* (sophistiquer).

Ces cas, parmi beaucoup d'autres, confirmeraient l'hypothèse de Claire Blanche-Benveniste, selon laquelle il y a, dans l'usage réel, une défektivité verbale qui atteint beaucoup plus de verbes que ceux traditionnellement présentés dans les grammaires et dans les dictionnaires.

En plus, nous avons vérifié que pour les emplois du Participe Passé, il s'agissait souvent des emplois adjectivaux et que le verbe, en tant que tel, n'était jamais utilisé. Ces informations seraient très pertinentes en lexicographie et devraient être fournies de façon explicite et aussi implicite, c'est à dire, à travers les exemples fabriqués par les lexicographes que nous trouvons parfois si artificiels.

Tableau 2 – Occurrence des formes verbales à l'oral et à l'écrit

VERBO	Frequência		ORAL		ESCRITO	
	oral	escrito	Tempo Verbal	Forma	Tempo Verbal	Forma
<b>ARREIGAR-SE</b>	7	28	Particípio Passado	arreigado 3	Particípio Passado	arreigada 11
				arreigados 2		arreigado 6
				arreigado 2		arreigados 6
						arreigadas 2
					Infinitivo	arreigar-se 1
					Pret. Imperfeito	arreigava-se 1
<b>GRADUAR</b>	3	75	Particípio Passado	graduado 2	Particípio Passado	graduado 28
				graduados 1		graduados 21
						graduada 18
					Infinitivo	graduar 3
					Pret. Perfeito	graduou 2
<b>SOFISTICAR</b>	3	203	Particípio Passado	sofisticada 1	Particípio Passado	sofisticados 60
				sofisticado 1		sofisticado 56
				sofisticados 1		sofisticada 51
						sofisticadas 35
					Infinitivo	sofisticar 1

Les faits d'association lexicale sont aussi très négligés par les dictionnaires portugais et, pourtant, ils contribuent fortement à la connaissance du fonctionnement et de la signification des mots. Les dictionnaires présentent surtout les phraséologies, c'est à dire, les expressions figées où chacun de leurs éléments a perdu son sens propre, mais ils présentent très peu de fois, ou très assystématiquement, des associations fréquentes ( cooccurrences préférentielles) ou des idiosyncrasies combinatoires. Là, l'intuition du lexicographe n'est pas du tout suffisante, soit pour les choisir, soit pour établir leur degré de figement ou leur fréquence dans les divers usages. Le recours au corpus est, dans ce cas, absolument indispensable.

En comparant les premières 25 paires contigües du corpus oral et du corpus écrit ( tableau 3 ) nous remarquons déjà des différences assez nettes.

Les premières paires de l'écrit sont constituées presque uniquement par des mots grammaticaux, tandis que dans les premières paires de l'oral on voit des syntagmes qui reflètent des phénomènes caractéristiques de l'oral, comme l'utilisation des marqueurs conversationnels não é ( n'est ce pas ) et quer dizer ( c'est à dire ) ou l'emploi pronominal de a gente ( on ) que la norme puriste refuse, surtout à l'écrit. La paire não não nous a vraiment intéressé parce que les répétitions ou les séquences de répétitions d'un mot (verbe, adverbe, pronom) constituent, en portugais, une forme de réponse très courante dans les dialogues, pour exprimer l'assentiment, la confirmation, la négation. Nous avons donc cherché les mots

Tableau 3 - Les 25 paires contigües plus fréquentes

ORAL		ESCRITO	
6585	não é	22394	que o
3786	quer dizer	19750	que se
3537	é que	19562	que a
3336	que é	17952	o que
1966	o que	16985	com a
1830	que eu	16205	de um
1732	que não	16192	e o
1699	não sei	15688	para o
1648	que se	15661	com o
1636	é uma	15550	para a
1510	é o	13754	a sua
1510	a gente	13737	de uma
1350	é um	13001	e a
1335	e depois	12872	que não
1289	por exemplo	11127	o seu
1186	que a	10490	em que
1185	para a	9853	e de
1164	para o	8934	que os

1147	é a	8626	é que
1085	acho que	8371	de que
1050	e a	8078	do que
1024	que o	7997	que é
1003	eu não	7679	e que
982	não não	7604	e os
964	uma coisa	7507	não se

répétés 2, 3 4 ou 5 fois, dont les plus fréquents sont présentés dans le tableau 4 avec leur fréquence.

Il faut dire que au moment de l'informatisation des transcriptions du corpus de langue parlée du Portugais Fondamental, nous n'avons gardé que les répétitions que le transcripneur considérait

Tableau 4 - Formes répétées 2, 3, 4 et 5 fois

<b>Palavra repetida</b>	<b>X 2</b>	<b>X 3</b>	<b>X 4</b>	<b>X 5</b>
não	982	58	26	15
é	699	36	10	
pois	241	13		
sim	195	12	8	11
há	102		2	
muito	102	16	4	
tudo	81	12	2	5
nada	62	10		
já	60	7		
tem	49	3		
gosto	47	2		
sempre	42	9	3	
tenho	38	7	2	
foi	37			
era	36	3		
são	33			
isso	32			
nunca	26			
também	24	2		
temos	21			
exactamente	20	4	2	
aqui	18			
está	16			
têm	14			
talvez	13			
vêm	13	5		

dá	13			
todos	12	5		
muita	12			
muitos	12			
bastante	11			
tou	10			
fazem	10			
vai	10			
imenso	9			
agora	9			
vem	8			
diga	8			
acontece	8			
tive	7			
vou	7			
vão	7			
claro	7			
havia	7			
ia	7			
absolutamente	7			
pode	6			
tava	6			
gostei	6			

intentionnelles et pas les répétitions par hésitation. Ces textes constituent la presque totalité du corpus oral dépouillé pour l'étude de ces associations. C'est pour ça que nous avons posé l'hypothèse, qui est encore à vérifier, que ces formes 2, 3, 4 et 5 fois répétées appartenaient à des prises de parole qui constituent des réponses.

Les concordances des paires nous laissent voir les liens entre les associations de mots, les formes de ces mots et le genre de discours. C'est le cas de l'exemple du tableau 5 où nous voyons que l'adverbe adiante ( après ) avec un sens temporel s'associe fréquemment, dans le discours didactique, à des verbes comme estudar ( étudier ), referir ( mentionner ) et ver ( voir ) employés toujours au Futur. Selon les indices combinatoires relevés, ces associations avec le mot adiante ont été significatives dans le corpus écrit.

La corrélation entre les associations lexicales et les différentes acceptions d'un mot est très visible dans le tableau 6.

Tableau 5 – Associations de adiante avec des formes verbales au Futur

<p>Exemplo: ADIANTE <b>Frequência do Lema no Corpus</b> Escrito: 496  <b>Frequência da combinatória:</b>7 <b>Lema:</b> ESTUDAR  estudará adiante -- <b>Índice de Combinatória:</b> 9.352  <b>Freq de Combin</b> 3 estudará adiante <b>Dist entre os elem</b> 1  <b>general Gomes da Costa, em 28 de Maio de 1926, como</b>  <b>»@i. Em 1929, porém, estalara uma grave crise (que</b>  <b>apoio económico-financeiro; - a Revolução Russa, que</b>  estudaremos adiante -- <b>Índice de Combinatória:</b> 8.659  <b>Freq de Combin</b> 2 estudaremos adiante <b>Dist entre os elem</b> 2  <b>como sendo a idade, o sexo, a profissão, clima, etc., como</b>  <b>objectos em @brepresentação cavaleira@b, fig. 9, que</b>  estudar adiante -- <b>Índice de Combinatória:</b> 4.283  <b>Freq de Combin</b> 2 estudar adiante <b>Dist entre os elem</b> 2  <b>6 b). Nos circuitos constituídos por condensadores (a</b>  <b>eléctrica produz diversos efeitos, os quais iremos</b>  <b>Frequência da combinatória:</b> 4 <b>Lema:</b> REFERIR  referiremos adiante --<b>Índice de Combinatória:</b> 8.659  <b>Freq de Combin</b> 4 referiremos adiante <b>Dist entre os elem</b> 2  <b>solamento eléctrico dos condutores deve ser diferente.</b>  <b>la constituição de classes com intervalos diferentes nos</b>  <b>lo método dos mínimos quadrados), a que nos</b>  <b>o problema do "envelhecimento da base", a que nos</b>  <b>Frequência da combinatória:</b> 21 <b>Lema:</b> VER  veremos adiante --<b>Índice de Combinatória:</b> 7.840  <b>Freq.de Combin</b> 10 veremos adiante <b>Dist entre os elem</b> 1  <b>individualizado como sub-região mesmo quando, como</b>  <b>afirmar que sejam ligados em série ou em paralelo.</b>  <b>anos, outros países se envolverão no mesmo, como</b>  <b>Matemática, Filosofia, Teologia e Medicina [como</b></p>	<p>estudará adiante. estudará adiante). estudará adiante,</p> <p>estudaremosmais adiante. estudaremos mais adiante</p> <p>estudar mais adiante), estudar mais adiante;</p> <p>Referiremos mais adiante referiremos mais adiante. referiremos mais adiante,</p> <p>veremos adiante, Veremos, adiante, veremos adiante. veremos adiante,</p>	<p>@tA educação@t O @bcombate @t7. 1.5 @bA Revolução Socialista enfraqueceu a frente oriental, acabam</p> <p>O Homem tem hábitos alimentares na página 64. Esta representação</p> <p>a tensão está em atraso relativamente de entre eles destacamos desde já os</p> <p>quais os tipos de canalizações eléct @t@bA REPRESENTAÇÃO DAS 1.º - @bMétodo gráfico@b (ou de poderíamos igualmente ter operado</p> <p>a análise das suas formas de povo a razão desta afirmação, através da Assim começava a 1ª @b@sGuerra com excepção da Teologia, ter-se-ia</p>
---	--	---

Tableau 6 – Associations lexicales du mot pressão

<p><b>ESCRITO</b>  <b>FT 1413 PRESSÃO</b>  <b>FC 6276 PRESSÃO</b>  <b>9 SUBTROPICAL (real:9)</b>  <b>pressões subtropicais 8.731</b>  <b>9 pressões subtropicais 1</b></p>		
<p>201065274 mais próximo das áreas de altas 201065253 am. O ar polar alimenta as altas 201065218 westerlies que provêm das altas 201065232 ssões no Mediterrâneo. Das altas 201065260 entos dos jet streams e às altas 201065239 s e, à direita, origina as altas 201065267 devido às influências das altas 201065225 nvolvidos por centros de @baltas</p> <p><b>44 ATMOSFÉRICO (real:46)</b>  <b>atmosférica pressão 8.643</b>  <b>40 pressão atmosférica 1</b></p> <p>201038714 resão exercida pelo ar chama-se 201038896 ormenor em capítulos anteriores. 201038756 s sob a acção da pressão do ar - 201038763 em que é corrente exprimirmos a 201038784 es, melhor equipadas, analisam a 201038791 a temperatura, a precipitação, a 201038805 diminuindo. b) O @bdecrécimo da 201038798 itude, da composição química, da 201038686 li não havia qualquer mudança de 201038875 emperatura sobe ligeiramente e a 201038931 e admissão é P-I- (normalmente a 201038952 do cilindro cai para o valor da 201038833 altitude está relacionada com a 201038742 erna dos pulmões fica inferior à 201038924 ias de temperatura e humidade, a 201038812 eso. Ou, mais simplesmente, se a</p>	<p>pressões subtropicais, pressões subtropicais pressões subtropicais. pressões subtropicais, pressões subtropicais@b. pressões subtropicais pressões subtropicais pressões subtropicais@b,</p> <p>@bpressão atmosférica. @bPressão atmosférica@b pressão atmosférica pressão atmosférica; pressão atmosférica, pressão atmosférica, pressão atmosférica pressão atmosférica, pressão atmosférica. pressão atmosférica. pressão atmosférica pressão atmosférica pressão atmosférica pressão atmosférica, pressão atmosférica pressão atmosférica pressão atmosférica</p>	<p>funcionan juntamente Na região nomeadame Os fenó pela subsi que migram sendo o</p> <p>- A press - diminu - porque est 1 atm= 760 a estrutura a nebulosid com a altitu da fuga mol De resto, a diminui. d) e a pressão e a força da e com a temp e o ar entr é de 1, 03kg é o @bpeso q</p>

201038882 porque se verifica um aumento da  
 201038721 ma-se @bpressão atmosférica. - A  
 201038945 ar comprimido introduzido até à  
 201038826 s variações da temperatura com a  
 201038679 nações: barómetros, para medir a  
 201038840 de de superfície dá-se o nome de  
 201038917 lugar terrestre que determina a  
 201038777 em-se à temperatura ambiente e à  
 201038770 neus de trás, de 27 lb/pol+2+. A  
 201038735 A compressão do ar é superior à  
 201038910 grandes altitudes por redução da  
 201038938 xpansão para as condições livres  
 201038903 de vista climatológico, a baixa  
 201038728 m influência sobre a variação da  
 201038693 como provar a sua existência?@b  
 201038700 ressonância exercida pelo ar chama-se  
 201038819 sobre nós; daí, a diminuição da  
 201038889 to maior quanto mais baixa for a  
 201038707 s a esta permitem concluir que a  
 201038868 ecebeu ao longo do dia; assim, a  
 201038749 sobre todos os corpos originam a  
 201038854 qual modo, podemos afirmar que a  
 201038861 enor a pressão. Concluímos que a  
 201038847 do; assim, podemos afirmar que a

**81 EXERCER (real:90)  
 pressões exercidas 8.521**

**4 pressões exercidas1**

201061860 Mundo de 1998, em França. Daí as  
 201061874 atribuiu a atitude tailandesa a  
 201061867 de opinião, mas sim as ameaças e  
 201061853 facto, como aponta, "das enormes

**25 pressão exercida 1**

201047234 , do que um casamento falhado. A  
 201047115 ão. Tento rodar o carro, mas a  
 201047108 cado abandonado. Recentemente,

201047255 aranda dos Windsor, não? DJOB- A  
 201047101 encontro. E nem quando baixou a  
 201047129 ocupar o lugar que ele deixou. À  
 201047136 xerce pressão sobre os corpos. À  
 201047143 rce-se em todos os sentidos@b. A  
 201047199 essão. Podemos assim dizer que a  
 201047206 essão exercida pelo ar é igual à  
 201047185 a Polónia, tem sido importante a  
 201047150 reia, é maior na situação I. @iA  
 201047094 em mais 70 milhões de dólares. A  
 201047213 eco (sem humidade) adicionada da  
 201047087 o entanto, alguns problemas pela  
 201047164 s frequentemente tensão.|| @t@bA  
 201047248 randa dos Windsor, não? DJOB - A  
 201047192 a-se cada vez mais em virtude da  
 201047227 Franco considerou inaceitável a  
 201047178 elo capital e, por outro lado, a  
 201047122 encontra a massa magmática e da  
 201047171 líquido resulta na diminuição da  
 201047220 vamente, um termo próprio para a  
 201047157 o devida à acção de uma força de  
 201047241 a da filha. É difícil imaginar a

**8 pressão exercidas1**

201047481 re as intensidades das forças de  
 201047488 e as intensidades das forças de  
 201047446 evidentemente fechada, as forças de  
 201047460 ncia mostra-nos que as forças de  
 201047439 e@i. Anteriormente, as forças de  
 201047453 inantemente sujeita às forças de  
 201047467 explicar o facto de as forças de  
 201047474 estão representadas as forças de

**exerceram pressão 7.287**

**2 exerceram pressão 2**

116064858 vitória dos Estados Unidos, que  
 116064851 perações a meio-campo, zona onde  
**exercendo pressão 7.142**

**4 exercendo pressão 1**

116057229 actuarem em todas as direcções,  
 116057222 sangue entra nessas cavidades e,

pressão atmosférica.  
 pressão atmosférica  
 pressão atmosférica  
 pressão atmosférica.  
 pressão atmosférica;  
 @bpressão atmosférica@b;  
 pressão atmosférica  
 pressão atmosférica  
 pressão atmosférica  
 pressão atmosférica.  
 pressão atmosférica.  
 (pressão atmosférica),  
 pressão atmosférica  
 pressão atmosférica.  
 @bPressão atmosférica@b  
 @bpressão atmosférica@b.  
 pressão atmosférica.  
 pressão atmosférica.  
 @bpressão atmosférica  
 pressão atmosférica  
 pressão atmosférica.  
 @bpressão atmosférica  
 pressão atmosférica  
 pressão atmosférica

Está-se, po  
 exerce-se em  
 exterior. Um  
 Fácil é con  
 manómetros,  
 mede-se  
 nesse mesmo  
 normal.|| @t  
 normal tem o  
 O ar compri  
 O diagnósti  
 o que rep  
 ocasiona uma  
 Os aparelho  
 Os resul  
 Pois cl  
 Por curiosi  
 Se a evapora  
 se exerce  
 será mais al  
 @t@bConclus  
 varia na r  
 varia na raz  
 varia segund

pressões exercidas  
 pressões exercidas  
 pressões exercidas  
 pressões exercidas

pela embaixad  
 pela embaixad  
 publicamente  
 sobre a Ingla

pressão exercida  
 pressão exercida  
 pressão exercida

contra esta mul  
 lá em baixo tir  
 pela edibilidade

pressão exercida  
 pressão exercida  
 pressão exercida  
 pressão exercida  
 pressão exercida  
 pressão exercida  
 pressão exercida  
 pressão exercida  
 pressão exercida  
 pressão exercida  
 pressão exercida  
 pressão exercida  
 pressão exercida  
 pressão exercida  
 pressão exercida  
 pressão exercida  
 pressão exercida  
 pressão exercida  
 pressão exercida  
 pressão exercida

pela mídia, tin  
 pela selecção p  
 pelo ar chama-s  
 pelo ar chama-s  
 pelo ar sobre o  
 pelo ar é igual  
 pelo ar seco (s  
 pelo movimento  
 pelo peso do ti  
 pelo poder para  
 pelo vapor de á  
 pelos adversári  
 pelos líquidos  
 pelos "media" t  
 pelos @imovimen  
 pelos sindicali  
 pelos trabalhad  
 por essa massa  
 por esse gás na  
 por grupos que  
 sobre uma dada  
 sobre uma mãe,

pressão exercidas  
 pressão exercidas  
 pressão exercidas  
 pressão exercidas  
 pressão exercidas  
 pressão exercidas  
 pressão exercidas  
 pressão exercidas

na face superi  
 na face inferi  
 na sua superfi  
 pela água sob  
 pelo ar e vapo  
 pelo ar atmosf  
 por um fluido  
 sobre o corpo

exerceram forte pressão  
 exerceram grande pressão.

para que  
 Quando

exercendo pressão  
 exercendo pressão

nas superfície  
 nas válvulas a







36405039 das. Se tivéssemos sido fracos e  
**cedendo pressão 6.737**  
**3 cedendo pressão 2**  
 36397120 privada dos sistemas municipais,  
 36397127 caminho-de-ferro, nalguns casos,  
 36397134 caminho-de-ferro, nalguns casos

**cedido às pressões** socialistas,  
**cedendo à pressão** das multinacio  
**cedendo à pressão** dos operadores  
**cedendo à pressão** dos operadores

Ainsi, comme nous le voyons, le mot pressão (pression), avec le sens de “ force exercé par l’atmosphère “ a les cooccurrents subtropical (subtropicale), atmosférica (atmosphérique), altas ( hautes ) et baixas (basses). Avec le sens de “pression du sang “ il a le cooccurrent arterial (artérielle), et avec le sens de "influence" il a les cooccurrents exercer (exercer) et ceder (céder).

Soit dans les usages courants, soit dans les langues de spécialité, ces types d’associations se présentent trop systématiquement pour qu’on puisse penser qu’elles sont accidentelles et, pourtant, on ne réussit pas à les atteindre par intuition. C’est donc dans le corpus qu’on relève soit des associations libres mais très usuelles qui montrent des tendances fortes dans plusieurs usages de la langue, soit des associations vraiment idiosyncratiques.

Ces phénomènes ont aussi des incidences sur la synonymie. Si nous prenons comme exemple les mots célebre (célèbre), famoso (fameux) et notável (remarquable) que les dictionnaires portugais présentent comme des synonymes, nous pouvons constater qu’ils ont des cooccurrents différents (tableau 7) et que les différences se situent surtout au niveau des noms et des adverbés terminés en -mente (tableau 8).

Tableau 7 – Cooccurrents des mots synonymes célebre, famoso, notável

FT 454 CÉLEBRE	FT 686 FAMOSO	FT 433 NOTÁVEL
10 TRISTEMENTE (real:10)	14 TORNAR (real:14)	6 CONJUNTO (real:6)
4 CRIMINOSO (real:4)	11 NOME (real:11)	6 QUALIDADE (real:6)

14 FICAR (real:14)	4 COLECÇÃO (real:4)	6 VERDADEIRAMENTE (real:6)
7 FRASE (real:7)	9 FICAR (real:9)	254 DE (real:328)
12 TORNAR (real:12)	4 AMERICANO (real:4)	8 ESFORÇO (real:8)
6 AUTOR (real:8)	5 GENTE (real:5)	7 OBRA (real:7)
10 MUITO (real:10)	433 DE (real:559)	6 ÉPOCA (real:6)
325 DE (real:422)	55 POR (real:62)	68 SER (real:76)
37 MAIS (real:37)	67 SER (real:77)	5 EXEMPLO (real:5)
45 SER (real:47)	77 MAIS (real:77)	48 MAIS (real:48)
10 TÃO (real:10)	4 MENOS (real:4)	7 FAZER (real:9)
11 SUA (real:20)	4 GRUPO (real:4)	16 MUITO (real:16)
73 EM (real:96)	8 TÃO (real:8)	7 TRABALHO (real:7)
6 SEUS (real:6)	22 SEU (real:22)	8 TER (real:13)
31 POR (real:36)	18 SUA (real:18)	29 COM (real:38)
4 DIA (real:4)	4 CASA (real:4)	77 E (real:112)
6 JÁ (real:6)	253 A (real:331)	56 EM (real:74)
42 QUE (real:67)	23 JÁ (real:23)	25 PARA (real:27)
46 E (real:77)	64 EM (real:83)	
14 COM (real:18)	9 MUITO (real:9)	
	16 COMO (real:16)	
	119 E (real:146)	
	30 COM (real:36)	
	36 NÃO (real:38)	

Les données de fréquence nous fournissent aussi des indications utiles à propos de la synonymie. Les exemples présentés dans les tableaux 9 et 10 ont été recueillis dans un corpus oral et dans un corpus écrit qui ont des dimensions semblables.

Nous avons pris dans un Dictionnaire de synonymes ceux qui étaient présentés pour les adverbes absolutamente ( absolument ), tableau 9, et completamente ( complètement ), tableau 10, et nous

Tableau 8 – Noms, adverbes et verbes cooccurrents de célebre, famoso, notável

FT 454 CÉLEBRE	FT 686 FAMOSO	FT 433 NOTÁVEL
CO-OCORRÊNCIA:	NOMES	

4 CRIMINOSO (real:4)	11 NOME (real:11)	6 CONJUNTO (real:6)
7 FRASE (real:7)	4 COLEÇÃO (real:4)	6 QUALIDADE (real:6)
6 AUTOR (real:8)	4 AMERICANO (real:4)	8 ESFORÇO (real:8)
4 DIA (real:4)	5 GENTE (real:5)	7 OBRA (real:7)
	4 GRUPO (real:4)	6 ÉPOCA (real:6)
	4 CASA (real:4)	5 EXEMPLO (real:5)
		7 TRABALHO (real:7)

<b>FT 454 CÉLEBRE</b>	<b>FT 686 FAMOSO</b>	<b>FT 433 NOTÁVEL</b>
CO-OCORRÊNCIA:	ADVÉRBIOS	
10 TRISTEMENTE (real:10)	77 MAIS (real:77)	6 VERDADEIRAMENTE (real:6)
10 MUITO (real:10)	4 MENOS (real:4)	48 MAIS(real:48)
37 MAIS (real:37)	8 TÃO (real:8)	16 MUITO (real:16)
10 TÃO (real:10)	23 JÁ (real:23)	
6 JÁ (real:6)	9 MUITO (real:9)	
	36 NÃO (real:36)	

<b>FT 454 CÉLEBRE</b>	<b>FT 686 FAMOSO</b>	<b>FT 433 NOTÁVEL</b>
CO-OCORRÊNCIA:	VERBOS	
14 FICAR (real:14)	14 TORNAR (real:14)	68 SER (real:76)
12 TORNAR (real:12)	9 FICAR (real:9)	7 FAZER (real:9)
45 SER (real:45)	67 SER (real:67)	8 TER (real:13)

avons observé leurs fréquences totales dans le corpus orale et dans le corpus écrit, ainsi que leurs fréquences en position pré-adjectivale en considérant leur fonction d' intensificateur.

Tableau 9 – Fréquences et position pré-adjectivale des synonymes du mot absolutamente

<b>ABSOLUTAMENTE</b>				
<b>Corpus oral</b>		<b>Sinónimos no dicionário</b>		
Frequência	212	decididamente	Freq. no corpus oral	0
ocorr. c/Adjectivo	45		Freq. no corpus escrito	2
<b>Corpus escrito</b>			ocorr. c/Adjectivo	0
Frequência	16	despoticamente	Freq. no corpus oral	0
ocorr. c/Adjectivo	13		Freq. no corpus escrito	0
		determinantemente	Freq. no corpus oral	0
			Freq. no corpus escrito	0
		geralmente	Freq. no corpus oral	212
			ocorr. c/Adjectivo	0
			Freq. no corpus escrito	26
			ocorr. c/Adjectivo	6
		ilimitadamente	Freq. no corpus oral	0
			Freq. no corpus escrito	0
		imperiosamente	Freq. no corpus oral	0
			Freq. no corpus escrito	1
			ocorr. c/Adjectivo	0
		independentemente	Freq. no corpus oral	18
			ocorr. c/Adjectivo	0
			Freq. no corpus escrito	22
			ocorr. c/Adjectivo	0
		indispensavelmente	Freq. no corpus oral	0
			Freq. no corpus escrito	0
		inteiramente	Freq. no corpus oral	11
			ocorr. c/Adjectivo	3
			Freq. no corpus escrito	25
			ocorr. c/Adjectivo	13
		necessariamente	Freq. no corpus oral	14
			ocorr. c/Adjectivo	1
			Freq. no corpus escrito	11
ocorr. c/Adjectivo	1			
peremptoriamente	Freq. no corpus oral	0		
	Freq. no corpus escrito	2		
	ocorr. c/Adjectivo	0		
precisamente	Freq. no corpus oral	295		
	ocorr. c/Adjectivo	3		
	Freq. no corpus escrito	46		
	ocorr. c/Adjectivo	0		
soberanamente	Freq. no corpus oral	0		
	Freq. no corpus escrito	0		

Tableau 10

<b>COMPLETAMENTE</b>
----------------------

<b>Corpus oral</b>		<b>Sinónimos no dicionário</b>		
Frequência	165	cabalmente	Freq. no corpus oral	0
ocorr. c/Adjectivo	102		Freq. no corpus escrito	1
<b>Corpus escrito</b>			ocorr. c/Adjectivo	0
Frequência	52	compridamente	Freq. no corpus oral	0
ocorr. c/Adjectivo	26		Freq. no corpus escrito	0
		exactamente	Freq. no corpus oral	510
			ocorr. c/Adjectivo	8
			Freq. no corpus escrito	49
			ocorr. c/Adjectivo	2
		integralmente	Freq. no corpus oral	4
			ocorr. c/Adjectivo	6
			Freq. no corpus escrito	5
			ocorr. c/Adjectivo	1
		imperiosamente	Freq. no corpus oral	0
			Freq. no corpus escrito	1
			ocorr. c/Adjectivo	0
		inteiramente	Freq. no corpus oral	11
			ocorr. c/Adjectivo	3
			Freq. no corpus escrito	25
			ocorr. c/Adjectivo	13
		juntamente	Freq. no corpus oral	15
			ocorr. c/Adjectivo	0
			Freq. no corpus escrito	30
			ocorr. c/Adjectivo	0
		perfeitamente	Freq. no corpus oral	150
			ocorr. c/Adjectivo	46
			Freq. no corpus escrito	51
			ocorr. c/Adjectivo	23
		plenamente	Freq. no corpus oral	9
			ocorr. c/Adjectivo	5
			Freq. no corpus escrito	7
			ocorr. c/Adjectivo	4
		radicalmente	Freq. no corpus oral	3
			ocorr. c/Adjectivo	1
		rasamente	Freq. no corpus oral	0
			Freq. no corpus escrito	0
		redondamente	Freq. no corpus oral	0
			Freq. no corpus escrito	1
			ocorr. c/Adjectivo	0
		totalmente	Freq. no corpus oral	61
			ocorr. c/Adjectivo	46 <sup>1</sup>
			Freq. no corpus escrito	52
			ocorr. c/Adjectivo	24

<sup>1</sup> 27 = diferente(s)

Si les différences entre les valeurs de la Fréquence à l'écrit et à l'oral de quelques-uns de ces adverbes sont déjà assez frappantes, on peut remarquer aussi les énormes différences qui s'observent entre les mots présentés comme synonymes, soit du seul point de vue quantitatif, soit du point de vue de leur fonctionnement auprès des adjectifs.

Ces quelques exemples montrent que les phénomènes linguistiques observables, soit du point de vue quantitatif, soit du point de vue qualitatif, rendent beaucoup plus ouvertes les frontières établies entre les phénomènes lexicaux et les phénomènes grammaticaux et qu'ils doivent, tous les deux, être intégrés à la description lexicographique.

### **Bibliographie**

- BACELAR DO NASCIMENTO, M. F. (1997), "A exploração de corpora linguísticos no ensino/aprendizagem do português ", Actas do Seminário Internacional de Português como Língua Estrangeira, Macau, pp.373-407.
- BACELAR DO NASCIMENTO, M. F. (1997), "Contribuição de análise de corpora para a descrição lexicográfica", Sentido que a vida faz, Estudos para Óscar Lopes, Porto, Ed. Campo das Letras, pp. 734-744.
- BACELAR DO NASCIMENTO, M. F. e L. A. S. PEREIRA (1996), "Dicionário de Combinatórias do Português: associações lexicais frequentes observadas num corpus de português contemporâneo", Actas do XI Encontro Nacional da Associação Portuguesa de Linguística, vol II – Dicionários, FARIA, J. H. e M. CORREIA (orgs.), APL, Lisboa, pp 43-54.
- BIBER, D. (1993) "Co-occurrence patterns among collocations. A tool for corpus-based lexical knowledge acquisition", Computational Linguistics, 19-3 –

Cambridge, M. A., ACL, pp. 531-538.

- CHURCH, K. W. e P. HANKS (1990), "Word association norms, mutual information and lexicography", Computational Linguistics, 16-1, Cambridge, M.A., ACL, pp. 22-29.
- SINCLAIR, J. M. (1991), Corpus, Concordance, Collocation, Oxford, Oxford University Press.
- SINCLAIR, J. M. (ed.) (1995) Cobuild English Collocations on CD-ROM, London, Harper Collins Publishers.
- SMADJA, F. (1993) "Retrieving Collocation from Text: Xtract", Computational Linguistics, 19-1, Cambridge, M.A., ACL, pp. 143-177.